

SEVEN

PUBLICAÇÕES ACADÊMICAS
2023



INTELIGÊNCIA ARTIFICIAL

HISTÓRICO, CONCEITOS E GESTÃO

Getulio K. Akabane

SEVEN

PUBLICAÇÕES ACADÊMICAS
2023



INTELIGÊNCIA ARTIFICIAL

HISTÓRICO, CONCEITOS E GESTÃO

Getulio K. Akabane

EDITORA CHEFE

Prof^o Me. Isabele de Souza Carvalho

EDITOR EXECUTIVO

Nathan Albano Valente

AUTOR DO LIVRO

Getulio K. Akabane

PRODUÇÃO EDITORIAL

Seven Publicações Ltda

EDIÇÃO DE ARTE

Alan Ferreira de Moraes

EDIÇÃO DE TEXTO

Natan Bones Petitemberte

BIBLIOTECÁRIA

Bruna Heller

IMAGENS DE CAPA

AdobeStok

2024 by Seven Editora

Copyright © Seven Editora

Copyright do Texto © 2024 Os Autores

Copyright da Edição © 2024 Seven Editora

O conteúdo do texto e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva do autor, inclusive não representam necessariamente a posição oficial da Seven Publicações Ltda. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos ao autor, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Todos os manuscritos foram previamente submetidos à avaliação cega pelos pares, membros do Conselho Editorial desta Editora, tendo sido aprovados para a publicação com base em critérios de neutralidade e imparcialidade acadêmica.

A Seven Publicações Ltda é comprometida em garantir a integridade editorial em todas as etapas do processo de publicação, evitando plágio, dados ou resultados fraudulentos e impedindo que interesses financeiros comprometam os padrões éticos da publicação.

Situações suspeitas de má conduta científica serão investigadas sob o mais alto padrão de rigor acadêmico e ético.



O conteúdo deste Livro foi enviado pelo autor para publicação de acesso aberto, sob os termos e condições da Licença de Atribuição Creative Commons 4.0 Internacional

CORPO EDITORIAL

EDITORA-CHEFE

Profº Me. Isabele de Souza Carvalho

CORPO EDITORIAL

Pedro Henrique Ferreira Marçal - Vale do Rio Doce University
Adriana Barni Truccolo - Universidade Estadual do Rio Grande do Sul
Marcos Garcia Costa Morais - Universidade Estadual da Paraíba
Mônica Maria de Almeida Brainer - Instituto Federal de Goiás Campus Ceres
Caio Vinicius Efigenio Formiga - Pontifícia Universidade Católica de Goiás
Egas José Armando - Universidade Eduardo Mondlane de Moçambique
Ariane Fernandes da Conceição - Universidade Federal do Triângulo Mineiro
Wanderson Santos de Farias - Universidade de Desenvolvimento Sustentável
Maria Gorete Valus - Universidade de Campinas
Luiz Gonzaga Lapa Junior - Universidade de Brasília
Janyel Trevisol - Universidade Federal de Santa Maria
Irlane Maia de Oliveira - Universidade Federal de Mato Grosso
Paulo Roberto Duailibe Monteiro - Universidade Federal Fluminense
Luiz Gonzaga Lapa Junior - Universidade de Brasília
Yuni Saputri M.A - Universidade de Nalanda, Índia
Arnaldo Oliveira Souza Júnior – Universidade Federal do Piauí, CEAD
Anderson Nunes Da Silva - Universidade Federal do Norte do Tocantins
Adriana Barretta Almeida - Universidade Federal do Paraná
Jorge Luís Pereira Cavalcante - Fundação Universitária Iberoamericana
Jorge Fernando Silva de Menezes - Universidade de Aveiro
Antonio da Costa Cardoso Neto - Universidade de Flores Buenos Aires
Antônio Alves de Fontes-Júnior - Universidade Cruzeiro do Sul
Alessandre Gomes de Lima - Faculdade de Medicina da Universidade do Porto
Moacir Silva de Castro - Pontifícia Universidade Católica de São Paulo
Marcelo Silva de Carvalho- Universidade Federal de Alfnas
Charles Henrique Andrade de Oliveira - Universidade de Pernambuco
Telma Regina Stroparo - Universidade Estadual de Ponta Grossa
Valéria Raquel Alcantara Barbosa - Fundação Oswaldo Cruz
Kleber Farinazo Borges - Universidade de Brasília
Rafael Braga Esteves - Universidade de São Paulo
Inaldo Kley do Nascimento Moraes - Universidade Estadual do Sudoeste da Bahia
Mara Lucia da Silva Ribeiro - Universidade Federal de São Paulo

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

A313i

Akabane, Getulio K.

Inteligência Artificial [recurso eletrônico] : Histórico, Conceitos e Gestão / Getulio K. Akabane. – São José dos Pinhais, PR: Seven Editora, 2024.

Dados eletrônicos (1 PDF).

Inclui bibliografia.

ISBN 978-65-6109-097-1

1. Tecnologias digitais. 2. Inteligência artificial. I. Título.

CDU 004.434

Índices para catálogo sistemático:

1. CDU: Inteligência artificial 004.434

Bruna Heller - Bibliotecária - CRB10/2348

DOI: 10.56238/livrosindi202465-001

Seven Publicações Ltda
CNPJ: 43.789.355/0001-14
editora@sevenevents.com.br
São José dos Pinhais/PR

DECLARAÇÃO DO(A) AUTOR(A)

O(a) autor(a) deste trabalho DECLARA, para os seguintes fins, que:

Não possui nenhum interesse comercial que gere conflito de interesse em relação ao conteúdo publicado;

Declara ter participado ativamente da construção dos respectivos manuscritos, preferencialmente nas seguintes condições: "a) Desenho do estudo, e/ou aquisição de dados, e/ou análise e interpretação dos dados; b) Elaboração do artigo ou revisão para tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão";

Certifica que o texto publicado está completamente livre de dados e/ou resultados fraudulentos e defeitos de autoria;

Confirma a citação correta e referência de todos os dados e interpretações de dados de outras pesquisas;

Reconhece ter informado todas as fontes de financiamento recebidas para realizar a pesquisa;

Autoriza a edição do trabalho, incluindo registros de catálogo, ISBN, DOI e outros indexadores, design visual e criação de capa, layout interno, bem como seu lançamento e divulgação de acordo com os critérios da Seven Eventos Acadêmicos e Editora.

DECLARAÇÃO DA EDITORA

A Seven Publicações DECLARA, para fins de direitos, deveres e quaisquer significados metodológicos ou legais, que:

Esta publicação constitui apenas uma transferência temporária de direitos autorais, constituindo um direito à publicação e reprodução dos materiais. A Editora não é co-responsável pela criação dos manuscritos publicados, nos termos estabelecidos na Lei de Direitos Autorais (Lei 9610/98), no art. 184 do Código Penal e no art. 927 do Código Civil; O(s) autor(es) é(são) exclusivamente responsável(eis) por verificar tais questões de direitos autorais e outros, isentando a Editora de quaisquer danos civis, administrativos e criminais que possam surgir.

Autoriza a **DIVULGAÇÃO DO TRABALHO** pelo(s) autor(es) em palestras, cursos, eventos, shows, mídia e televisão, desde que haja o devido reconhecimento da autoria e edição e sem qualquer finalidade comercial, com a apresentação dos devidos **CRÉDITOS** à **SEVEN PUBLICAÇÕES**, sendo o(s) autor(es) e editora(es) responsáveis pela omissão/exclusão dessas informações;

Todos os e-books são de acesso aberto, portanto, não os venda em seu site, sites parceiros, plataformas de comércio eletrônico ou qualquer outro meio virtual ou físico. Portanto, está isento de transferências de direitos autorais para autores, uma vez que o formato não gera outros direitos além dos fins didáticos e publicitários da obra, que pode ser consultada a qualquer momento.

Todos os membros do conselho editorial são doutores e vinculados a instituições públicas de ensino superior, conforme recomendado pela CAPES para obtenção do Qualis livro;

A Seven Eventos Acadêmicos não atribui, vende ou autoriza o uso dos nomes e e-mails dos autores, bem como de quaisquer outros dados deles, para qualquer finalidade que não seja a divulgação desta obra, de acordo com o Marco Civil da Internet, a Lei Geral de Proteção de Dados e a Constituição da República Federativa.

GETULIO K. AKABANE

Pós-Doutorado (FEA-USP e TUMST-Tokyo University of Marine Science and Technology, Tokyo-Japão); Doutorado em Administração de Empresas (EAESP-FGV); Mestrado em Administração de Empresas (EAESP-FGV); Especialização em Administração (CEAG-FGV-SP); Pesquisador visitante (The University of Tokyo e TUMST-Tokyo University of Marine Science and Technology, ambos em Tokyo-Japão). Atuação como executivo em TI, Logística e Produção em empresas nacionais e multi-nacionais concomitante com a atividade de Docência (Direção de Faculdade, Coordenação de curso e de ensino) e Pesquisa e orientação na Graduação e Programa de Mestrado e Doutorado em IES pública e privada na área de TI, IA, Produção e Logística. Autor de Livros, capítulo de livros, inúmeros artigos científicos publicados em revistas científicas e Anais de congressos nacionais e internacionais.

APRESENTAÇÃO

CONHECIMENTO ALIMENTADO POR SOBRIEDADE E PRUDÊNCIA

Desde a publicação, em 2010, do livro, hoje clássico, sobre Artificial Intelligence: A Modern Approach, de Stuart Russel e Peter Norvig, títulos sobre Inteligência Artificial (IA) vêm crescendo a olhos vistos. Em sua maioria, durante algum tempo publicados em inglês, eles começaram a chamar a atenção de editoras brasileiras que passaram a providenciar as suas traduções para o português. Basta uma consulta à Amazon ou outra fonte de informação para que isso seja constatado. Diante desse quadro, é promissor que títulos acadêmicos sejam publicados em língua portuguesa por brasileiros, ou seja, por aqueles que conhecem não apenas o tema de que tratam, mas, sobretudo, as condições pelas quais a IA tem entrado no panorama brasileiro, em especial no panorama universitário. Este é o caso do livro de Getúlio Akabane que aqui apresento.

Graças ao gigantismo dos dados, à escalabilidade dos computadores, à sofisticação das CPUs e à computação na nuvem em grande escala, os sistemas de IA tornaram-se capazes de realizar uma série de tarefas que há pouco tempo eram ainda precárias ou então impensáveis: jogos, traduções de idiomas, conversão da fala em texto ou imagens e vídeos, planejamento, diagnóstico médico, previsões de várias ordens, e a fila se estende a se perder de vista.

Portanto, o crescimento visível das publicações sobre IA não é casual. Não deverá haver qualquer atividade humana, seja de que campo for, que não venha a ter a interferência da IA em suas práticas. Aceleradamente a IA está se tornando onipresente em todas as nervuras da vida humana. Quaisquer investigações sobre tecnologia hoje devem se tornar investigações sobre IA. Esses são os prognósticos daquelas cujas afirmações estão bem plantadas em conhecimento de causa, e esses não são apenas os desenvolvedores que têm a mão na massa, mas também todos aqueles que, de uma forma ou de outra, estão enfiados na pesquisa sobre IA, em especial nos impactos e consequências que ela traz para as sociedades, em todos os seus aspectos, desde as relações cotidianas até o poder, riqueza, guerra e trabalho,

Dada a relevância da questão que aí se desenha, informar-se sobre a IA em fontes confiáveis tornou-se madatório. A emergência de novas tecnologias é sempre capaz de provocar abalos que transformam a economia, as injunções políticas e as esferas da cultura. Isso não é novo. O que se perfila como novidade na IA é sua capacidade de transmutação técnica e penetração social a uma velocidade jamais vista. O que é novo também são as condições sócio-técnicas digitais dataficadas que a IA encontra para a sua repercussão e disseminação nos discursos sociais de que resulta um alarido de ideias que complicam e confundem o cenário.

Em qualquer campo em que estivermos, desde aqueles mais colados às questões da IA até aqueles que parecem mais longínquos, todos hoje sofreremos as interferências da IA em nossas vidas. Diante disso, informar-se é preciso. É necessário ter uma ideia do nosso grau de conhecimento sobre o assunto, pelo menos para que sejamos capazes de separar o joio do trigo, ou seja, para que possamos discriminar a informação válida daquela em que temores infundados apenas disfarçam a alienação.

Não resta dúvida de que a IA se constitui hoje em um recurso indispensável para atender a uma miríade de necessidades humanas, mas, ao mesmo tempo, isso não pode ocultar os riscos de várias ordens que a IA abriga. Vale lembrar que a IA tem dois Ps (Poder e Perigo) e dois Rs (Recurso e Risco). Ela é um prolongamento do humano e, como tal, carrega em si e, mais do que isso, expande as contradições do humano. Conseqüentemente, é preciso estar alerta. O primeiro caminho para isso é livrar-se da caixa de ressonância, que beira a histeria, dos eufóricos, de um lado, e dos disfóricos, de outro. Só o conhecimento nos dá a sobriedade e a prudência necessárias a um necessário engajamento sadio. Para isso, o leitor encontrará neste livro um auxílio inestimável.

Lucia Santaella Kassel

Livre Docente da PUC SP
Junho 2024

PREFÁCIO

O motor a vapor no século XVIII, as ferrovias no século XIX e os computadores do século XX transformaram dramaticamente o comportamento das economias e da sociedade.

Ao longo da história, a humanidade sempre utilizou a capacidade de aprender com a inteligência com as quais nascemos. Basta praticar o que aprendemos, adquirir novos conhecimentos, acumular experiências e... viver.

Desta forma, somos capazes de adquirir novas habilidades ou modificar aquelas que já temos com a faculdade de pensar, imaginar, criar nossas próprias idéias e... sonhar.

Neste século XXI a humanidade está transferindo estas virtudes para as máquinas. Ou seja, as máquinas capazes de aprender e que podemos ensiná-los...

Observamos assim, avanços consideráveis como veículos que podem reconhecer pedestres e outros objetos, sistemas que distinguem os seres vivos e até mesmo inteligências capazes de sonhar, pintar e compor música de forma autônoma.

Surge desta forma a questão: as máquinas podem pensar, aprender e superar os humanos? Uma máquina que está dialogando com uma pessoa a distância faz acreditar que está conversando com outro humano? Essa é uma dúvida que esteve presente desde que Alan Mathison Turing desafiou e que ainda não foi completamente resolvida.

Na ciência da computação, a Inteligência Artificial (IA) é a inteligência das máquinas em contraste com a inteligência natural exibida por humanos. A IA não se limita a sistemas construídos por seres humanos. É um termo coletivo que abrange uma ampla variedade de tecnologias cognitivas.

O campo da IA apresenta tecnologias bem conhecidas como a robótica, os veículos e drones autônomos, o reconhecimento facial e de imagens, o processamento da linguagem natural, e agora o metaverso com uma rede de mundos virtuais tridimensionais (3D) focados na conexão social híbrida entre o real e a imersão virtual em que o processo de geração de cena tem sido significativamente acelerado para quase em tempo real dentro da IA generativa.

No entanto, o grande desafio está no convívio entre os humanos e máquinas inteligentes sobretudo no equilíbrio e na responsabilidade no apoio a tomada de decisões.

Observa-se assim, a gradativa presença da IA no nosso cotidiano sobretudo na automatização das atividades rotineiras provocando aumento exponencial na disponibilidade de dados e com novas técnicas de aprendizado de máquina permitem que as organizações se beneficiem em tarefas gerenciais.

Como tal, pode-se esperar que no futuro, os humanos e máquinas desempenhem um papel combinado nos processos de tomada de decisão no nosso cotidiano.

Essas abrangências e visões múltiplas requer investigação e estudos aprofundados sobre como os seres humanos podem coexistir de forma harmônica com a IA e como suavizar o impacto desta tecnologia nas suas atividades diárias.

Desta forma, é uma tarefa desafiadora a definição única e comumente aceita da IA. Normalmente referida como a capacidade de aprendizado das máquinas e com acúmulo de experiências, ajustar-se a novas demandas e passa a executar tarefas semelhantes às dos humanos.

Os termos IA e sistemas IA foram introduzidos pela primeira vez na década de 1950.

Após anos de esperança e promessa, a IA está ganhando tração dentro das principais corporações, instituições de pesquisa e nas academias para aprimorar a tomada de decisões e fazer previsões com os algoritmos inteligentes, reinventar modelos de negócios e criar novos ecossistemas, aprimorar as redes sociais, e refazer a experiência do cliente e da própria sociedade.

No entanto, a IA enfrenta os obstáculos usuais de qualquer tecnologia em evolução sobretudo nas questões não comprovadas e desconhecidas. Ou seja, quando os sistemas inteligentes estão prestes a reproduzir fielmente os comportamentos humanos, “que têm inteligência cognitiva, emocional e social?”.

Para uma melhor compreensão e implementação da IA, o mundo deve considerar os requisitos e expectativas da IA, ou seja, nos campos da aplicação, emprego, ética, educação, entendimento e a dinâmica da sua evolução.

A presente obra tem no seu capítulo inicial o histórico da Inteligência Artificial e suas evoluções, diferentes visões conceituais e fundamentações, suas variadas formas, representações e vertentes.

No capítulo II são tratadas questões que envolvem o Machine Learning, como uma vertente do IA, as estruturas das Redes Neurais e o Deep Learning. As formas da Aprendizagem Supervisionada, Não Supervisionada e por Reforço também são aqui tratadas.

No capítulo III são apresentados os conceitos e a relação dos Algoritmos que formam a base do Machine Learning. Os Sistemas Especialistas, Reconhecimento de voz e discurso, o Processamento da Linguagem Natural e a Visão Computacional são abordadas no Capítulo IV. O universo da Ciência de Dados é abordado no Capítulo V.

No Capítulo VI estão reservados a estrutura e aplicações das Plataformas digitais, e no capítulo VII os conceitos da Robótica, sua Arquitetura e o Espaço da configuração.

O capítulo VIII trata da família de analíticos e a estrutura do Big Data e as Redes Neurais e por último as Considerações Finais

Assim, uma nova era impulsionada pela IA está surgindo. Está se movendo em direção ao pensamento colaborativo, que é um amálgama da capacidade cognitiva dos humanos e do poder computacional excepcional das máquinas. Essa nova onda de IA garante que os sistemas de IA sejam desenvolvidos para capacitar os seres humanos sem substituí-los ou ameaçá-los. Humanos e máquinas devem trabalhar como parceiros colaborativos para atingir objetivos. Os humanos estão projetando, treinando, implantando e testando esses sistemas com o uso das habilidades cognitivas. Simultaneamente, as máquinas fornecem aos humanos um poder computacional excepcional, permitindo o processamento e análise de dados volumosos em tempo real.

Desta forma, os seres humanos desempenham um papel crucial no sucesso do sistema de IA. Diferentes usuários satisfazem diferentes propósitos com sistemas de IA. Cientistas de dados, pesquisadores e desenvolvedores lidam com o design, desenvolvimento e melhoria contínua do sistema aumentando o desempenho, corrigindo erros e adicionando novas funcionalidades. Outra categoria são os usuários que são diretamente ou indiretamente afetados pelas decisões do sistema de IA. Esses usuários devem ser capazes de entender o raciocínio que levou a uma determinada decisão. O terceiro tipo são os especialistas do domínio, como médicos que usam sistemas de IA para auxiliar na tomada de decisões sobre seus pacientes. Todos os usuários devem conviver com o sistema de IA para poder usá-lo de maneira eficaz. Outro grupo são os formuladores de políticas e órgãos reguladores, garantindo que o sistema de IA cumpra os princípios éticos e morais por meio de auditoria e testes adequados juntamente com os gestores e membros do conselho da empresa que supervisionam o desenvolvimento do sistema de IA e avaliam seu uso e aplicação no mundo real. Todos esses usuários estão envolvidos em diferentes níveis do ciclo de vida da IA.

O humano digital produzido pela IA aumentará dramaticamente, e uma segunda vida digital está surgindo - uma sociedade combinada com humanos reais e digitais. É provável que outras espécies ou organismos renascerá e que um “verso-criador” surgirá. Tecnologia imortal-viver para sempre - eventualmente se tornará verdade. No estágio atual, é mais do lado inorgânico onde tudo é gerado com computadores digitais; contudo, a sociedade continuará, e as pessoas eventualmente alcançarão o terceiro verso ou até mesmo o multiverso em que o mundo orgânico e o mundo inorgânico podem se mesclarem. A IA acelerará essa revolução.

Boa Leitura a todos,

O Autor

SUMÁRIO

INTRODUÇÃO	12
CAPÍTULO 1 INTELIGÊNCIA ARTIFICIAL (IA)	18
CAPÍTULO 2 MACHINE LEARNING	55
CAPÍTULO 3 ALGORITMOS	88
CAPÍTULO 4 SISTEMAS ESPECIALISTAS	121
CAPÍTULO 5 CIÊNCIA DOS DADOS	142
CAPÍTULO 6 PLATAFORMAS	180
CAPÍTULO 7 ROBÓTICA	202
CAPÍTULO 8 ANALÍTICOS	231
CONSIDERAÇÕES FINAIS	245
REFERÊNCIAS	247

Desde os primórdios, os seres vivos precisavam desenvolver alguma habilidade física para mudar suas vidas cuja experiência foi passada de geração para geração. A mudança das estações era tão estável que proporcionou à espécie a possibilidade única de sobrevivência.

Os humanos aprenderam a cultivar a terra e continuam a inová-la e explorar conforme as estações para produzir alimentos.

Assim evoluiu a civilização ao longo de gerações a desenvolver e a progredir, cujos vestígios encontrados nos registros fósseis. E a biodiversidade ajudou trazer a estabilidade ao planeta terra.

Por mais de 2000 anos, os filósofos tentaram entender os desdobramentos da arte de imaginar, observar, aprender, raciocinar e recordar.

A primeira Revolução Industrial do século XVIII introduziu novos meios de mecanização, os teares e dispositivos motores a vapor que libertaram os trabalhadores das tarefas eminentemente físicas e as fábricas liberadas das limitações dos trabalhos braçais e a força dos animais.

No século XIX, a segunda revolução trouxe técnicas de produção em massa em que os produtos fossem eficientemente fabricados em escala até então sem precedentes.

Após a segunda guerra mundial (1945) a tecnologia passou a facilitar as nossas vidas. As mudanças foram se acelerando e parecia que nada deteria o nosso progresso. O futuro se configurava emocionante! Proporcionava tudo o que sempre sonhávamos!

A transformação para a tecnologia digital liderada pelo desenvolvimento dos computadores onde a terceira revolução industrial dominou o final do século XX.

Estas eras industriais compartilham em comum é que o crescimento econômico seguido de cada avanço tecnológico, não era meramente aditivo, mas exponencial. Pois, cada ciclo industrial introduz saltos na velocidade, na eficiência e na utilidade dos processos e traz um efeito multiplicador sobre o crescimento econômico.

Assim, as novas formas de energia ampliaram a capacidade física do ser humano e possibilitou criar uma ampla gama de produtos e capazes de replicá-los num ritmo mais veloz e assim estimular o desejo pelo consumo.

Resolvemos assim a produção de alimentos em grande escala e nada poderia nos deter a não ser por nós mesmos. Pois, os humanos continuam a consumir os recursos da terra até observar a sua descontinuidade. Observou o declínio das espécies vivas, o desequilíbrio do ecossistema e o

surgimento das intemperes da natureza antes não observadas como encontradas na terra de Serengeti¹.

Pois, as florestas tropicais são *habitats* especialmente preciosos. São lugares em que o talento e a inteligência da criação transbordam!

Esta verdade sempre existiu na pré-história, com a agricultura de sobrevivência; e gradativamente os humanos se distanciaram e os predadores naturais gradativamente estão sendo eliminados.

Neste século XXI vivenciamos a transformação social impulsionado pelas redes sociais e a tecnologia inteligente.

MAS, DO QUE SE TRATA A TECNOLOGIA INTELIGENTE?

O termo refere-se a qualquer sistema ou dispositivo que adota a combinação do aprendizado de máquina, a inteligência artificial, a robótica e a análise de dados para aumentar a produtividade com menos recursos.

Até então, cada ciclo tecnológico permitiu que os processos superassem as limitações existentes. Com o advento da tecnologia inteligente vislumbra-se avanços ainda maiores com possibilidades de superar as capacidades humanas.

Pois, os humanos são criaturas biológicas com emoções que influenciam nos vários processos de julgamento. Robôs e máquinas não compartilham estas características biológicas.

Muitos acreditam no "supercérebro eletrônico" dos computadores que agregou um potencial ilimitado da Inteligência Artificial (IA).

Então vejamos como realizamos as atividades diárias: “ No período do despertar matinal até quando vamos dormir, tomamos decisões de grande ou de pequenas proporções.” Mas, quais decisões são tomadas com pequena base de dados que conduz a imperfeições? Tomadas de decisões do cotidiano como qual *podcast* ouvir, o que e onde devo comprar, a melhor rota que conduz ao trabalho?...

Estas e outras questões com base nas intuições ou conhecimento prévio passou a ser um processo desafiador para o nosso cérebro biológico; mas com o apoio do cérebro eletrônico podemos ampliar a capacidade de executar as funções de perceber, entender, prever e manipular informações num mundo cada vez mais integrado e mais complexo de hoje.

¹ Serengeti significa em Maa (língua Maasai), "o lugar onde a terra continua para sempre" ou "planícies sem fim". O ecossistema Serengeti é uma região geográfica na África, abrangendo o norte da Tanzânia. Com seus rios, florestas, pântanos, savanas e extensões áridas abrigam cerca de 70 grandes mamíferos e mais de 500 espécies de pássaros onde os Maasai pastoreiam seu gado ao lado de animais selvagens. A área protegida na região inclui aproximadamente 30.000 km² de terra, incluindo o Parque Nacional do Serengeti e várias reservas de caça e hospeda a segunda maior migração de mamíferos terrestres do mundo; o que ajuda a protegê-lo como uma das Sete Maravilhas Naturais da África, e como uma das dez maravilhas naturais de viagens do mundo.

Estamos na era digital ou na quarta revolução industrial.

A digitalização é a conversão de dados e processos analógicos num formato legível pela máquina; ou seja, em caracteres 1 e 0 que podem ser interpretados e manipulados por computadores. É a conversão do conteúdo analógico em código binário que tornam os dados muito mais produtivos que o seu equivalente analógico. Isso se deve ao código digital (algoritmos que podem interpretar, manipular, processar e transformar qualquer forma de dados digitais) ativado pelo software em forma de instruções executáveis.

Os padrões para categorizar, estruturar, vincular e mover dados digitais tornaram possíveis com os algoritmos para serem mais significativo e utilizável e, portanto, mais valioso. A expansão da coleta, o armazenamento, processamento e a análise de dados se ampliam com a proliferação dos sensores e dispositivos em toda a economia e na própria sociedade. Esse fenômeno é descrito como "big data" e "Internet de tudo".

Nesse ambiente altamente conectado, os algoritmos criam valor a partir dos dados, e os dados, por sua vez, aprimoram os algoritmos levando ao aprendizado de máquina e ao desenvolvimento de inteligência artificial.

A abrangência da inteligência artificial permite que os dados passem a ter sentido, razão, que possibilita ação e adaptação para se transformar num importante ativo estratégico.

O relatório do comitê de política da economia digital da OECD (Organization for Economic Cooperation and Development) cita três fatores em que a transformação digital afeta a economia e a sociedade: a escala, o escopo e a velocidade que atuam em conjunto para criar efeitos econômicos e sociais complexos capaz de estimular mudanças fundamentais no sistema de produção.

O fator escala nas operações e na economia atende a demanda (aumento do valor à medida em que os usuários aprimoram a utilidade do produto). O escopo atua na riqueza e na integridade das ofertas de produtos). Enquanto que a alta velocidade de comercialização eletrônica combinada com economias de escala e o escopo contribuem para agregar valor aos produtos e serviços.

Os vetores da propriedade, dos ativos e do valor econômico fornecem vantagem competitiva e direcionam o investimento na direção das fontes de valores intangíveis no formato digitalizados. Os valores intangíveis como software e dados são combinados com investimentos tangíveis como infraestruturas, ferramentas, máquinas, equipamentos e edifícios - criando assim novos ativos flexíveis e escaláveis para atender à dinâmica das condições ambientais e explorar as novas oportunidades de negócios.

O elemento-chave da transformação digital da infraestrutura e as estruturas de capital é o fluxo de dados gerados por sensores incorporados ao sistema que permite monitorar o desempenho de toda a operação de forma remota em tempo real. O compartilhamento dos dados permite a melhoria

contínua do uso dos ativos com *loops de feedback* mediante os recursos do aprendizado de máquina, além da provisão de novos serviços com base na análise do fluxo de informações.

Outra mudança na natureza do investimento do capital tangível é o advento das plataformas digitais que conectam os provedores aos usuários facilitando acesso a informações proporcionando relacionamentos, mercados e a formação de ecossistemas.

A Internet como a infraestrutura comum para conectar usuários, ativos digitais, objetos físicos e serviços permitiu novos modelos e a natureza dos relacionamentos.

Os mercados se tornam maiores e mais bem informados, eficientes e completos e gerenciados pelas grandes empresas privadas que reúnem interesses e objetivos díspares com plataformas tecnologicamente sofisticadas em *design* e configurações para fins específicos, seja para interatividade ou na criação e compartilhamentos de novos conhecimentos.

O aprimoramento na coleta de dados robotizados pode estimular e sustentar relacionamentos de longo prazo entre os componentes que, por sua vez, reduzem ainda mais os custos de transações individuais. Embora a natureza virtual das conexões digitais possa corroer a confiança em algumas transações, sistemas de reputação *on-line*, análise de fraude em tempo real e outras análises de *Big Data* podem gerar e aumentar a confiança dos relacionamentos.

Assim, a IA é uma área da ciência da computação que explora a capacidade das máquinas de executar tarefas que normalmente requerem inteligência humana como raciocínio, o planejamento, o aprendizado e a o entendimento da linguagem escrita e falada. Trata-se da inteligência que não surge com processos naturais, ou onde a inteligência humana pode ser simulada ou associada. Da forma objetiva, refere-se a programas de computador que tomam decisões e procedem julgamentos semelhantes aos dos humanos como a arte de reconhecer objetos, animais ou até indivíduos em imagens que requerem certa dose de raciocínio.

Por outro lado, a inteligência humana é geralmente usada como contraponto da IA, embora na literatura, a inteligência animal também surja como uma forma de comparação.

Por outro lado, a prática do aprendizado de máquina requer três tipos de recursos: modelos de dados mais inteligentes, acesso a volume ilimitados de dados e uma poderosa computação na nuvem de custo acessível, os quais são cada vez mais viáveis no mundo de hoje.

O campo da IA procura assim dialogar com as entidades inteligentes. Diferentemente da filosofia e da psicologia que também se ocupam com a inteligência, a IA procura entender e construir a utilidade das entidades inteligentes.

Para o sucesso da IA requerem a combinação de dois elementos: a inteligência e o artefato como sistemas computadorizados que melhor representam a inteligência.

A IA também representa uma das faces do software conjugados com os sistemas operacionais, linguagens de programação e conjunto de ferramentas necessárias para escrever os programas e as respectivas documentações.

Outro recurso que compõe os recursos da IA é a rede neural projetado para simular a estrutura física e a lógica dos neurônios análogos a estrutura do cérebro humano.

Na tomada de decisões em estados binários simples de ligar ou desligar, as redes neurais repassam coletivamente os "pesos" das decisões entre os neurônios estabelecendo as melhores suposições para modelagem dos dados após os processos biológicos.

Outro conceito relacionado com a história da informação é a da lógica difusa (fuzzy) compostos de ferramentas de pesquisa prototipadas com técnicas de IA com base na correspondência e na similaridade ou na análise semântica. A lógica difusa refere-se a operações lógicas que não possui valores booleanos de verdadeiro ou falso, mas possui uma classificação de confiabilidade expressa em 0 e 1. Esses valores permitem tomados de decisões lógicas de maneira semelhante às redes neurais.

Esses recursos que formam as tecnologias cognitivas capazes de executar as tarefas na forma de realidade aumentada que apoiam na tomada das melhores decisões; e criar interações que tradicionalmente exigiam a inteligência dos recursos humanos.

Os desafios técnicos na construção de modelos complexos para fornecer dados para alimentar, treinar os modelos e as poderosas plataformas de banco de dados que estão em vias de viabilização gradativa.

Mas os dados e os recursos da tecnologia não são os maiores desafios. Nelas englobam a visão, a cultura e mudança na gestão que a sociedade e as organizações podem enfrentar.

Assim a IA contribui na análise de forma humanizada em conformidade com os tipos de inteligência que exhibe (cognitiva, emocional e social) ou no seu estágio mais evoluído como a inteligência artificial específica, generativa ou a superinteligência.

O termo IA universalmente adotado é a execução de tarefas geralmente reservadas para a cognição humana como reconhecendo padrões, previsão dos resultados nebulosos com certo grau de incerteza e na tomada de decisões complexas. Os algoritmos inteligentes podem perceber, interpretar e agir em ambientes formados de emoção, a compaixão e a criatividade que aproximam da inteligência humana.

A abundância dos dados, as mudanças na sociedade e na estrutura econômica criam desafios como explorar e proteger os dados² para estimular a inovação e a produtividade, novas posturas políticas, como proteger a privacidade e os direitos de propriedade intelectual e garantir a segurança pessoal e das organizações.

Encontrar esse equilíbrio será um desafio, mas reconhecer os custos e benefícios relativos ao compartilhamento e reutilização aprimorados de dados é essencial para fazer a transformação digital funcionar para o crescimento e o bem-estar da humanidade.

Cada um dos elementos apresentam a perspectiva única na forma de influenciar o nosso futuro impulsionado pela inteligência, automação, biotecnologia, nanotecnologia, veículos autônomos, dispositivos de Internet das Coisas (IoT), drones inteligentes e autônomos, saúde inteligente, sensores inteligentes, tecno cidades inteligentes e muito mais.

Envolvem também questões mais amplas como mudança climática, privacidade e segurança on-line em que as preocupações que nos afetarão e às futuras tecnologias com a qual conviveremos.

O autor

² LGPD: lei nº 13.709, aprovada em agosto de 2018 e com vigência a partir de agosto de 2020 para criar um cenário de segurança jurídica, com a padronização de normas e práticas, para promover a proteção, de forma igualitária e protocolar, no país e no mundo de compartilhar os dados pessoais de todo cidadão que esteja no Brasil. São compostos de elementos como Consentimento; Automatização com autorização do cidadão; ANPD (Autoridade Nacional de Proteção de Dados Pessoais) e agentes de tratamento; Gestão em foco: administração de riscos e falhas.

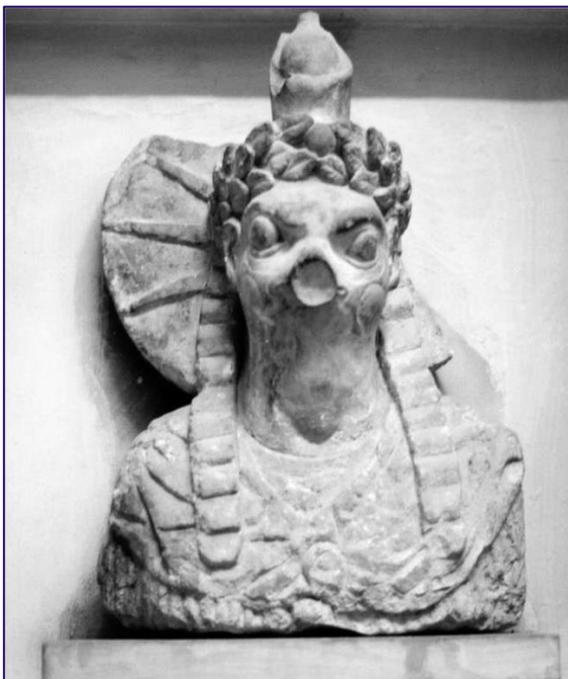
INTELIGÊNCIA ARTIFICIAL (IA)

“Se o conhecimento pode criar problemas, não é com a ignorância que podemos solucioná-los.” Isaac Asimov (1919-1992)

HISTÓRIA DA IA: DUPLICANDO A MENTE HUMANA

A busca pela inteligência artificial é tão moderna tanto quanto as fronteiras da ciência da computação, e é tão antiga quanto a própria antiguidade.

O conceito da "máquina pensante" começou cerca de 2500 a.C., quando os egípcios adornavam as estátuas falantes para obter conselhos místicos do Deus Re-Harmakhis¹. Tratava-se da figura de um gênio em que na região do seu dorso havia uma acomodação onde um sacerdote simulava a palavra divina que ecoava numa trombeta escondida nas cavidades que desce até a boca para impressionar os fiéis.



Estátua falante - Museu Egípcio no Cairo.

Antiga civilização egípcia

Com o amplo conhecimento tecnológico evocava a imaginação coletiva com uso dos dispositivos automatizados para afastar os predadores dos tesouros enterrados nos túmulos; ou como apoio nas reuniões da complexa liturgia religiosa, ou nas seções civis na produção de efeitos vocais especiais. A estátua de Re-Harmakhis do Baixo Egito é um artefato construído de pedra calcária branca com espaço no dorso em que a fala dos sacerdotes podem ser amplificadas e/ou alteração da tonalidade para simular a voz de divina. A imagem de Deus simbolizava não apenas o sol nascente, mas como o repositório da mais profunda sabedoria para a consulta, aconselhamentos ou orientação dos fiéis. Essas estátuas não tinham movimentos e nem ações, mas eram motivos de adorno e venerações segundo o egiptólogo Loukianoff (1936).

De fato, o sonho de desenvolver máquinas que possam imitar a cognição humana remonta há séculos.

¹ Harmakhis era considerado não apenas o Deus do sol como significa "Hórus no horizonte", o repositório da mais profunda sabedoria.

Os autômatos encontrados na tumba de Tutancâmon² são os antecessores dos robôs de hoje representados por figuras egípcias antigas com membros móveis.

A palavra “autômato” é a latinização do grego “αὐτόματον”, autômato, (neutro), “agir com a vontade de alguém”. Pode ser uma máquina autônoma, ou um mecanismo de controle projetado para seguir automaticamente uma sequência predeterminada de operações, ou responder às instruções predeterminadas. O relógio mecânico bellstrickers proporciona uma ilusão ao observador casual como se estivesse sendo operado com o poder próprio.

Os autômatos no mundo helenístico (período da história da Grécia e da parte do Oriente Médio compreendido entre a morte de Alexandre, o Grande em 323 a.C. e a anexação da península e das ilhas por Roma em 146 a.C.) foram concebidos como artefatos e ferramentas, brinquedos, adornos religiosos e protótipos para demonstrar os princípios científicos básicos.

Para o filósofo grego Aristóteles (384-324 a.C) é creditado a definição de raciocínio dedutivo³ e do silogismo⁴. Mais tarde, no século XV d.C., os ursos de tambor e figuras dançantes nos relógios eram os autômatos mais apreciados e a figura dos jogos como Maelzel Chess Automaton (1836)⁵, de Wolfgang von Kempelen em 1769, que predominara no século dezoito.



"The Turk", o jogador de xadrez automatizado de von Kempelen provou ser uma farsa; um mestre de xadrez sem pernas estava escondido no seu interior.

² Faraó da XVIII Dinastia egípcia Tutankamon, nasceu no Egito em 1341 a. C. que reinou apenas nove anos, de 1332 a 1323 a. C., e tornou-se famoso após sua tumba ter sido encontrada intacta em 1922, repleta de tesouros.

³ O raciocínio dedutivo é um processo lógico em que uma conclusão é baseada na concordância de múltiplas premissas que geralmente são consideradas verdadeiras e é provavelmente o processo mais utilizado em toda a matemática. O filósofo grego Aristóteles, considerado o pai do raciocínio dedutivo, escreveu o clássico: “Todos os homens são mortais. Sócrates é um homem. Portanto, Sócrates é mortal”.

⁴ Raciocínio dedutivo estruturado formalmente a partir de duas proposições (premissas) das quais se obtém por inferência uma terceira (conclusão). Ex.: "todos os homens são mortais; os gregos são homens; logo, os gregos são mortais".

⁵ "Maelzel's Chess Player" (1836) é um ensaio de Edgar Allan Poe que expõe um jogador de xadrez autômato fraudador chamado *The Turk* que se tornou famoso na turnê na Europa e nos Estados Unidos. O autômato falso foi inventado por Wolfgang von Kempelen em 1769 e foi trazido para os EUA em 1825 por Johann Nepomuk Mälzel após a morte de von Kempelen.

Com a invenção do mecanismo analítico de Charles Babbage em 1833, a inteligência artificial tornou-se uma possibilidade real. O escritor de ficção científica *H.G. Wells*⁶ com a obra *The Time Machine* (1895) começou a explorar o conceito de robôs e outras máquinas que pensam e agem como seres humanos.

Embora seja difícil identificar as raízes da IA pode-se rastrear a década de 1940 quando o escritor russo-americano de ficção científica Isaac Asimov⁷ publicou seu conto *Runaround* – uma história sobre um robô desenvolvido pelos engenheiros Gregory Powell e Mike Donovan⁸ que evoluiu para a publicação das “Três Leis da Robótica”:

1. o robô não pode ferir por inação ou permitir que um ser humano seja prejudicado;
2. um robô deve obedecer às ordens dadas pelos seres humanos, exceto onde as ordens estejam em conflito com a primeira Lei; e
3. um robô deve proteger a sua própria existência, desde que tal proteção não entre em conflito com a primeira nem a segunda lei.

As atividades das questões da IA se tornaram densas no meado do século XX, com Vannevar Bush⁹ com a obra "*As We May Think*" (1945) oferecendo uma visão do futuro onde os computadores poderiam ajudar os humanos de várias maneiras; John Von Neumann¹⁰ com a afirmativa: "*Diga-me exatamente o que uma máquina não possa fazer e eu vou programá-la para fazê-lo*".

Quase ao mesmo tempo, o matemático inglês Alan Turing¹¹ trabalhou em questões muito menos ficcionais e desenvolveu uma máquina chamada *The Bombe* para o governo britânico com o objetivo de decifrar o código *Enigma* usado pelo exército alemão na Segunda Guerra Mundial. *The Bombe* tinha um peso de cerca de uma tonelada e é considerado o primeiro computador eletromecânico em funcionamento. Pois, a poderosa maneira pela qual conseguiu quebrar o código *Enigma*, uma tarefa considerada impossível até mesmo para os melhores matemáticos humanos, fez Turing questionar sobre a inteligência de tais máquinas.

⁶ Romancista, jornalista, sociólogo e historiador inglês e conhecido por romances de ficção científica como “A Máquina do Tempo”, “A Guerra dos Mundos” e romances cômicos como “Tono-Bungay” e “A História do Mr. Polly”.

⁷ foi um escritor e bioquímico norte-americano, nascido na Rússia, autor de obras de ficção e divulgação científica.

⁸ Gregory Powell e Mike Donovan são personagens fictícios dos contos de Isaac Asimov “I Robot”. Eles são engenheiros práticos especialistas e não roboticistas teóricos que são empregados principalmente no teste de novos robôs experimentais em situações práticas - em planetas ou estações espaciais (os robôs são proibidos na Terra). Eles entram regularmente em situações complexas e potencialmente perigosas ao tentar resolver problemas de robôs em campo.

⁹1890-1974, engenheiro, inventor e político estadunidense conhecido pelo seu papel político no desenvolvimento da bomba atômica e pela ideia do *memex* — visto como um conceito pioneiro, precursor da world wide web.

¹⁰ 1903-1957, foi um matemático húngaro de origem judaica, naturalizado estadunidense, é conhecido por ter formalizado o projeto lógico de um computador sugerindo que as instruções fossem armazenadas na memória que tornaria mais rápido. A maioria dos computadores de hoje em dia segue ainda o modelo proposto por Von Neumann.

¹¹ Alan Mathison Turing (1912-1954) britânico, foi um matemático, lógico, criptoanalista e influente cientista de computação. Formalizador conceitual de algoritmo da máquina de Turing, e precursor da inteligência artificial.

Em 1950, ele publicou seu artigo “*Computing Machinery and Intelligence 3*”, onde descreveu como criar máquinas inteligentes e, em particular, como testar a sua inteligência.

O teste de Turing é considerado hoje como uma referência para identificar a inteligência de um sistema: “*se um humano está interagindo com outro humano e uma máquina, e se o humano é incapaz de distinguir a máquina do humano, diz-se que a máquina é inteligente*”.

A IA, desta forma foi introduzida como uma área da ciência em meados de 1950 e evoluiu rapidamente, e em 1956 os pesquisadores John McCarthy (Dartmouth College), Marvin Minsky (MIT), Claude Shannon (Bell Labs- que formou a teoria da informação) e Nathaniel Rochester (IBM- mais tarde projetou o IBM 701, o primeiro computador científico comercial) reunidos no evento “*The Dartmouth College Artificial Intelligence Conference: the Next Fifty Years*” no Dartmouth College (Hanover, N.H).

O *workshop Dartmouth Summer Research Project on Artificial Intelligence* (DSRPAI- finalidade de criar uma área de pesquisa voltada para a construção de máquinas capazes de simular a inteligência) financiado pela Fundação Rockefeller com três objetivos: celebrar o projeto “*Dartmouth Summer Research*” para avaliar o progresso, seu estágio e o rumo do projeto da IA.

Assim, McCarthy cunhou o termo “inteligência artificial” para solidificar a orientação no campo da pesquisa.

Desta forma, a inteligência artificial (IA) foi definida como a “*capacidade do sistema de interpretar e aprender corretamente os dados externos para usá-los, interpretar e executar as tarefas com adaptação flexível para atingir metas e objetivos específicos*”. Foi estabelecida como uma disciplina acadêmica nos anos de 1950.

Outro fato é o programa de computador ELIZA, criado entre 1964 e 1966 por Joseph Weizenbaum no MIT (Massachusetts Institute of Technology) como uma ferramenta de processamento de linguagem natural capaz de simular a conversa humana e passar no teste de Turing.

Outro marco é o surgimento de *Expert Systems*, uma coleção de regras onde a inteligência humana pode ser formalizada e reconstruída de cima para baixo com a série de declarações “se-então”.

Na década de 1940, o psicólogo canadense Donald Hebb desenvolveu a teoria da aprendizagem conhecida como *Hebbian Learning* que replica o processo dos neurônios no cérebro humano. Isso levou à criação de pesquisas sobre as Redes Neurais Artificiais. Mas o trabalho estagnou em 1969 quando Marvin Minsky e Seymour Papert mostraram que os computadores não tinham capacidade de processamento suficiente exigido pelas redes neurais artificiais.

As redes de aprendizado profundo podem ter milhares de camadas com bilhões de parâmetros. Ao contrário do cérebro humano, são “treinadas” com grande quantidade de dados conhecidos e aplicam o que “aprenderam” para escolher matematicamente e reconhecer os padrões de forma

precisa num universo de dados não estruturados em multimídia. São tomadas de decisão ou uma previsão relacionada a qualquer pergunta como “de quem é o rosto na imagem? Quais palavras foram utilizadas no segmento de áudio? A compra é fraudulenta?”

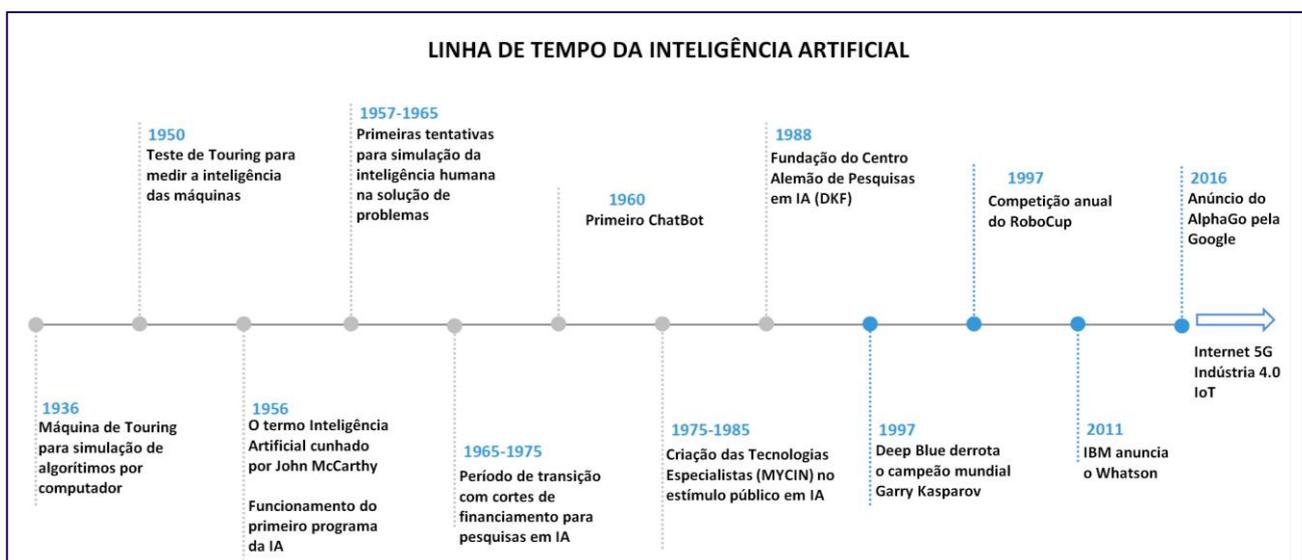
O surgimento do *Deep Blue* em 1997 desenvolvido pela IBM que venceu o campeão mundial de xadrez da época, o soviético Garry Kasparov também foi um fato importante na história da IA.

As redes neurais artificiais voltaram na forma de *Deep Learning* (aprendizado profundo) em 2015 com o *AlphaGo*, o programa desenvolvido pela *Google* que conseguiu superar o campeão do mundo de Go (jogo de tabuleiro japonês), o chinês Ke Jie numa partida melhor de três.

O Go é substancialmente mais complexo do que xadrez (na abertura, o xadrez há 20 movimentos possíveis e 361 no Go) onde havia a crença de que os computadores nunca seriam capazes de derrotar os humanos.

Na rede neural *Deep Learning*, o aprendizado ocorre com base nos algoritmos de reconhecimento de imagem e da fala, cujas aplicações vão desde os alto-falantes inteligentes até os veículos autônomos. As milhares de camadas de profundidade com bilhões de parâmetros são “treinadas” com grande quantidade de dados conhecidos e a aplicação do “aprendizado” se destinam para reconhecimento de padrões no universo de dados não estruturados.

Na rede neural *Deep Learning*, o aprendizado ocorre com base nos algoritmos de reconhecimento de imagem e da fala, cujas aplicações vão desde os alto-falantes inteligentes até os veículos autônomos. As milhares de camadas de profundidade com bilhões de parâmetros são “treinadas” com grande quantidade de dados conhecidos e a aplicação do “aprendizado” se destinam para reconhecimento de padrões no universo de dados não estruturados.



Fonte: Wehle, 2019

Para os fins deste estudo, podemos definir a IA como: “o sistema com a capacidade de interpretar e aprender com os dados externos para atingir objetivos e tarefas específicas mediante adaptação flexível”.

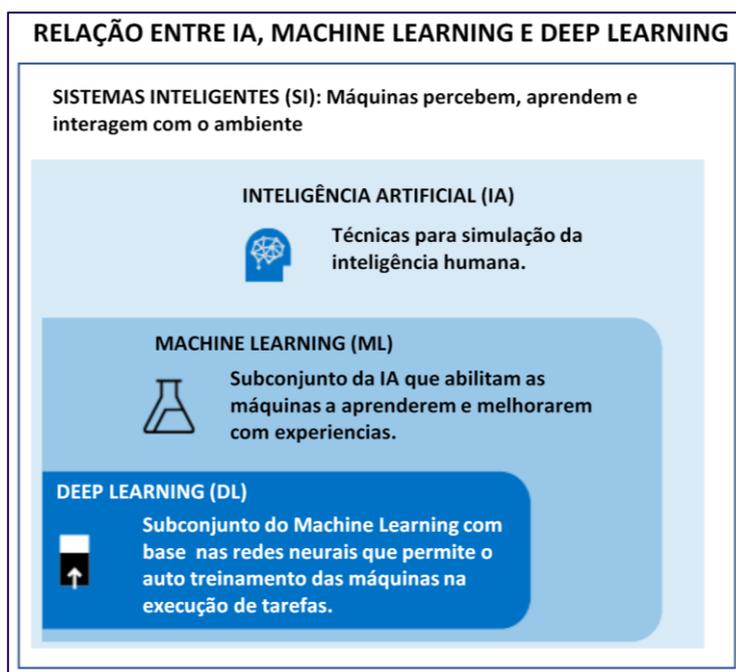
Pois a tecnologia da IA opera no domínio da automação e do aprendizado contínuo que impulsiona a coleta, o armazenamento, a manipulação, a análise, a gestão e a recuperação de informações focada na tomada de decisões.

As tecnologias como aprendizado profundo, algoritmos genéticos, e processamento de linguagem natural, a IA pode treinar máquinas para reconhecer padrões com grande quantidade de dados.

As ferramentas mais disponíveis da IA de uso pessoal incluem assistente pessoal (Alexa, Siri e Cortana), planejamento de viagens (Mezi), música (Pandora), planejamento financeiro (Olivia), tradução de idiomas (Liv) e soluções para *smarhome* (Nest). As ferramentas populares de IA para uso comercial incluem plug-and-play soluções para necessidades de negócios (por exemplo, Fluid AI), comércio eletrônico e marketing digital (Sentient), automação de processos (Amazon MTurk), reconhecimento facial (Haystack), assistente de linguagem jurídica (Legal Robot) e pontuação de crédito (Lenddo).

A partir de 2015, o Google Cloud, a Amazon Web Services e o Microsoft Azure e outros começaram a intensificar pesquisas para aprimorar os recursos de processamento de linguagem natural, visão computacional e ferramentas de análise dos dados.

Desta forma, os recursos do aprendizado de máquina e o aprendizado profundo que aprendem, pensam, raciocinam, decidem e novamente aprendem com as experiências acumuladas se aprimoram na tomada de decisões e se aproximam dos humanos.



Fonte: Adaptado de Microsoft, 2020

IA ASSUME DIFERENTES FORMAS

A Inteligência artificial é um termo genérico que se aplica em diferentes áreas, seja de caráter geral ou específicas tanto no campo da pesquisa quanto nas aplicações sociais e corporativas de forma sobreposta, que incluem:

- **IA de caráter geral:** Sistemas que interagem com o ambiente adjacente com base nos dados conhecidos. No caso do DeepMind do Google que se conecta com a rede neural no aprendizado do videogame simulando os humanos.
- **Processamento de linguagem natural (PNL):** Permite a leitura, interpretação e a comunicação da linguagem humana. Com os métodos estatísticos e programação semântica entende a gramática e o sintaxe e, em alguns casos, percebe as emoções do agente que interage com o sistema, como os robôs de bate-papo.
- **Percepção da máquina:** Sensores - câmeras, microfones, acelerômetros, GPS, radar que estimula os recursos como o reconhecimento de fala, da face, de objetos, do ambiente (luminosidade, temperatura, pressão, velocidade do vento) e a visão computacional entre outros.
- **Robótica:** Dispositivos robóticos utilizados em fábricas, hospitais e outras aplicações como *drones* em atividades como sensoriamento e mapeamento remotos, processos estocásticos ou programação complexa.

- **Inteligência social:** Veículos autônomos, robôs e assistentes digitais como *Siri* e *Alexa* que exigem coordenação e orquestração apropriadas. Estão associados com a compreensão do comportamento humano além do reconhecimento das normas e dos protocolos sociais.

O QUE É IA?

Inteligência Artificial (IA) é a parte da ciência da computação destinada a projetar sistemas de computadores inteligentes, isto é, sistemas que exibem as características que são associadas com a inteligência no comportamento humano - como a linguagem falada, o aprendizado, raciocínio, resolução de problemas entre outros (IBarr e Feigenbaum, 1981).

Para o pai da inteligência artificial, McCarthy (2001) é a "*ciência e a engenharia de fabricação de máquinas inteligentes, especialmente programas de computadores inteligentes*".

A Comissão Europeia sobre IA define como "*sistemas que exibem o comportamento inteligente pela análise do ambiente e realiza ações com certo grau de autonomia para atingir objetivos específicos*".

Semelhante à tecnologia, a IA trata do uso de técnicas (pesquisa heurística, proposições) que formam o universo da ciência dos computadores (T. Winograd apud Bobrow e Hayes 1985, p. 380).

Na sua origem, a palavra artificial vem da raiz latina "*facere arte*" que significa "fazer arte", desta forma se traduz livremente em inteligência artificial.

Já como software, a IA é um "*ser humano que se comporta de forma limitada*" com o propósito de replicar a inteligência humana na sua totalidade.

Winston (1984) sugere como "*estudo de ideias que permitem aos computadores serem inteligentes*". Rich e Knight (1991) definem a IA como o "*estudo para que os computadores façam coisas que, no momento, as pessoas se saem melhor*".

A seguir algumas definições e/ou descrições mais comuns da IA:

- A IA é inteligente porque aprende;
- IA transforma dados em conhecimento;
- IA é uma solução inteligente de um problema;
- A IA incorpora a capacidade de se adaptar ao ambiente, de lidar com o conhecimento incompleto ou incorreto.

Na academia as definições mais comuns:

- A capacidade do computador digital ou um robô de executar tarefas comumente associadas a seres inteligentes.

- Máquina que executa tarefas com certo grau de inteligência que antes era realizado por seres humanos
- Simulação dos processos de inteligência humana por máquinas, especialmente sistemas de computador que incluem aprendizado, raciocínio e autocorreção.
- Capacidade da máquina imitar de forma inteligente o comportamento humano

Observa-se assim que a IA é um termo amplo abordando os aspectos técnicos pelos quais os computadores simulam a inteligência humana com recursos da lógica, das regras de programação *if-then*, das árvores de decisão e do aprendizado de máquina (incluindo a aprendizagem profunda).

Os propósitos da Inteligência artificial é dotar o computador da capacidade de realizar o pensamento abstrato, a capacidade de análise dentro de certo contexto e ser criativo sem ser inteligente. Ou seja, as máquinas com a capacidade de solução com software e algoritmos que normalmente são resolvidos por humanos com a inteligência natural.

Para visualizar mais claramente o conceito de IA, você pode pensar num *chatbot* cuja tarefa é auxiliar os frequentadores do restaurante a reservar uma mesa. Por natureza, este *chatbot* é um programa treinado com milhares de perguntas e respostas relevantes sobre reservas. Desta forma ele aprende com o fluxo de diálogo sobre reservas e após o treinamento, o *chatbot* pode conversar com os clientes. No entanto, se um cliente desvia do tópico principal da reserva da mesa e solicitam recomendações da comida, o *chatbot* torna-se inoperante, pois não foi treinado para executar esta tarefa.

Assim, a inteligência artificial é um “*software ou programa de computador provido de mecanismos com capacidade de aprendizado*”. No seu sentido prático é a “*capacidade de tomada de decisão semelhante aos humanos*”.

Kaplan e Haenlein (2019, p. 17) define IA como “*a capacidade do sistema de interpretar e aprender corretamente com os dados externos e usar essas aprendizagens para realizar tarefas e atingir metas e objetivos específicos através de adaptação flexível*”.

Em outras palavras, pode-se definir a IA como a “*habilidade com que as máquinas utilizam os algoritmos para aprender a partir dos dados e aplicar na tomada de decisões semelhante aos humanos*”.

A vantagem sobre os humanos está na capacidade de analisar grande volume de dados e informações num curto espaço de tempo com acurácia e precisão e de forma contínua, sem interrupções ou intervalos.

Pode se assim apontar a inteligência artificial (IA) como o ramo da ciência da computação para a pesquisa e desenvolvimento do comportamento humano simulado em máquinas inteligentes (Hans-Dieter Wehle, 2019).

Trata se desta forma de examinar métodos que permitem um comportamento inteligente dos computadores para trabalhar de forma autônoma na solução de problemas. Softwares capazes de processar imagens, textos, vídeos e áudios ou outros formatos de dados como sinais elétricos e de comunicação com variadas frequências que representam diversos fenômenos da natureza como velocidade do vento, temperatura, pressão, umidade com capacidade de aprender continuamente a partir destas mensagens.

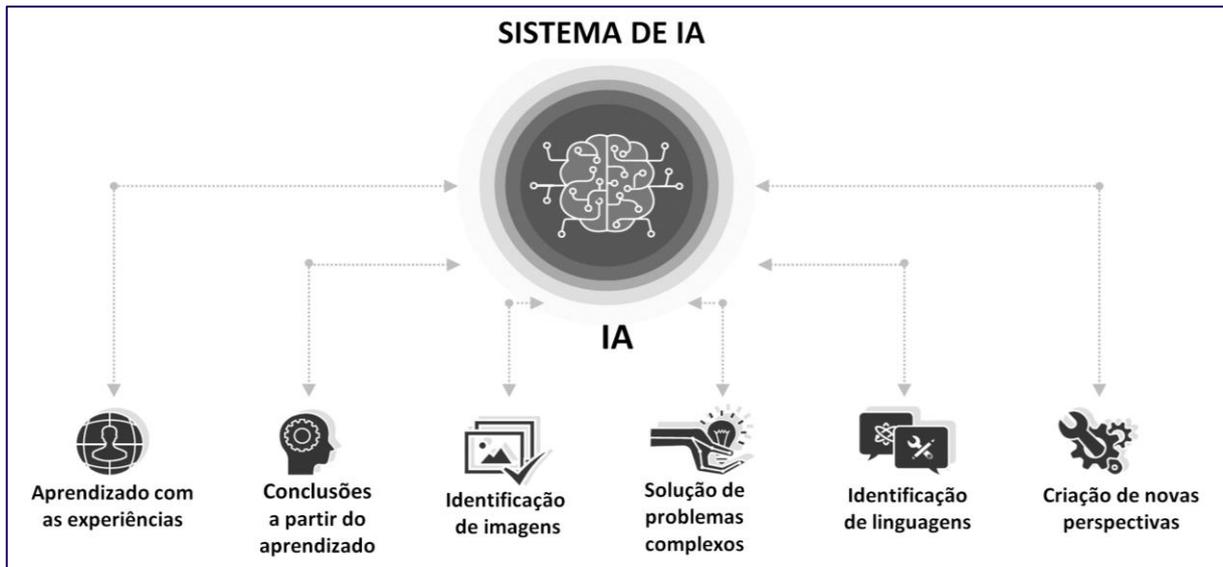
Mediante estes aprendizados, a máquina pode aplicar os conhecimentos adquiridos em outras situações nas diversas áreas do conhecimento.

A taxa de erro é também significativamente menor em relação às outras máquinas e os próprios humanos.

Nas outras áreas do conhecimento, a inteligência (das máquinas e dos humanos) é objeto de estudo dos psicólogos, biólogos e neurocientistas que tem como base a racionalidade e da emotividade na escolha da melhor ação para atingir um determinado objetivo por meio de critérios a serem otimizados e dos recursos disponíveis.

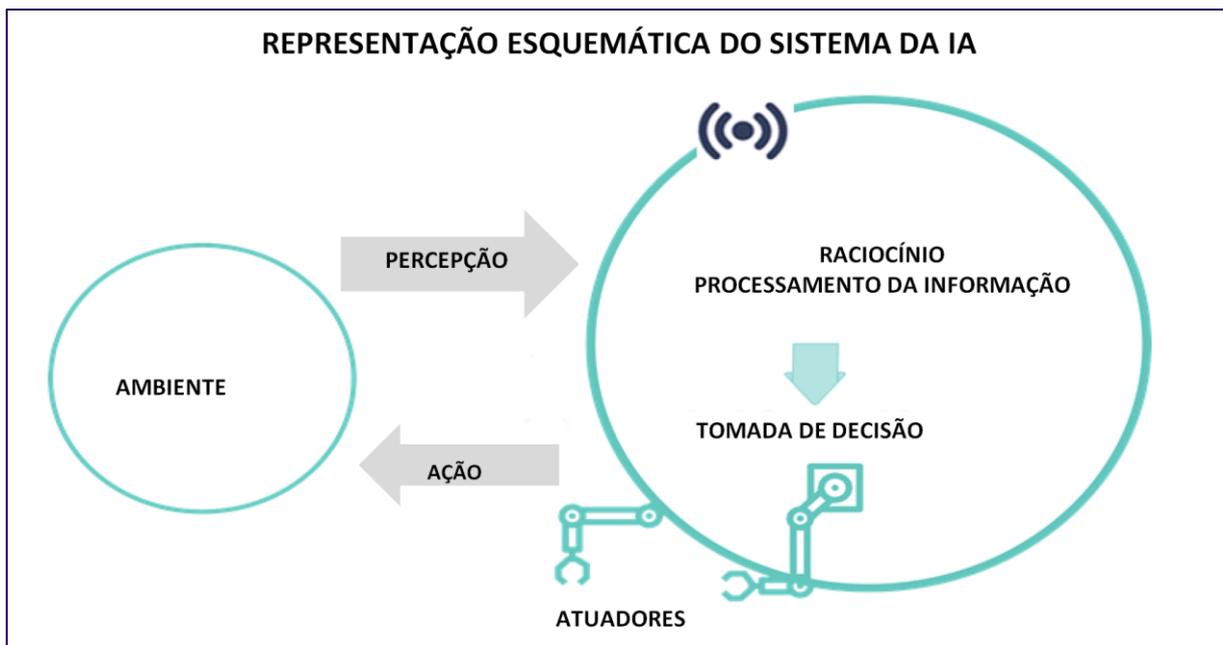
No campo das aplicações, a IA começa a ser adotadas no apoio aos processos mecânicos e repetitivos agregando significativos benefícios com o aumento da eficiência no trabalho, nas tarefas e no próprio estilo de vida dos humanos.

Assim, o Sistema de IA é composto por máquinas configuradas com diferentes níveis de autonomia e para um determinado objetivo definidos pelos humanos que permitem fazer análises, interpretações, previsões, recomendações ou tomadas de decisões que influenciam os ambientes reais ou virtuais.



Fonte: Wehle, 2019

Assim, o termo sistema de IA pode significar qualquer componente, software e/ou hardware baseado em IA. De fato, são subsistemas de um sistema maior, ou seja, são sistemas dependentes. Antes de tudo é racional (Russell e Norvig, 2009) que é capaz de interagir com o ambiente em que está imerso coletando e interpretando dados com sensores, processando os sinais coletados e tomando as decisões conforme as percepções e escolhendo a melhor ação com os atuadores e modificando o ambiente. Os sistemas de IA podem usar regras simbólicas ou aprender um modelo numérico e podem adaptar seu comportamento segundo o ambiente analisado conforme figura a seguir:



Fonte: Russel e Norvig (2009)

Sensores e percepções: representados pelo símbolo *wifi* e podem ser câmeras, microfones, teclado, *sites* ou outros dispositivos de entrada como sensores físicos (temperatura, pressão, proximidade, força/torque, sensores táteis).

Em geral, os sensores captam dados relevantes do ambiente conforme atribuído pelo *designer* humano e podem ser estruturados e não estruturados.

Os dados estruturados são organizados de acordo com modelos predefinidos (como dados de transação de uma empresa), enquanto dados não estruturados não possuem uma estrutura definida ou conhecida (imagem de uma câmera de segurança ou trecho qualquer de um texto). (Será objeto de descrição detalhada no capítulo 5)

COMPREENDENDO A INTELIGÊNCIA

Os recursos da IA e os sistemas inteligentes permitem que o cérebro humano possa construir outros sistemas semelhantes ao do comportamento humano.

Mas, o que é inteligência?

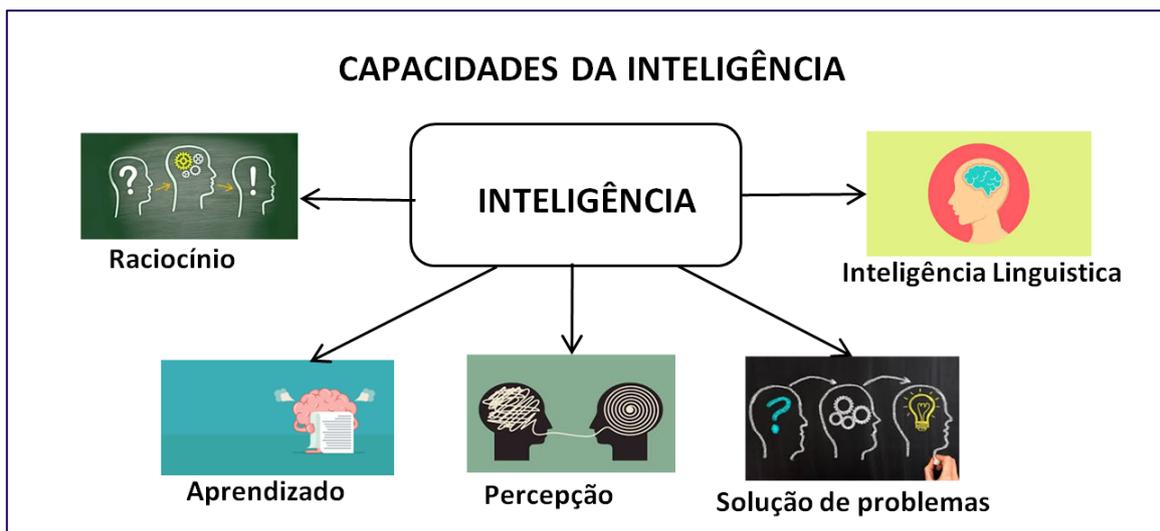
É a capacidade do sistema para calcular, raciocinar, perceber relacionamentos e analogias, aprender com a experiência, armazenar e recuperar informações da memória, resolver problemas, compreender ideias complexas, usar a linguagem natural com fluência, classificar, generalizar e adaptar novas situações (McCarthy, 2007).

Desta forma, uma máquina ou um sistema é artificialmente inteligente quando está equipado com pelo menos uma ou todas estas inteligências.

COMPOSIÇÃO DA INTELIGÊNCIA

A inteligência é intangível, e é composto pelo:

- Raciocínio
- Aprendizagem
- Capacidade de solução de problemas
- Percepção
- Inteligência Linguística



INTELIGÊNCIA MULTIFUNCIONAL

Nº	Inteligência e a Descrição	Exemplos
1	Inteligência Linguística Capacidade de falar, reconhecer e usar mecanismos de fonologia (sons da fala), sintaxe (gramática) e semântica (significado).	Locutores, Oradores
2	Inteligência Musical Capacidade de criar, comunicar e entender significados feitos de som, compreensão de tom, ritmo.	Músicos, Cantores, Compositores
3	Inteligência Lógica-matemática Capacidade de usar e entender relacionamentos na ausência de ação ou objetos. É também a capacidade de entender ideias complexas e abstratas.	Matemáticos, Cientistas
4	Inteligência Espacial Capacidade de perceber informações visuais ou espaciais, alterá-las e recriar imagens visuais sem referência aos objetos, construir imagens 3D e movê-las e girá-las.	Projetistas, Astronautas, Fisiatras
5	Inteligência Corporal-Cinemática Capacidade de usar todo ou parte do corpo para resolver problemas ou produtos da moda, controlar as habilidades motoras com precisão e habilidade e/ou com força e manipular os objetos.	Esportistas, dançarinos
6	Inteligência Intrapessoal Capacidade de distinguir os próprios sentimentos, intenções e motivações.	Buda Gautama
7	Inteligência Interpessoal Capacidade de reconhecer e fazer distinções entre os sentimentos, crenças e intenções de outras pessoas.	Comunicadores de massa, entrevistadores

Fonte: Gardner (2011)

RACIOCÍNIO

É o conjunto de processos que nos permite fornecer base para julgamento, tomada de decisões e previsão. De forma ampla é composta por dois tipos - Raciocínio indutivo e o Raciocínio dedutivo.

Raciocínio Indutivo	Raciocínio dedutivo.
Conduz a observações específicas para fazer declarações gerais e amplas.	Começa com uma declaração geral e examina as possibilidades de chegar a uma conclusão lógica e específica.
Exemplo - "Nita é professora. Nita é estudiosa. Portanto, todos os professores são estudiosos".	Exemplo - "Todas as mulheres acima de 60 anos são avós. Shalini tem 65 anos. Portanto, Shalini é avó."

APRENDIZAGEM

A capacidade de aprender é propriedade dos seres humanos, espécies particulares de animais e sistemas habilitados para IA. Assim, o processo de aprendizagem é categorizado da seguinte forma:

- **Aprendizagem Auditiva:** Aprendizado pela audição. Ex. alunos ouvindo as palestras em áudio gravado.
- **Aprendizagem Episódica:** Aprendizado pela memorização das sequências de eventos que alguém testemunhou ou experimentou de forma linear e ordenado.
- **Aprendizagem Motora:** Aprendizado pelo movimento preciso dos músculos. Ex. escolhendo objetos, escrevendo etc.
- **Aprendizagem por observação:** Aprendizado pela observação e pela imitação. Ex. a criança tenta aprender imitando os pais.
- **Aprendizagem Perceptiva:** Aprendizado pelo reconhecimento dos estímulos que já se viu antes. Ex. identificando e classificando os objetos e/ou situações conhecidas.
- **Aprendizagem relacional:** Aprendizado pela diferenciação entre vários estímulos com base em propriedades relacionais, em vez de propriedades absolutas. Ex. adicionar "pouco menos" de sal na hora de cozinhar batatas que ficaram salgadas da última vez, utilizando uma colher de sopa como dosador.
- **Aprendizado espacial:** Aprendizado por meio de estímulos visuais como imagens, cores, mapas, etc. Ex. uma pessoa pode criar um roteiro mental antes de seguir o caminho.
- **Aprendizagem com resposta a estímulos:** Aprendizado pelo comportamento específico diante de um determinado estímulo. Ex. um cão levanta a orelha ao sonar da campainha.

CAPACIDADE DE SOLUÇÃO DE PROBLEMAS

É o processo em que se percebe e/ou tenta chegar a uma solução desejada a partir de uma situação inicial, seguindo algum caminho que é bloqueado por obstáculos conhecidos ou desconhecidos. Isto inclui processo de tomada de decisões, que é o processo de selecionar a melhor alternativa dentre várias alternativas para atingir a meta desejada.

PERCEPÇÃO

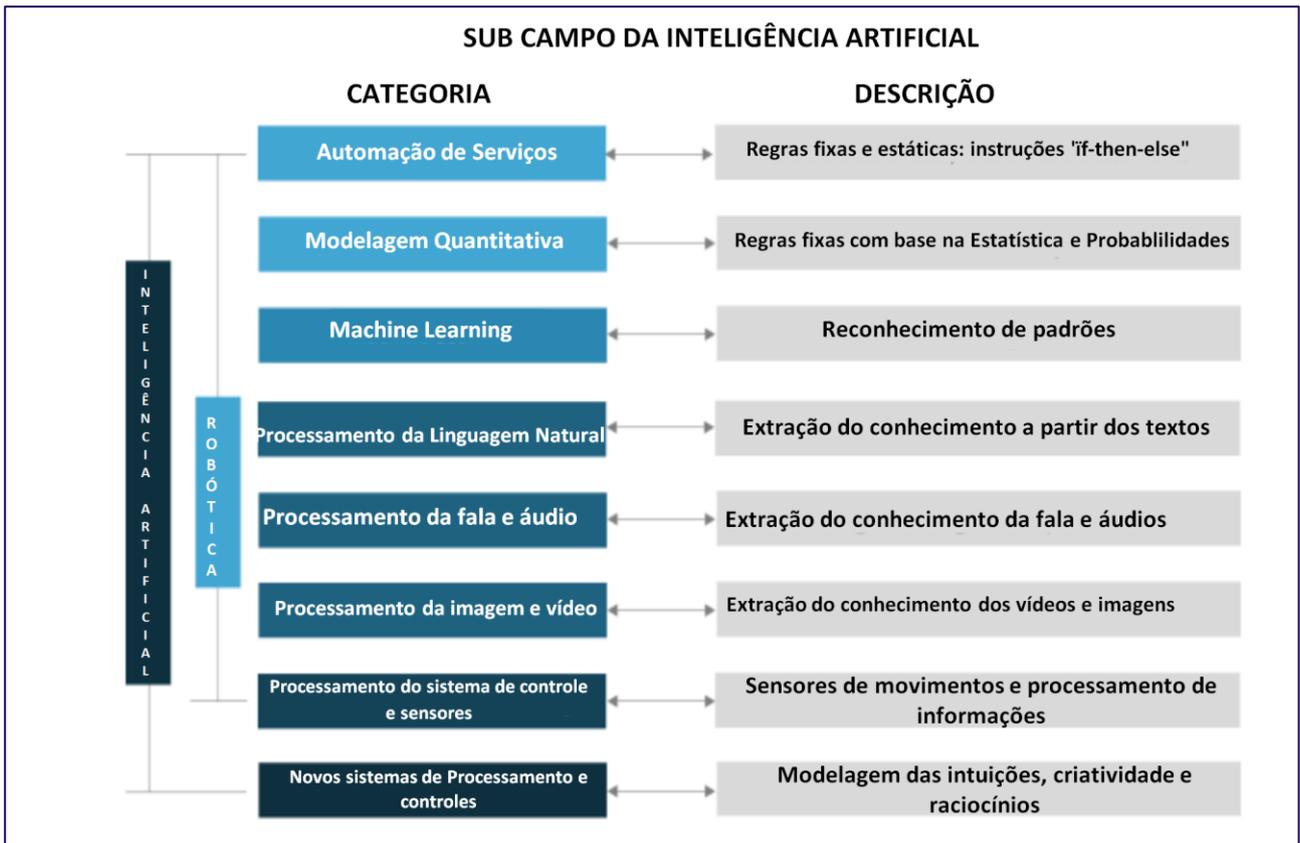
É o processo de aquisição, interpretação, seleção e organização de informações que pressupõe a capacidade de sensoriamento. Nos seres humanos, a percepção é auxiliada por órgãos sensoriais como visão, audição, olfato, sensação térmica entre outros. No domínio da IA, o mecanismo de percepção reúne os dados adquiridos de maneira significativa pelos diversos tipos de sensores existentes como câmeras, detectores de fumaça, calor, luminosidade, presença, entre outros.

INTELIGÊNCIA LINGUÍSTICA

É a capacidade de interpretar, compreender, falar e escrever a linguagem verbal e escrita. É um recurso importante na comunicação interpessoal.

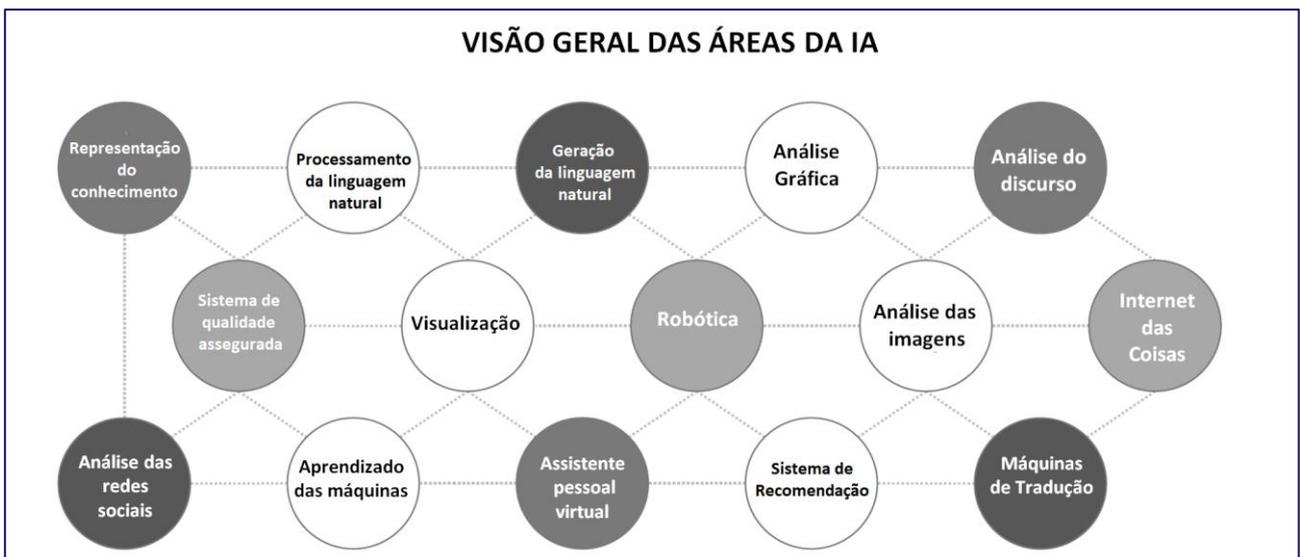
O QUE ESTÁ ENVOLVIDO NA IA

Entre 1956 e 1982 deu origem a vários subcampos da IA e grande parte deste trabalho levou aos primeiros protótipos da teoria moderna da IA.



Fonte: Gartner, 2020

A inteligência artificial é um campo vasto de estudo e de pesquisa de forma a encontrar soluções para problemas do mundo real conforme a figura a seguir:



Fonte: Wehle, 2019

MACHINE LEARNING

É um dos campos mais difundidos em IA. O conceito básico deste recurso é fazer com que a máquina aprenda com dados à medida que os seres humanos possam aprender com a experiência. Ele contém modelos de aprendizado com base nos quais as previsões podem ser feitas com dados desconhecidos que será objeto de detalhamento posterior.

LÓGICA

Tem como base o uso da lógica matemática para executar os programas de computador. Nela contém regras e fatos para realizar a correspondência de padrões, análise semântica, etc.

BUSCA

Este campo de estudo é basicamente usado em jogos como xadrez, jogo da velha, Go. Os algoritmos de pesquisa fornecem a solução ideal depois de pesquisar todo o espaço de pesquisa que serão tratadas neste estudo.

REDES NEURAIAS ARTIFICIAIS (RNA)

Essa é uma rede de sistemas computacionais eficientes, cujo tema central tem como base na analogia das redes neurais biológicas. A RNA pode ser usada em robótica, reconhecimento de voz, processamento de fala, etc.

ALGORITMO GENÉTICO

Os algoritmos genéticos ajudam na resolução de problemas com a assistência de mais de um programa. O resultado seria baseado na seleção do mais apto.

REPRESENTAÇÃO DO CONHECIMENTO

É o campo de estudo com o qual podemos representar os fatos de uma maneira que a máquina é compreensível para a máquina. O conhecimento mais eficiente é representado; quanto mais o sistema seria inteligente.

APLICAÇÃO DE IA

Nesta seção, veremos os diferentes campos suportados pelo AI -

JOGOS

A IA desempenha papel crucial em jogos estratégicos como xadrez, pôquer, jogo da velha, etc., onde a máquina pode pensar em um grande número de posições possíveis com base no conhecimento heurístico.

PROCESSAMENTO DE LINGUAGEM NATURAL

É possível interagir com o computador que entende a linguagem natural falada pelos seres humanos.

SISTEMAS ESPECIALISTAS

Existem alguns aplicativos que integram máquina, software e informações especiais para transmitir raciocínio e aconselhamento. Eles fornecem explicações e conselhos aos usuários.

SISTEMAS DE VISÃO

Esses sistemas entendem, interpretam e compreendem a entrada visual no computador. Por exemplo:

- Um avião espião tirar fotografias, que são usadas para descobrir informações espaciais ou mapas das áreas.
- Os médicos usam o sistema especialista clínico para diagnosticar o paciente.
- A polícia usa um software de computador capaz de reconhecer o rosto do criminoso com o retrato armazenado feito pelo artista forense.

RECONHECIMENTO DE FALA

Alguns sistemas inteligentes são capazes de ouvir e compreender a linguagem em termos de sentenças e seus significados, enquanto um humano fala com ela. Ele pode lidar com diferentes sotaques, gírias, ruído em segundo plano, alteração no ruído humano devido ao frio etc.

RECONHECIMENTO DE CALIGRAFIA

O software de reconhecimento de manuscrito lê o texto escrito em papel por uma caneta ou na tela por uma caneta. Ele pode reconhecer as formas das letras e convertê-las em texto editável.

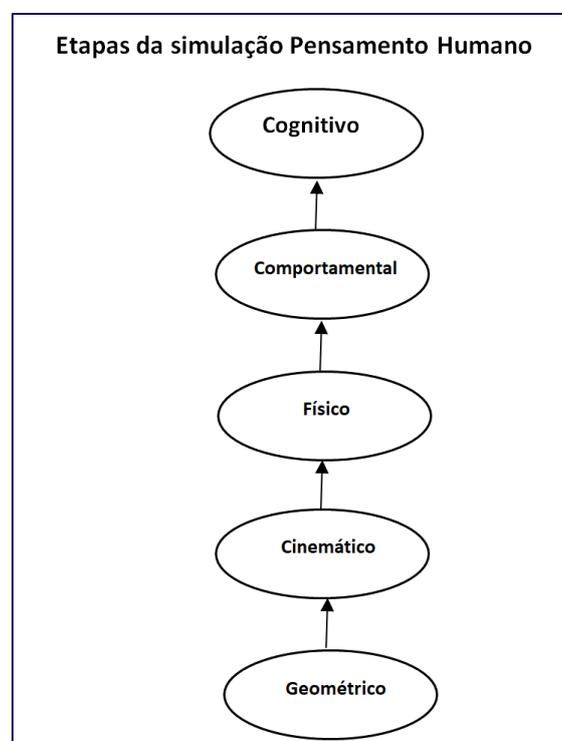
ROBÔS INTELIGENTES

Os robôs são capazes de executar as tarefas dadas por um ser humano. Eles possuem sensores para detectar dados físicos do mundo real, como luz, calor, temperatura, movimento, som, impacto e pressão. Eles possuem processadores eficientes, múltiplos sensores e memória enorme, para exibir inteligência. Além disso, eles são capazes de aprender com seus erros e podem se adaptar ao novo ambiente.

MODELAGEM COGNITIVA: SIMULANDO O PROCEDIMENTO DE PENSAMENTO HUMANO

A modelagem cognitiva é basicamente o campo de estudo da ciência da computação que lida com o estudo e simula o processo de pensamento dos seres humanos. A principal tarefa da IA é fazer a máquina pensar como os humanos. Pois, a característica mais importante do processo de pensamento humano é a solução de problemas. É por isso que a modelagem cognitiva procura entender como os humanos podem resolver os problemas por meio da construção de modelos que são usados para várias aplicações em IA, como aprendizado de máquina, robótica, processamento de linguagem natural etc.

A seguir, é apresentado o diagrama dos diferentes níveis de pensamento do cérebro humano:



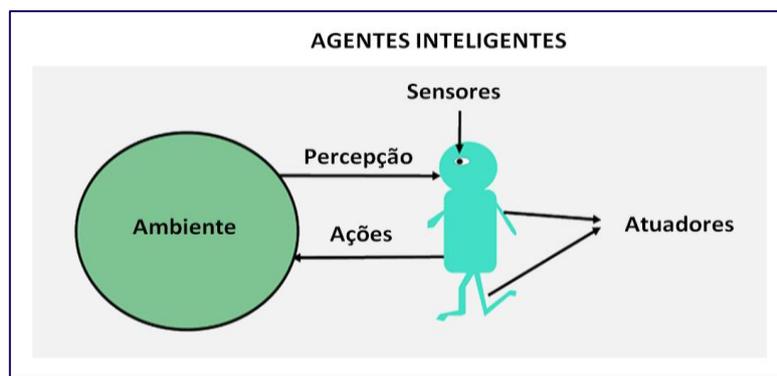
Fonte: Tutorials point, 2019

AGENTES EM INTELIGÊNCIA ARTIFICIAL

Um sistema de IA é caracterizado como o estudo do agente que percebe o ambiente com os sensores e age com os atuadores. Um agente de IA pode ter propriedades mentais, como conhecimento, crença, intenção etc.

O QUE É UM AGENTE?

Um agente pode ser qualquer coisa que perceba o ambiente por meio de sensores e atue nesse ambiente por meio de atuadores. Um agente corre no ciclo de perceber, pensar e agir conforme a seguinte ilustração.



Antes de avançar, devemos primeiro saber sobre sensores e atuadores.

- **Sensor:** é um dispositivo que detecta a mudança no ambiente e envia as informações para outros dispositivos eletrônicos. O agente observa seu ambiente através de sensores.
- **Atuadores:** são componentes de máquinas que convertem a energia em movimento. Os atuadores são responsáveis apenas por mover e controlar o sistema. Pode ser um motor elétrico, engrenagens, trilhos etc. Ou dispositivos que afetam o meio ambiente como pernas, rodas, braços, dedos, asas, nadadeiras e tela de exibição.

Um agente pode ser:

- **Agente humano:** Possui olhos, ouvidos e outros órgãos que trabalham para sensores como mãos, pernas, tato, boca para atuadores.
- **Agente robótico:** Pode ter câmeras, telêmetro infravermelho, PLN como sensores e vários motores como atuadores.
- **Agente de software:** Trata se de ações como pressionar as teclas, arquivar o conteúdo da entrada sensorial e atuar nessas entradas e exibir a saída na tela. Por outro lado, alguns agentes de software (para robôs ou softbots) ou simulador que atua num ambiente muito detalhado e complexo seja ela real e/ou artificial.

ESTRUTURA DO AGENTE DE IA

A tarefa da AI é projetar um programa que implemente a função do agente. A estrutura de um agente inteligente é uma combinação de arquitetura e programa de agentes.

Pode ser visto como:

$$\text{Agente} = \text{Arquitetura} + \text{Programa do Agente}$$

Onde,

Arquitetura: mecanismo que um agente de IA executa.

Função do agente: usada para mapear uma percepção para uma ação.

$$f: P^* \rightarrow A$$

Programa do agente: é uma implementação da função do agente e é executado na arquitetura física para produzir a função f .

MEIO AMBIENTE

Alguns programas operam em um ambiente totalmente artificial confinado aos dispositivos de entrada, como teclado, banco de dados e outros sistemas de arquivos de computador e saída de caracteres em uma tela.

REPRESENTAÇÃO PEAS

PEAS é um tipo de modelo no qual um agente de IA trabalha e podemos agrupar suas propriedades no modelo de representação composto por quatro palavras:

P (Performance): Desempenho

E (Environment): Ambiente

A (Actuator): Atuadores

S (Sensor): Sensores

Assim, o desempenho é o objetivo pelo qual se avalia o comportamento de um agente como representado na tabela a seguir:

AGENTE	MEDIDA DE DESEMPENHO	AMBIENTE	ATUADOR	SENSOR
1- Diagnósticos médicos	<ul style="list-style-type: none"> • Paciente sadio • Redução de custo 	<ul style="list-style-type: none"> • Paciente • Hospital • Staff 	<ul style="list-style-type: none"> • Teste • Tratamento 	<ul style="list-style-type: none"> • Teclado para entrada do sintoma
2- Aspirador de pó	<ul style="list-style-type: none"> • Limpeza • Eficiência • Vida da bateria • Segurança 	<ul style="list-style-type: none"> • Sala • Mesa • Piso de madeira • Carpet • Obstáculos diversos 	<ul style="list-style-type: none"> • Rodas • Escovas • Vácuo • Sugador 	<ul style="list-style-type: none"> • Câmera • Sensores de detecção de detritos • Sensor de degraus • Sensor de colisão • Sensor de parede Infravermelho
3- Robô de seleção de peças	<ul style="list-style-type: none"> • Porcentagem de peças em posições corretas. 	<ul style="list-style-type: none"> • Correia transportadora com peças • Caixas 	<ul style="list-style-type: none"> • Braços articulados • Mãos 	<ul style="list-style-type: none"> • Câmera • Sensores de ângulo comum.
4- Veículo autônomo	<ul style="list-style-type: none"> • Segurança • tempo • condução legal, • conforto 	<ul style="list-style-type: none"> • Estradas • Outros veículos • Sinais de trânsito • Pedestres 	<ul style="list-style-type: none"> • Direção • Acelerador • Freio • Sinal • Buzina 	<ul style="list-style-type: none"> • Câmera • GPS, • Velocímetro • Odômetro • Acelerômetro • Sonar.

TIPOS DE INTELIGÊNCIA ARTIFICIAL (IA)

O campo da IA permanece em grande parte inexplorada, o que significa que todos os notáveis programas de IA que vemos hoje representam apenas o início desta vasta área da ciência.

O nível em que o sistema de IA pode replicar as capacidades humanas é usado como critério para determinar os tipos de IA. Assim, a forma em que a máquina é comparada com os humanos no tocante à flexibilidade e funcionalidade, a IA pode ser categorizado entre os vários tipos.

Ou seja, a IA mais desenvolvida possa executar mais funções semelhantes aos humanos enquanto a IA com funcionalidade mais restritas é considerada o tipo menos evoluído.

Com base nesse padrão, a IA é categorizada em:

- **Tipo 1:** IA fraca ou estreita- é um sistema desenvolvido, programado e treinado para executar uma tarefa específica e funciona em um contexto limitado. Destinado a trabalhos físicos e cognitivos rotineiros. Baseia-se na semelhança com a mente humana na capacidade de "acreditar" e possivelmente "sentir", como indivíduos. Todo e qualquer cenário possível precisa ser previamente inserido manualmente assim permite identificar padrões e correlações de dados com mais eficiência do que os humanos.

Aplicado em previsões de vendas, sugestões de compra, perfil do cliente ou mecanismo de tradução do *Google*.

Cada IA fraca contribui na construção da IA forte como na indústria automotiva, nos carros autônomos, mecanismos de busca são resultados da coordenação de várias IA estreitas.

- **Tipo 2:** IA Geral ou Forte- possui habilidades cognitivas generalizadas que procura soluções de forma autônoma para uma tarefa desconhecida. Sistemas capazes de entender e raciocinar interagindo com o ambiente como um ser humano. Tem a vantagem de cálculos rápidos e em grande volume, podem pensar abstratamente, planejar e resolver problemas sem entrar em detalhes. Tem capacidade limitada em inovar, suscitar pensamentos e ideias sem um *script* prévio.
- **Tipo 3:** Super IA- Refere-se à posição em que computadores/máquinas superam os seres humanos e seriam capazes de imitar os pensamentos humanos e em determinada situação a capacidade cognitiva das máquinas será superior aos seres humanos.
No passado, houve desenvolvimentos como o supercomputador Watson da IBM derrotando enxadristas humanos ou em dispositivos auxiliares como o Siri da Google envolvendo-se em conversa com pessoas, mas ainda não existe uma máquina desenvolvida capaz de processar a profundidade do conhecimento e da capacidade cognitiva como a de um sistema totalmente humano.

COM BASE NAS FUNCIONALIDADES

1. **Máquinas reativas:** É a forma básica de IA com capacidades extremamente limitadas. Não resguarda e não utiliza memórias do passado para ações futuras. Simula de forma limitada a capacidade da mente humana pois não possuem funções baseadas na memória. Pois respondem a um conjunto restrito ou combinação de entradas. Um exemplo de uma máquina de IA é um sistema que venceu o Grandmaster Garry Kasparov em 1997, o Deep Blue da IBM.
2. **Capacidade de memorização limitada:** Sistemas de IA capaz de usar experiências passadas ou aprender com os dados históricos para tomada de decisões. No caso da IA de reconhecimento de imagem é treinado com milhares de rótulos e imagens identificar os objetos. Utilizam estas imagens como referências para compreender o conteúdo com base na "experiência de aprendizado" e rotula as imagens novas com precisão crescente.
Trata se dos *chatbots*, assistentes virtuais e veículos autônomos, entre outros.
3. **Teoria da mente:** capaz de entender as emoções, crenças, pensamentos, expectativas dos humanos para interagir socialmente. Embora existam muitas melhorias nesse campo se encontra em evolução.
4. **Autoconsciência:** IA superinteligente, autoconsciência e consciência (ser humano completo). Obviamente, esse tipo de *bot* ainda não existe e, se for alcançado, será um dos marcos no campo da IA.

- 5. IA Analítica:** Alimentada com os recursos do aprendizado da máquina (aprendizado profundo) analisa toneladas de dados para obter dependências e padrões e produzir recomendações, fornecer informações para tomada de decisões dos negócios orientada pelos dados. A análise de opinião e a avaliação de risco do fornecedor, otimização de inventário e previsão de demanda são aplicações da IA analítica.

Dados capturados pela IoT e armazenadas na nuvem pode detectar mau funcionamento de máquinas e acionar o comando para desligar o sistema. Robôs do Centro de Distribuição que conduz o operador da empilhadeira às prateleiras para identificar as mercadorias agilizando o processo de coleta.

- 6. IA interativa:** Melhoria nos processos internos com uso de *chatbots* e assistentes pessoais inteligentes para facilitar atividades corporativas como tomada de decisão e previsões dos resultados esperados.
- 7. IA de texto:** Visa potencializar os recursos de reconhecimento de texto, conversão de fala em texto, tradução automática e geração de conteúdo.

Substitui a pesquisa por palavras-chave e hipertextos, Tem como base a pesquisa semântica e ao processamento de linguagem natural, que permitem criar mapas semânticas e reconhecer sinônimos para entender o contexto da pergunta do usuário.

- 8. IA visual:** Permite identificar, reconhecer, classificar objetos ou converter imagens e vídeos para serem interpretados e/ou opinados posteriormente.

Pode auxiliar uma empresa seguradora a estimar os danos com base em fotos de objetos danificados ou máquinas. Abrange técnicas como a visão computacional, campos de realidade aumentada. Permite ainda reconhecimento facial ou inspeções visuais automatizadas que permitiam ao fabricante controlar imediatamente a qualidade dos detalhes na linha de produção.

- 9. Autoconsciência:** Provavelmente é o estágio final do desenvolvimento da IA, que agora existe apenas hipoteticamente. Também conhecida como autoexplicativa trata do estágio de evolução para assemelhar ao cérebro humano. Requer tempo de materialização, e será o objetivo supremo da pesquisa em IA. Provavelmente não será capaz de entender e evocar emoções com quem interage, porém possui sentimentos, necessidades, crenças e, possivelmente, desejos próprios.

Outra forma de classificação que é mais comumente utilizado é a classificação da Inteligência Artificial Estreita (IAE), Inteligência Artificial Geral (IAG) e a Superinteligência Artificial (SIA).

1. Inteligência Artificial Estreita (IAE)

Representa toda a IA disponível que foi criada até o momento. Essas máquinas se limitam a executar o que estão programadas e, portanto, possuem variedades estreitas ou muito restrita de competências de respostas. São sistemas de caráter geral que executam funções específicas e limitadas realizadas pelos humanos.

2. Inteligência Artificial Geral (IAG)

Trata-se da capacidade de aprender, perceber, entender e funcionar como os humanos. São sistemas com a capacidade de construir diversas competências de forma independente e estabelecer generalizações e conexões entre os domínios, reduzindo assim o tempo necessário para o treinamento. A replicação dos recursos multifuncionais dos humanos tornará os sistemas de IA tão capazes quanto as pessoas.

3. Superinteligência Artificial (SIA)

O seu desenvolvimento provavelmente indicará o auge da pesquisa em IA, já que a IAG está se tornando viáveis e aplicáveis no ambiente humano.

Para replicar a inteligência multifacetada dos humanos, o SIA requer capacidade maior da memória, investigação e processamento mais rápido de dados e a capacidade ampliada na tomada de decisão conhecida como de singularidade.

Para que o sistema seja treinado e que ser capaz de aplicar os conhecimentos adquiridos com as novas entradas baseia-se nos recursos da matemática, métodos quantitativos e algoritmos inteligentes, e é aplicável quando o sistema for treinado para executar estas tarefas.

Davenport e Ronanki (2018) destacam três importantes necessidades corporativas: automatizar processos de negócios, obter informações com a análise de dados e do envolvimento com clientes, fornecedores e funcionários para adquirir habilidades e capacidades cognitivas para alcançar seus objetivos de negócios.

- 1. Automação do processo.** Trata da automação das tarefas físicas repetitivas desenvolvidos pelos humanos que são substituídas pelas tecnologias de automação sobretudo os RPAs (*Robotic Process Automation*), onde "robôs" agem como ser humano e processam informações de vários sistemas de TI.

Estas tarefas incluem:

- sistemas de transferência e registro de dados de sistemas de *e-mail* e *call center* em processos para atualizar os arquivos de clientes com alterações de endereço ou adições de serviços;
- pedidos de substituição ou perdas de cartões de crédito via ATM envolvem vários sistemas para atualizar registros e manipular comunicações com clientes;
- reconciliar falhas na cobrança de serviços extraíndo informações de diferentes tipos de documentos; e
- leitura de documentos legais e contratuais para extrair as disposições com processamento de linguagem natural.

O RPA é uma das tecnologias cognitivas de menor custo e mais adequado na implementação e normalmente com retorno rápido do investimento. São IAs de caráter geral que não estão programados para aprender e melhorar, embora os desenvolvedores estejam adicionando inteligência e capacidade de aprendizado.

2. Visão cognitiva. Utiliza algoritmos para detectar padrões, interpretar o significado nos grandes volumes de dados. São aplicações para:

- prever a compra provável de um cliente
- identificar ou detectar reclamação fraudes de crédito em tempo real;
- identificar e analisar dados segurança ou problemas de garantia da qualidade de produtos e serviços;
- automatizar a segmentação personalizada de anúncios digitais;
- fornece informações mais precisas e modelagem atuarial detalhadas às seguradoras.
- análises futuras das Informações cognitivas quando: os dados são muito mais intensos e detalhados; modelos são treinados para melhorar o aprendizado; capacidade de usar dados novos dados para fazer previsões ou ajustes para aplicações em outras uma categoria amis precisa.

Versões de aprendizado de máquina (aprendizado profundo, que tenta imitar a atividade do cérebro humano para reconhecer padrões) como reconhecer imagens e da fala.

Os aplicativos de *insight* cognitivo geralmente são usados para melhorar o desempenho em trabalhos que apenas máquinas podem fazer - tarefas como a compra programática de anúncios que

envolvem trituração e automação de dados em alta velocidade que há muito que estão além da capacidade humana - então eles estão geralmente não é uma ameaça para empregos humanos.

Engajamento cognitivo. Projetos que envolvem funcionários e clientes que usam processamento de linguagem natural como *chatbots*, agentes inteligentes e máquinas que aprendem cuja categoria incluem:

- agentes inteligentes que oferecem atendimento ao cliente abordando uma ampla e crescente variedade de questões solicitações de senha para perguntas de suporte técnico - tudo na linguagem natural do cliente; sites internos para responder perguntas de funcionários sobre tópicos incluindo TI, benefícios dos funcionários e política de RH;
- sistemas de recomendação de produtos e serviços para varejistas que aumentam a personalização, o engajamento, e vendas - normalmente incluindo idioma avançado ou imagens; e
- sistemas de recomendação de tratamento de saúde que ajudar os provedores a criar planos de atendimento personalizados que levar em consideração o estado de saúde de cada paciente e tratamentos anteriores.

METODOLOGIAS DE IA

O termo inteligência artificial (IA) refere-se a sistemas de computação que executam tarefas que normalmente são consideradas no âmbito da tomada de decisão humana. Esses sistemas controlados por software e agentes inteligentes incorporam análises avançadas de dados e aplicativos de Big Data. Os sistemas de IA utilizam esse repositório de conhecimento para tomar decisões e tomar ações que se aproximam das funções cognitivas, incluindo aprendizado e solução de problemas.

Existem várias abordagens usadas para desenvolver e construir sistemas de IA que incluem:

Aprendizado de Máquina (ML). Ramo da IA usa métodos e algoritmos estatísticos para descobrir padrões e "treinar" sistemas para fazer previsões ou decisões sem programação explícita. São formados por ML supervisionado, semi-supervisionado (inclui classificações e rotulagens) e não supervisionado (entradas de dados sem rótulos aplicados por humanos).

Aprendizado Profundo (DL). Abordagem baseada em redes neurais artificiais (RNAs) para aproximar as vias neurais do cérebro humano. Os sistemas de aprendizado profundo são particularmente valiosos para o desenvolvimento da visão computacional, reconhecimento de fala, tradução automática, filtragem de redes sociais, videogames e diagnóstico médico.

Redes Bayesianas. Esses sistemas dependem de modelos gráficos probabilísticos que usam variáveis aleatórias e independência condicional para entender e agir melhor nas relações entre as coisas, como uma droga e efeitos colaterais ou escuridão e um interruptor de luz aceso.

Algoritmos genéticos. Utilizam abordagem heurística modelada pela seleção natural por meio da mutação e técnicas de cruzamento para resolver desafios biológicos complexos e demais questões associadas.

A Inteligência Artificial (IA) preocupa-se em entender a natureza da inteligência humana e projetar artefatos inteligentes que podem executar as tarefas que, quando executadas por seres humanos, dizem exigir inteligência. Qualquer grande avanço na tecnologia traz consigo uma variedade de oportunidades e desafios.

Os aplicativos em IA podem ser classificados em quatro ondas, que estão acontecendo simultaneamente, mas com diferentes pontos de partida e velocidade:

A PRIMEIRA ONDA: CONHECIMENTO ARTESANAL

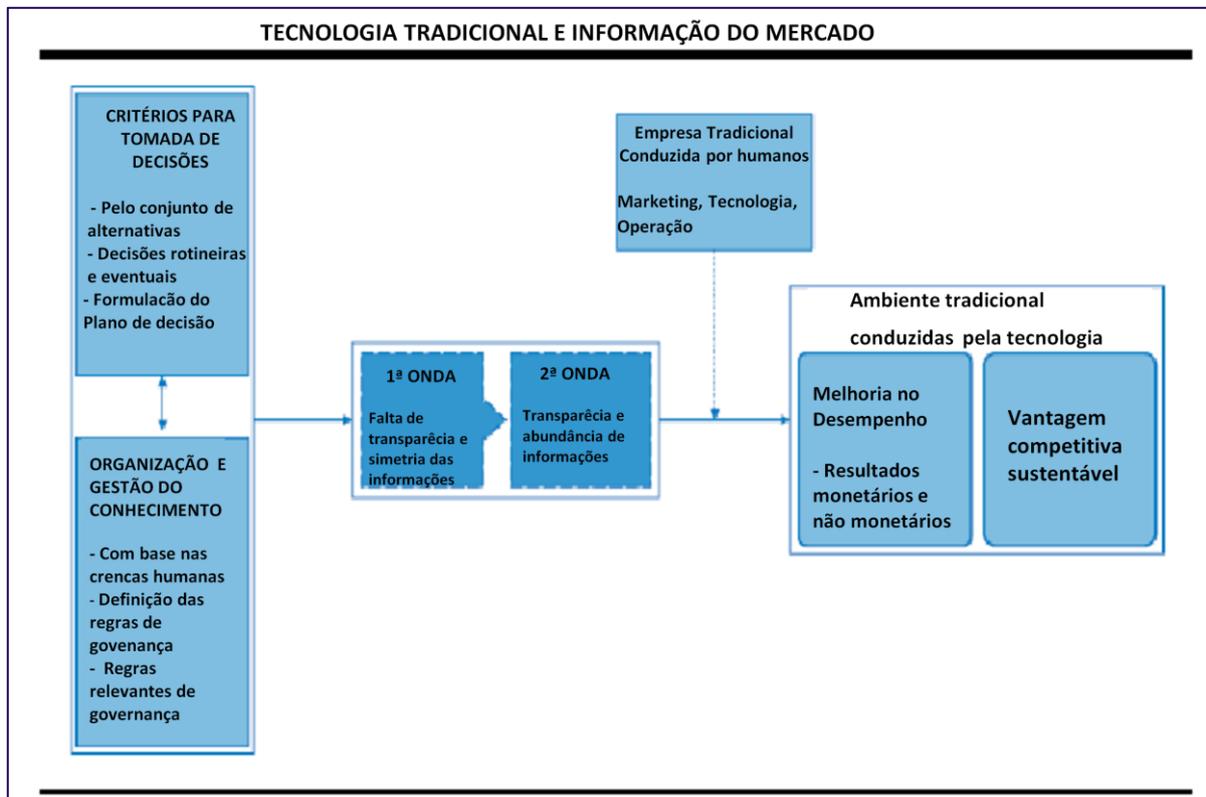
Conhecida como a "*Internet IA*" potencializada pela enorme quantidade de dados que fluem pela web. Aproveita o fato de que os usuários rotulam automaticamente os dados enquanto navegam na web: as ações de comprar vs não comprar; clicar vs não clicar.

Estas cascatas de dados rotulados criam um perfil detalhado das personalidades, hábitos, demandas e desejos dos navegadores: trata-se da receita perfeita para estruturar conteúdo personalizado para manter conectado numa determinada plataforma ou para maximizar receita ou lucro de um provedor de conteúdo.

Trata-se de sistemas especializados com base em regras que representam conhecimentos humanos disponíveis em domínios bem definidos. Observa-se neste contexto a limitação de capacidade dinâmica do aprendizado e do conhecimento tácito – as habilidades, ideias e experiências que não são codificados ou facilmente expressos.

Nos sistemas de diagnósticos para apoio aos médicos, as pesquisas indicaram que as informações limitadas no diagnóstico final necessitam ser complementadas com o conhecimento pessoal dos médicos que inclui conhecimentos explícitos e tácitos (Groopman 2007). De fato, os médicos tomam decisões usando a combinação dos princípios gerais dos protocolos clínicos aceitos associando com base em experiências e raciocínios pessoais para confirmar ou não confirmar os diagnósticos incluindo os sintomas dos pacientes. Assim trabalham com conhecimentos explícitos, mas seu conhecimento tácito é fundamental na tomada de decisão do diagnóstico. Desta forma, os sistemas especialistas tinham lacunas em seus conhecimentos e capacidades da primeira versão que não tinham capacidade de aprender com conhecimento tácito e que não pôde ser expresso com regras definidas.

A figura a seguir apresenta as informações que conduzem a resultados com os recursos gerando recomendações gerais conforme as configurações de produtos e de serviços.



Fonte: Adaptado de Kumar et al, 2019

A SEGUNDA ONDA: APRENDIZADO DE MÁQUINA- IA DE NEGÓCIOS.

Os processos nesta onda compreendem os conhecimentos específicos que aprendem com as informações dos dados que são alimentadas com uma variedade de dispositivos, sensores, coletores e instrumentos que constroem o conhecimento tácito. O aprendizado profundo, que fornece subsídios para construir e treinar redes neurais na tomada de decisões que é a marca registrada desta onda. Fatos importantes impulsionam os avanços a segunda onda: o crescente surgimento de computadores e coletores com baixo consumo de energia e a big data em alta escala e disponibilidade. Pois, o aprendizado profundo requer volume grande de dados que apoia o modelo de treinamento com base no aprendizado supervisionado. Por outro lado, à medida em que a segunda onda avança, o progresso e a evolução do aprendizado não supervisionado e a aprendizagem por reforço se tornará mais central.

Os algoritmos são treinados de forma que o conjunto de dados proprietários que variam conforme o comportamento de compra dos clientes e conduzir os gerentes a melhor tomada de decisão. Ou pode estudar milhares empréstimos bancários e taxas de reembolso para saber se um tipo de mutuário é um potencial risco oculto por inadimplência ou, surpreendentemente um bom potencial para tomada de novos empréstimos, mas que estava esquecido pela administração.

Da mesma forma, os médicos podem usar os algoritmos de aprendizado profundo para digerir enormes quantidades de dados sobre o diagnóstico do paciente, perfis genômicas, terapias resultantes e consequentes resultadas de saúde e talvez descubra um protocolo de tratamento personalizado digno. Pela exploração de correlações ocultas que escapam da nossa lógica linear de causa e efeito, os recursos da IA pode superar até o mais experiente dos especialistas.

A segunda onda inclui também a computação perceptiva - a capacidade do computador reconhecer o que está presente ao seu redor. A tecnologia perceptiva está cada vez mais sendo usada em serviços de segurança e nas indústrias nas aplicações de cognição visual ou visão computacional. Muitas indústrias – na assistência médica, petróleo e gás, serviços públicos, agricultura, na área de seguros, a visão computacional com dispositivos assistidos por aprendizes para diagnosticar pacientes, operações de pesquisa, monitorar plataformas de petróleo ou ativos segurados, fornecendo acesso rápido a informações críticas com maior precisão e menos risco para a vida humana.

A segunda onda impulsionou avanços no campo dos negócios e a sociedade no apoio para encontrar informações, encomendar produtos, tocar música, realizar traduções e interpretações de idiomas do cotidiano.

Reconhecimento de imagem, processamento de linguagem natural, segurança cibernética, plataformas autônomas também são resultadas da segunda onda. Trouxe também os recursos diferenciados de classificação e de previsão, porém ainda sem a capacidade contextual ou a capacidade mínima de raciocínio para completar e apoiar a inteligência completa.

TERCEIRA ONDA

Chamada de “IA de percepção” proporciona um novo recurso à visão, audição e outros sensores para coleta de novos dados que permite novos usos e criar novos aplicativos.

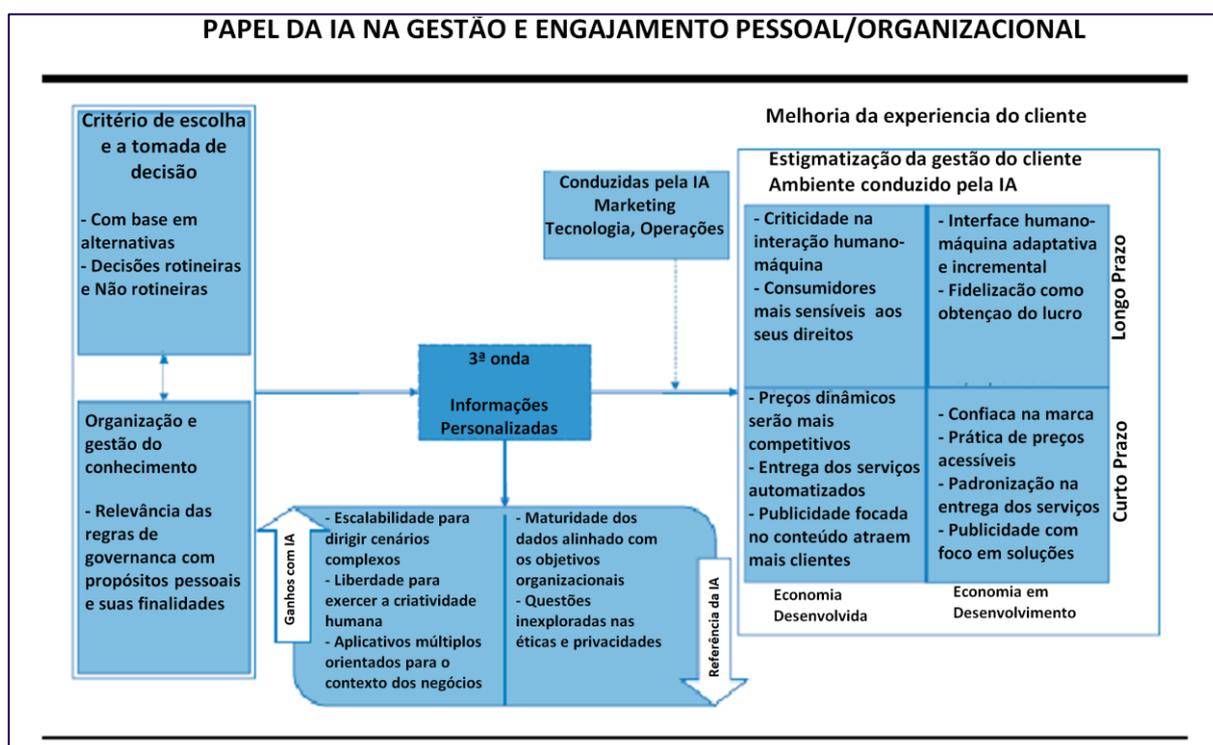
Com a proliferação de dispositivos inteligentes nas residências, zona urbana e rural que inclui interfaces de fala, da visão computacional, reconhecimento facial até na inspeção de qualidade nas fábricas.

Os seres humanos usam uma variedade de ferramentas para se adaptar incluindo regras e modelos, bem como os valores como a ética e moral que trazemos numa situação. Assim, ensinar um tigre pode levar milhares ou milhões de repetições, mas uma criança que vê um tigre uma vez e na próxima diz: *"isso é um tigre!"*, ela conseguiu o aprendizado, pois não precisamos ensinar repetidamente. Trata se de aprendizado único que recorda as instâncias do passado e aplicar posteriormente numa situação diferente.

O campo da sinergia cognitiva em que os sistemas com a capacidade de reconhecer por semelhança um novo episódio com situações do passado trata se da adaptação contextual.

A terceira onda começará para abordar esses componentes importantes da verdadeira IA onde:

1. constroem modelos explicativos para classes ou fenômenos do mundo real,
2. aprendem e raciocinam à medida que encontram novas tarefas e situações, e
3. permitem a comunicação natural entre máquinas e pessoas. É ativado com a computação em nuvem; combina o pensamento completo sobre o gerenciamento de dados, o processamento da linguagem natural, a representação de dados e o conhecimento; lojas virtuais com ferramentas analíticas robustas de aprendizado de máquina; e diálogo humano/máquina por meio de plataformas de comunicação robustas.

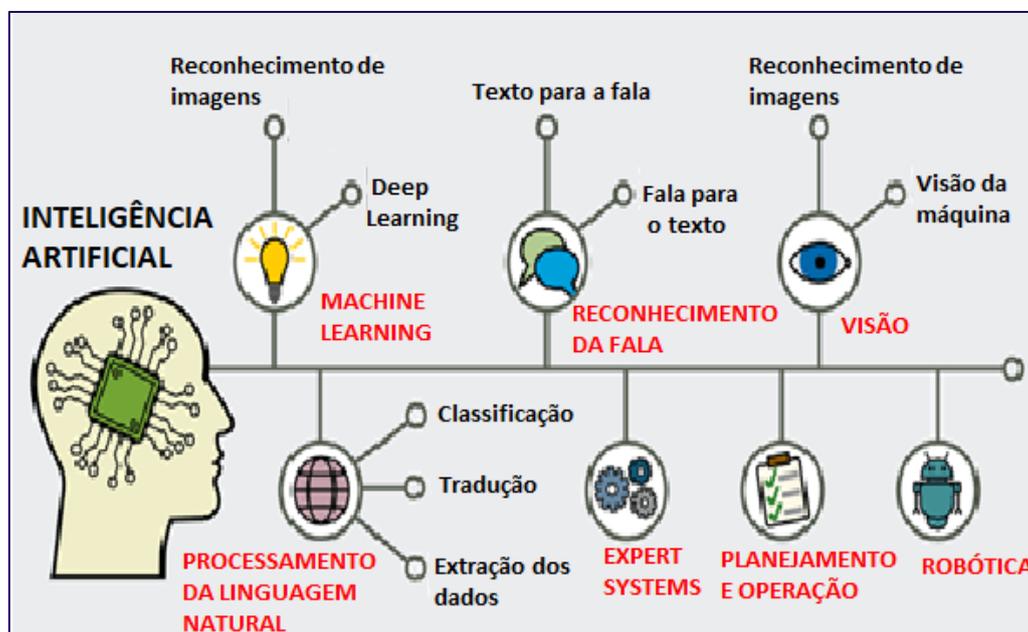


Fonte: Adaptado de Kumar et al, 2019

Trata-se da representação mais próxima da inteligência completa com a tomada de decisão de ponta a ponta que incluem sistemas de aprendizagem e raciocínio contextualmente adaptáveis; sistemas que aprendem *online* e *offline* a partir dos dados complexos e dados esparsos; interação humano máquina e máquina a máquina com capacidades e funções para ser dirigido, de forma colaborativa e delegatório.

A quarta onda é a mais ampla, mas também de maior desafio: a "IA autônoma" integrando todas as ondas anteriores IA autônoma permite às máquinas a capacidade de sentir e responder para o ambiente ao seu redor, para se mover intuitivamente e manipular os objetos com facilidade quanto um ser humano. São incluídos nesta onda os veículos autônomos que podem "ver" o meio ambiente

à sua volta, reconhecendo os padrões de *pixels* da câmera, descobrir as cores com o que eles se correlacionam (sinais de parada) entre outros; e depois usar essas informações para tomar decisões (aplicando pressão no freio para parar lentamente o veículo).



Na área da robótica, essa etapa avançada os algoritmos industriais da IA serão aplicadas em linhas de montagem automatizadas e operações de armazéns, em tarefas comerciais (lavagem de louça e colheita de frutas) e também eventualmente nas aplicações nas áreas de consumo em geral como lojas, bares e padarias.

A IA pode ser programado para maximizar a lucratividade ou substituir o trabalho humano, e desta forma agrega valor imediato à economia, pois é rápido, preciso, funcionam 24 horas por dia, e não reclama,

Desta forma, a contribuição da IA em primeiro e mais importante é o deslocamento do trabalho: Desde que a IA pode executar tarefas únicas com precisão sobre a humana, pois a maioria dos trabalhos humanos é de tarefa única e de rotina será substituído por esta tecnologia de última geração. Isso inclui ambos os trabalhos de colarinho branco e colarinho azul. A IA também será questionado nos quesitos de segurança, privacidade, viés de dados e manutenção de monopólio. Todos são problemas significativos sem solução conhecida, para que governos e as empresas devem voltar atenção nestas questões.

Mas uma preocupação que ainda não temos que enfrentar é a que pode ser mais comum hoje em dia, expressada na imagem dos filmes de ficção científica em que as máquinas atingirão a

verdadeira inteligência em nível humano (ou mesmo de nível sobre-humano) tornando-os capazes de presumivelmente de ameaçar a humanidade.

O aumento acentuado atual no número de aplicativos de IA deve-se a vários fatores.

Em primeiro lugar, o século XXI trouxe aumento considerável na capacidade de processamento e de armazenamento dos computadores enquanto que o custo vem diminuindo devido aos avanços das tecnologias dos circuitos eletrônicos usados em sistemas digitais complexos tornando assim a computação cada vez mais acessível. São os principais requisitos para o crescimento das aplicações da IA complementadas pela tecnologia da nuvem e o aprendizado das máquinas.

O segundo fator importante são os avanços nas técnicas de IA, como redes neurais profundas e a sofisticação dos algoritmos aumentou substancialmente e causou um impacto significativo nas aplicações práticas da IA.

Outra área da IA em que houve progresso significativo no processamento das linguagens naturais usado na tradução automática, na interface do usuário, os chatbots que são utilizados para tarefas como responder às perguntas dos clientes ou realizam tarefas em nome dos usuários que são conhecidos como assistentes virtuais.

O terceiro fator é a geração, armazenamento, transmissão e processamento de alto volume, em alta velocidade e alta variedade de dados devido aos avanços tecnológicos incluindo Internet das Coisas (IoT), sensores de baixo custo, redes de comunicação de alta velocidade e dispositivos móveis.

O alto volume de dados pode fornecer vários tipos de informações, que não foram possíveis até então.

Em geral a IA requer recursos avançados como raciocínio, aprendizado conceitual, senso comum, planejamento, pensamento multidisciplinar, criatividade e até autoconsciência e emoções que permanecem além do nosso alcance. Não existem caminhos de engenharia para evoluir em direção a estes recursos gerais. A que distância estamos da IA geral?

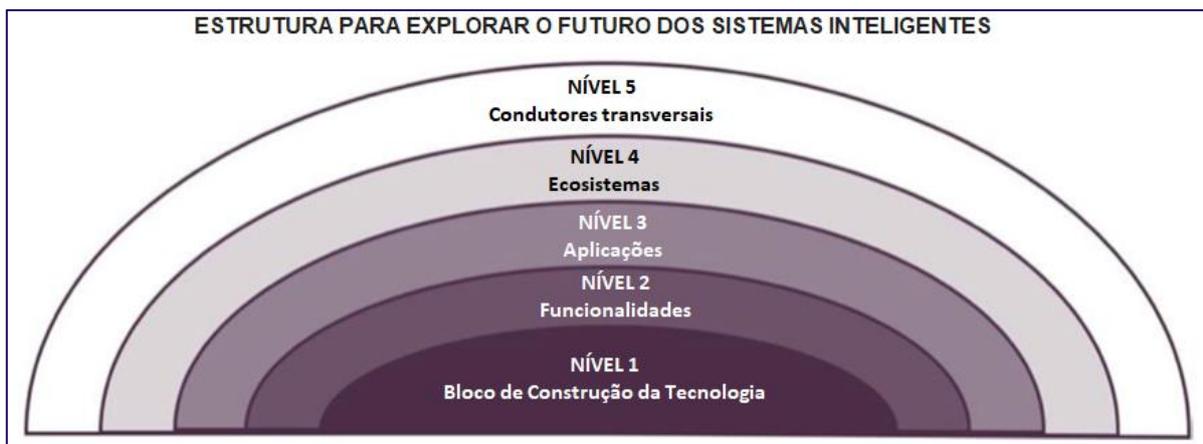
Requer dezenas de grandes avanços para chegar nestas etapas que pode trazer uma revolução tecnológica de magnitude transcendental que está acontecendo muito mais rápido.

Cabe a nós entender seu impacto monumental, benefícios generalizados e sérios desafios.

SISTEMAS INTELIGENTES

Um sistema inteligente é uma máquina ou conjunto conectado de máquinas tecnologicamente avançadas que percebem e respondem ao mundo físico ou digital, de aprender com a experiência e apoiar pessoas ou concluir tarefas que podem variar do simples e repetitivo ao muito difícil e complexo. Os sistemas inteligentes podem também incluir tecnologias como a de aprendizado ou da inteligência artificial juntamente com os sensores avançados, a robótica entre outros.

Em face das dimensões, a complexidade e a velocidade da evolução do domínio dos sistemas inteligentes, foram propostas um mecanismo estruturante para orientar a pesquisa e facilitar a coleta e síntese dos sinais, tendências e fatores relevantes detas mudanças. Estão organizadas a partir daa partes da micro-visão do componente do sistema inteligente para a visão macro de drivers globais moldar os sistemas inteligentes. Tem a finalidade orientativa de examinar os sinais e tendências em cada uma dessas múltiplas escalas porque os principais fatores de mudança podem ser encontrados dentro e em todas essas escalas conforme a figura a seguir:



Fonte: Culver et. AI, 2019

- 1. Blocos de construção da tecnologia** - Exploração dentro do nível da estrutura que procura entender o componente e partes de vários sistemas inteligentes, como sensores, algoritmos e dados. Ao explorar sinais e tendências dos componentes da tecnologia, pode se examinar melhor o que pode ser possível no futuro próximo como os avanços dos componentes se expandem os recursos do sistema.
- 2. Funcionalidades** - os componentes básicos da tecnologia podem ser combinados para criar poderosas funcionalidades polivalentes incluindo a visão de máquina e o processamento da linguagem natural.
- 3. Aplicativos e casos de uso** - As manchetes das notícias estão repletas de histórias das empresas que aplicam IA a diferentes segmentos das indústrias e para várias aplicações.

A exploração dentro desse nível procura entender como os desenvolvedores de tecnologia e as empresas adotam sistemas inteligentes para realizar tarefas específicas. O objetivo é identificar e entender a situação atual e inferir as possíveis realizações futuras.

- 4. O Ecosistema** - Exploração dentro deste nível da estrutura busca entender os fatores que moldam as atividades das organizações em todo o sistema inteligente do ecossistema mais

amplo. Esses fatores incluem grandes investimentos, importantes parcerias, modelos de negócios emergentes, disponibilidade de talentos, ritmo e coordenação do desenvolvimento em setores relevantes.

- 5. Condutores proeminentes** - Este nível da estrutura procura entender as tendências e os fatores transversais que moldam o mundo em que os sistemas inteligentes são residentes.

A análise nesse nível baseia-se nos aspectos sociais, tecnológicos, estrutura econômica, ética e política que são ferramentas usadas em exercícios de prospecção estratégica para explorar sistematicamente as tendências e fatores que moldarão um determinado domínio (Bishop e Hines 2007, p. 88). Também considera se os fatores ambientais nesse nível da estrutura.

Considerando os influenciadores de alto nível nos permitiu pensar mais profundamente sobre as forças em jogo e como elas moldam o curso dos outros níveis no futuro.

Tendo estabelecido uma estrutura para o trabalho, a equipe então passou a digitalizar e pesquisar em cada nível da estrutura para desenvolver uma compreensão mais profunda dos

problemas. Permite que se examine mais extenso do conjunto de fontes secundárias, uso de métodos analíticos simples para considerar a importância relativa e influência de diferentes tendências para identificar e examinar diferentes fatores de mudança que, de forma direta e combinada, molda o futuro dos sistemas inteligentes que foram condensados e combinados na figura anterior. Desta forma, o mapeamento da estrutura de orientação para os principais fatores de mudança, fatores de mudança relacionados aos níveis 1, 2 e 3 são capturados em Tecnologia, o nível 4 é capturado no ecossistema e o restante do mapa no nível 5.

IA como disciplina científica

A descrição do tópico anterior trata da descrição abstrata muito simples de um sistema de IA por meio de três recursos principais: percepção, raciocínio/tomada de decisão e atuação que formam as capacidades de raciocínio e de aprendizado. A robótica é outra disciplina muito relevante.

Raciocínio e tomada de decisão. Esse grupo de técnicas inclui representação e raciocínio, planejamento, programação, pesquisa e otimização do conhecimento. Essas técnicas permitem executar o raciocínio sobre os dados provenientes dos sensores. Para poder fazer isso, é preciso transformar dados em conhecimento; portanto, uma área da IA tem a ver com a melhor forma de modelar esse conhecimento (representação do conhecimento). Uma vez modelado o conhecimento, o próximo passo é argumentar com ele (raciocínio do conhecimento), o que inclui fazer inferências por meio de regras simbólicas, atividades de planejamento e programação, pesquisar um grande conjunto de soluções e otimizar entre todas as soluções possíveis para um problema. O passo final é

decidir que ação tomar. O raciocínio/tomada de decisão em um sistema de IA geralmente é muito complexo e requer uma combinação de várias das técnicas mencionadas.

Aprendizado. Esse grupo de técnicas inclui aprendizado de máquina, redes neurais, aprendizado profundo, árvores de decisão e muitas outras técnicas de aprendizado. Essas técnicas permitem que um sistema de IA aprenda a resolver problemas que não podem ser especificados com precisão ou cujo método de solução não pode ser descrito por regras de raciocínio simbólico.

Como aplicação associados aos recursos de percepção, como compreensão de fala e linguagem, a visão computacional ou previsão de comportamento. Observe que esses problemas são aparentemente fáceis para os humanos. No entanto, não são tão fáceis para os sistemas de IA, pois não podem confiar no raciocínio de senso comum (por enquanto) e são especialmente difíceis quando o sistema precisa interpretar dados não estruturados. É aqui que as técnicas que seguem a abordagem de aprendizado de máquina são úteis. No entanto, as técnicas de aprendizado de máquina podem ser usadas para muito mais tarefas do que apenas a percepção. As técnicas de aprendizado de máquina produzem um modelo numérico (ou seja, uma fórmula matemática) usada para calcular a decisão a partir dos dados que serão abordados ao longo dos capítulos seguintes.

Onde e como a personalização e a IA convergem?

Em termos de marketing, a personalização é frequentemente apresentada e é estudada juntamente com a personalização, pois estão relacionados no conceito, mas diferem no aplicativo.

A personalização ocorre quando a empresa decide, geralmente com base em informações coletadas anteriormente; dados do cliente, qual mix de marketing é adequado para o indivíduo, enquanto a personalização ocorre quando o cliente especifica proativamente um ou mais elementos de seu mix de marketing. Uma boa quantidade de consenso prevalece em entender a personalização como um processo amplamente controlado pela empresa que é alimentado usando dados e personalização no nível do cliente como uma decisão amplamente decidida pelo cliente processo focado no design e na entrega da oferta. A personalização demonstrou funcionar em ambientes digitais e não digitais. No domínio dos serviços, chamadas serviços inteligentes de roteamento também oferecem personalização combinando clientes com serviços representantes com habilidades e personalidade apropriadas.

A personalização como um processo interliga clientes e profissionais de marketing⁹ e solidifica o relacionamento entre eles. Os relacionamentos com os clientes que também têm emoções vincular o progresso a um estado de engajamento e que as relações positivas desempenham um papel na influência do comportamento do envolvimento do cliente (EC) .

O engajamento foi definido como a atitude, o comportamento, o nível de conexão entre clientes; entre clientes e funcionários; e de clientes e funcionários de uma empresa. Além disso, quanto mais positiva a atitude e compreendendo o papel da inteligência artificial no marketing de envolvimento personalizado comportamento, e quanto maior o nível de conexão, maior o nível de noivado.

O alto grau de personalização na IA é considerado um fator importante por trás de sua popularidade. A IA mudou o paradigma de sistemas especialistas baseados em regras abordagem a uma abordagem baseada em aprendizado profundo e orientada por dados. A IA é tão discreta que os usuários geralmente não sabem que interagiram com a tecnologia.

Quando a tecnologia funciona em nível pessoal, cria um vínculo com os usuários. Além disso, quando os profissionais de marketing utilizam esse vínculo, o potencial de criação de valor para o cliente é enorme. No entanto, o sucesso da personalização iniciativas é restringido pelo volume e qualidade do cliente em formação; a capacidade das empresas de gerar insights a partir dos dados dos clientes; e a implementação eficaz de *insights*. Para superar essas três restrições e Além do nível atual de ofertas personalizadas, as empresas agora estão recorrendo para soluções com inteligência artificial.

Estas ferramentas podemos ser capazes de aplicar a experiência de um a muitos como sistemas do conhecimento que se tornam mais inteligentes a cada experiência. Um dos propósitos corporativos que se pode obter é a análise individualizada e centrada no cliente para aprimorar na tomada de decisão.

A capacidade de ver conexões esparsas entre pessoas, lugares e coisas num ambiente de enormes quantidades dados heterogêneos oriundos de diversas fontes, e entender o seu contexto torna se viável no contexto da IA.

Ao mesmo tempo, o progresso nas capacidades perceptivas como reconhecimento de imagem e da fala continua a avançar com novos métodos de observação, senso de criação e de previsão entre aplicativos que surgem e se aprimoram a cada instante.

Desta forma, a IA tornou-se uma ferramenta valiosa e essencial para orquestrar tecnologias digitais e gerenciar operações de negócios e atitudes pessoais.

Hoje, a robótica, o reconhecimento de imagem, o processamento de linguagem natural, as ferramentas de análise em tempo real e vários sistemas conectados na Internet das Coisas (IoT) utilizam a IA para oferecer recursos e capacidades mais avançados.

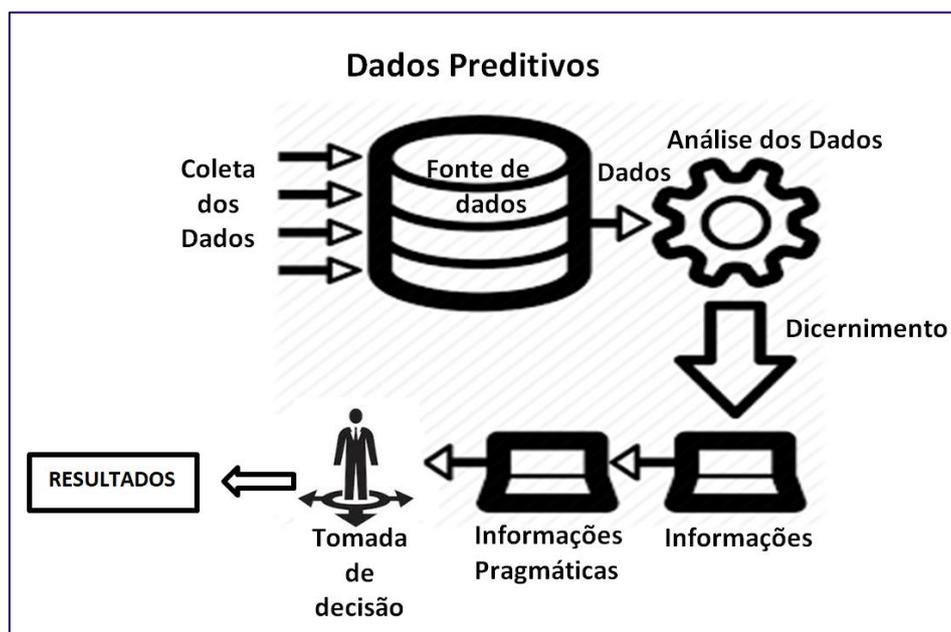
Assim, a IA está incorporado um número crescente de aplicativos e ferramentas. Desde programas de análise corporativa e assistentes digitais como Siri e Alexa a veículos autônomos e reconhecimento facial.

APRENDIZADO DE MÁQUINA (Machine Learning)

INTRODUÇÃO

A proliferação de dados e os avanços na velocidade de processamento e do aumento da capacidade de armazenamento de dados e, sobretudo pelo aumento do uso da Internet tornou o aprendizado das máquinas um componente significativo dos pesquisadores e gestores em negócios.

Modernas organizações coletam volumes massivos de dados e para que tenham valor devem ser analisadas e extraídas os discernimentos que proporcionem uma boa tomada de decisão ilustrada na figura a seguir:



Fonte: Adaptado de Kelleher et al, 2018

O aprendizado de máquina é um subconjunto essencial da inteligência artificial (IA) em que máquinas são instruídas a aprender de forma semelhante aos humanos. Enquanto os humanos estão na fase inicial dos estudos, a compreender seus recursos extremamente dinâmicos, o conceito existe há décadas.

Exemplos típicos de aplicação do aprendizado de máquina são os filtros de spam dos e-mails, o software aplicativo de análise e concessão de créditos financeiros, o *feed* de notícias, as publicidades personalizadas e direcionadas pelas empresas de tecnologia, entre outros.

Você já foi fazer compras no mercado? Normalmente parte da lista prévia de itens com base nas necessidades e experiências anteriores e um orçamento orientativo.

Com o aumento da inflação e outros fatores do mercado, o lojista pode alterar o preço e os itens ofertados com certa frequência. Fatos estes que requer ajustar a lista de compras e esforços adicionais para novas pesquisas e de tempo para nova atualização. É aqui que o Machine Learning pode ajudá-lo.

APRENDIZADO DA MÁQUINA

O aprendizado de máquina (ML) é o processo de usar modelos matemáticos para tratamento de dados para ajudar o computador a aprender sem instruções diretas. O aprendizado de máquina usa algoritmos para identificar padrões dentro dos dados que são usados para criar um modelo para fazer previsões. Com o aumento do volume de dados e a experiência, os resultados do aprendizado de máquina são mais precisos - assim como os humanos tendem a melhorar suas intuições com o acúmulo da prática.

A adaptabilidade do aprendizado de máquina o torna uma ótima opção em cenários em que os dados estão sempre mudando, a natureza da solicitação ou tarefa está sempre mudando ou a codificação de uma solução seria efetivamente impossível.

Os recursos do Machine Learning revolucionaram a maneira como percebemos, classificamos e estruturamos as informações e fazer vários *insights* que permita prever os fatos futuros. Trata-se de um processo automatizado que extrai padrões a partir dos dados para a construção de modelos analíticos que permite entendimentos e fazer previsões.

As máquinas são programadas para aprender de forma automatizada e interativa a partir dos dados analisados tornando o processo de aprendizado de forma contínua; como a máquina é exposta a volume crescentes de dados requer padrões robustos e reconhecidos para serem ajustados de forma que suas ações sejam automatizadas.

Assim, o processo de aprendizado começa com observações, experiência direta ou instrução e exemplos na formação de padrões dos dados que permita a melhor tomada de decisões no futuro. O objetivo principal é permitir que os computadores aprendam automaticamente sem intervenção ou assistência humana e que se auto ajustem as ações futuras.

O aprendizado de máquina e a tradicional análise estatística são semelhantes em muitos aspectos, mas diferentes na sua execução. Embora a análise estatística seja baseada na teoria e distribuições de probabilidade, o aprendizado de máquina é projetado para que encontre a combinação das equações matemáticas que melhor preveja um resultado.

Portanto, aprendizado de máquina é adequado para análise de uma ampla gama de problemas que envolvem a classificação, regressão linear e análise de agrupamentos com grandes quantidades de dados.

O QUE É APRENDIZADO DE MÁQUINA?

Pode ser definido como o campo da ciência da computação, mais especificamente uma aplicação de inteligência artificial, que fornece aos sistemas de computação a capacidade de aprender com os dados e melhorar a experiência sem serem explicitamente programados.

O aprendizado de máquina ("ML") é o estudo científico de algoritmos e modelos estatísticos que os sistemas de computador usam para executar uma tarefa específica sem usar instruções explícitas, baseando-se em padrões e inferência derivados de dados.

ML foi cunhado pela primeira vez em 1959 por Arthur Samuel, pioneiro em jogos de computador e IA.

Mitchell (1977), outro cientista da computação, mais tarde forneceu uma definição formal amplamente citada de ML:

“Diz-se que um programa de computador aprende com a experiência E com relação a alguma classe de tarefas T e com a medida de desempenho P , se seu desempenho nas tarefas em T , medido por P , melhorar com a experiência E ”.

Em termos práticos, isso significa projetar algoritmos que:

Para um dado de consumo " E " acima; aplicar técnicas estatísticas para detectar padrões; e assim, “aprende” como executar alguma tarefa “ T ” acima com o desempenho P acima.

Esse processo é iterativo. A tentativa de cada iteração na tarefa específica é medida em relação ao resultado correto para essa tarefa (o " P " acima). Se melhorar, o algoritmo se ajustará ainda mais na direção dos ajustes que levaram a essa melhoria de desempenho. Se piorar, a ação oposta é realizada pelo algoritmo.

O método pelo qual muitas técnicas de ML atingem o processo de otimização mencionado acima é chamado de descida em gradiente, que abordaremos posteriormente, pois por si só trata-se de um tópico significativo.

Por fim, dependendo do progresso do processo acima, o algoritmo é ajustado, alimentando-o com mais dados e/ou ajustes específicos nas configurações do algoritmo até que o desempenho desejado seja alcançado.

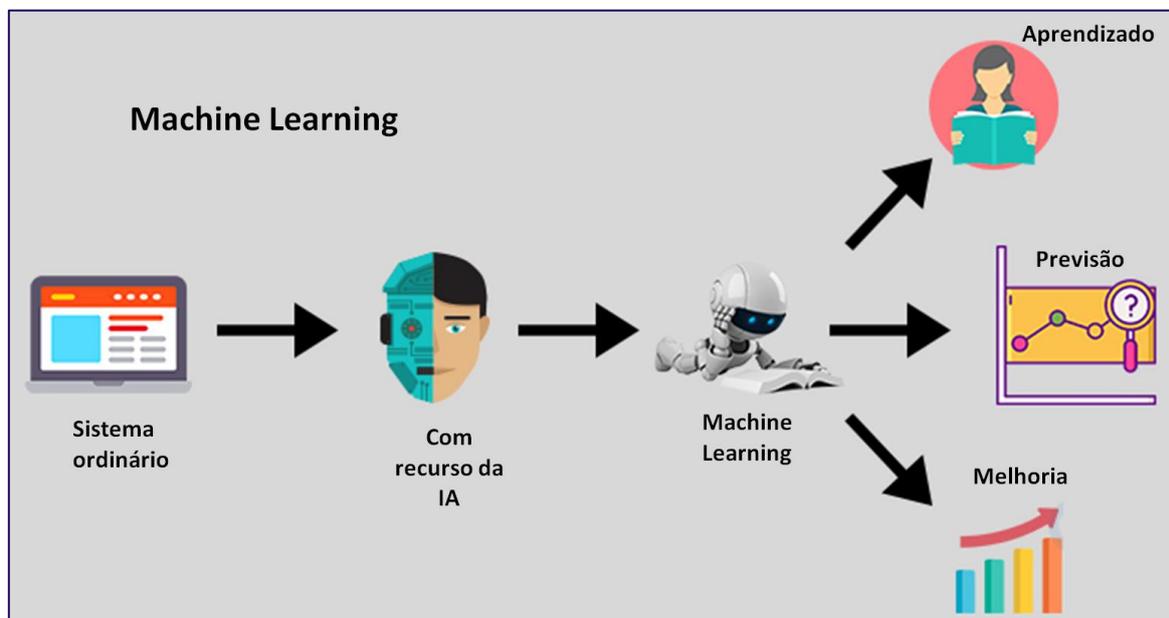
Basicamente, o foco principal do aprendizado de máquina é permitir que os computadores aprendam automaticamente sem intervenção humana.

Assim, os sistemas tomam decisões autonomamente sem nenhum suporte externo. Essas decisões são tomadas quando a máquina é capaz de aprender com os dados e entender os padrões subjacentes contidos nela. Então, através da correspondência de padrões e de análises adicionais cujos resultados pode ser uma classificação ou uma previsão.

Agora, surge a questão de como esse aprendizado pode ser iniciado e realizado?

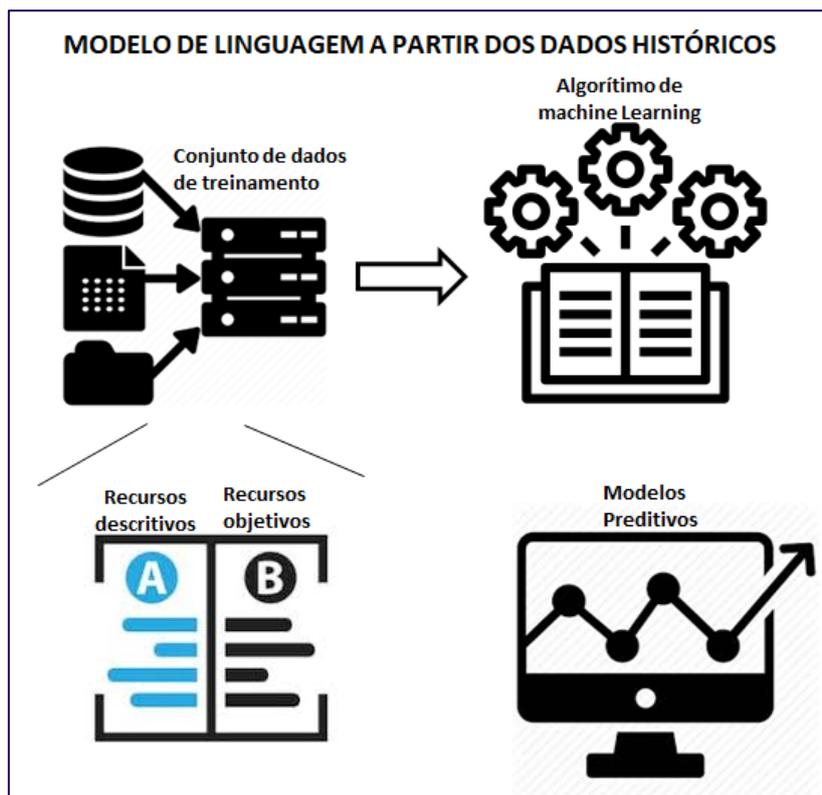
Pode ser iniciado com as observações dos dados com amostras e/ou exemplos, instruções ou também com algumas experiências diretas. Então, com base nessas entradas, a máquina toma melhor decisão procurando alguns padrões nos dados.

Trata-se, portanto, do subconjunto da IA que inclui técnicas estatísticas abstratas que permitem que as máquinas melhorem na execução de tarefas por meio de experiências. Nesta categoria inclui aprendizado profundo (deep learning) que veremos adiante.



Fonte: DataFly, 2019

Com os recursos do machine learning permite identificar numa operação empresarial, as oportunidades mais lucrativas ou identificar os altos riscos com maior rapidez e precisão. Caso contrário, exigiria maior tempo e recursos adicionais para que possa ser treinado de forma adequada seguindo a curva de aprendizado.



Fonte: Adaptado de Kelleher et al, 2018

Assim a combinação de aprendizado de máquina com a IA e as tecnologias cognitivas torna o sistema mais eficaz no processamento de grandes volumes de informações.

Métodos de aprendizado de máquina: Algoritmos.

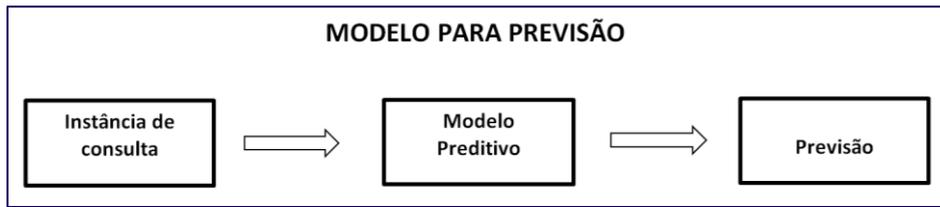
No conceito básico do aprendizado de máquinas utilizam-se algoritmos que possuem a capacidade de aprender a partir de experiências, que é comumente associada com o processo de automação.

Desta forma, os dados formam a espinha dorsal para que os algoritmos de aprendizado de máquina possam ser treinados.

Assim, os algoritmos formam os blocos de construção para permitir que os computadores aprendam a se comportar de maneira mais inteligente.

Assim, os algoritmos são treinados com os dados históricos, e são capazes de criar mais dados para melhorar a previsão dos resultados mediante a combinação dos recursos da ciência da computação, matemática e estatística.

Pois, a estatística é essencial para extrair inferências dos dados; a matemática é útil para o desenvolvimento de modelos de aprendizado de máquina e, finalmente, a ciência da computação é usada para implementar os algoritmos.



Fonte: Adaptado de Kelleher et al, 2018

Ao serem treinados de forma repetida com os dados experimentais, os algoritmos aprendem e formam padrões correspondentes com base nas análises estatísticas para identificar, classificar e melhorar as previsões sobre o futuro.

Sempre que um dado novo é introduzido no modelo de ML, ela associa com os padrões aprendidos para aprimorar as previsões futuras.

O modelo é sempre aperfeiçoado com novas experiências elevando assim a precisão de análise e aumentando o grau de adaptabilidade para novos dados e assim produzir melhores resultados.

Os cientistas de dados usam diferentes e variados tipos de algoritmos de aprendizado de máquina para descobrir novos padrões no universo do *Big Data* que proporcionam novos *insights* aumentando assim o grau de precisão na tomada de decisão.

As implementações práticas de tais estratégias de extração de recursos serão apresentadas nas últimas seções deste tutorial.

Com base na maneira com os algoritmos “aprendem” sobre os dados e assim fazer previsões são classificados em três grandes grupos: o aprendizado supervisionado, o não supervisionado e o de reforço (Castle, 2017) conforme figura a seguir:



Fonte: DataFly, 2019

APRENDIZADO SUPERVISIONADO

Trata-se do tipo de algoritmo mais usado. É chamado de supervisionado porque o processo de aprendizado dos algoritmos a partir do conjunto de dados de treinamento pode ser considerado como um professor que supervisiona este processo. Nesse tipo de algoritmo de ML, os possíveis resultados já são conhecidos e os dados de treinamento também são rotulados com respostas corretas. Pode ser entendido da seguinte forma:

Suponha que temos variáveis de entrada x e uma variável de saída y e aplicamos um algoritmo para aprender a função de mapeamento da entrada para a saída, como -

$$Y = f(x)$$

O objetivo principal é aproximar tão bem a função de mapeamento que, quando temos novos dados de entrada (x), podemos prever a variável de saída (Y) correspondente para esses dados.

Podem ser divididos nos dois tipos de problemas a seguir:

Classificação – Quando temos a saída categorizada como "preto", "ensino", "não ensino", etc.

Regressão - Um problema é chamado de problema de regressão quando temos o valor real de saída, como "distância", "quilograma" etc.

Árvore de decisão, floresta aleatória, knn, regressão logística são exemplos de algoritmos de aprendizado de máquina supervisionados.

Portanto, em vez de regras e comportamentos, são exemplos de entrada e saída com intuito de que possa generalizar a partir dos exemplos (normalmente descrevendo o passado) a se comportarem em situações não mostradas nos exemplos. Há um mapeamento claro e distinto entre os dados de entrada e de saída. Com base nas entradas, o modelo pode ser treinado nas diversas instâncias.

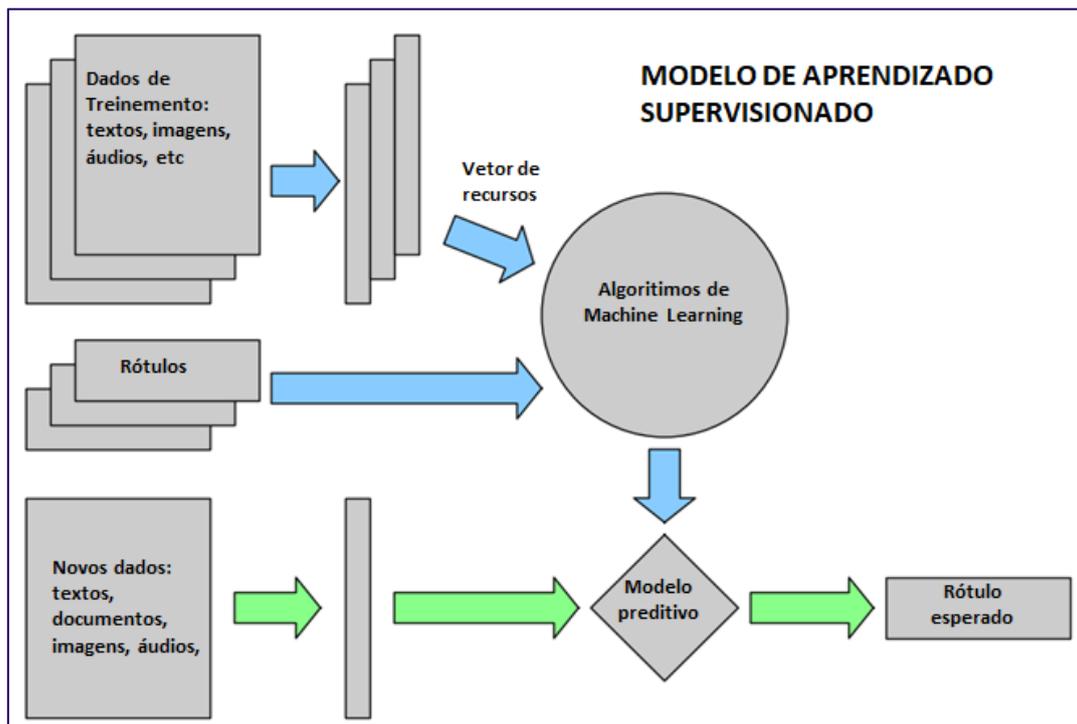
Um exemplo de aprendizado supervisionado é a filtragem de spam. Com base nos dados rotulados, o modelo é capaz de determinar se os dados são spam ou coisas significativas. Esta é uma forma mais usual de treinamento.

No nosso exemplo consiste na figura de um piso e a interpretação correspondente (isto é, se o piso está limpo ou não). Se fornecermos exemplos suficientes que são suficiente da maioria das situações, o sistema, por meio de seu algoritmo de aprendizado de máquina, será capaz de generalizar para saber também como interpretar boas imagens de pisos nunca antes vistos pelo sistema.

Pois, os algoritmos de aprendizado de máquina supervisionados podem aplicar o que foi aprendido no passado, e com a adição de novos dados rotulados podem prever eventos futuros. O

sistema torna se capaz de fornecer metas para qualquer nova entrada após o treinamento adequado e suficiente. O algoritmo de aprendizado também pode comparar uma saída com a saída correta e pretendida e encontrar erros para modificar e ajustar o modelo de forma adequada.

São utilizados em situações em que os dados históricos são conhecidos e pode ser usado para prever resultados futuros, como identificar quais clientes é mais provável para a inadimplência de uma dívida.



Fonte: Adaptado de scikit-learn.org, 2019

Assim, o aprendizado supervisionado pode responder perguntas como:

- **Quantos clientes solicitarão um empréstimo no próximo mês?**
Dados de treinamento: com base dos pedidos de empréstimo de meses anteriores
- **Essas células são cancerígenas?**
Dados de treinamento: exemplos de células cancerígenas e não cancerígenas
- **Este e-mail é spam?**
Dados de treinamento: e-mails anteriores conhecidos como spam ou não spam
- **Esta transação é fraudulenta?**
Dados de treinamento: transações anteriores conhecidas que são fraudulentas ou não fraudulentas

APRENDIZADO NÃO SUPERVISIONADO

Como o nome sugere, esses tipos de algoritmos não têm nenhum supervisor para fornecer qualquer tipo de orientação, ou seja, estão intimamente alinhados com a verdadeira inteligência artificial.

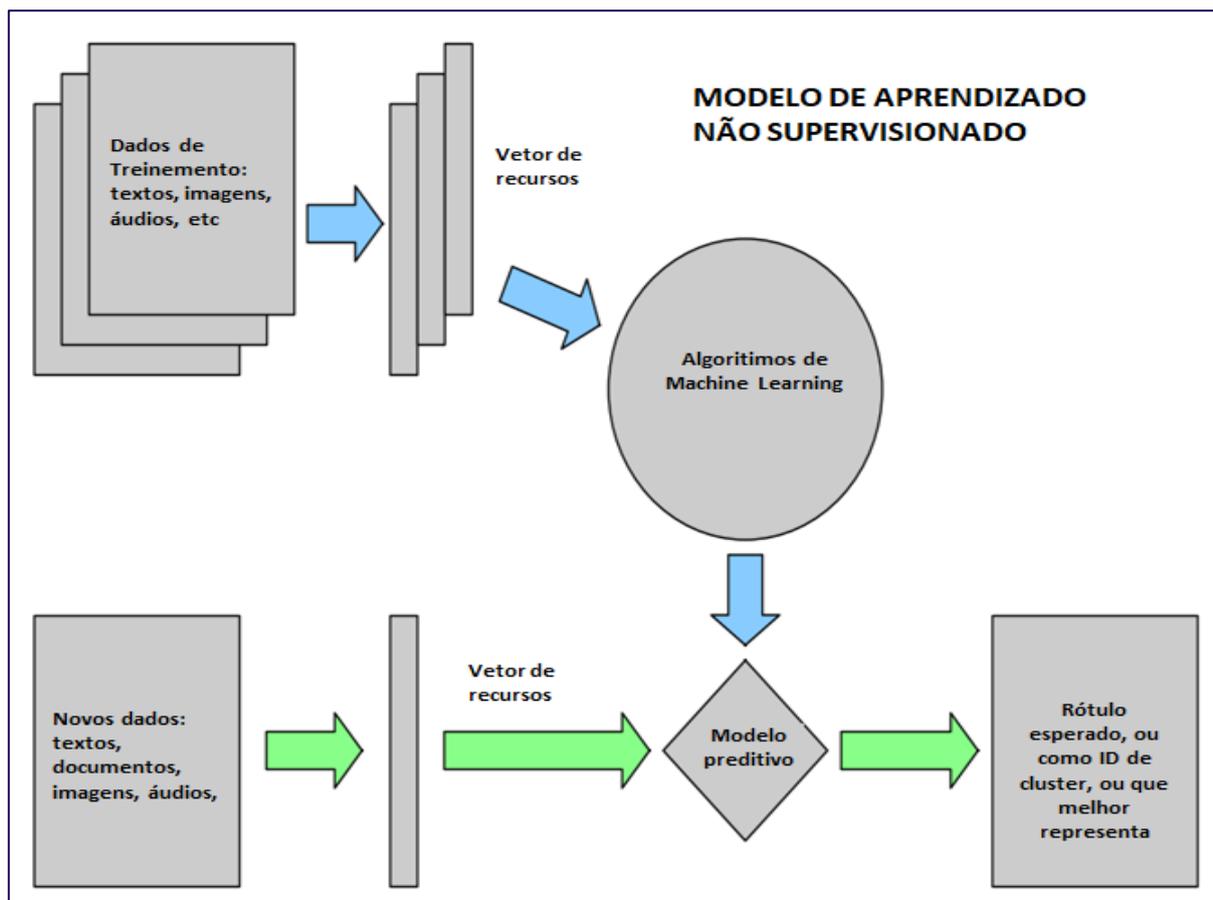
Pode ser entendido da seguinte forma:

Para uma variável de entrada x , então não haverá variáveis de saída correspondentes, como existem nos algoritmos de aprendizado supervisionado. Ou seja, na aprendizagem não supervisionada não haverá resposta correta nem professor para orientação pois os próprios algoritmos ajudam a descobrir padrões interessantes nos dados.

Os problemas de aprendizado não supervisionado podem ser divididos nos dois tipos de problemas a seguir:

- **Clustering:** Em problemas de clustering, precisamos descobrir os agrupamentos inerentes nos dados. Por exemplo, agrupando clientes por seu comportamento de compra.
- **Associação:** Nesse tipo de problema exige a descoberta das regras que descrevem grandes partes dos dados do cotidiano. Por exemplo, localizando os clientes que compram produtos x e y .

No aprendizado não supervisionado, não há dados rotulados, ou seja, os dados não são conhecidos. Nesse caso, o algoritmo aprende a identificar os padrões no conjunto de dados existentes. Desta forma, o algoritmo agrupa os dados em vários subconjuntos com base na sua densidade e visualiza em alta dimensão. E depois de aprender e extrair os padrões, o modelo tira conclusões. Não há rótulos nas variáveis de saída; o sistema não está preparado para uma resposta definida, mas auto descobre os padrões de dados. Possui diferentes técnicas que podem ser usadas nas transações com os dados (análise de cluster) e pode ser benéfico se usado previamente como parte do processo de avaliação do risco para descobrir os riscos imprevistos.



Fonte: Adaptado de scikit-learn.org, 2019

O aprendizado não supervisionado estuda como os sistemas podem inferir numa função para descrever uma estrutura oculta a partir de dados não rotulados. O sistema não descobre uma saída correta, mas explora os dados e extrair as inferências desse conjunto para descrever estruturas ocultas de dados não rotulados.

O modelo de aprendizado não supervisionado tenta ajustar os parâmetros para resumir melhor as regularidades encontradas nos dados.

A diferença entre dados supervisionados e não supervisionados é que os dados de treinamento fornecem exemplos, mas não existem rótulos específicos. O objetivo do aprendizado supervisionado é criar um modelo que possa prever o resultado para novas instâncias com base nos exemplos anteriores. Com o aprendizado não supervisionado visa criar um modelo para fazer uma descoberta em vez de fazer previsões.

Considere os seguintes exemplos de alguns usos possíveis de aprendizado não supervisionado:

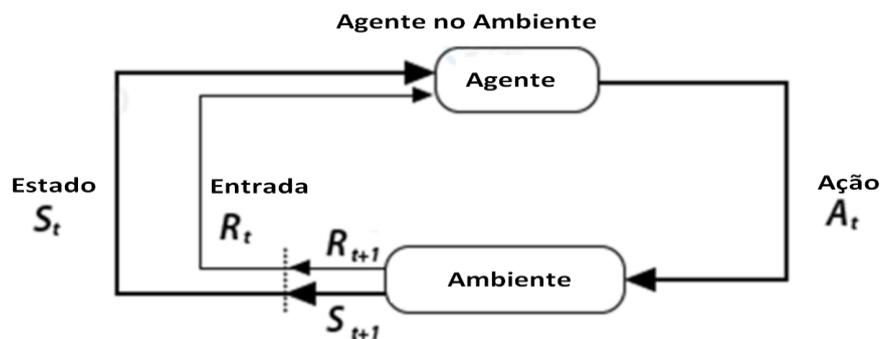
Descobrir o que os clientes compram junto com outros itens de compra. Por exemplo, você pode achar que os clientes que compram café também tendem a comprar leite. Pode se concluir dividindo a base de usuários em grupos com gostos, locais ou informações demográficas semelhantes.

A categoria dos algoritmos de aprendizado semi-supervisionado está entre o aprendizado supervisionado e o não supervisionado, pois eles usam dados rotulados e não rotulados para treinamento - geralmente uma pequena quantidade de dados rotulados e uma grande quantidade de dados não rotulados. Os sistemas que usam esse método são capazes de melhorar consideravelmente a precisão do aprendizado. Geralmente, a aprendizagem semi-supervisionada é escolhida quando os dados rotulados adquiridos requerem recursos qualificados e relevantes para treiná-los/aprender com eles. Caso contrário, a aquisição de dados não rotulados geralmente não requer recursos adicionais.

APRENDIZADO POR REFORÇO

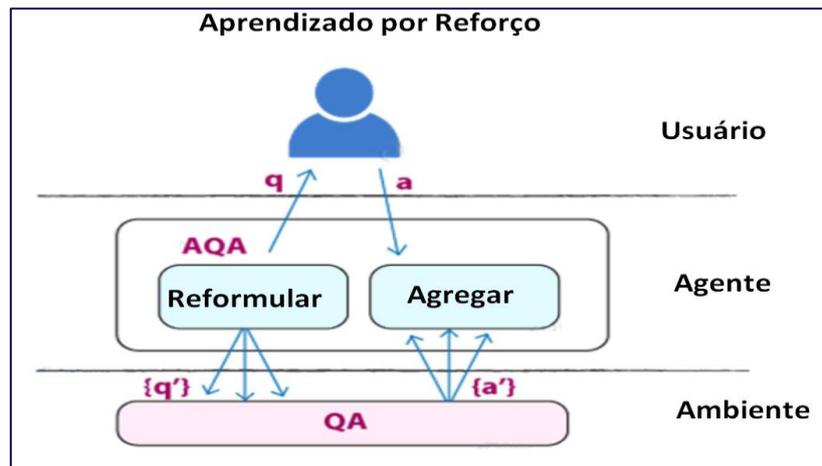
O Aprendizado por Reforço abrange mais áreas da Inteligência Artificial, o que permite que as máquinas interajam com seu ambiente dinâmico para atingir seus objetivos.

Assim, as máquinas e agentes de software são capazes de avaliar o comportamento ideal em um contexto específico e com a ajuda do *feedback* de recompensa (sinal de reforço), são capazes de aprender o comportamento e melhorá-lo com o tempo.



Os algoritmos que aprendem com a experiência passada e capturam os melhores conhecimento e treinam os sistemas para tomar decisões específicas de forma precisa. Basicamente, a máquina é exposta a um ambiente em que se treina continuamente usando o método de tentativa e erro.

Quando não há um conjunto de dados de treinamento, ele aprende com sua própria experiência.



É aplicado em sistemas autônomos como veículos autônomos inteligentes e na programação da robótica industrial nas execuções de ações que requer certa dose de autonomia.

Pois, a eficiência é melhorada com o acúmulo de experiências junto ao agente do ambiente de atuação. Esse é o princípio básico do aprendizado por reforço que interage com seu ambiente produzindo ações para descobrir erros e faz as compensações com ajustes constantes. Esta pesquisa por meio da tentativa e erro são as características mais relevantes do aprendizado por reforço. Trata-se do método em que as máquinas e os agentes de software determinem automaticamente o comportamento ideal para um contexto específico no sentido de maximização do desempenho. É necessário *feedback* de compensação para que o agente aprenda qual é a melhor ação; isso é conhecido como sinal de reforço.

O Processo de Decisão de Markov¹ é um exemplo de algoritmos de aprendizado de máquina de reforço.

COMO O APRENDIZADO DE MÁQUINA RESOLVE PROBLEMAS

Aqui está uma visão geral dos estágios envolvidos ao usar o aprendizado de máquina para resolver problemas:

¹ MDP (Markov Decision Process) é uma forma de modelar processos de forma em que as transições entre estados são probabilísticas, e é possível observar e interferir periodicamente (épocas de decisão) em que estado o processo se encontra executando ações. Cada ação tem uma recompensa (ou custo), que depende do estado em que o processo se encontra. Alternativamente, pode-se definir recompensas por estado apenas, sem que estas dependam da ação executada. Trata-se de uma tupla (S, A, T, R), onde:

- S é um conjunto de estados em que o processo pode estar;
- A é um conjunto de ações que podem ser executadas em diferentes épocas de decisão;
- $T : S \times A \times S \rightarrow [0, 1]$ é uma função que dá a probabilidade do sistema passar para um estado $s' \in S$, dado que o processo estava em um estado $s \in S$ e o agente decidiu executar uma ação $a \in A$ (denotada $T(s'|s, a)$);
- $R : S \times A \rightarrow R$ é uma função que dá o custo (ou recompensa) por tomar uma decisão $a \in A$ quando o processo está em um estado $s \in S$.

Pode-se também definir para cada estado $s \in S$, um conjunto de ações possíveis naquele estado (A_s). Assim, o conjunto de todas as ações poderia ser $A = \bigcup_{s \in S} A_s$. Usaremos apenas "A", a fim de simplificar a notação. Os conjuntos S e A podem ser finitos ou infinitos.

Etapa 1: coletar e preparar os dados

Depois que as fontes de dados são identificadas, os dados disponíveis e coletados são compilados. O tipo de dados obtidos direciona na escolha dos algoritmos de aprendizado de máquina a ser utilizado. À medida que se analisa os dados, as anomalias são identificadas, a estrutura é desenvolvida e os problemas de integridade dos dados são resolvidos.

Etapa 2: Treinamento do modelo

Os dados preparados são divididos em dois grupos: o conjunto de treinamento e o conjunto de testes. O conjunto de treinamento forma a grande parte dos dados para ajustar os modelos com a maior precisão.

Etapa 3: validar o modelo

Nesta etapa para selecionar o modelo de dados final, o conjunto de testes é usado para avaliar o desempenho e a precisão.

Etapa 4: interpretar os resultados

Analise o resultado para encontrar idéias, tirar conclusões e prever resultados.



Prever valores

São úteis na identificação de causa e efeito entre variáveis como os algoritmos de regressão criam um modelo a partir dos valores que são então usados para fazer previsões. Os estudos de regressão ajudam a prever o futuro, o que pode ajudar a antecipar a demanda do produto, prever números de vendas ou estimar os resultados da campanha.



Identificar ocorrências incomuns

Frequentemente usados para identificar riscos em potencial, os algoritmos de detecção de anomalias identificam dados fora da norma prevista. Mau funcionamento do equipamento, defeito estrutural, erros de texto e casos de fraude são exemplos de como o aprendizado de máquina pode ser usado para resolver preocupações.



Localizar a estrutura

Os algoritmos de agrupamento são frequentemente o primeiro passo no aprendizado de máquina, revelando a estrutura subjacente no conjunto de dados. Categorizando itens comuns, o agrupamento é comumente usado na segmentação de mercado, oferecendo informações que podem ajudar a selecionar o preço e antecipar as preferências do cliente.



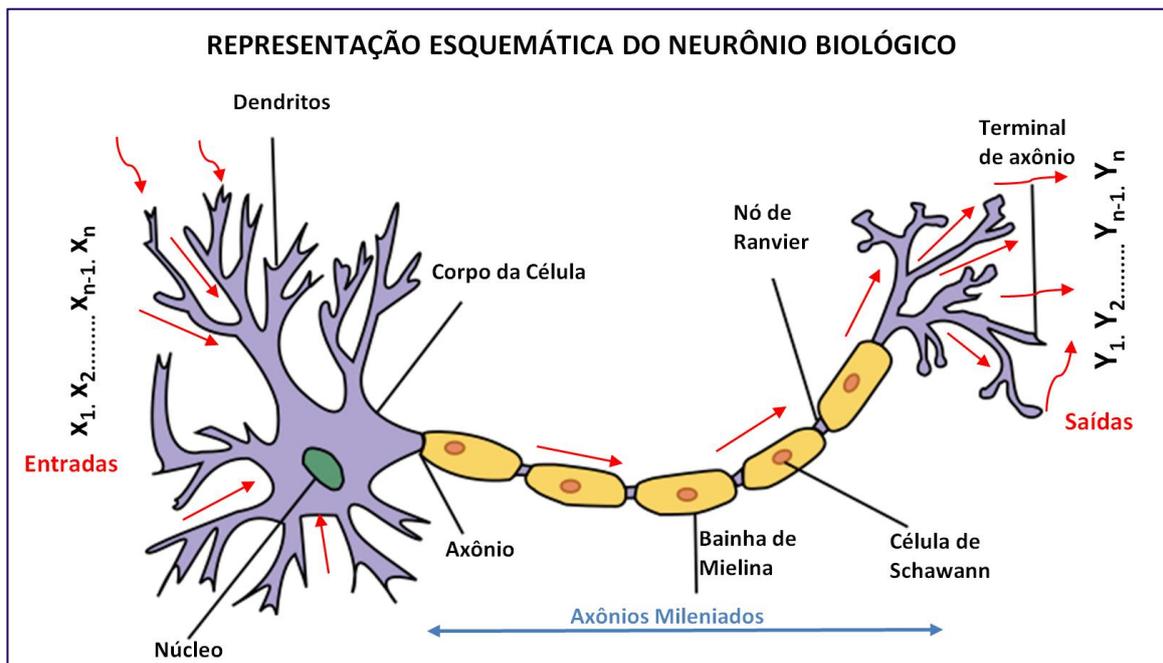
Previsão por categorias

Os algoritmos de classificação ajudam a determinar a categoria correta para informações. Tendo semelhança com o agrupamento, a classificação difere na aplicação no aprendizado supervisionado, onde são atribuídos rótulos predefinidos.

REDES NEURAIS

NEURÔNIOS BIOLÓGICOS

O cérebro humano é estruturado, conectado e integrado por bilhões de neurônios que constitui a rede de comunicação chamada de rede neural. Cada neurônio possui um corpo central, diversos dendritos e axônios conforme a figura a seguir:



Fonte: Towards Data Science, 2019

Os neurônios são células eletricamente excitáveis que se comunicam com outras células por meio de conexões especializadas formando os blocos de construção do cérebro.

- Dendritos ou dendrites (do grego: *déndron*: árvore) são os numerosos prolongamentos dos neurônios que atuam na recepção de estímulos nervosos do ambiente ou de outros neurônios e na transmissão desses estímulos para o corpo da célula, também chamado de pericário.
- Soma é onde os sinais recebidos dos dendritos são unidos e transmitidos. Junto com os dendritos formam a superfície de entrada do neurônio
- Axônios é uma parte especializada da soma que se conecta ao axônio que controla o disparo do neurônio formando a superfície de saída do fluxo de informação.
- Seu papel é transmitir o sinal neural para outros neurônios através de suas sinapses.
- Sinapses são pequenas lacunas localizadas no terminal do axônio como ponto de contato de um neurônio e o dendrito que conecta o neurônio a outras células nervosas formando as redes neurais. Lá, neurotransmissores são usados para transportar o sinal através da sinapse para outros neurônios.

O processo de comunicação entre os neurônios ocorre a partir dos dendritos que recebem os impulsos em forma de sinais elétricos dos neurônios por meio das sinapses captados pelos sensores biológicos como visão, olfato, audição, tato, paladar, entre outros. Em seguida o corpo celular processa a informação e envia para outro neurônio, e assim sucessivamente.

Os sinais elétricos acima do limiar do disparo (*threshold*²) caminham pelos axônios pode ser amplificado ou atenuado, dependendo do dendrito de origem, pois a cada condutor está associado um peso pelo qual o sinal é multiplicado. Os pesos são o que chamamos de memória que são estabelecidos por meio de treinamento recebido pelo cérebro durante a vida útil que é a memorização.

O princípio operacional de um neurônio biológico pode ser resumido da seguinte forma. Primeiro, ele recebe informações de seus dendritos (ou seja, de outros neurônios).

Numa segunda etapa, uma soma ponderada dessas entradas é realizada dentro do soma. O resultado é então repassado para a colina do axônio. Se essa soma ponderada for maior que o limite, o neurônio será acionado.

Caso contrário, ele permanece em repouso. O estado do neurônio (ativado ou desativado) então se propaga através de seu axônio e é passado para os outros neurônios conectados através de suas sinapses. Embora muito simples essa descrição de alto nível do princípio operacional de um neurônio

² Magnitude ou intensidade que deve ser excedida para que uma determinada reação, fenômeno, resultado ou condição ocorra ou se manifeste.

biológico é suficiente para entender o modelo matemático de um neurônio artificial proposto por McCulloch & Pitts³ em 1943.

Com base no modelo do neurônio biológico cuja rede neural forma a inteligência humana, os pesquisadores desenvolveram um modelo de neurônio matemático que se tornou a base da Inteligência Artificial.

Algumas abordagens de aprendizado de máquina adotam algoritmos baseados no conceito de redes neurais que são inspiradas pelo cérebro humano onde as unidades de processamento (analogamente aos nossos neurônios) com muitas conexões ponderadas entre elas formando uma rede.

A entrada são os dados provenientes dos sensores (em nosso exemplo, a imagem do piso) e tem como saída a interpretação da imagem (em nosso exemplo, se o piso está limpo ou não). Durante a análise dos exemplos (a fase de treinamento da rede), os pesos das conexões são ajustados para corresponder o máximo possível ao que dizem os exemplos disponíveis (ou seja, para minimizar o erro entre a saída esperada e a saída calculada pela rede) . No final da fase de treinamento há uma fase de teste do comportamento da rede neural onde é avaliada com os exemplos nunca vistos antes para comprovar se a tarefa foi bem aprendida.

É importante notar que essa abordagem (como todas as técnicas de aprendizado de máquina) sempre tem uma certa porcentagem de erro), embora geralmente seja pequena. Portanto, uma noção essencial nesta fase é o grau de precisão, ou seja, uma métrica para avaliar a porcentagem de respostas corretas.

Existem vários tipos de redes neurais e abordagens de aprendizado de máquina, dos quais atualmente um dos mais bem-sucedidos é o aprendizado profundo (deep learning) cuja abordagem refere-se ao fato de que a rede neural possui várias camadas entre a entrada e a saída que permitem aprender a relação geral de entrada e saída em etapas sucessivas que serão objeto tópicos subsequentes.

REDES NEURAIS ARTIFICIAIS (RNA)

O aprendizado de máquina envolve mecanismos adaptativos que permitem aos computadores aprender com a experiência, aprender pelo exemplo e aprender por analogia. Assim, os recursos de aprendizado podem melhorar o desempenho de um sistema inteligente ao longo do tempo.

³ Em 1943, o neurocientista Warren S. McCulloch, e o lógico Walter Pitts publicaram no Boletim de Biofísica Matemática, a pesquisa "Um cálculo lógico das ideias iminentes à atividade nervosa". Neste artigo tentaram entender como o cérebro poderia produzir padrões altamente complexos usando muitas células básicas chamadas neurônios que estão conectadas, que deu uma contribuição importante ao desenvolvimento de redes neurais artificiais - que modelam as características-chave dos neurônios biológicos.

As origens das redes neurais artificiais (RNA) estão no campo da biologia. Pois o cérebro biológico consiste em bilhões de neurônios altamente interconectados formando uma rede neural. O processamento de informações humanas depende do sistema que conecta as células nervosas.

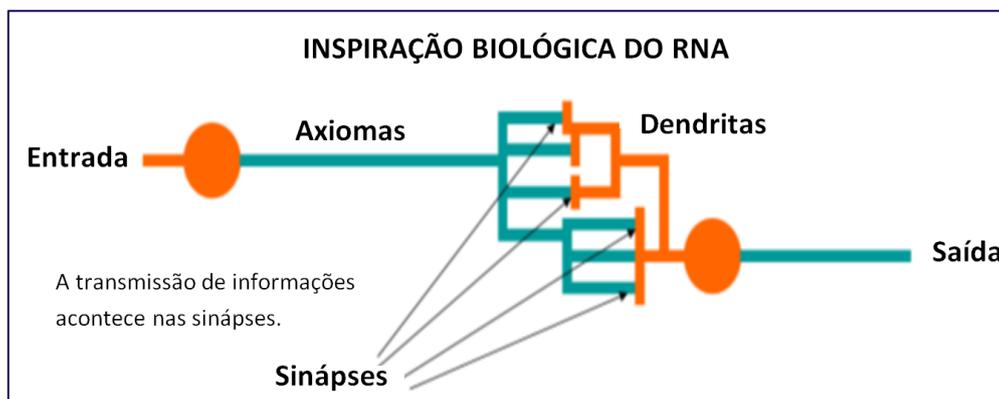
Uma comparação clássica entre o processamento de informações por um humano e um computador concentra-se na capacidade de reconhecimento e o aprendizado de padrões. O computador pode calcular grandes números com alta velocidade, mas possui limitações em reconhecer algo como um problema de classificação, texto escrito, compactação de dados e um algoritmo de aprendizado.

Pelo contrário, um ser humano reconhece e lida com estes desafios mencionados processando informações distribuídas e transformações através de milhares de neurônios interconectados no cérebro (Jeng et al, 2008).

PROCESSAMENTO

As abordagens mais usuais do aprendizado de máquina são redes neurais artificiais e algoritmos genéticos. A RNA simula a rede neural biológica do cérebro humano para replicar os processos de aprendizado. Desta forma emula a estrutura do cérebro humano.

Tem como base o mecanismo semelhante ao sistema nervoso de um organismo vivo que tem como princípio de que as tarefas complexas sejam executadas instintivamente.



Fonte: Mehdi Akhoondzadeh, Kami Azizi, 2019

A unidade central de processamento desse sistema nervoso é conhecida como "neurônio". O cérebro humano tem cerca de 10 a 100 bilhões de neurônios, onde cada um é conectado a muitos outros por meio de "sinapses". O cérebro humano tem cerca de 100 trilhões de sinapses e pode ser considerado como um sistema de processamento de informações altamente complexo, não linear e paralelo.

As informações são armazenadas e processadas em uma rede neural simultaneamente em toda a rede, e não em locais específicos. Em outras palavras, nas redes neurais, os dados e seu processamento são globais e não locais.

O aprendizado é uma característica fundamental e essencial das redes neurais biológicas.

Essas conexões controlam o corpo humano e os processos do pensamento.

A facilidade com que eles podem aprender levou a tentativas de emular uma rede neural biológica em um computador.

As primeiras teorias da RNA foram expostas por pesquisadores com formação em psicologia que tentavam explicar o comportamento humano e o processo de pensamento modelando o cérebro humano.

Uma rede neural artificial consiste em um número de processadores muito simples, também chamados neurônios, que são análogos aos neurônios biológicos no cérebro.

Os neurônios são conectados por links ponderados que passam sinais de um neurônio para outro.

O sinal de saída é transmitido através da conexão de saída do neurônio. A conexão de saída se divide em vários ramos que transmitem o mesmo sinal. Os ramos de saída terminam nas conexões de entrada de outros neurônios na rede.

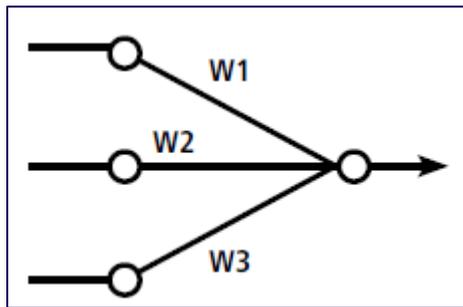
ANALOGIA ENTRE REDES NEURAIS BIOLÓGICAS E ARTIFICIAIS

REDE NEURAL BIOLÓGICA	REDE NEURAL ARTIFICIAL
Núcleo	Neurônio
Dendritos	Entrada
Axônio	Saída
Sinapse	Pêso
Lento	Rápido
Muitos neurônios (10^9)	Poucos neurônios (~100)

O trabalho começou na década de 1940 com o primeiro modelo matemático do neurônio artificial de McCulloch & Pitts e foi consolidado em 1957 pelo advento de Frank Rosenblatt⁴, o "Perceptron" que era um classificador linear, um tipo mais simples de "feed forward" da rede neural conforme o modelo simplificado a seguir:

⁴ Inspirado nos princípios biológicos de um neurônio desenvolveu o conceito de Perceptron no Laboratório Aeronáutico de Cornell-EUA em 1957.

- Um neurônio recebe "mensagens de comunicação" de outros neurônios na forma de impulsos elétricos de força diferente que podem ser excitatórios ou inibitórios.
- Um neurônio integra todos os impulsos recebidos de outros neurônios.
- Se a integração resultante for maior que um determinado limite, o neurônio 'dispara', acionando o potencial de ação que é transmitido a outros neurônios conectados.



O Perceptron original recebeu três entradas W1, W2 e W3 com pesos variáveis e gerou uma única saída conforme o modelo.

O neurônio é a unidade de estrutura básica da rede neural e recebe qualquer número de entradas, possuindo pesos com base na sua importância. Caso receba sinais de estímulo suficientes dispara e aciona todas as suas saídas.

Como um neurônio real, as entradas ponderadas são somadas e a saída é baseada numa função limitadora e envia para todos os neurônios subsequentes. Para que o sinal de saída seja traduzido em informações reais de trabalho, os impulsos transmitidos atinjam a camada de saída.

Embora este sistema funcionasse bem para problemas simples, nas classificações não lineares chamados de problemas de “ou exclusivo” são impossíveis de solução, porém no mundo real se apresenta como uma questão relativamente simples.

No caso do exemplo, “é possível ir às compras ou ao cinema, mas não é possível que faça as duas coisas ao mesmo tempo”.

Tabela de "ou exclusivo"

Entrada1	Entrada2	Resultado
V	V	F
V	F	V
F	V	V
F	F	F

V- Verdadeiro
F- Falso

Seleção e escolha

Tabela de “ou exclusivo” que simulam os cenários da vida real e são difíceis de modelar, porque só é capaz de processar problemas lineares.

Assim as redes neurais artificiais formam um tipo especial de algoritmo de aprendizado de máquina que é modelado a partir do funcionamento do cérebro humano. Análogo aos neurônios do nosso sistema nervoso que são capazes de aprender com os dados do passado, a RNA é capaz de aprender com os dados passados e fornecer respostas na forma de previsões ou estabelecer classificações.

Portanto, as RNAs são modelos estatísticos não lineares que exibem relacionamentos complexos entre as entradas e saídas para descobrir novos padrões. Trata uma variedade de tarefas

como reconhecimento de imagem, reconhecimento de fala, tradução simultânea e automática de linguagens, diagnósticos médicos, entre outros.

São compostos de quatro áreas distintas de pesquisa que são:

- **Modelagem das redes biológicas:** procura o entendimento do cérebro humano e de suas funções. Esta área é de particular interesse para psicólogos e pesquisadores em neuroanatomia;
- **Ferramenta educacional:** para entender como resolver tarefas complexas que as metodologias tradicionais de IA e os algoritmos de computador têm dificuldade de solução. Pesquisadores nessa área incluem cientistas da computação, engenheiros, matemáticos e estatísticos interessados em construir algoritmos de computador a partir do processo de resolução de problemas da RNA;
- **Problemas reais do mundo real:** Solução nas áreas comerciais e de negócios. Trata-se do campo multidisciplinar que requer experiências em outras áreas além daquelas relacionadas à RNA. Existem pacotes de software da RNA que são fáceis de usar o suficiente para que os novos usuários comecem a usar sem exigir que eles tenham um conhecimento profundo dos algoritmos da RNA. Isso é diferente das técnicas convencionais de computador, que exigem que o usuário compreenda completamente o algoritmo antes de escrever o programa para aplicá-lo. No caso de RNAs, tudo que um usuário precisa saber é como apresentar o problema em questão de uma forma que uma RNA possa entender.

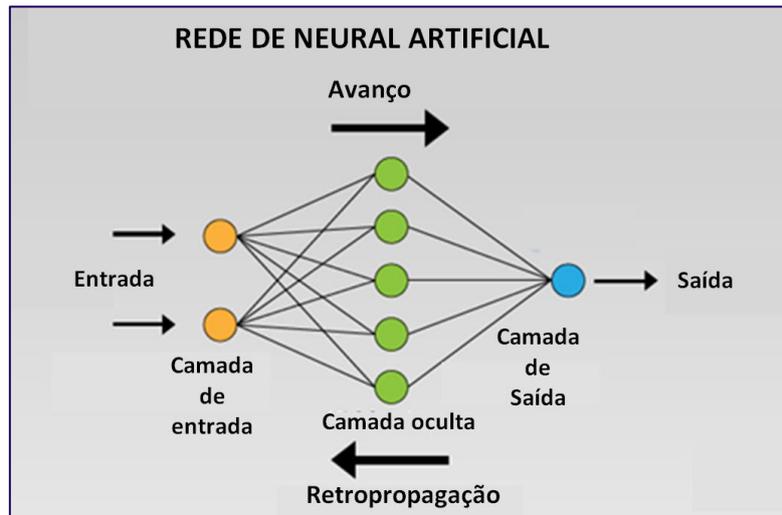
Estas ferramentas podem-se obter um método econômico de chegar às soluções que definem os modelos de distribuição mediante o aprimoramento de técnicas de análise de dados devido às suas capacidades preditivas avançadas a partir de amostras para fornecer dados de saída.

ARQUITETURA DE REDES NEURAIS ARTIFICIAIS

As redes neurais remontam ao início dos anos 70, quando Warren S. McCulloch e Walter Pitts cunharam esse termo. O funcionamento das redes neurais artificiais é semelhante ao modo como os neurônios funcionam em nosso sistema nervoso. Para entender o funcionamento das RNAs, primeiro entendemos que está estruturada em três camadas essenciais:

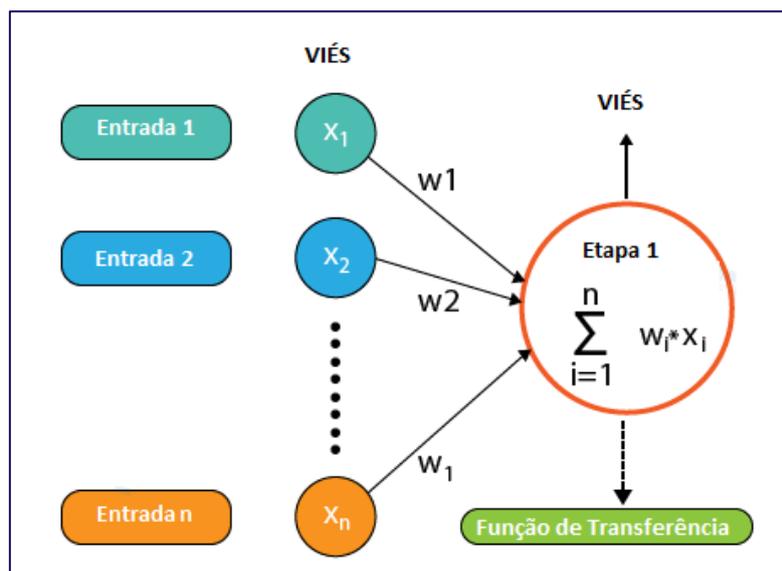
- **Camadas de entrada:** é a primeira camada de uma RNA que recebe as informações de entrada na forma de multimídia- textos, números, arquivos de áudio, pixels de imagem, fotos, gráficos, sinais elétricos e frequências, etc.

- **Camadas ocultas:** trata se das camadas intermediárias formadas por uma única camada oculta, como no caso de um perceptron ou várias camadas ocultas. Essas camadas executam vários tipos de cálculos matemáticos dos dados de entrada e reconhecem os padrões dos quais fazem parte.
- **Camada de saída:** são os resultados obtidos através de cálculos rigorosos realizados pela camada do meio, ou seja, a oculta.



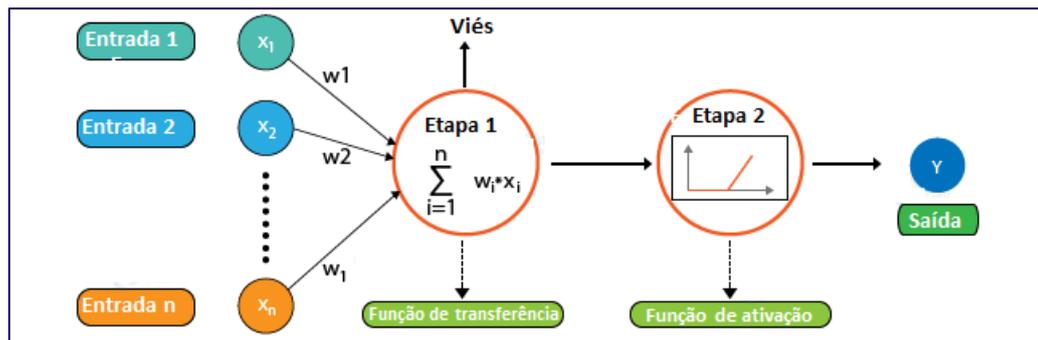
Fonte: Data Flair, 2019

Possui também vários parâmetros e hiperparâmetros que afetam o desempenho do modelo, pois, a saída das RNAs depende principalmente desses parâmetros. São formados por pesos, vieses, taxa de aprendizado, tamanho do lote etc, onde cada nó na RNA possui um determinado peso.



Após a soma efetuada pela função de transferência, a função de ativação apresenta o resultado. Com base na saída recebida, as funções de ativação disparam o resultado apropriado do nó. Por exemplo, se a saída recebida estiver acima de 0,5, a função de ativação disparará 1, caso contrário, permanecerá 0.

Algumas das funções de ativação populares usadas nas redes neurais artificiais são Sigmoid, RELU, Softmax, tanh etc.



Com base no valor que o nó disparou, obtemos a saída final. Em seguida, usando as funções de erro, calculamos as discrepâncias entre a saída prevista e a saída resultante e ajustamos os pesos da rede neural através de um processo conhecido como retropropagação.

As RNAs fazem parte de uma área emergente no Machine Learning, conhecida como Deep Learning.

RETRO PROPAGAÇÃO DAS REDES NEURAS ARTIFICIAIS

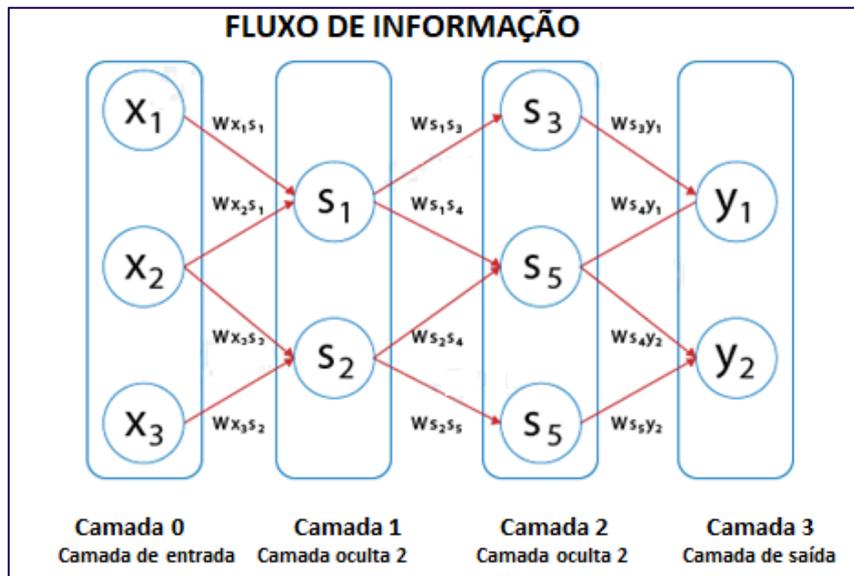
Para treinar uma rede neural por meio de exemplos de mapeamentos de entrada e saída e após a conclusão são testados com dados reais. A rede neural prediz a saída e são avaliados como a saída está correta usando as várias funções de erro. Finalmente, com base no resultado, o modelo será otimizado com ajustes com os pesos e regras da cadeia.

Tipos de redes neurais artificiais

- Existem dois tipos importantes de redes neurais artificiais
 - Rede Neural FeedForward
 - Rede Neural de FeedBack
 - Redes Neurais Artificiais Avançadas

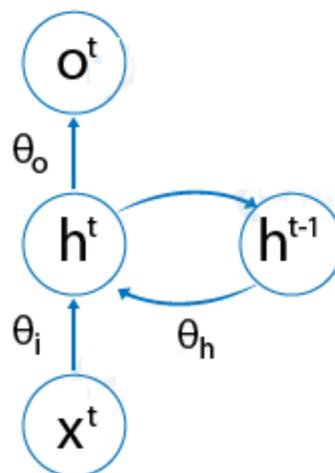
Nas ANNs de feedforward, o fluxo de informações ocorre apenas em uma direção. Ou seja, o fluxo de informações é da camada de entrada para a camada oculta e, finalmente, para a saída. Não há loops de feedback presentes nessa rede neural. Esse tipo de rede neural é usado principalmente no

aprendizado supervisionado para instâncias como classificação, reconhecimento de imagem etc. Nós as usamos em casos em que os dados não são sequenciais por natureza.



REDES NEURAIIS ARTIFICIAIS

Nas RNAs de feedback, os loops de feedback fazem parte dele. Esse tipo de rede neural é principalmente para retenção de memória, como no caso de redes neurais recorrentes. Esses tipos de redes são mais adequados para áreas em que os dados são sequenciais ou dependem do tempo.



REDES BAYESIANAS

Esse tipo de rede neural possui um modelo gráfico probabilístico que utiliza a inferência bayesiana para calcular a probabilidade. Esse tipo de rede bayesiana também é conhecido como rede de crenças. Nestas redes bayesianas, existem arestas que conectam os nós que representam as

dependências probabilísticas presentes entre esses tipos de variáveis aleatórias. A direção do efeito é tal que, se um nó está afetando o outro, eles caem na mesma linha de efeito. A probabilidade associada a cada nó quantifica a força do relacionamento. Com base no relacionamento, é possível inferir a partir das variáveis aleatórias no gráfico com a ajuda de vários fatores.

A única restrição que essas redes precisam seguir é que elas não podem retornar ao nó pelos arcos direcionados. Portanto, as redes bayesianas são chamadas de gráficos acíclicos direcionados (DAGs).

Essas redes bayesianas podem lidar com variáveis de vários valores e são compostas por duas dimensões:

- Gama de preposições
- Probabilidade de atribuir cada preposição.

Suponha que exista um conjunto finito de variáveis aleatórias, de modo que cada variável do conjunto finito seja denotada por $X = \{x_1, x_2, \dots, x_n\}$, em que cada variável X retira dos valores presentes no conjunto finito, como Valor $\{x_1\}$. Se houver um link direcionado da variável X_i para a variável X_j , X_i será o pai de X_j que mostra as dependências diretas entre essas variáveis.

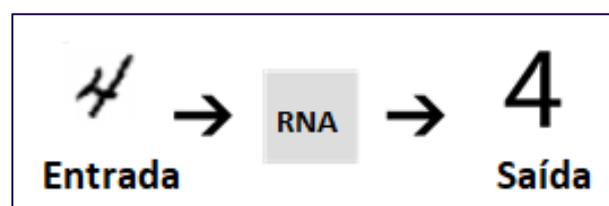
Com a ajuda das redes bayesianas, pode-se combinar o conhecimento prévio e os dados observados. As redes bayesianas são principalmente para aprender as relações causais e também entender o conhecimento do domínio para prever o evento futuro. Isso ocorre mesmo no caso de dados ausentes.

APLICAÇÕES DE REDES NEURAIS ARTIFICIAIS

A seguir, estão as importantes aplicações de Redes Neurais Artificiais:

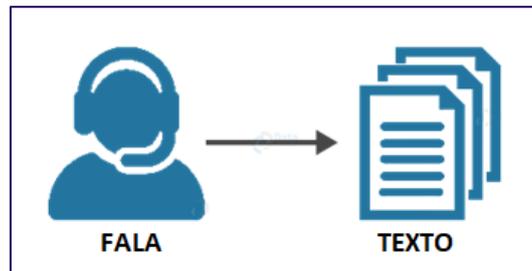
Reconhecimento de caracteres manuscritos

As RNAs são usadas para reconhecimento de caracteres manuscritos. As redes neurais são treinadas para reconhecer os caracteres manuscritos que podem estar na forma de letras ou dígitos.



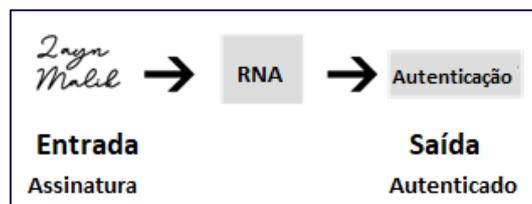
Reconhecimento de fala

As RNAs desempenham um papel importante no reconhecimento de fala. Os modelos anteriores de reconhecimento de fala foram baseados em modelos estatísticos, como os modelos Hidden Markov. Com o advento da aprendizagem profunda, vários tipos de redes neurais são a escolha absoluta para obter uma classificação precisa.



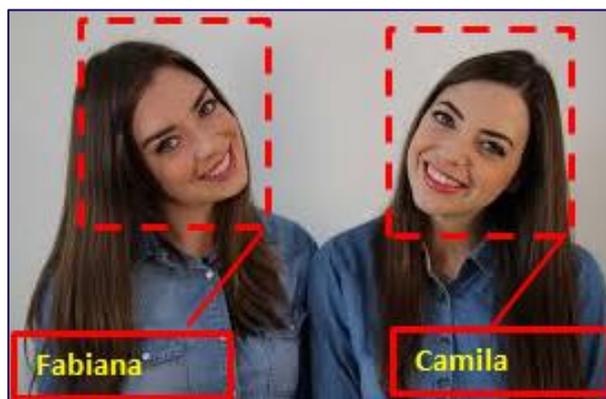
Classificação de Assinaturas

Para reconhecer assinaturas e categorizá-las para a classe da pessoa, usamos redes neurais artificiais para construir esses sistemas para autenticação. Além disso, as redes neurais também podem classificar se a assinatura é falsa ou não.



Reconhecimento Facial

Para reconhecer os rostos com base na identidade da pessoa, usamos redes neurais. Eles são mais comumente usados em áreas onde os usuários exigem acesso de segurança. Redes neurais convolucionais são o tipo mais popular de RNA usado neste campo.



Sumário

Então, você viu o uso de redes neurais artificiais através de diferentes aplicações. Espero que o DataFlair seja o melhor para explicar a introdução às redes neurais artificiais. Além disso, adicionamos vários exemplos de RNA entre o blog para que você possa relacionar o conceito de redes neurais facilmente. Estudamos como as redes neurais são capazes de prever com precisão usando o processo de retropropagação. Também passamos pelas Redes Bayesianas e, finalmente, analisamos as diversas aplicações das RNAs.

ALGORITMO GENÉTICO (GA)

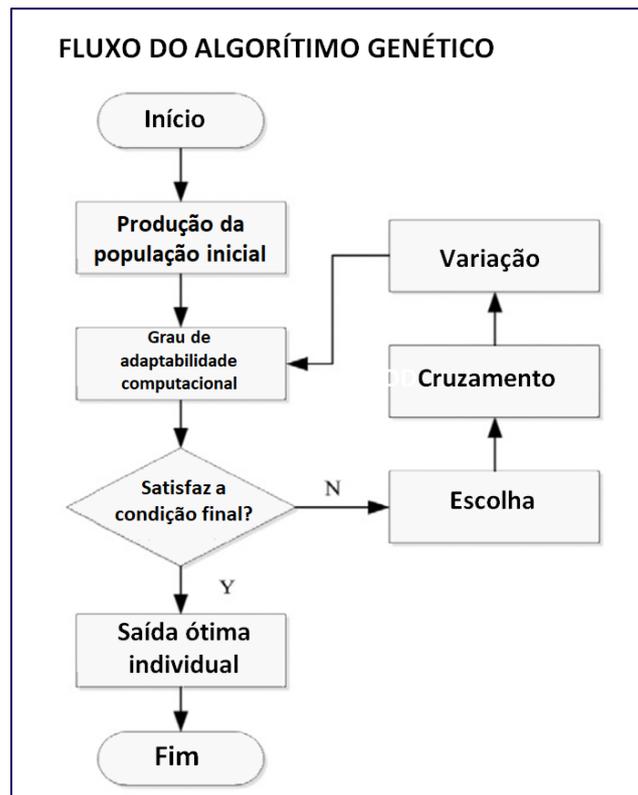
O GA é uma tecnologia do sistema de inteligência artificial auto-organizado e adaptável que simula os procedimentos e mecanismos de evolução da natureza para resolver problemas extremos.

A idéia básica do GA é formular um algoritmo de processo que busca a solução ideal com base na simulação natural do valor extremo genético e do evolucionismo biológico, que tem como base biológica firme. O GA é um algoritmo de pesquisa colateral e global que adota estratégia probabilística de busca aleatória capaz de obter automaticamente as informações relacionadas no âmbito da pesquisa e se ajusta e se adapta às direções da busca de forma robusta.

O GA tem sido amplamente utilizado nos campos do controle e otimização dos processos, processamento de sinais e vida artificial, etc. (Nguyen G, Dang TT, Hluchy L., 2007).

Foi proposto pelo Prof. Holland da Universidade de Michigan em 1975 com base na teoria da evolução de Darwin, da teoria da seleção natural de Weizmann e a Teoria da genética em populações de Mendel.

Na simulação da recombinação e da evolução genética natural, o GA inicialmente codifica os problemas a serem solucionados. Cada unidade no código é chamada de "gene"; um grupo de genes pode formar um "cromossomo", conhecido como indivíduo; vários cromossomos serão processados repetidamente pelos operadores como seleção, cruzamento e mutação até chegar um valor com solução ótima ou subótima. O processo básico de operação com o uso do GA para resolver problemas é mostrado na figura a seguir:



Fonte: fei e Xiu, 2015

1. **Código:** o código é a questão principal a ser resolvida pelo GA, o que pode tornar numa solução prática solução dos dados da estrutura genética;
2. **Geração das espécies iniciais:** Geração aleatória das espécies iniciais formados por genes K;
3. **Avaliação individual:** Avaliação das vantagens e desvantagens de cada cromossomo com funções adaptativas;
4. **Cálculo da seleção:** O objetivo da seleção é transmitir diretamente os resultados para a próxima geração com base na avaliação adaptativa individual.
5. **Cálculo cruzado:** Parte de duas estruturas individuais pai são substituídas e recombinada para gerar um novo indivíduo. Cruzamento é a operação de troca de informações entre os pais;
6. **Cálculo da variação:** altera os genes de alguns cromossomos individuais na espécie para aumentar a variedade da espécie e melhorar a habilidade de pesquisa local do algoritmo;
7. **Determinação das condições de terminação:** O algoritmo termina quando a adaptabilidade do indivíduo ideal atinge o valor limite determinado ou a adaptabilidade do indivíduo ideal e a espécie para de subir (Li L, Horrocks I., 2004).

APRENDIZADO PROFUNDO (Deep learning)

O aprendizado profundo ("DL") é um subtipo de aprendizado de máquina. O DL pode processar uma gama mais ampla de recursos de dados, requer menos pré-processamento de dados por

humanos (por exemplo, rotulagem de recursos) e às vezes pode produzir resultados mais precisos do que as abordagens tradicionais de ML (embora exija uma quantidade maior de dados para fazê-lo).

No entanto, é computacionalmente mais caro na hora de executar, custos de hardware e quantidades de dados. O subconjunto de aprendizado de máquina composto por algoritmos que permitem ao software para treinar-se para executar tarefas, como fala e reconhecimento imagem expondo várias camadas de redes neurais a vastas quantidades de dados.

O deep learning combina o poder da computação das máquinas com os padrões de conexão das redes neurais para entender os relacionamentos complexos como os diagnósticos médicos e o reconhecimento da localização física.

Um sistema de aprendizado profundo é autodidata, aprendendo ao filtrar informações através de múltiplas camadas ocultas, de maneira semelhante aos seres humanos. Sem redes neurais, não haveria aprendizado profundo. E isso significa apenas uma coisa - se você deseja dominar o aprendizado profundo, começa com redes neurais!

A confiabilidade preditiva da aprendizagem máquina depende da qualidade dos dados históricos que foram inseridos.

Eventos novos e imprevistos podem criar resultados inválidos se não forem identificados ou ponderados inadequadamente. Como um resultado, os vieses humanos podem desempenhar um papel importante papel no uso do aprendizado de máquina.

Tais vieses podem afetar quais conjuntos de dados são escolhidos para o treinamento da IA, os métodos escolhidos para o processo, e a interpretação da saída.

Finalmente, embora o aprendizado de máquina tenha grande potencial, seus modelos ainda estão atualmente limitados por muitos fatores, incluindo armazenamento e recuperação de dados, poder de processamento, modelagem algorítmica suposições e compreensão humana e julgamento.

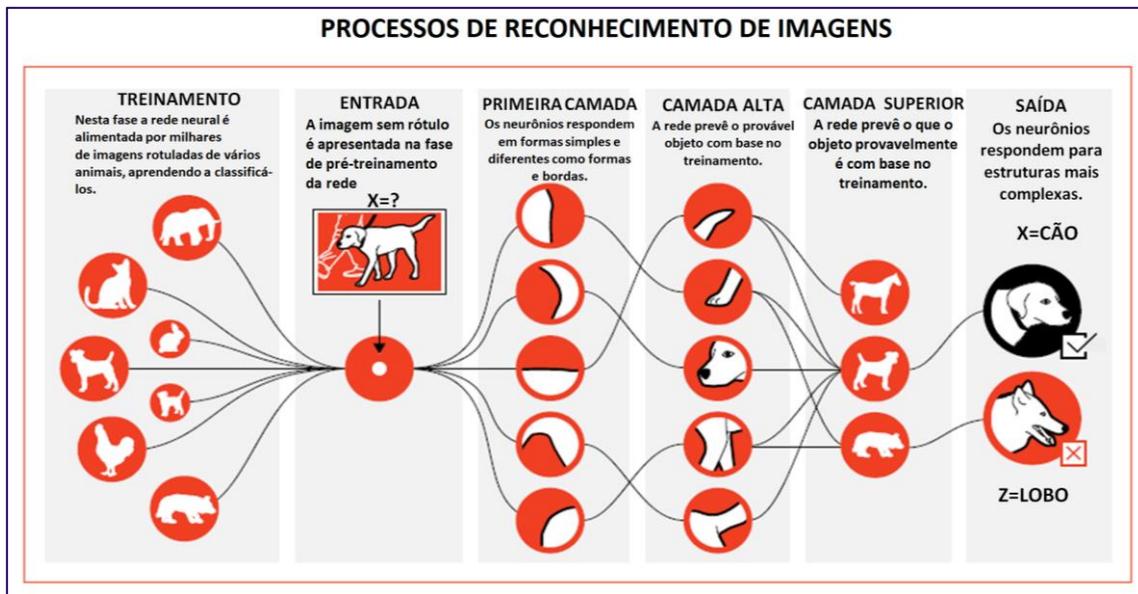
O principal componente dos sistemas de DL - redes neurais artificiais - começou a tomar forma nos anos 40, observando avanços significativos nos anos 60 e em cada década a partir de então.

O uso da DL acelerou nas últimas décadas. Isso é devido:

- a disponibilidade de hardware de computador mais barato, mas cada vez mais poderoso; e
- o crowdsourcing de ricos conjuntos de dados via Internet, o que ajuda a criar, capturar e selecionar os conjuntos de dados rotulados necessários em grande escala.

Na aprendizagem profunda, as camadas interconectadas de calculadoras baseadas em software conhecidas como "neurônios" formam uma rede neural. A idéia é replicar um entendimento abstrato de como acreditamos que o cérebro humano possa processar informações semelhantes e aprender com os arredores e com as informações sensoriais.

A rede ingere grandes quantidades de dados de entrada, processando-os através de várias camadas de neurônios que aprendem recursos cada vez mais complexos dos dados em cada camada:



Fonte: Parloff, 2018

A rede pode então:

- Fazer uma determinação sobre os dados, por exemplo foto X = cão; aprende se sua determinação está correta, isto é, a foto X é de fato cão; e usa o que aprendeu para fazer determinações sobre novos dados, ou seja, foto X = cachorro, mas foto Z = não cachorro.
- Redes neurais formam um tópico enorme e complexo.
- Não é tão difícil criar uma a partir do zero e ainda mais fácil criar uma usando bibliotecas de códigos pré-criadas, mas os conceitos matemáticos subjacentes podem levar algum tempo para total entendimento.

Como a máquina e o aprendizado profundo diferem?

DL é um subcampo de ML. Nesse sentido, ML e DL compartilham muitas características, incluindo sua capacidade de detectar padrões a partir de dados e aprender a executar uma tarefa específica com a melhoria do desempenho ao longo do tempo, com mais informações.

No entanto, existem muitas diferenças entre aprendizado de máquina e aprendizado profundo, tanto em termos quanto em aplicação.

As principais diferenças que distinguem o aprendizado de máquina do aprendizado profundo são quatro:

- consumo de dados;
- hardware dedicado;
- extração de recursos; e
- casos de uso.

1. Consumo de dados

O DL requer um grande número de amostras rotuladas para ter sucesso. Por esse motivo, a explosão de dados nas últimas décadas permitiu o DL como uma técnica viável (junto com um hardware mais barato e mais poderoso, dos quais veja abaixo).

No entanto, a quantidade de dados não é suficiente: ela precisa ter a qualidade certa, ou seja, rotulada. Nem todos os dados coletados são rotulados, rotulados corretamente ou de maneira adequada para DL. Esses dados também não são sempre acessíveis ao público.

Onde for esse o caso, você (ou alguém em seu nome) deve realizar um exercício de rotulagem de dados, que é oneroso em termos de tempo e dinheiro e geralmente requer um conjunto definido e rigoroso de procedimentos, controles de qualidade e experiência no domínio.

Esse fato e seu impacto na utilidade da DL para problemas do mundo real são frequentemente subestimados nas discussões sobre a DL (e em um grau semelhante nas discussões sobre a ML).

2. Hardware dedicado

A fase de treinamento dos sistemas de DL geralmente requer hardware dedicado, como GPUs (unidades de processamento gráfico) para reduzir o tempo de execução a algo gerenciável, ou seja, horas, dias ou semanas versus anos. Esses sistemas, embora cada vez mais baratos, ainda são caros versus as necessidades de configurações mais simples de ML.

3. Extração de Recursos

A extração de recursos (também conhecida como engenharia de recursos) é o processo de colocar o conhecimento do domínio na criação de extratores de recursos para reduzir a complexidade dos dados e tornar os padrões mais visíveis aos algoritmos de aprendizado.

Esse processo é difícil e caro em termos de tempo e experiência.

Isso é melhor explicado através de um exemplo:

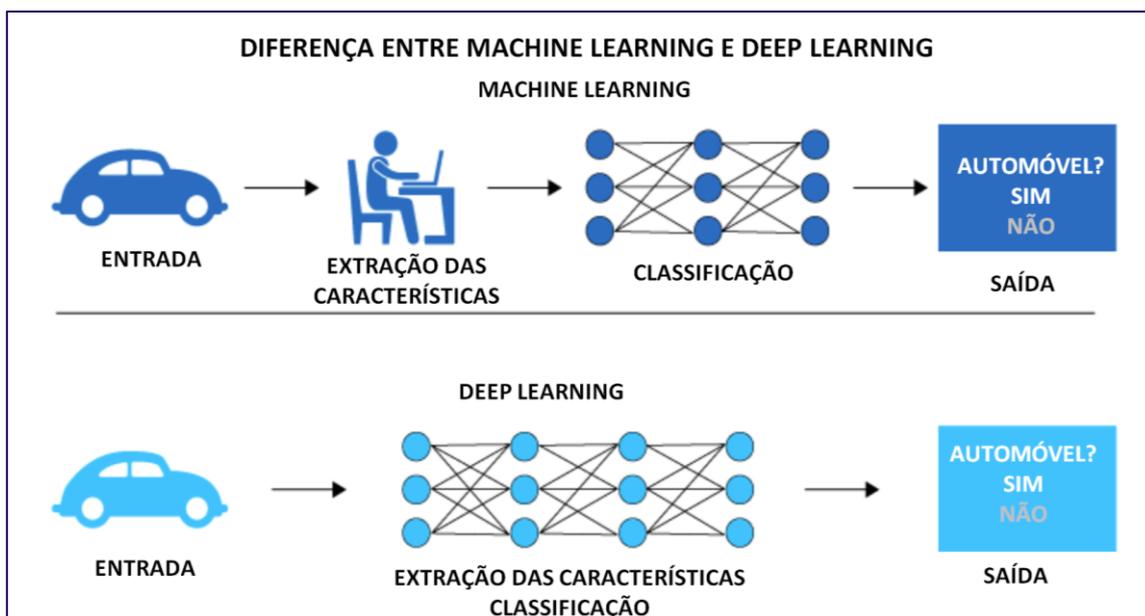
Suponha que você esteja criando um sistema que aprenderá a classificar as imagens como Automóvel ou Não Automóvel.

No ML clássico, a abordagem algorítmica usará os dados para saber se a imagem é Carro ou Não Carro. Para ajudar isso, pode ter sido possível para um ser humano rotular características constituintes indicativas de Carro (por exemplo, rodas) nas imagens, fornecendo recursos extras com os quais o sistema pode avaliar.

Por outro lado, uma solução DL também tentará determinar quais partes da imagem compõem o carro, por exemplo, rodas, espelhos retrovisores, faróis, pára-brisas, etc.

Como resultado, o DL pode reduzir a quantidade de códigos que os humanos precisam aplicar para definir recursos nos conjuntos de dados. Essa é a diferença entre rotular a imagem como Carro vs. Carro e rotular outros dados sobre a imagem, como rodas, pára-brisas, espelho retrovisor, etc., que indicam Carro ou Não.

Este diagrama é um resumo útil da distinção principal entre aprendizado de máquina e aprendizado profundo:



Casos de Uso

O custo do DL é maior em termos de:

- tempo - tanto em relação ao tempo de configuração quanto ao tempo de execução;
- hardware; e
- dados.

Como resultado, podemos generalizar e dizer que DL e ML são melhor usados da seguinte maneira:

	MACHINE LEARNING	DEEP LEARNING
Volume de dados	Para volume pequeno de dados	Para volume grande de dados
Custo Computacional	Menor custo de hardware e operação	Maior custo de hardware e operação
Adaptabilidade	Técnicas de ML requer domínio, aplicativo e recursos de engenharia específicos são necessárias para criar modelos de ML de alto desempenho. Portanto, os modelos resultantes são menos adaptáveis, mesmo em domínios semelhantes.	As técnicas de DL e aplicativos podem ser adaptadas a diferentes domínios com muito mais facilidade do que os algoritmos de ML. Por exemplo, uma vez que se entende a teoria subjacente do aprendizado profundo para o domínio do reconhecimento de fala, aprender como aplicar redes profundas ao processamento de linguagem natural não é muito desafiador, pois o conhecimento da linha de base é bastante semelhante.
Característica de Engenharia	O ML geralmente requer engenharia de recursos complexos, que é dispendiosa em termos de tempo e de experiência no domínio da contratação ou contratação.	As técnicas de DL podem eliminar ou reduzir a necessidade de extração complexa de recursos, reduzindo assim o tempo e o custo dessa etapa, embora potencialmente em troca de mais despesa em tempo e dinheiro em relação ao hardware e tempo de execução das abordagens de DL.
Interpretabilidade	Devido à engenharia de recursos e aos modelos mais simples, os sistemas ML são geralmente mais fáceis de interpretar do que o DL. Em outras palavras, é mais fácil entender como e por que o algoritmo ML chegou a um resultado. Isso pode ser incrivelmente útil, sem mencionar o necessário (por exemplo, em ambientes regulamentados), para desenrolar e corrigir um sistema que produz resultados incorretos em circunstâncias inesperadas.	Menor capacidade de interpretação. Frequentemente vistos como sistemas de "caixa preta", os pesquisadores lutam para relaxar e explicar como e por que o sistema alcançou um resultado específico. Dito isto, continuam a haver desenvolvimentos significativos nessa área que estão abrindo a caixa preta, portanto essa distinção pode diminuir com o tempo.

Qual é a relação de ML e DL com IA?

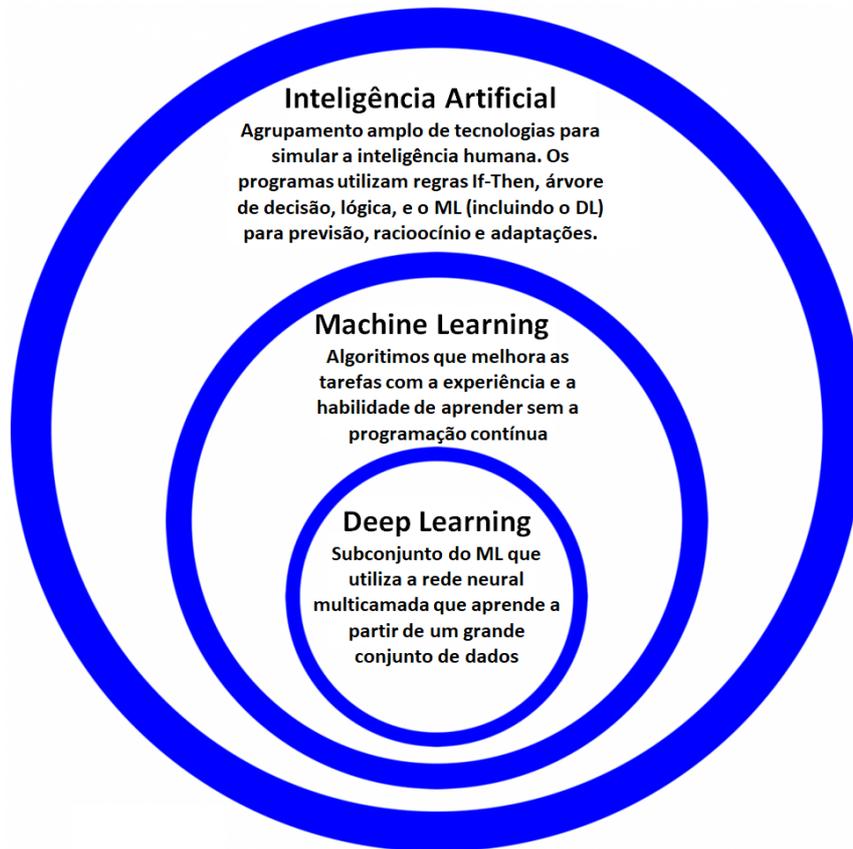
O DL é um subtipo de ML e cada um é um subtipo de IA onde o ML e DL são as principais técnicas (juntamente com regras e pesquisa) por meio das quais a IA de hoje acontece. Não existe "outro" I.A. para falar em termos práticos, embora haja muitos meios teóricos pelos quais podemos criar IA. (incluindo meios biológicos ou biológicos híbridos, conforme explorado no excelente livro de Nick Bostrom, Superintelligence).

Com isso, entendemos a capacidade de um IA para aprender a executar qualquer tarefa específica característica da inteligência humana versus ser especialista em uma tarefa e apenas uma tarefa. No caso do primeiro, dizemos o IA. é capaz de generalizar sua capacidade de um domínio para outro, enquanto no último isso simplesmente não é possível.

Através dessa lente (ou seja, as extremidades), todo o IA de hoje. - se ML ou DL é usado (isto é, os meios) - são Inteligência Artificial Estreita ("ANI").

Isso ocorre porque eles fazem uma coisa bem e apenas uma coisa. O Alpha Go do Google DeepMind só pode jogar Go, mas não pode jogar xadrez e, portanto, é um ANI, apesar de seus feitos técnicos incrivelmente impressionantes.

Uma boa maneira de juntar isso é o diagrama abaixo:



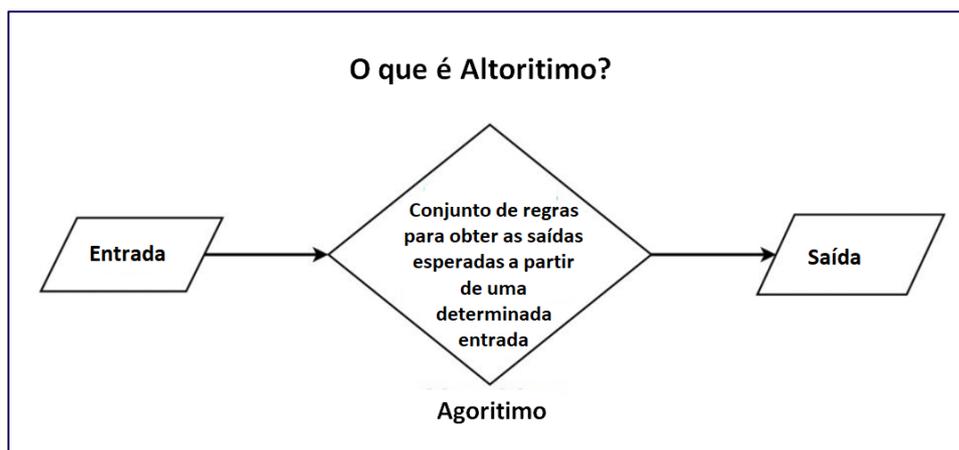
ALGORITMOS

Algoritmo é um conjunto ou arranjo de instruções implementadas por um humano ou um computador para executar um processo. Essas instruções ajudam na resolução de um problema complexo ou na computação de dados. Os computadores devem seguir estas instruções para fazer um cálculo ou executar operações na solução de problemas.

O nome algoritmo foi concebido a partir do intelectual de Bagdá, Al Khwarizmi¹ que foi o primeiro a introduzir os algoritmos precisos e inequívocos no mundo mecânico.

Em termos didáticos, um algoritmo é uma solução passo a passo bem definida ou uma série de instruções para resolver um problema.

A palavra Algoritmo significa "um processo ou conjunto de regras a serem seguidas em cálculos ou outras operações de solução de problemas".



Portanto, o algoritmo refere-se a um conjunto de regras/instruções que definem passo a passo como um trabalho deve ser executado para obter os resultados esperados (Mittal, 2020).

Algoritmos são necessários para processos simples, como multiplicação de dois números ou problemas complexos, procurando o conteúdo de um arquivo compactado. Os programadores tendem a desenvolver bons algoritmos para que os problemas sejam solucionados da melhor forma.

Um algoritmo deve:

¹ Mohammad ibn Musa al-Khwarizmi (780–850 AC), foi um matemático, astrônomo e geógrafo persa e um estudioso da Casa da Sabedoria em Bagdá durante o califado de Abbasid. Algumas palavras refletem a importância das contribuições de al-Khwarizmi à matemática como "álgebra" que é derivada de al-jabr, uma das duas operações que ele usou para resolver equações quadráticas. Assim os termos *Algorism* e *algorithm* derivaram do *Algoritmi*, que é a forma latina do seu nome.

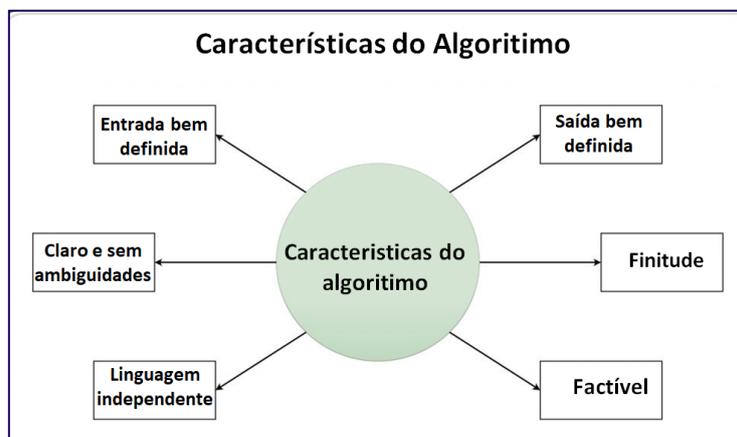
- Ser bem definido e ordenado - instruções fornecidas para o algoritmo devem ser compreensíveis e bem definidas.
- Ter operações inequívocas, ou seja, cada uma das etapas deve ser simples o suficiente para que não exija mais simplificação.
- Ter operações efetivamente computáveis.

Um algoritmo de computador é um procedimento computacional que transforma um conjunto de entradas finitas em saídas com a aplicação dos recursos da matemática e da lógica.

O algoritmo projetado é independente do idioma, ou seja, são apenas instruções simples que podem ser implementadas em qualquer idioma e, no entanto, a saída será a mesma, conforme o esperado.

QUAIS SÃO AS CARACTERÍSTICAS DE UM ALGORITMO?

Como não segue nenhuma instrução escrita como receita, mas apenas um padrão pois nem todas as instruções escritas para programação representam um algoritmo. Para que algumas instruções sejam um algoritmo, ele deve ter as seguintes características:



Claro e inequívoco: Cada uma das etapas deve ser clara em todos os aspectos e deve levar a apenas um significado.

Entradas bem definidas: Cada entrada que o algoritmo recebe deve ser entradas definidas.

Saídas bem definidas: O algoritmo deve definir claramente qual saída será gerada.

Finitude: O algoritmo deve ser finito, ou seja, não deve terminar em um loop infinito ou similar.

Factível: O algoritmo deve ser simples, genérico e prático, de modo que possa ser executado conforme os recursos disponíveis. Não deve conter tecnologia futura e sim do presente.

Independente de idiomas: o algoritmo projetado deve ser independente de idioma, ou seja, deve ser apenas instruções simples que podem ser implementadas em qualquer idioma e, no entanto, a saída será a mesma, conforme o esperado.

Um algoritmo de programação possui várias etapas como:

- Definição do problema - O que deve ser feito?
- Coleta de dados - O que temos para resolver o problema? Ou entradas.
- Processamento de dados - Entendendo e transformando-os numa planilha utilizável.
- Abordagem lógica – conjunto lógico para formular uma solução a partir dos dados.
- Solução – apresentada de maneira desejada num terminal GUI(graphic user interface), diagrama ou gráfico.

Para resumir, para um dado de entrada de valor finito x , o algoritmo transforma num valor de saída efetivo y , onde y é $f(x)$ para alguma função bem definida f .

Outro aspecto importante é que os algoritmos não estão estritamente vinculados a nenhuma linguagem de programação. São soluções genéricas como tais.

COMPONENTES DO ALGORITMO

Geralmente, os algoritmos compreendem quatro componentes:

- 1. Entrada ou feedback:** como os algoritmos são sequências de etapas a serem seguidas para obter alguma saída, precisamos fornecer algumas entradas para o algoritmo.
- 2. Uma sequência de algoritmos:** seqüenciamento das etapas presentes no algoritmo de acordo com a saída que queremos para organizar as etapas presentes no algoritmo.
- 3. Condição:** precisamos verificar a condição de acordo com a entrada.
- 4. Repetição:** às vezes, existem etapas que precisamos repetir até atingir os resultados desejados. Por exemplo, se eu quiser pesquisar um registro em arquivos, temos que comparar nossa entrada com arquivos diferentes até descobrirmos a saída desejada. Nesse caso, as mesmas etapas são repetidas até a saída (Arquivos).

CARACTERÍSTICAS DO ALGORITMO

Ajustabilidade: Qualquer algoritmo deve ser concluído num determinado momento, caso contrário entrará em estado infinito e nunca será concluído.

Definição: cada algoritmo deve ser claro e inequívoco.

Eficácia: Cada linha de um algoritmo deve ser essencial. Deve ser simples para que seja fundamentada.

Entrada: Todo algoritmo deve receber entrada, pode inclusive ser zero ou de acordo com as condições que o usuário pode usar.

Resultado: é necessário para qualquer trabalho que venha a ser produzido; caso contrário, não há significado para nenhum trabalho pois deve gerar uma ou mais saídas.

Generalidade: O algoritmo deve ser comum para o conjunto de entradas e requisitos, para que o mesmo algoritmo possa ser aplicado em vários locais, de acordo com os requisitos do usuário.

VANTAGENS DOS ALGORITMOS

- Os algoritmos são muito fáceis de entender e podem ser escritos em linguagem simples, que pode ser entendida por qualquer pessoa
- Os algoritmos podem ser divididos em partes diferentes, que serão fáceis de implementar na prática.
- Usando algoritmos, podemos entender facilmente a sequência a ser seguida no processamento.

DESVANTAGENS DOS ALGORITMOS

- Dificuldade em converter a tarefa complexa em algoritmos adequados.
- Demora no processo, porque precisamos gastar tempo adequado para escrever algoritmos e, posteriormente, precisamos implementar em uma linguagem de programação.
- É complicado mostrar funcionalidades para cada etapa da introdução aos algoritmos e é difícil entender cada fluxo no termo loop e ramificação.

O QUE É A COMPLEXIDADE DO ALGORITMO E COMO ENCONTRÁ-LO?

Um algoritmo é definido como complexo com base na quantidade de espaço e tempo que consome na sua execução que são caracterizados pelo fator eficiência.

Refere-se à medida do tempo consumido para executar e obter a saída esperada e o espaço necessário para armazenar todos os dados (entrada, dados temporários e saída).

Fator de tempo: o tempo é medido contando o número das principais operações comparados no algoritmo de classificação.

Fator de espaço: O espaço é medido contando o espaço máximo de memória exigido pelo algoritmo que pode ser dividida em dois tipos:

- A complexidade do espaço de um algoritmo é calculada determinando os seguintes 2 componentes:
 - **Parte fixa:** refere-se ao espaço que é definitivamente requerido pelo algoritmo. Por exemplo, variáveis de entrada, variáveis de saída, tamanho do programa etc.
 - **Parte variável:** refere-se ao espaço que pode ser diferente com base na implementação do algoritmo. Por exemplo, variáveis temporárias, alocação dinâmica de memória, espaço de pilha de recursão etc.

Portanto, a complexidade espacial $S(P)$ de qualquer algoritmo P é $S(P) = C + SP(I)$, onde C é a parte fixa e $S(I)$ é a parte variável do algoritmo, que depende da característica da instância I .

ALGORITMOS DE APRENDIZADO DE MÁQUINA

Os algoritmos são utilizados para modelar um problema de diferentes maneiras de acordo com os dados de entrada onde são definidos os algoritmos adequados conforme os estilos de aprendizado.

Desta forma é muito importante examinar as funções dos dados de entrada e no processo de preparação do modelo para cada problema e obter desta forma o melhor resultado.

Existem duas maneiras de categorizar os algoritmos de Machine Learning.

1. O primeiro é um agrupamento pelo estilo de aprendizado.
2. O segundo é um agrupamento por uma similaridade na forma ou função.

Geralmente, ambas as abordagens são úteis.

Vamos examinar os algoritmos conforme os estilos diferentes de aprendizado da máquina detalhados no capítulo anterior:

APRENDIZADO DE MÁQUINA: TIPOS DE ALGORITMOS

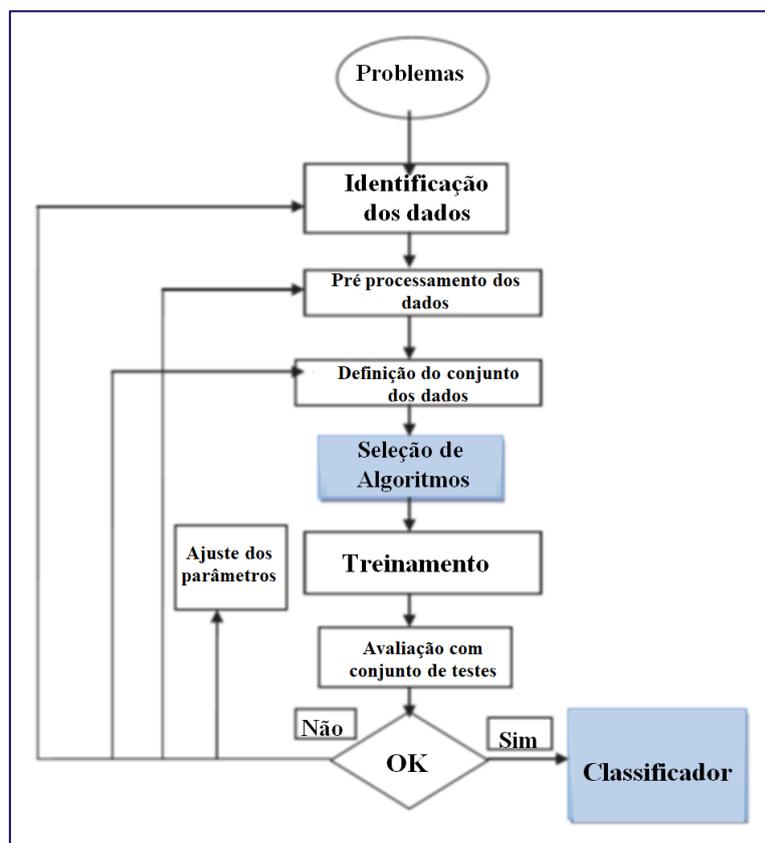
Os algoritmos de aprendizado de máquina são organizados em taxonomia, com base no resultado desejado do algoritmo. Os tipos de algoritmo comuns incluem:

- **Aprendizado supervisionado:** o algoritmo gera uma função que mapeia entradas conforme as saídas desejadas. Uma formulação padrão da tarefa de aprendizado supervisionado é o problema de classificação: o treinando deve aprender (para aproximar ao comportamento) de uma função que mapeia um vetor em uma das várias classes observando várias exemplos de entrada e saída da função.

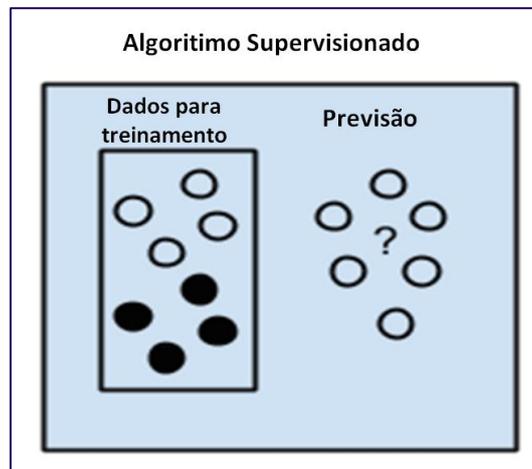
- **Aprendizado não supervisionado:** modela um conjunto de entradas: exemplos rotulados (conhecidos) não são acessíveis.
- **Aprendizado semi-supervisionado:** combina exemplos rotulados (conhecido) e não rotulados (conhecido) para gerar ou classificar apropriadamente uma função.
- **Aprendizado por reforço:** o algoritmo aprende uma política (jeito) de como agir a partir da observação do ambiente. Toda a ação tem algum impacto no meio ambiente e o ambiente fornece *feedback* que orienta o algoritmo de aprendizado.
- **Transdução:** semelhante ao aprendizado supervisionado, mas não constrói explicitamente uma função mas tenta prever novos resultados com base nos dados de treinamento de entrada.
- **Aprendendo a aprender:** o algoritmo aprende seu próprio viés indutivo com base na experiência anterior.

1. Aprendizado supervisionado

Neste estilo, os dados de entrada são chamados de dados de treinamento e têm um rótulo ou resultado conhecido como spam/não spam ou preço das ações ou as vendas do dia. Desta forma, o modelo é preparado pelo processo de treinamento.



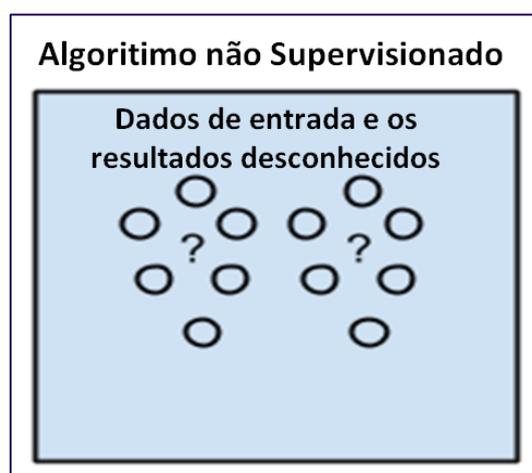
São geralmente utilizados para fazer previsões com treinamentos sucessivos até que modelo atinja o nível desejado.



- Aplicações mais usuais são problemas de classificação e regressão.
- Os algoritmos utilizados são a regressão logística e rede neural de propagação reversa.

2. Aprendizado não supervisionado

Os dados de entrada não são rotulados (conhecidos) e os resultados também não são conhecidos. Os modelos são preparados por dedução das estruturas presentes nos dados de entrada para extrair regras gerais. Geralmente são adotados processos matemáticos para reduzir a redundância.



- Exemplos de problemas são clustering, redução de dimensionalidades e aprendizado de regras por associação.
- Os algoritmos incluem o Apriori e k-Means.

3. Aprendizado por reforço

Num problema de previsão onde o resultado é desejado, o modelo deve aprender as estruturas para organizar os dados e fazer estas previsões.

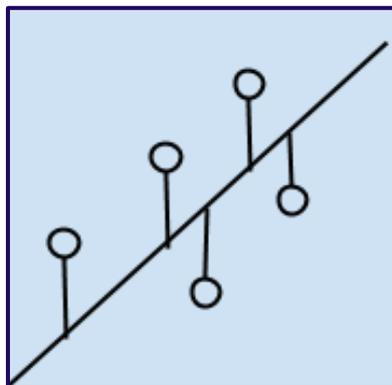
- Geralmente estes problemas são de classificação e regressão.
- Os algoritmos são as extensões para outros métodos flexíveis com base nas suposições sobre como modelar os dados não rotulados.

Algoritmos de exemplo são extensões para outros métodos flexíveis. Isso faz suposições sobre como modelar os dados não rotulados.

Algoritmos de aprendizado de máquina agrupados por similaridade

Algoritmos de ML são frequentemente agrupados por similaridades em termos da sua função.

1. Algoritmo de regressão



Algoritmos de regressão procura modelar o relacionamento entre variáveis para refinar usando a medida de erro nas previsões do modelo.

O método reforça o trabalho estatístico por meio da classe do problema e à classe do algoritmo.

A técnica estatística utilizada é a regressão linear, na qual o valor da variável dependente é previsto por meio de variáveis independentes. O relacionamento é formado pelo mapeamento da variável dependente e independente numa linha que é chamada linha de regressão representada por:

$$Y = a * X + b.$$

Onde:

Y = variável dependente (por exemplo, peso).

X = variável independente (por exemplo, altura)

a = inclinação

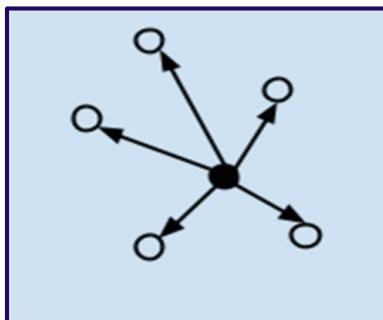
b = Interceptação.

Os algoritmos de regressão mais populares do Machine Learning são:

- Ordinary Least Squares Regression (OLSR)
- Linear Regression
- Logistic Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)

2. Algoritmos baseados em instâncias

Este modelo é um problema de decisão com instâncias com dados de treinamento nos bancos de dados que são consideradas recursos importantes ou necessário para sua realização.



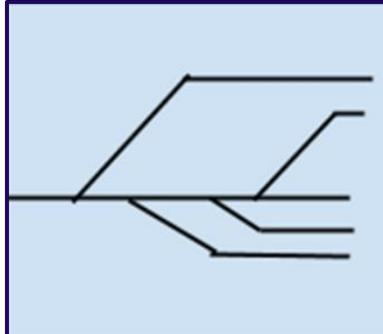
Trata-se de métodos que constroem banco de dados que comparam novos dados com os dados do banco de dados por similaridade para encontrar a melhor correspondência e fazer uma previsão. Por esse motivo, os métodos baseados em instâncias também são chamados de métodos “vencedor-leva-tudo” e aprendizado baseado em memorização. O foco é colocado na representação das instâncias armazenadas e assim, as medidas de similaridade usadas entre as instâncias.

Os algoritmos baseados em instâncias mais populares do Machine Learning são:

- k-Nearest Neighbor (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)

3. Algoritmos de Regularização

Trata-se de uma extensão para outro método devido à sua complexidade além de favorecer os modelos mais simples que são melhores para generalização.

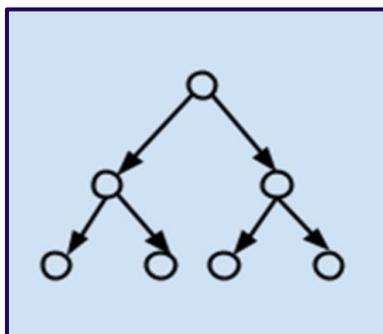


Os algoritmos de regularização mais populares e poderosos e geralmente são modificações simples feitas em outros no Machine Learning que são:

- Ridge Regression
- Least Absolute Shrinkage and Selection Operator (LASSO)
- Elastic Net
- Least-Angle Regression (LARS)

4. Os métodos da árvore de decisão

Trata-se da construção do modelo de decisão com base nos valores reais dos atributos nos dados. As decisões bifurcam-se nas estruturas das árvores até que uma decisão sobre a previsão seja tomada para um determinado registro. As árvores de decisão são treinadas em dados para problemas de classificação e regressão e geralmente são rápidas e precisas e um dos algoritmos preferidos no aprendizado de máquina.

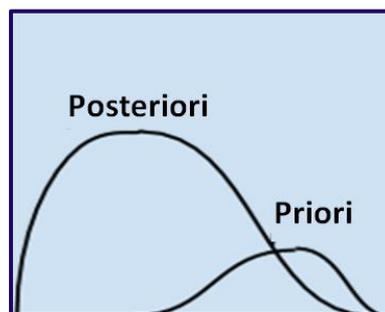


Os algoritmos mais populares da árvore de decisão no Machine Learning são:

- Classification and Regression Tree (CART)
- Iterative Dichotomies 3 (ID3)
- C4.5 and C5.0 (different versions of a powerful approach)
- Chi-squared Automatic Interaction Detection (CHAID)
- Decision Stump
- M5
- Conditional Decision Trees

5. Algoritmos Bayesianos

Esses métodos são aqueles que aplicam o Teorema de Bayes para problemas como classificação e regressão.

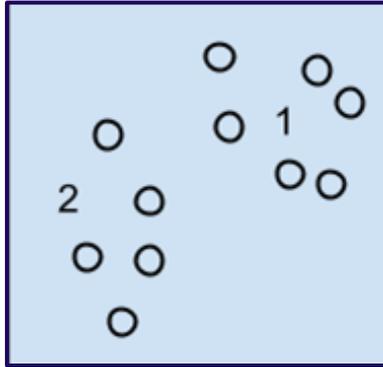


Os algoritmos bayesianos mais populares no Machine Learning são:

- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Averaged One-Dependence Estimators (AODE)
- Bayesian Belief Network (BBN)
- Bayesian Network (BN)

6. Algoritmos de Clustering

O agrupamento como a regressão descreve a classe do problema e a classe dos métodos. E são organizados pelas abordagens de modelagem como centróide e hierarquia. Todos os métodos se preocupam em usar as estruturas inerentes aos dados. Essa é uma necessidade de organizar melhor os dados em grupos de máxima semelhança.



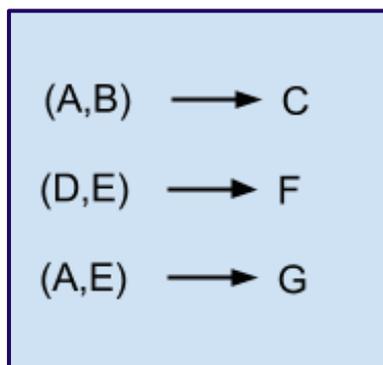
Os algoritmos de cluster mais populares no Machine Learning são:

- k-Means
- k-Medians
- Expectation Maximization (EM)
- Hierarchical Clustering

7. Algoritmos de aprendizado por regras de associação

Os métodos de aprendizado de regras de associação extraem regras pois isto explica melhor as relações observadas entre variáveis nos dados.

Essas regras podem descobrir associações importantes e úteis em grandes conjuntos de dados multidimensionais. Isso pode ser explorado por uma organização.

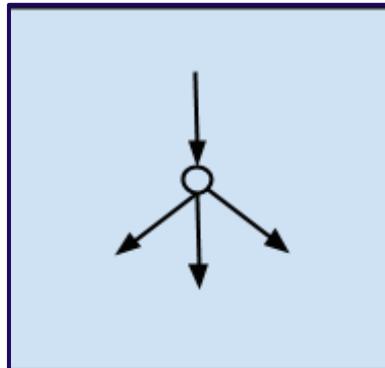


Os algoritmos mais populares de aprendizado de regras de associação no Machine Learning são:

- Apriorist algorithm
- Eclat algorithm

8. Algoritmos de redes neurais artificiais

Estes modelos são inspirados na estrutura das redes neurais biológicas.



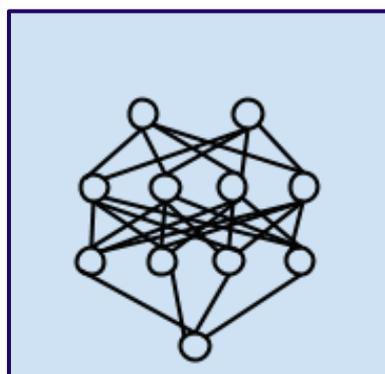
Formam uma classe de correspondência de padrões semelhantes aos que usamos para problemas de regressão e classificação. Embora exista subcampos enormes e também combina centenas de algoritmos e variações.

Os algoritmos mais populares de aprendizado de redes neurais artificiais no Machine Learning são:

- Perceptron
- Back-Propagation
- Hopfield Network
- Radial Basis Function Network (RBFN)

9. Algoritmos Deep Learning

Os métodos de aprendizagem profunda são uma atualização moderna das redes neurais artificiais que procura explorar a computação abundante de baixo custo e redes neurais muito maiores e mais complexas.



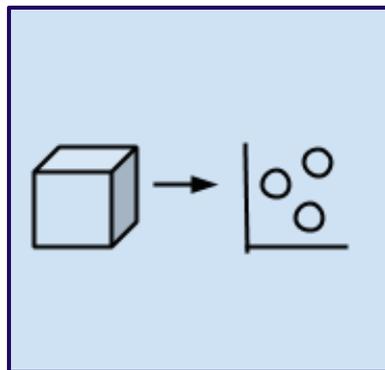
Os algoritmos mais populares de aprendizado de deep learning no Machine Learning são:

- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)
- Convolutional Neural Network (CNN)
- Stacked Auto-Encoders

10. Algoritmos de redução de dimensionalidade

Como nos métodos de agrupamento, a redução de dimensionalidade busca uma estrutura inerente nos dados mediante seus resumos.

Geralmente, pode ser útil visualizar dados dimensionais para utilização num método de aprendizado supervisionado. Muitos desses métodos são adotados em classificação e regressão.



Geralmente, pode ser útil visualizar dados dimensionais. Além disso, podemos usá-lo em um método de aprendizado supervisionado. Muitos desses métodos adotamos para uso em classificação e regressão.

Os algoritmos mais populares de aprendizado de redução de dimensionalidade no Machine Learning são:

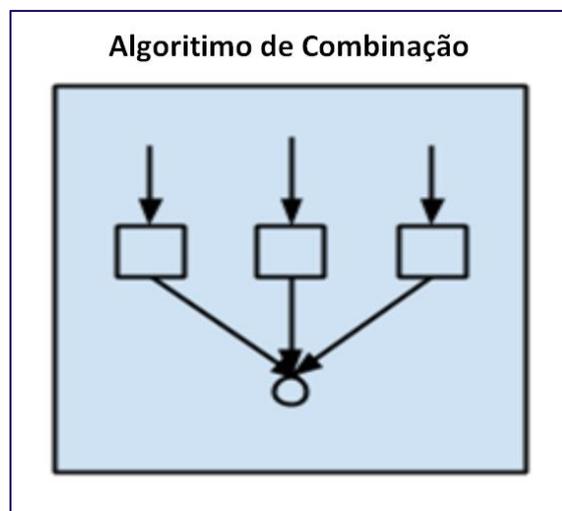
- Principal Component Analysis (PCA)
- Principal Component Regression (PCR)
- Partial Least Squares Regression (PLSR)
- Summon Mapping
- Multidimensional Scaling (MDS)
- Projection Pursuit
- Linear Discriminant Analysis (LDA)
- Mixture Discriminant Analysis (MDA)

- Quadratic Discriminant Analysis (QDA)
- Flexible Discriminant Analysis (FDA)

O que é aprendizagem por combinação?

Basicamente, esses métodos são compostos por modelos mais fracos. Além disso, à medida que são treinados são combinadas de alguma forma para fazer a previsão.

Além disso, são necessárias muitas tentativas para estabelecer a combinação dos tipos de treinandos fracos como eles podem ser combinados. Portanto, esta é uma classe muito poderosa de técnicas e, como tal, a sua adoção é bastante frequente.



O mundo do aprendizado de máquina onde os tipos de dados são variados (multimídia), heterogêneos, cheio de viés e acompanhados de ruídos onde requer também níveis variados e diferentes tipos de algoritmos.

Ao usamos o método tradicional de algoritmo de aprendizado de máquina para treinar um modelo, na maioria das vezes descobrimos que não temos uma boa precisão e encontramos um nível mais alto de variação ou perda de precisão no resultado. Requer desta forma estabelecer a meta, que é que uma combinação dos modelos de algoritmos. Nesta primeira fase são apresentados uma série de algoritmos de aprendizado de máquina por combinação, e também uma prévia para o aprendizado profundo.

No grupo dos algoritmos, vários modelos são simples e têm menos poder computacional e outros são complexos e exigem mais recursos da computação.

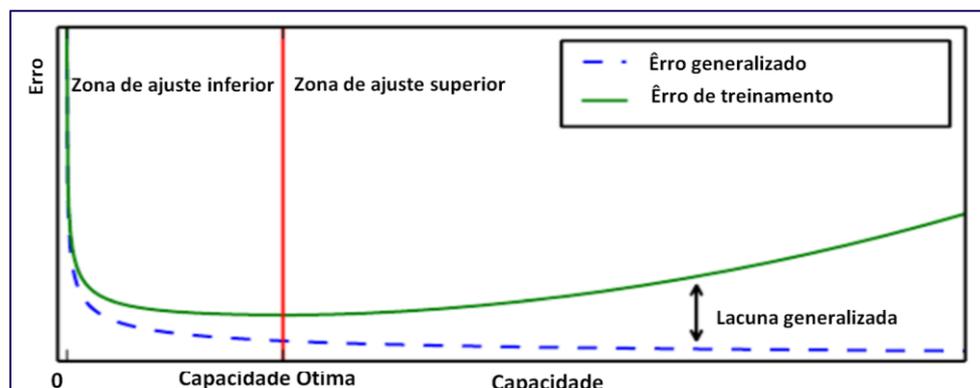
Num processo de treinamento de um modelo em que requer alta precisão nos resultados, o uso de um modelo simples (algoritmo simples) pode ocorrer resultados não tão precisos por não se ajustar corretamente aos dados. Nesses casos, comprometemos a precisão e o tempo de computação.

Para resolver esse problema, podemos treinar um modelo com diferentes algoritmos (aprendizado fraco) para obter um resultado médio (índice de confiança) ou implementar aplicativos em tempo real com uma precisão muito boa.

Os métodos de combinação utilizam vários algoritmos da base da árvore para construir um modelo que proporcione um melhor desempenho preditivo.

O princípio básico por trás do modelo por combinação é que um grupo de treinandos fracos são reunidos para formar um treinando forte, aumentando assim a precisão do modelo. Quando tentamos prever a variável de destino usando qualquer técnica de aprendizado de máquina, as principais causas de diferença nos valores reais e previstos são os ruídos, a variação e o viés. O métodos de combinação ajudam a reduzir esses fatores (exceto ruídos, que são erros irreduzíveis).

Se obtivermos a variância, ruído e o viés nos dados brutos na imagem ou em qualquer outro formato dos dados podemos ver que o modelo está bem ou mal ajustado. Esses são os motivos de se criar um grande impacto diretamente no modelo. Pois, o erro causado no treinamento e o erro de generalizado têm uma lacuna que é representada como a lacuna de generalização, onde mostra que o modelo está com ou sem ajuste.

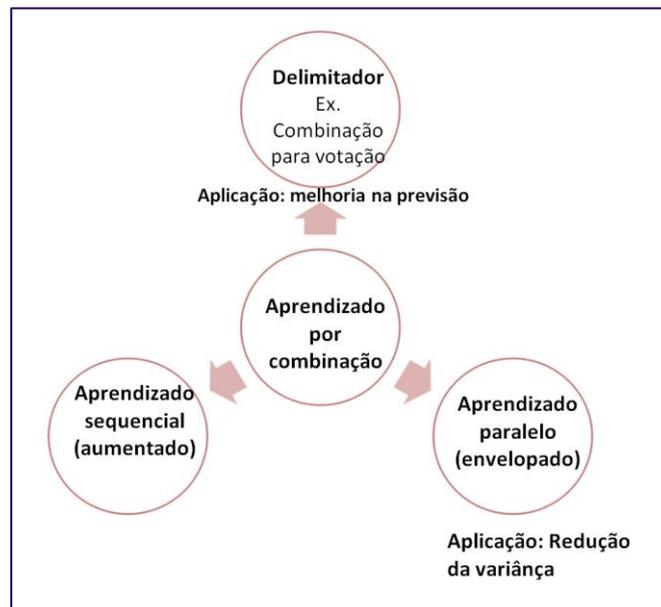


Razões para usar a combinação

- **O conjunto de dados é muito grande ou pequeno** – Requer o uso da amostragem para escolher o tamanho da amostra para obter a média do resultado.
- **Dados complexos (não lineares)** - O conjunto de dados em tempo real é não linear, portanto, quando treinamos um único modelo que não pode definir claramente o limite da classe e o modelo fica mal ajustado. Nesse caso, temos que pegar uma subamostra diferente e tirar uma média do modelo diferente.

- **Alta Confiabilidade** - quando treinamos um modelo com várias classes e obtemos alta produção correlacionada, e essa situação leva à alta confiabilidade, pois a maioria do modelo prevê a mesma classe de confiança.

TIPO DE APRENDIZADO POR COMBINAÇÃO



Os algoritmos por combinação são divididos em três categorias principais de base para o seu uso:

	Envolvamento	Potencialização	Empilhamento
Posicionamento dos dados no subconjunto	Aleatório	Fornecer testes sem classificação como prioridade alta	Vários
Objetivo	Minimização da variância	Aumento da força preditiva	Ambos
Área de aplicação	Subespaço aleatório	Gradiente descendente	Mistura
Combinação com modelos simples	Média ponderada	Ponderada de maior peso	Regressão logística

APRENDIZAGEM SEQUENCIAL POR COMBINAÇÃO (IMPULSO)

Impulso: é um meta-algoritmo do conjunto de aprendizado de máquina para reduzir principalmente o viés e, além disso, a variância no aprendizado supervisionado, e um grupo de algoritmos de aprendizado de máquina que convertem treinandos fracos em treinandos da cadeia. São métodos sequenciais de conjuntos em que os treinandos da base são gerados sequencialmente.

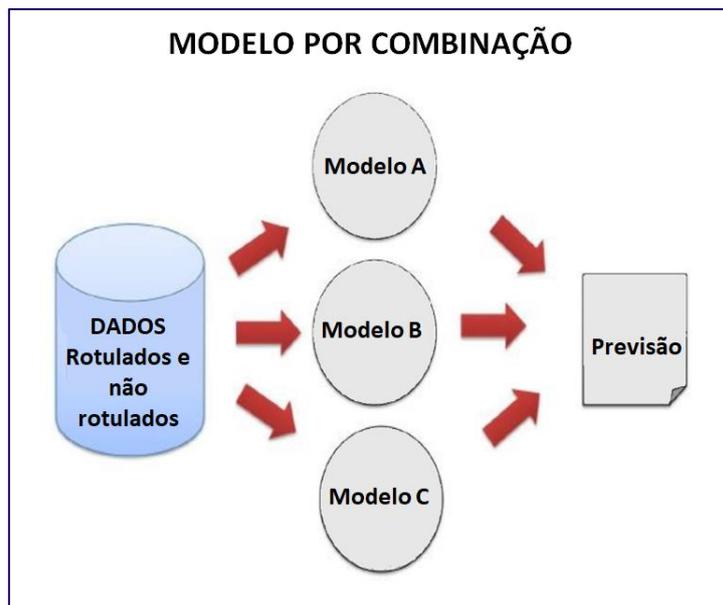
Exemplo: Adaboost, reforço estocástico de gradiente

Como criar um sistema por combinação

Todos os modelos devem ter uma diferença no nível da população. Assim temos que dividir o conjunto de dados em subconjuntos de maneira que cada subconjunto possui a menor correlação entre si. Assim, o modelo de combinação criará resultados diferentes e independentes.

Todo modelo deve fornecer resultados de hipóteses diferentes criando desta forma uma combinação generalizada.

Com base na categoria do modelo, podemos visualizar os dados para obter uma perspectiva diferente. Ele deve ser (linear ou não linear) ou (supervisionado ou não supervisionado).



Fonte: Patel, 2019

Quantificação de desempenho

O desempenho do modelo é calculado pela diferença entre a Entrada e a Saída.

$$Err(x) = (E[\bar{f}(x)] - f(x))^2 + E[\bar{f}(x) - E[\bar{f}(x)]]^2 + \sigma_e^2$$

↓ ↓ ↓ ↓ ↓

Previsto Verdadeiro Previsto Valor médio Erro irreduzível

Viés²: Diferença entre o valor real

Variância: Como a previsão varia para diferentes realizações do modelo

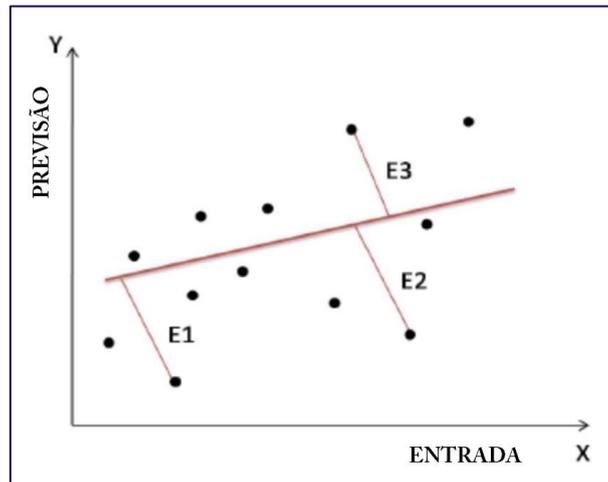
Erro: representa os três componentes básicos: Viés, Variância, e o Erro irreduzível

OBS: Não se pode lidar com erros irreduzíveis, mas pode se manipular as variâncias e os vieses.

Compromisso de variação de polarização

- **Erro de polarização: (baixo ajuste de polarização alta)**

Viés é a média da diferença entre o resultado previsto e o real. Alta polarização significa que estamos obtendo baixo desempenho.



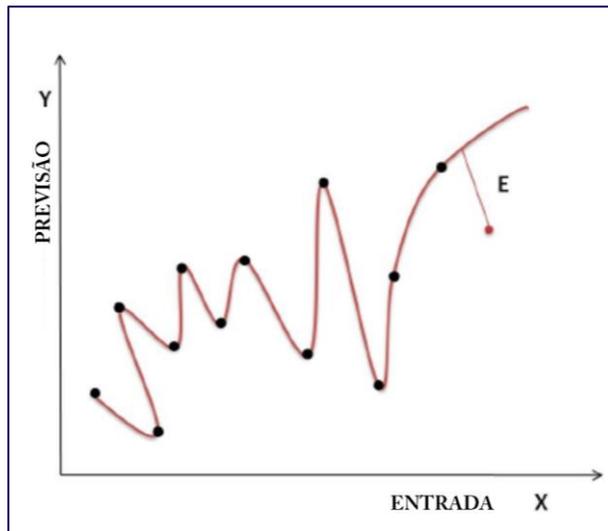
Quando treinamos dados lineares e complexos, podemos identificar o erro com os pontos pretos e as respectivas distância da linha do erro E1, E2, E3 em cor vermelha. O significado simples da situação é que o modelo não está se encaixando adequadamente e, portanto, apresenta baixo desempenho; nesse caso, temos que usar o modelo complexo (polinomial) para minimizar esse cenário.

- **Erro de variância: (Ajuste acima da previsão)**

- Quantifica a diferença do valor previsto na mesma observação em que o modelo é ajustada para cima.
- Treinar o modelo que apresenta a variância próximo de 100% de precisão nos dados de treinamento. Ao verificamos o modelo com os dados atuais, observa se falha ao prever o resultado correto.

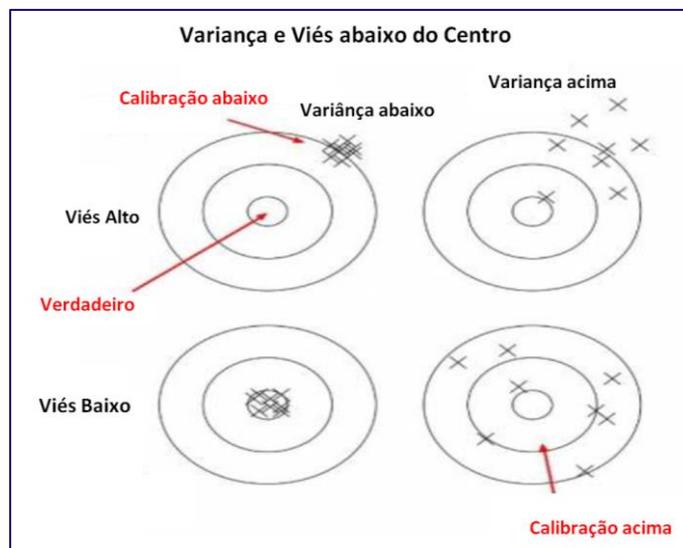
Nestas duas condições, podem ocorrer:

1. Necessário aumentar os dados de treinamento.
2. Uso do modelo complexo para dados simples.



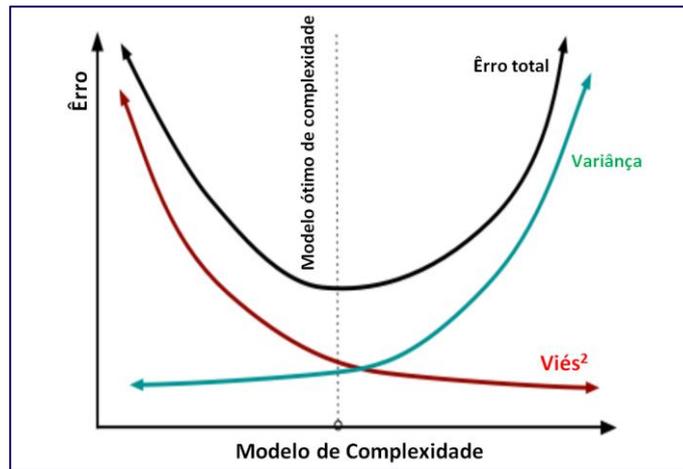
Na figura acima, podemos ver que nosso modelo tem um bom desempenho com os dados de treinamento dos pontos em preto. Mas, quando prevemos o valor do ponto em vermelho, vemos a diferença na forma da curva, que é um erro de variação.

O problema de ajuste excessivo pode ser resolvido aumentando o número de instâncias de treinamento ou escolhendo o classificador correto para previsão.



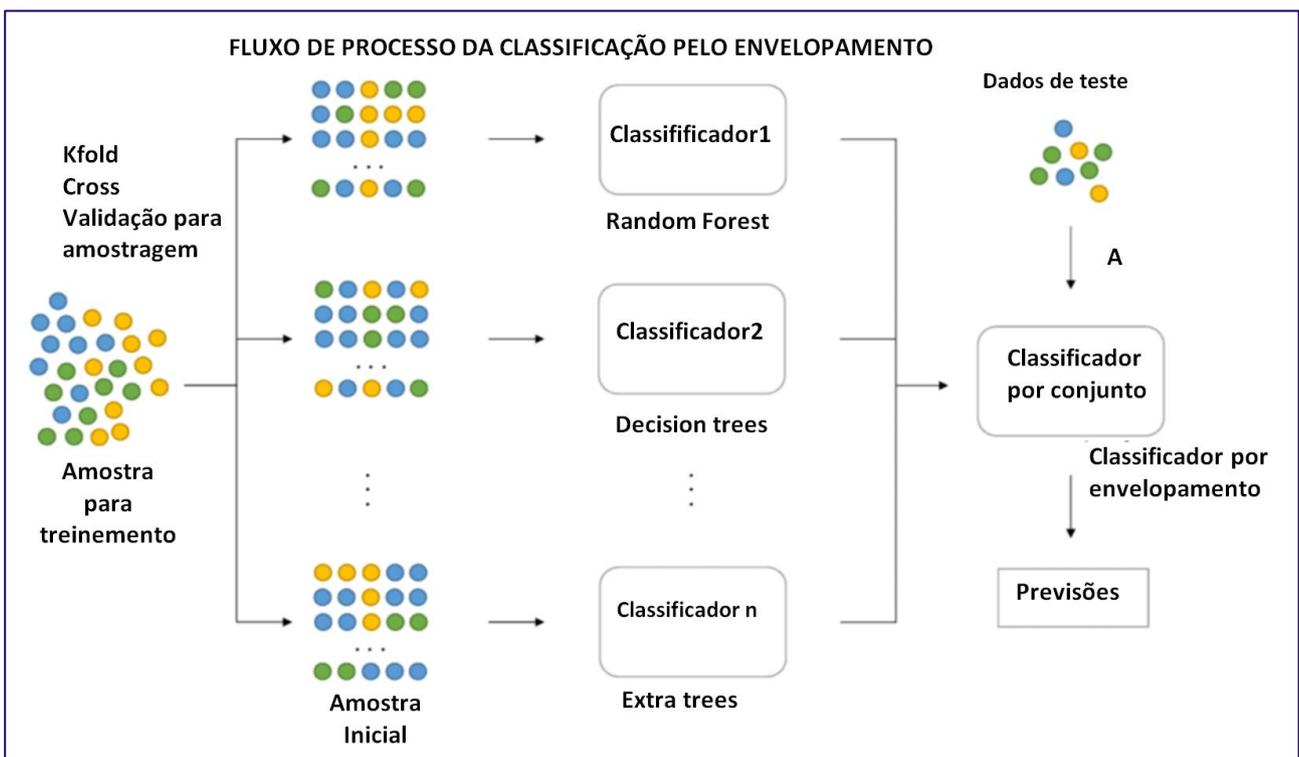
Na figura acima, imagine que o centro do círculo (O) contenha o valor real e a cruz (X) representa o valor previsto.

Quando temos um viés alto, aumentamos a complexidade do modelo, obtendo um viés baixo, portanto esperamos o resultado de baixo viés e baixa variação.



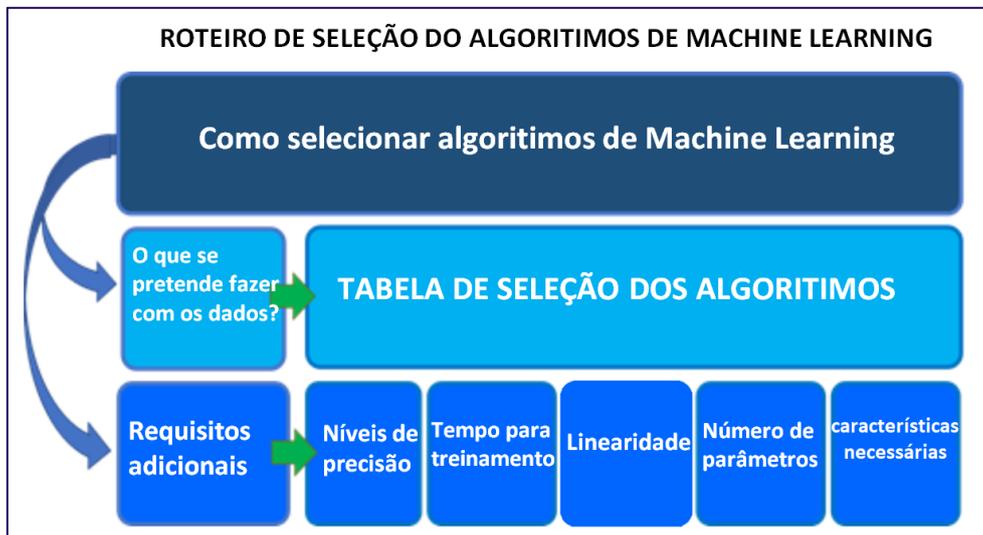
Acima temos o gerenciamento do viés e a variação de maneira equilibrada, onde se observa que o aprendizado do conjunto está ajustado.

Envolvimento-Meta-algoritmo do conjunto para reduzir a variância



Como selecionar o algoritmo de machine learning

Uma pergunta comum é "Qual algoritmo de aprendizado de máquina devo usar?"



Fonte: Adaptado de Microsoft, 2020

O algoritmo a ser selecionado depende principalmente de dois aspectos diferentes do cenário de ciência de dados:

O que se quer fazer com os dados?

Especificamente, qual é a pergunta comercial que você deseja responder aprendendo com os dados anteriores?

Quais são os requisitos do cenário de ciência de dados?

Especificamente, qual é a precisão, o tempo de treinamento, a linearidade, o número de parâmetros e o número de recursos suportados pela sua solução?

- **Precisão**

A precisão no aprendizado de máquina mede a eficácia de um modelo como a proporção de resultados reais em relação ao total de casos. Para tanto requer estabelecer um conjunto de métricas de avaliação padrão do setor.

Obter a resposta mais precisa nem sempre é possível e necessário. Às vezes, uma aproximação é adequada dependendo do caso desejado. Pois assim poderá reduzir drasticamente o tempo de processamento e também tendem naturalmente a evitar ajustes excessivos.

Existem três maneiras de avaliar o modelo:

- Gerar pontuações nos dados de treinamento
- Gerar pontuações no modelo comparando com as pontuações do conjunto reservado de testes
- Compare pontuações para dois modelos diferentes, mas relacionados, usando o mesmo conjunto de dados

- **Tempo de treino**

No aprendizado supervisionado, o treinamento significa usar dados históricos para criar um modelo de aprendizado de máquina capaz de minimizar os erros. O número de minutos ou horas necessárias para treinar um modelo varia bastante entre os algoritmos. O tempo de treinamento geralmente está intimamente ligado à precisão; um normalmente acompanha o outro.

Além disso, alguns algoritmos são mais sensíveis ao número de pontos de dados do que outros. Você pode escolher um algoritmo específico porque possui uma limitação de tempo, especialmente quando o conjunto de dados é grande.

O processo de criação e o uso do modelo de aprendizado de máquina geralmente ocorre em três etapas:

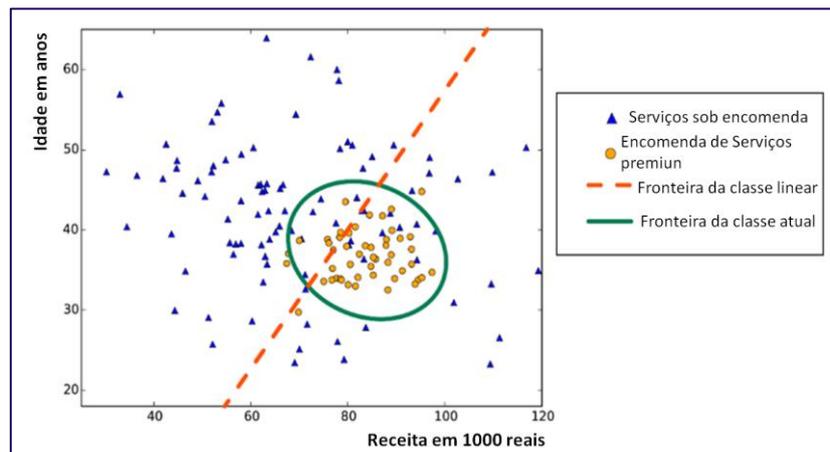
- **Configuração do modelo:** após a escolha do tipo específico de algoritmo para definir os parâmetros ou hiperparâmetros.
- **Fornecimento do conjunto de dados rotulados:** para ser compatíveis com o algoritmo.
- **Conclusão do treinamento:** o modelo treinado é utilizado com um dos módulos de pontuação para fazer previsões sobre novos dados.

- **Linearidade**

A linearidade nas estatísticas e no aprendizado de máquina significa que há um relacionamento linear entre uma variável e uma constante no seu conjunto de dados. Por exemplo, algoritmos de classificação linear assumem que as classes podem ser separadas por uma linha reta (ou seu análogo de maior dimensão).

Muitos algoritmos de aprendizado de máquina fazem uso da linearidade como:

Regressão logística multiclasse

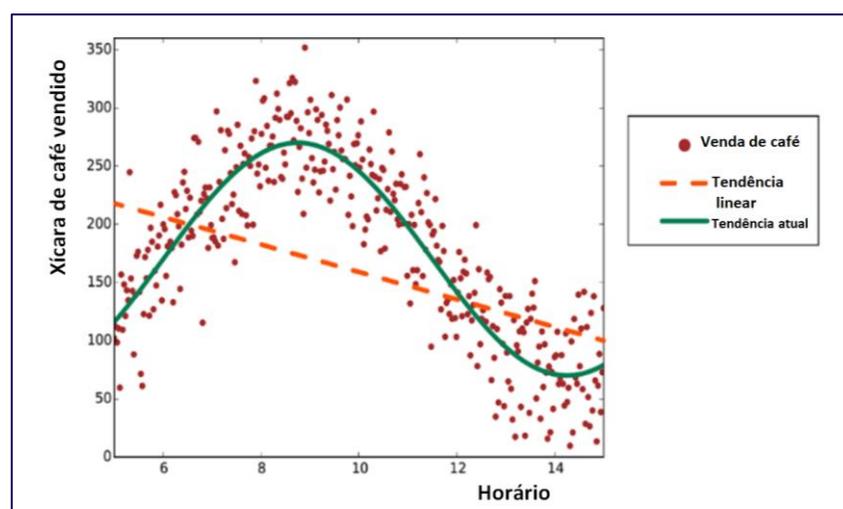


Regressão logística de duas classes

Máquinas de vetores de suporte

Os algoritmos de regressão linear assumem que as tendências de dados seguem uma linha reta. Essa suposição não é ruim para alguns problemas, mas para outros ela reduz a precisão. Apesar de suas desvantagens, algoritmos lineares são populares como primeira estratégia. Eles tendem a ser algoritmicamente simples e rápidos de treinar.

Limite de classe não linear: confiar num algoritmo de classificação linear resultaria em baixa precisão.



Dados com tendência não linear: o uso do método de regressão linear geraria erros muito maiores que o necessário.

- **Número de parâmetros**

Os parâmetros são os botões que um cientista de dados aciona ao configurar um algoritmo. São números que afetam o comportamento do algoritmo, como tolerância a erros ou número de iterações ou opções entre variantes de como o algoritmo se comporta. Às vezes, o tempo de treinamento e a precisão do algoritmo podem ser sensíveis à obtenção das configurações corretas. Normalmente, algoritmos com grande número de parâmetros exigem mais tentativa e erro para encontrar uma boa combinação.

Como alternativa, existe o módulo Tune Model Hyperparameters no designer de Machine Learning: O objetivo deste módulo é determinar os melhores hiperparâmetros para um modelo de machine learning. O módulo cria e testa vários modelos usando diferentes combinações de configurações. Ele compara métricas sobre todos os modelos para obter as combinações de configurações.

Embora essa seja uma ótima maneira de garantir a abrangência do espaço dos parâmetros, o tempo necessário para treinar um modelo aumenta exponencialmente com o número de parâmetros. A vantagem é que ter muitos parâmetros normalmente indica que um algoritmo tem maior flexibilidade. Muitas vezes, é possível obter uma precisão muito boa, desde que você encontre a combinação certa de configurações de parâmetros.

- **Número de funções**

No aprendizado de máquina, um recurso é uma variável quantificável do fenômeno que você está tentando analisar. Para certos tipos de dados, o número de recursos pode ser muito grande comparado ao número de pontos de dados. Geralmente, esse é o caso de dados genéticos ou de texto.

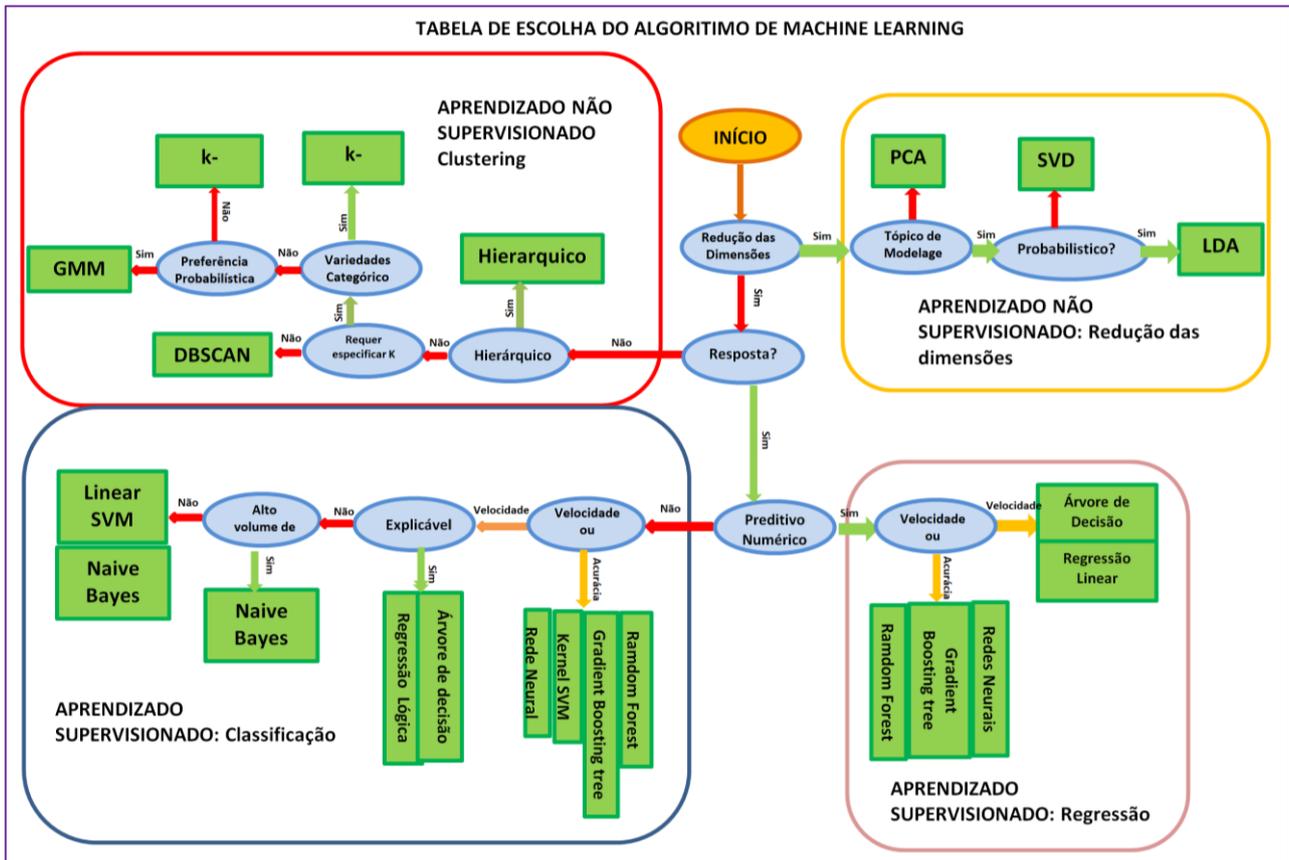
Um grande número de recursos pode atrapalhar alguns algoritmos de aprendizado, tornando o tempo de treinamento imprevisivelmente longo. As máquinas de vetores de suporte são particularmente adequadas para cenários com um grande número de recursos. Por esse motivo, eles foram usados em muitos aplicativos, desde recuperação de informações até classificação de texto e imagem. Máquinas de vetores de suporte podem ser usadas para tarefas de classificação e regressão.

A seleção de recursos refere-se ao processo de aplicação de testes estatísticos às entradas, dada uma saída especificada. O objetivo é determinar quais colunas são mais preditivas da saída. O módulo Seleção de recurso com base em filtro no designer do Machine Learning fornece vários algoritmos de seleção de recurso para você escolher. O módulo inclui métodos de correlação, como correlação de Pearson e valores do qui-quadrado.

Você também pode usar o módulo Importância do recurso de permutação para calcular um conjunto de pontuações de importância do recurso para o seu conjunto de dados. Assim aproveitar essas pontuações para ajudar a determinar os melhores recursos a serem usados em um modelo.

ALGORITMOS DE ML MAIS COMUNS

Nesta seção, aprenderemos sobre os algoritmos de aprendizado de máquina mais comuns.



Fonte: Adaptado de Patel, 2018

Nota:

DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

SVD (singular Value Decomposition)

DBM (Deep Boltzmann Machine)

GMM (Gaussian Mixture Models)

PCA (Principal Component Analysis)

Linear SVM (support vector machine)

Os algoritmos são descritos abaixo:

- **Regressão linear**

É um dos algoritmos mais conhecidos em estatística e aprendizado de máquina.

Conceito básico – Trata-se de um modelo linear que assume uma relação linear entre as variáveis de entrada x , e a variável de saída única y . Em outras palavras, podemos dizer que

y pode ser calculado a partir de uma combinação linear das variáveis de entrada x. A relação entre variáveis pode ser estabelecida ajustando uma melhor linha.

- **Tipos de regressão linear**

A regressão linear temos dois seguintes tipos:

- **Regressão linear simples:** Possui apenas uma variável independente.
- **Regressão linear múltipla:** Possui mais de uma variável independente.

A regressão linear é usada principalmente para estimar os valores reais com base nas variáveis contínuas.

Por exemplo, a venda total diária estimada de uma loja pode ser estimada por regressão linear com base no histórico de vendas.

Regressão logística

O grande volume de dados que cuja classificação é feita com a construção da equação. Este método é usado para encontrar a variável dependente discreta no conjunto de variáveis independentes com o objetivo de encontrar o melhor conjunto de parâmetros. Nesse classificador, cada recurso é multiplicado por um peso e todos são totalizados. Em seguida, o resultado é passado para a função sigmóide, que produz a saída binária. É um algoritmo de classificação e também conhecido como regressão *logit*. É usado principalmente para estimar valores discretos como 0 ou 1, verdadeiro ou falso, sim ou não com base num determinado conjunto de variáveis independentes. Basicamente, ele prevê a probabilidade, portanto, seu resultado fica entre 0 e 1.

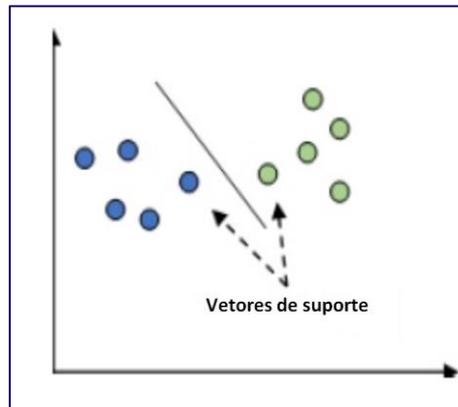
Árvore de decisão

A árvore de decisão é um algoritmo de aprendizado supervisionado usado principalmente para problemas de classificação.

Basicamente, é um classificador expresso como partição recursiva com base nas variáveis independentes. A árvore de decisão possui nós que formam a árvore raiz. Árvore enraizada é uma árvore direcionada com um nó chamado "raiz". A raiz não possui arestas de entrada e todos os outros nós têm uma aresta de entrada. Esses nós são chamados de folhas ou nós de decisão. Por exemplo, considere a seguinte árvore de decisão para ver se uma pessoa está em forma ou não.

Máquina de vetores de suporte (SVM)

É usado para problemas de regressão mas sobretudo para problemas de classificação. O conceito principal do SVM é plotar cada item de dados como um ponto no espaço n-dimensional, com o valor de cada recurso sendo o valor de uma coordenada específica. Aqui n seriam os recursos que teríamos. A seguir, é apresentada uma representação gráfica simples para entender o conceito de SVM:



No diagrama acima, temos dois recursos, portanto, em primeiro precisamos plotar essas duas variáveis no espaço bidimensional, em que cada ponto tem duas coordenadas chamadas de vetores de suporte. A linha divide os dados em dois grupos com diferentes classificações. Essa linha seria o classificador.

Naïve Bayes

Faz parte da técnica de classificação. A lógica por trás dessa técnica é o teorema de Bayes para construir classificadores. A suposição é que os preditores são independentes. Em palavras simples, assume que a presença de um recurso específico em uma classe não está relacionada à presença de qualquer outro recurso. Abaixo está a equação do teorema de Bayes -

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right)P\left(A\right)}{P\left(B\right)}$$

O modelo Naïve Bayes é fácil de construir e particularmente útil para grandes conjuntos de dados.

K-Nearest Neighbors (KNN)

É usado para regressão e é amplamente utilizado para resolver problemas de classificação. O principal conceito desse algoritmo é o armazenamento de todos os casos disponíveis e classifica novos casos por maioria de votos de seus vizinhos k . O caso é então atribuído à classe que é a mais comum entre os vizinhos K -mais próximos, medida por uma função de distância. A função de distância pode ser a distância euclidiana, Minkowski e Hamming. Considere o seguinte para usar KNN:

- Computacionalmente os KNN são mais custoso do que outros algoritmos usados para problemas de classificação.
- No processo de normalização das variáveis necessárias, as variáveis de maior intervalo podem influenciá-los.
- No KNN, precisamos trabalhar no estágio de pré-processamento, como remoção de ruído.

K-Means Clustering

Como o nome sugere, é usado para resolver os problemas de cluster. É basicamente um tipo de aprendizado não supervisionado. A sua principal lógica é classificar o conjunto de dados através de vários clusters como as etapas para formar clusters por meios K -

K-means seleciona k número de pontos para cada cluster conhecido como centróide

Cada ponto de dados forma um cluster com os centróides mais próximos, ou seja, k clusters.

Agora, ele encontrará os centróides de cada cluster com base nos membros existentes. Estas etapas são repetidas até que ocorra a convergência.

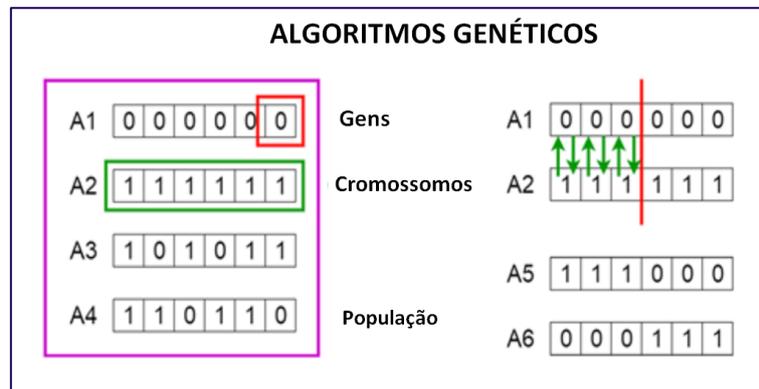
Random Forest

É um algoritmo de classificação supervisionado. A vantagem do algoritmo de floresta aleatória é que ele pode ser usado para problemas de classificação e regressão. Basicamente, é a coleção de árvores de decisão (ou seja, floresta) ou pode-se dizer um conjunto das árvores de decisão. O conceito básico de floresta aleatória é que cada árvore faz uma classificação e a floresta escolhe as melhores classificações a partir delas.

As vantagens do algoritmo Random Forest é o uso nas tarefas de classificação e regressão e podem lidar com os valores ausentes e não se encaixará no modelo, mesmo que tenhamos mais número de árvores na floresta.

Algoritmos Genéticos

Trata-se de uma heurística de pesquisa inspirada na teoria da evolução natural de Charles Darwin². Esse algoritmo reflete o processo de seleção natural, onde os indivíduos mais aptos são selecionados para reprodução, a fim de produzir descendentes da próxima geração.



Fonte: Mallawaarachchi, 2017

Noção de Seleção Natural

O processo de seleção natural começa com os indivíduos mais aptos de uma população. Pois, são eles que produzem filhos que herdam as características dos pais e serão adicionados à próxima geração. Se os pais tiverem melhor condicionamento físico, seus filhos serão melhores que os pais e terão melhores chances de sobrevivência. Esse processo continua se repetindo e, no final, será encontrada uma geração com os indivíduos mais aptos de continuidade.

Esta noção pode ser aplicada a um problema de pesquisa. Consideramos um conjunto de soluções para um problema e selecionamos o conjunto das melhores possibilidades em cinco fases do algoritmo genético.

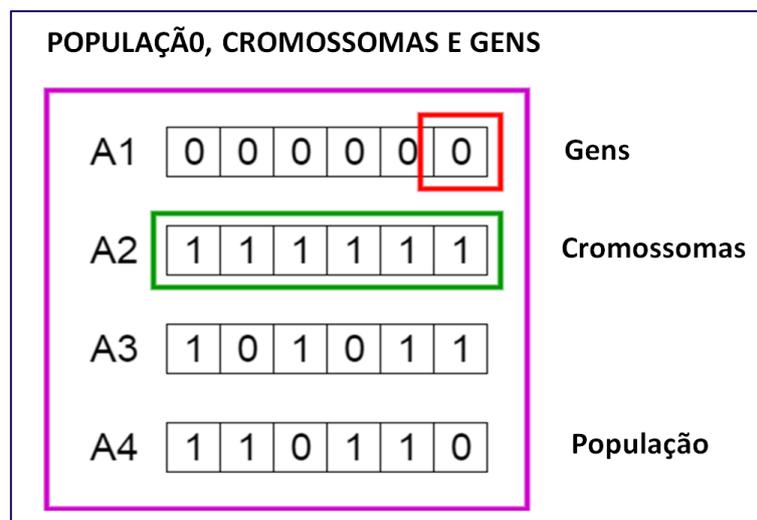
- População inicial
- Função de fitness
- Seleção
- Crossover
- Mutação

² **Charles Darwin** (1809-1882), naturalista inglês, desenvolveu uma teoria evolutiva que é a base da moderna teoria sintética: a teoria da **seleção natural**. “*Os organismos mais bem adaptados ao meio (selecionados para aquele ambiente) têm maiores chances de sobrevivência do que os menos adaptados, deixando um número maior de descendentes*”.

População Inicial

Trata-se do conjunto de indivíduos em que começa o processo onde cada indivíduo é uma solução para o problema que se deseja resolver. Um indivíduo é caracterizado por um conjunto de parâmetros (variáveis) conhecidos como Genes. Os genes são unidos por strings para formar um cromossomo (solução).

Assim, o conjunto de genes de um indivíduo é representado usando uma string, em termos de um alfabeto. Geralmente, são utilizados valores binários (sequência de 1s e 0s). Dizemos que codificamos os genes em um cromossomo.



Fonte: Mallawaarachchi, 2017

Função de Aptidão

Trata-se da função de condicionamento físico que determina a adequação do indivíduo (capacidade de competir) com a respectiva pontuação de aptidão. A probabilidade de um indivíduo ser selecionado para reprodução é baseada em sua pontuação de condicionamento físico.

Seleção

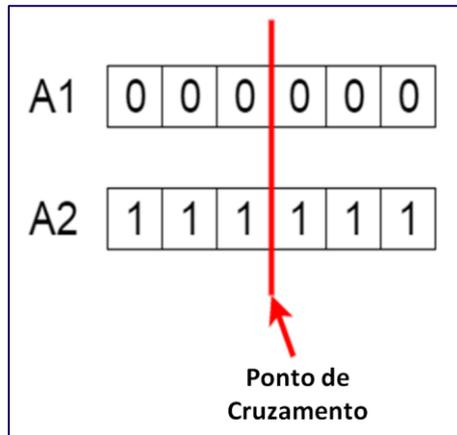
Nesta fase é a seleção dos indivíduos mais aptos e permitir que eles passem seus genes para a próxima geração.

Cada par de indivíduos (pais) são selecionados com base em suas pontuações de condicionamento físico. Indivíduos com alta aptidão têm mais chance de serem selecionados para reprodução.

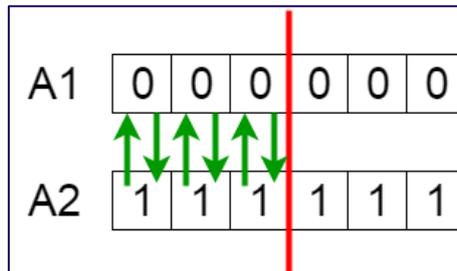
Cruzamento

O cruzamento é a fase mais significativa de um algoritmo genético. Para cada par de pais a serem acasalados é escolhido aleatoriamente um ponto de cruzamento dentre os genes.

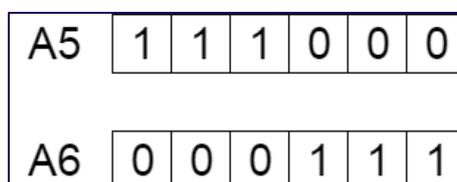
No exemplo a seguir são considerados 3 pontos de cruzamento.



Os filhos são criados trocando os genes dos pais entre si até que o ponto de cruzamento seja alcançado.



Os novos filhos são adicionados à população.



Mutação

No processo de novas criações, alguns genes podem ser submetidos a uma mutação com baixa probabilidade aleatória. Isso implica que alguns dos bits da sequência podem ser invertidos.



Terminação

O algoritmo termina se a população convergiu (quando não produz filhos significativamente diferentes da geração anterior). Diz-se então que o algoritmo genético forneceu um conjunto de soluções para o nosso problema.

Comentários

A população tem um tamanho fixo. À medida que novas gerações são formadas, indivíduos com menos condicionamento físico morrem, proporcionando espaço para novos filhos.

A sequência de fases é repetida para produzir indivíduos em cada nova geração que são melhores que a geração anterior.

Pseudo Código

```
START  
Generate the initial  
population  
Compute fitness  
REPEAT  
Selection  
Crossover  
Mutation  
Compute fitness  
UNTIL population has  
converged  
STOP
```

SISTEMAS ESPECIALISTAS

O campo de pesquisa dos Sistemas especialistas (ES) foi anterior à Inteligência Artificial (IA) e são baseados em software intensivos em conhecimentos na execução tarefas que requerem a experiência humana.

Trata da solução dos problemas específicos de domínio e raciocínio específicos que é determinado pelo especialista humano e se comportam como um sistema consultivo artificial para um domínio de problema específico.

Embora a IA seja predominante em várias aplicações comerciais hoje, o especialista das aplicações do sistema também é considerado IA.

Os sistemas especialistas saíram dos laboratórios de pesquisa no início dos anos 1980 e se expandiram em vários campos de aplicação como engenharia, química, medicina, indústria e muitos outros.

Os Sistemas especialistas procuram captar a capacidade de raciocínio dos humanos para responder aos usuários. Ao incorporar a experiência fornecem soluções nas questões de maneira semelhante aos tomadores de decisão usando as experiências pessoais anteriores ou recorrendo às experiências de outras pessoas (Sariyar & Ural, 2010)

Por ser um ramo da inteligência artificial, os sistemas especialistas apresentam características que os diferenciam de outros tipos de sistemas, que são:

- **Uso do raciocínio simbólico e heurísticas:** Integram o conhecimento existente com o raciocínio subjetivo do raciocínio simbólico utilizando métodos heurísticos de processamento (não algorítmicos) com o propósito de resolver problemas normalmente encontrados na interdependência.
- **Autoconhecimento:** São capazes de examinar o próprio raciocínio e explicar os motivos pelos quais chegou a um determinado nível de conclusões.
- **Inferência:** São capazes de combinar e vincular conhecimentos com o objetivo de obter novos conhecimentos.
- **Experiência:** Capacidade de encontrar soluções para um problema e determinar a competência necessária para a respectiva área.

Desta forma encontram aplicabilidade em diversas áreas especialmente para organizações onde existem pessoas com alto nível de conhecimento e experiência. (Giarranto & Riley, 2005).

O processo de construção de sistemas especialistas com conhecimento de domínio especializado é definido como engenharia do conhecimento. Sistemas especialistas baseados em conhecimentos contêm conhecimento adquirido em banco de dados de periódicos, livros, do conhecimento ou entrevistas no domínio com especialistas humanos.

Além dos sistemas especialistas clássicos, existem sistemas especialistas hoje usando técnicas como redes neurais artificiais e algoritmos genéticos.

Posteriormente os ESs passaram a compor um ramo da IA para gerenciar questões utilizando o conhecimento humano em vários campos da análise e consulta para controlar problemas complexos em vários domínios. São programas de computador que manipula fatos, conhecimento e raciocínio para resolver problemas com eficiência e eficácia numa área específica e restrita de problemas onde requer fontes multidisciplinares de conhecimento e especialistas humanos.

Semelhante aos especialistas humanos, os ESs usam a lógica simbólica e regras heurísticas para encontrar soluções. A perícia artificial desfruta de certas vantagens sobre a perícia humana; é consistente, permanente, replicável e transferível que preservam e disseminam o conhecimento de forma eficiente a custos razoáveis.

Pois, fornece uma avaliação matemática e logicamente robusta no tratamento de informações imprecisas que são particularmente relevantes em ecossistemas de plataformas, impacto ambiental e na gestão de recursos operacionais e corporativos.

Assim, a representação lógica é adequada quando os dados disponíveis são escassos ou o estado atual de conhecimento sobre um domínio de problema é muito impreciso para métodos matemáticos clássicos sobretudo para avaliar os sistemas mais complexos.

Além disso, o campo de dados do especialista é "heterogêneo" e contém muitos fatos procedimentais; portanto, o engenheiro do conhecimento deve ser um especialista no processo de segregação das informações onde o ES representa uma forma de capturar, codificar e reutilizar informações. Fundamentalmente, um ES compreende a representação da *expertise* ou um problema a ser resolvido com os mecanismos em forma de regras.

O QUE É UM ES?

Dentre as muitas definições são destacadas as mais adotadas como sendo o programa de computador que:

- Contém uma porção significativa do conhecimento especializado humano num domínio específico e estreito e emula a capacidade de tomada de decisão na área.
- Destinado a fundamentar os julgamentos ou prestar assistência numa área complexa na qual as habilidades humanas são limitadas ou escassas.

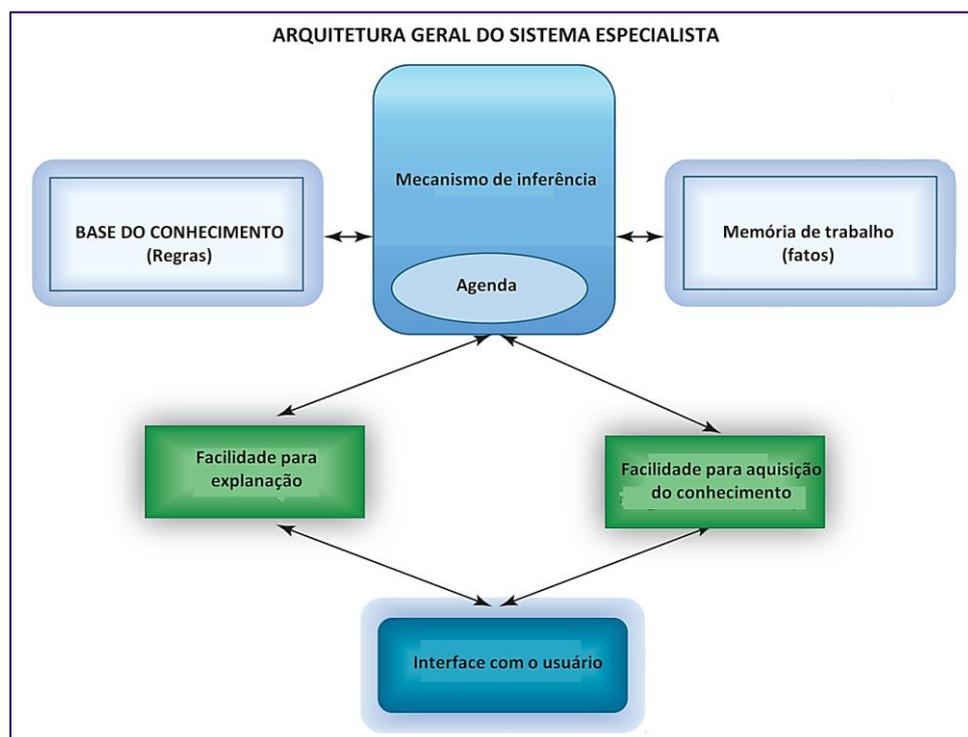
- Projetado para resolver problemas num nível comparável ao especialista humano em um determinado domínio.
- Opera um mecanismo de inferência a um corpo de perícia especializada representada na forma de "conhecimento".

Uma das definições mais amplamente aceitáveis foi dado por Edward Feigenbaum (Universidade de Stanford) como “um programa de computador inteligente que emprega o conhecimento e procedimentos de inferência na solução de problemas que são considerados difícil o suficiente para exigir experiência humana significativa.”

Alguns conceitos associados:

- **Idioma:** o tradutor do comando escrito numa sintaxe específica. O Expert System Language fornece um mecanismo de inferência para executar as declarações da Linguagem.
- **Ferramenta:** software utilitário para depuração, desenvolvimento e implantação de sistemas aplicativos. Exemplos: editores de texto, compiladores, geradores de código (XTRAN e outros semelhantes). Ferramentas integrando todos os programas utilitários numa única interface do usuário.

A arquitetura geral de um sistema especialista é apresentada na figura a seguir e os componentes são definidos da seguinte forma:



Fonte: Adaptação de Tolun e Oztoplak, 2016

- **Interface do usuário:** o mecanismo pelo qual o usuário e o sistema especialista se comunicam.
- **Recurso de explicação:** explica o raciocínio do sistema para o usuário.
- **Memória de trabalho:** um banco de dados de fatos usados pelas regras.
- **Mecanismo de inferência:** faz inferências ao decidir quais regras que satisfazem os fatos ou objetos, prioriza e executa as regras satisfeitas com prioridade máxima.
- **Agenda:** lista prioridade de regras criadas pelo mecanismo de inferência e transfere para a memória de trabalho.
- **Instalação de aquisição de conhecimento:** inserção automática do conhecimento no sistema pelo usuário sem a necessidade de codificação pelo engenheiro do conhecimento.

Os primeiros sistemas especialistas foram construídos por meio de entrevistas junto a especialistas no intuito de coletar conhecimentos, daí o termo "Sistemas Especializados (ES)". O ES é um programa de computador que é construído com base nas experiências de um especialista no domínio específico executando funções como “fazer perguntas e explicar o seu raciocínio com base no modelo “pergunta e resposta” junto ao usuário final.

O kernel (núcleo) do sistema especialista tem dois componentes principais:

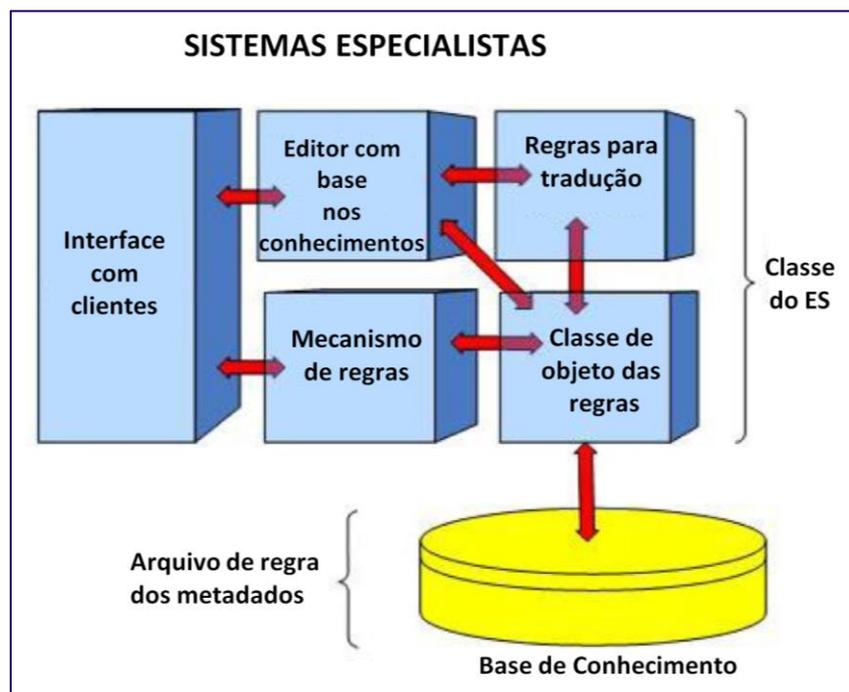
- **Base de conhecimento:** contém o conhecimento específico dentro do domínio do especialista. Mediante regras, pode ser representado por fatos simples, ou por representações complexas em forma de quadros.
- **Mecanismo de inferência:** possui várias funções como o sistema de razões usando as regras IF – THEN para a construção da base de conhecimento para fazer as inferências. As formas mais comuns de inferência é o encadeamento para frente e para trás. O processo de avançar com os fatos conhecidos para chegada nas conclusões é chamado de encadeamento direto. Alternativamente, o processo de retroceder a partir de uma hipótese com apoio dos fatos conhecidos é chamado de encadeamento reverso.

CARACTERÍSTICAS DO ES

Uma base de conhecimento é formado pelo conhecimento humano com fatos e heurísticos do especialista humano capturado e armazenamento em forma de regras, quadros lógicos para resolução de problemas complexos utilizando os conceitos de inteligência artificial. compartilhamentos. Conhecimento factual é um conhecimento amplamente compartilhado obtido de livros, periódicos,

banco de dados e sites de Internet e o aprendizado heurístico é raro são usados em procedimentos para resolução de problemas e para apoiar a aprendizagem humana e na tomada de decisão e ações.

- O ES resolve programas complicados num determinado domínio melhor do que os especialistas humanos.
- ES possui uma vasta quantidade de conhecimentos específicos de domínio nos seus mínimos detalhes.
- O quadro do ES aplica métodos representativos heurísticas para gerenciar os argumentos pelos quais reduzem o número de soluções na zona de revisão.
- A característica única do sistema especialista é a capacidade de explanação. Possibilita que os especialistas revisem seu próprio raciocínio para explicar suas decisões.
- A estrutura dos sistemas especialistas utiliza o pensamento representativo ao considerar um problema. Os símbolos são usados para simular vários tipos de aprendizagens como as realidades, ideias e diretrizes.



Fonte: Adaptado de Ambak e Mosa, 2019

MECANISMO DE INFERÊNCIA

Trata da identificação do raciocínio humano e a implementação do processo de raciocínio da inteligência artificial. Tem como objetivo de extrair os dados e informações disponíveis no banco de dados do sistema e do usuário para fornecer as respostas na forma de previsões e sugestões como um especialista humano.

Existem dois tipos de mecanismos de inferência:

- **Encadeamento reverso:** É direcionado ao objetivo onde a solução é conhecida; ou seja, o sistema retrocede para encontrar os fatos que suportam esta solução.
- **Encadeamento direto:** É orientado por dados; assim, é usado quando a solução absoluta está ausente. O sistema coleta os dados e são usados até que a solução seja conhecida.

AQUISIÇÃO DO CONHECIMENTO

Opera com o editor para inserir o conhecimento específico diretamente no domínio no banco de dados que serve como interface entre os especialistas e o ES.

O objetivo geral é fornecer um meio eficiente e conveniente de capturar e armazenar todos os componentes da base de conhecimento; e auxilia no processo de segregação do conhecimento do especialista e a codificação para a base de conhecimento.

A edição do conhecimento pode ser realizada de duas maneiras para gerar e modificar o arquivo de regras:

- Pelo engenheiro do conhecimento:
- Próprio sistema especialista

FACILIDADE NA EXPLANAÇÃO

A maioria dos ESs possui recursos de explanação onde indica o caminho da solução específica para o usuário chegar a uma determinada conclusão. Ou simplesmente responde como o ES chegou à solução solicitado pelo usuário, as razões da rejeição de algumas alternativas. Por exemplo, no caso de avaliação de proposta de empréstimo bancário, o ES esclarecerá as razões da aprovação ou rejeição do pedido.

INTERFACE COM O USUÁRIO

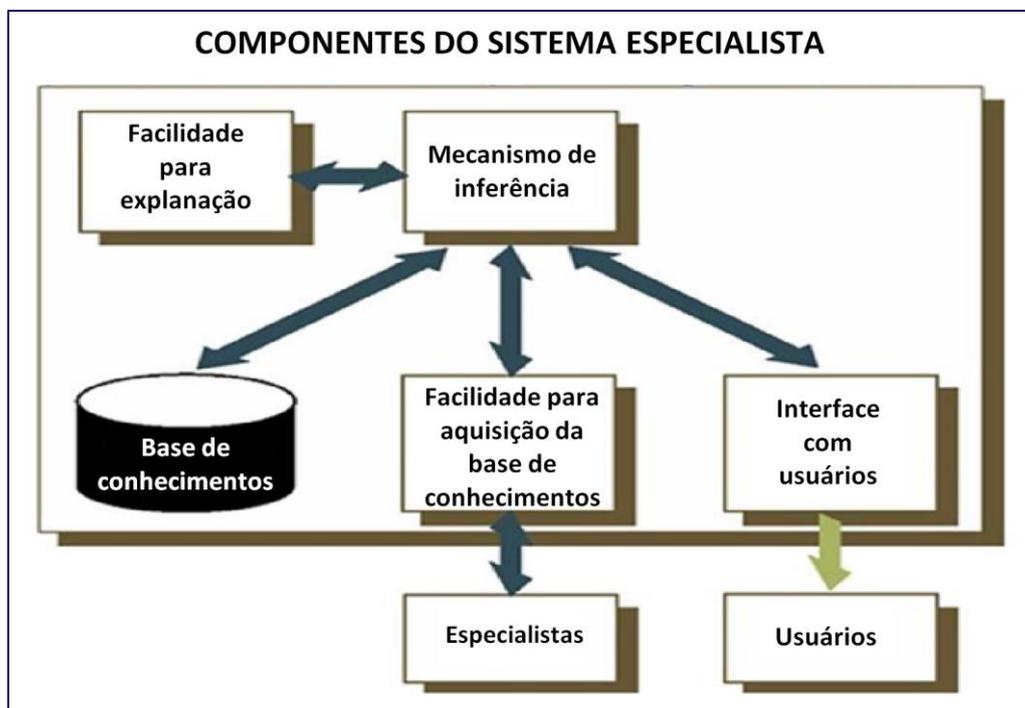
Trata se de uma ferramenta intermediária que gerencia o diálogo entre o usuário e o sistema com os recursos como menus e interface gráfica, áudio, vídeo, ícones, etc. O recurso do processamento de linguagem natural que é uma forma de inteligência artificial que permitem aos computadores a ler e responder, simulando a habilidade humana de entender a linguagem cotidiana. Possui a função de facilitar o uso para os desenvolvedores, usuários e administradores.

ENGENHARIA DO CONHECIMENTO

Trata do processo contribuem no projeto, desenvolvimento, construção, suporte e a manutenção do sistema do ES. Os recursos humanos como o especialista no domínio têm conhecimento, julgamento, experiência e métodos especiais para aconselhar e resolver problemas.

Assim, o engenheiro de conhecimento trata do desenvolvimento da base de conhecimento, no mecanismo de inferência e na interface que atende às necessidades do usuário. Finalmente, existem dois componentes no desenvolvimento de um sistema educacional eficaz e inteligente em qualquer domínio:

- **base de conhecimento:** técnicas de representação e gerenciamento do conhecimento, como listas, árvores, redes semânticas, frames, scripts e as regras de produção conforme mostrado na figura a seguir:



Fonte: Adaptado de Ambak e Mosa, 2019

VANTAGENS DO SISTEMA ESPECIALISTA

O ES possui muitas vantagens sobre os especialistas humanos como listadas em conformidade:

- **Acessibilidade:** Integração combinada e sistêmica do conhecimento de vários especialistas humanos em comparação com o conhecimento de uma única pessoa. Os Estão disponíveis para utilização e pode auxiliar o especialista humano na solução de problemas.
- **Eficiência:** Capacidade de revisar e estruturar todas as transações em comparação com o especialista humano que é limitado a revisar apenas uma amostra.

- **Restrições de tempo:** O tempo menor no treinamento de engenheiros em ES comparado com o treinamento do especialista humano.
- **Uniformidade:** Menores erros, pois o ES possui a representação organizada de conhecimentos onde os erros podem ser neutralizados sem esforço.

COMPONENTES DO ES

1. Sistema baseado em conhecimentos

O Knowledge-Based System (KBS) é o elemento essencial no grupo da IA para execução de tarefas mais exigentes com base na inteligência.

A sociedade e a indústria são orientadas na era do conhecimento contam com a capacidade de tomada de decisão de vários recursos de especialistas.

O KBS atuar como um especialista sob demanda em tempo real a qualquer hora e em qualquer lugar como uma ferramenta produtiva de forma em que os usuários funcionem num nível superior de segurança e consistência.

2. Sistema baseado em regras

Simulam a capacidade básica de liderança dos especialistas humanos. São projetados para resolver problemas como os humanos com base no conhecimento humano codificado ou em experiências representadas na forma de regras, como If-then. São utilizadas em operações de inferência para análise de tendência dos dados para se chegar à conclusão apropriada. Trata de metodologias para raciocinar sobre informações com base em regras ou de conhecimento para formular conclusões.

3. Redes neurais artificiais (ANNs)

É uma ferramenta de software projetada para estimar relacionamentos a partir do mapeamento entre dados brutos às suas características. A ubiquidade dos sistemas neurais depende da flexibilidade e da capacidade de processar informações contínuas ou em duplas e criar resultados em espaços complexos. No caso em que a saída é binária, a rede funciona como um classificador, mas quando a saída é contínua a rede pode resolver questões de previsão.

4. Sistema especialista difuso (Fuzzy)

Trata se da combinação das capacidades e dos padrões difusos em vez da lógica booleana para gerar informações. Os princípios do ES difuso seguem a seguinte forma: Se w é baixo e n é alto, então p = médio, onde os tipos são conhecidos com base em regras ou com base no conhecimento.

O processo de inferência dominante é formado de três (ou quatro) procedimentos.

- **na difusão (fuzzificação):** as capacidades de inscrição caracterizadas nos fatores de informação estão conectadas às suas qualidades reais, para decidir o nível de verdade para cada início da governança.
- **sob inferência:** o valor verdadeiro na premissa de cada regra é calculado e aplicado na etapa de conclusão de cada regra. O subconjunto difuso resultante deve ser distribuído para cada variável de rendimento de cada regra. Geralmente, apenas o produto secundário é utilizado como regras de inferência. Na inferência, a função de participação no rendimento é coletada numa proporção para comparação com as premissas de governança registradas em nível de verdade (lógica fuzzy).

Na inferência do produto, a função de participação no rendimento é dimensionada pelo nível de verdade registrado na abertura da gestão.

- **sob a estrutura:** a maioria dos subconjuntos fuzzy permitidos para cada variável de rendimento são reunidos para formar o subconjunto fuzzy solitário para cada variável de rendimento. São utilizadas as funções Max ou Sum. No arranjo Sum (Soma), o subconjunto fuzzy de saída unificado é construído tomando-se o todo pontual terminado a maioria dos subconjuntos fuzzy distribuídos para a variável de rendimento pela regra de inferência de dedução.
- **Defuzzificação (discricionária):** que é utilizada quando é valioso mudar a saída difusa definida para um novo número [24].

5. Raciocínio baseado em caso (CBR)

A ideia básica do CBR é adotar soluções que foram usadas na solução de questões anteriores e reutilizá-los para resolver novos problemas. As descrições das experiências anteriores dos especialistas humanos, representadas como casos são armazenadas no banco de dados para recuperação posterior para o uso num caso novo com parâmetros semelhantes.

As características comparáveis ao novo problema se procedem com o ajuste por aproximação e aplica se ao novo caso. As soluções bem-sucedidas são marcadas na base de conhecimento para utilização no caso novo. Para soluções ineficazes são armazenadas com com esta observação.

6. Modelagem

Trata-se da metodologia interdisciplinar do ES de modo a construir relações formais no design do modelo lógico em diferentes domínios do conhecimento/problema. Além disso, a tecnologia de modelagem pode fornecer técnicas quantitativas para analisar dados, para comentar ou obter conhecimento especializado com programação ou cálculos de lógica indutiva, de modo que a IA, ciência da construção e outros campos de pesquisa atinjam estágios mais avançados para mudança mais amplas do ES.

O naipe de regras em um sistema especialista difuso é conhecido como base de regra ou base de conhecimento. O processo de inferência abrangente prossegue em três (ou quatro) procedimentos. Em primeiro lugar, na fuzzificação, as capacidades de inscrição caracterizadas nos fatores de informação estão conectadas às suas qualidades reais, para decidir o nível de verdade para cada início de governo. Em segundo lugar, sob inferência, o valor verdade para a premissa de cada regra é calculado e aplicado à parte de conclusão de cada regra. Resulta em um subconjunto difuso a ser distribuído para cada variável de rendimento para cada regra. Geralmente, apenas um produto secundário é utilizado como regras de inferência. Na inferência, a função de participação de rendimento é cortada em uma estatura em comparação com as premissas do governo registradas em um nível de verdade (lógica fuzzy).

Na inferência do produto, a função de participação de rendimento é dimensionada pelo nível de verdade registrado do prefácio de gerenciamento. Em terceiro lugar, sob a estrutura, a maioria dos subconjuntos fuzzy permitidos para cada variável de rendimento são unidos o tempo todo para enquadrar um subconjunto fuzzy solitário para cada variável de rendimento. Novamente, geralmente, Max ou Sum são usados. Na composição máxima, o subconjunto fuzzy de saída combinada é construído tomando o máximo pontual sobre todos os subconjuntos fuzzy atribuídos a uma variável pela regra de inferência (lógica fuzzy). No arranjo Soma, o subconjunto fuzzy de saída unido é construído tomando-se o todo pontual terminado a maioria dos subconjuntos fuzzy distribuídos para a variável de rendimento pela regra de inferência de dedução. Por último, está a Defuzzificação (discricionária), que é utilizada quando é valioso mudar a saída difusa definida para um novo número.

DESENVOLVIMENTO HISTÓRICO DE SISTEMAS ESPECIALISTAS

No final dos anos 1950, vários projetos de programação como o General Problem Solver desenvolvido por Newell e Simon¹ para resolver uma variedade de problemas que vão desde integração simbólica até quebra-cabeças para encontrar o melhor caminho.

¹ O programa de computador projetado para simular a resolução de problemas humanos, introduzido em 1958 pelo cientista cognitivo americano Allen Newell (1927–92) e pelos economistas e teóricos da decisão dos EUA John Clark

Na década de 1960, os precursores dos sistemas especialistas começaram pelo desenvolvimento de programas práticos para os problemas do mundo real em diferentes

Campos. Em 1965, o DENDRAL² é criado como o primeiro sistema especialista baseado em conhecimentos. A aplicação é usada para encontrar estruturas de moléculas orgânicas. Trata a questão do espaço de busca aplicando o conhecimento heurístico de especialista em química, para encontrar a estrutura correta.

O MYCIN é desenvolvido para diagnóstico médico na detecção de infecções bacterianas em 1973 e recomenda uma lista de terapias para o paciente com a base de conhecimento inicial de cerca de 200 regras.

A base de conhecimento é separada do motor de inferência e este recurso de MYCIN oferece outros sistemas de diagnóstico para ser construído rapidamente. A base de conhecimento inicial do MYCIN continha cerca de 200 regras.

Na década de 1980, o valor dos sistemas especialistas estava bem estabelecido com o surgimento da linguagem de programação C. A ferramenta do sistema especialista Language Integrated Production System (CLIPS) é adotada na academia, indústria e instituições governamentais. O CLIPS suporta programação orientada a objetos, baseada em regras e procedural, e pode ser implementado em computadores que usam o compilador ANSI C, as versões estendidas do CLIPS (FuzzyCLIPS, COOL, ...) para o uso integrado de lógica difusa e orientada a objetos. Nos meados da década de 1990 surge a versão Java do CLIPS, chamada de JESS com todas propriedades do CLIPS que oferece correspondência de regras com fatos e o suporte ao processo de encadeamento reverso.

Shaw (nascido em 1933) e Herbert Alexander Simon (1916–2001), e desenvolvido em 1972 por Newell e Simon. Neste programa, o problema é representado numa tabela de conexões mostrando as distâncias entre todos os pares de estados (inicial, intermediário e final), e é modelada como uma pesquisa no espaço do problema usando operadores permitidos (ações), onde a tarefa é encontrar um caminho a partir do estado inicial para o estado objetivo. Num jogo de xadrez, o estado inicial típico é a posição inicial e o estado objetivo correspondente é o xeque-mate no oponente; ou o estado inicial mais comum é estar em casa com uma criança e um carro com bateria descarregada, e o estado da meta correspondente é a entrega da criança à creche. Em geral, o espaço do problema é muito grande para que a busca exaustiva seja viável e, portanto, a heurística cognitiva da análise de meios-fins é implementada. Outros problemas como encher jarros d'água, o anel da Torre de Hanói, travessia dos missionários e o canibal em que o programa é capaz obter a solução do quebra-cabeça. Representa os primeiros marcos na história da aplicação da inteligência artificial.

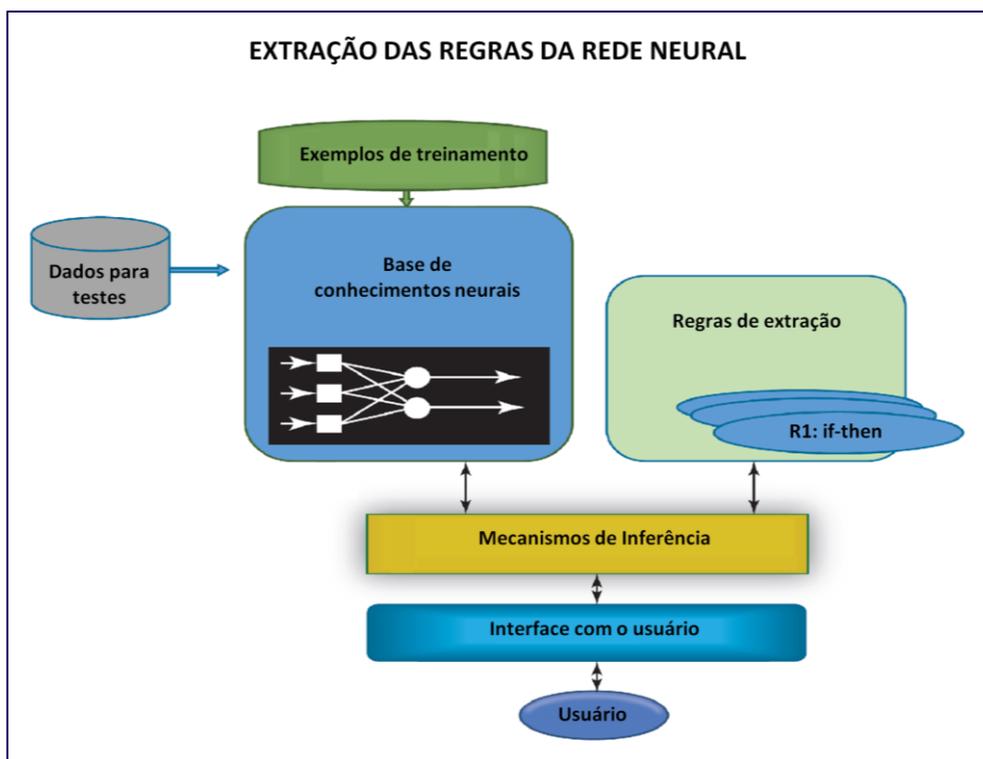
² Dendral é um sistema especialista e um projeto pioneiro em inteligência artificial desenvolvido na Universidade de Stanford. O programa Dendral é considerado o primeiro sistema especialista por causa do modo automático de tomar decisões e resolver problemas relativos a química orgânica. Ele é dividido em dois subprogramas, o Heuristic Dendral e o Meta-Dendral e foi inteiramente escrito em Lisp.

PROCESSO DE AQUISIÇÃO DE CONHECIMENTO

O processo de aquisição do conhecimento ocorre a partir do domínio do especialista e convertido em conjunto de regras ou quadros lógicos. São baseados em regras ou em quadros a Sistemas especialistas conectáveis e indutivas baseados nas redes neurais ofereceram abordagens para superar a barreira de aquisição de conhecimento. Tais abordagens são qualitativamente diferentes dos sistemas especialistas padrão na representação de conhecimento e de aprendizado.

Para um sistema especialista na rede neural, a representação do conhecimento consiste de uma rede, peso da conexão e as interpretações semânticas ligado a células e ativações. A principal vantagem desta abordagem é que o algoritmo de aprendizagem associado como Backpropagation (retropropagação) pode gerar exemplos de treinamento e gerar automaticamente os sistemas especialistas.

Este procedimento é ilustrado na figura a seguir:



Fonte: Adaptado de Tolun e Oztoplak, 2016

Os aplicativos típicos incluem nas áreas de gestão e administração (estimativa dos custos, programação), industrial (controle de processo, controle de qualidade de fabricação, diagnóstico de falha), saúde (diagnóstico médico em domínios especializados, identificação de bactérias), bancário (decisões de crédito e empréstimos) e outros campos (previsão).

O modelo neural desenvolvido por McCulloch e Pitts (1943) é considerado como ponto de partida dos sistemas especialistas conectáveis. E o processo de aprendizagem por neurônios é definido por Hebb em 1949, e a eficiência do modelo de Hebbian é determinado pela emulação e transferência de cada impulso de um neurônio para outro pelas regras de disparo entre cada neurônio.

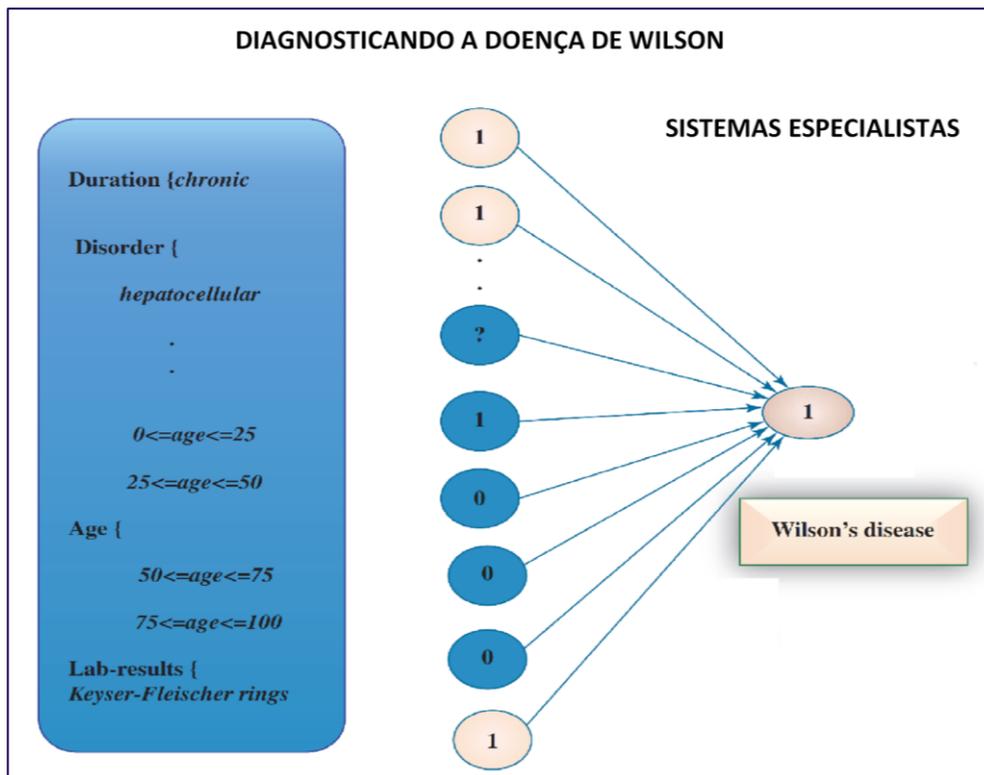
A aprendizagem indutiva é usada nos sistemas especialistas da rede neural conectável quando há muitos dados empíricos e também para prevenir gargalos na aquisição de conhecimento.

O aplicativo de regras do sistema especialista define o treinamento e o padrão de teste é representado no exemplo do sistema especialista médico.

O conhecimento base do sistema especialista consiste de várias regras IF–THEN que estão relacionadas com o diagnóstico das doenças para identificar o tratamento mais adequado e o sistema pode ser implementado como uma rede neural de três camadas do exemplo a seguir:

<p>IF duration (x, chronic) and disorder (x, hepatocellular) and age (x, <25) and lab-results (x, Kayser–Fleischer rings) THEN diagnostic (x, Wilson’s disease)</p>

A rede neural consiste nas células que correspondem aos sintomas na camada de entrada, e as doenças são apresentadas na camada intermediária ou oculta e os tratamentos definidos na camada de saída das células neurais (nós). Os padrões de treinamento são os 0's (falta de conhecimento sobre a presença ou ausência da doença), 1's (presença da doença).



Fonte: Adaptado de Tolun e Oztoplak, 2016

A topologia geral para este ES conectável consiste no sistema neural de duas camadas rede onde o diagnóstico da doença é ativado apenas nas situações sintomáticas que são identificadas como verdadeiras (1).

Ao mesmo tempo, o sistema é treinado e testado pelo algoritmo de aprendizagem supervisionado de retropropagação.

Portanto, os problemas mais adequados do sistema especialista com a abordagem da rede neural são aqueles que procuram classificar as entradas num pequeno número de grupos.

SISTEMAS BASEADOS EM INDUÇÃO

O principal gargalo para a construção de sistemas especialistas reside na obtenção de conhecimento de especialistas no domínio que são várias dificuldades:

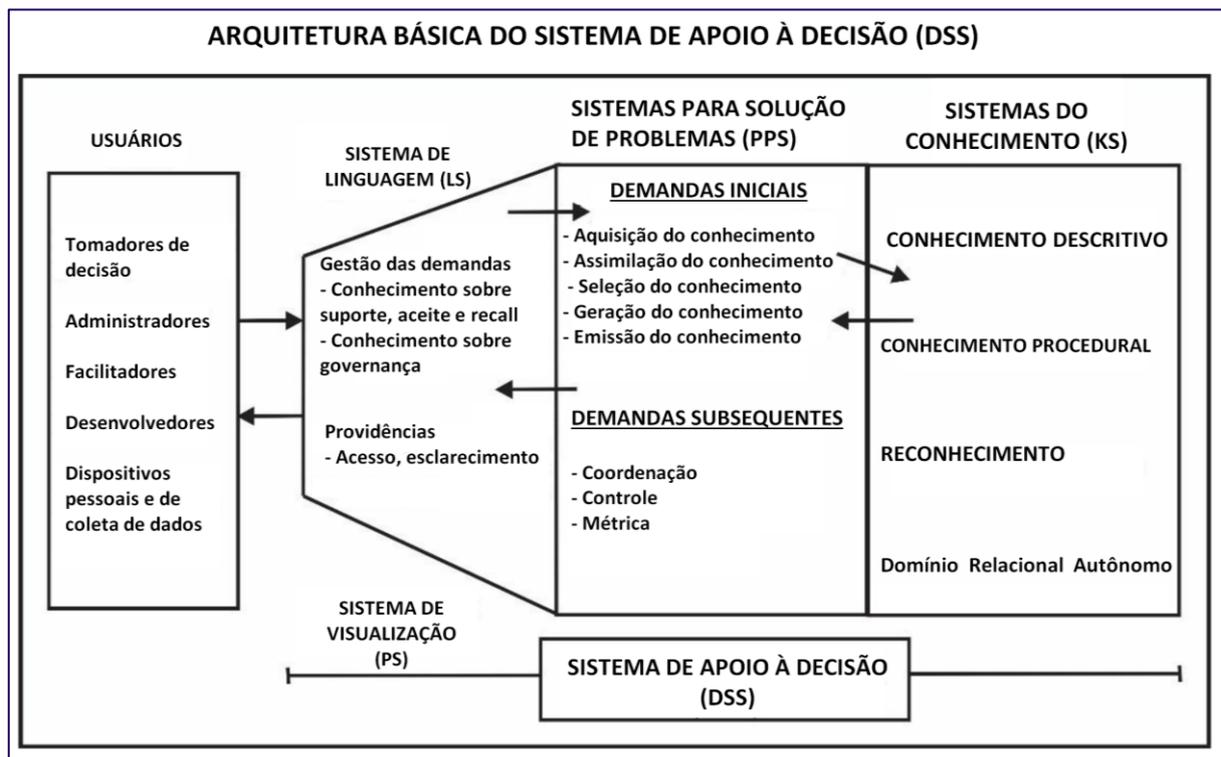
1. Incompatibilidade do conhecimento, a diferença entre o conhecimento do próprio e o especializado e como é apresentado no programa.
2. Incapacidade dos humanos de expressar o conhecimento que possuem e a natureza inerente do conhecimento (subconsciente, aproximado, incompleto, inconsistente, etc).
3. Problema de verificação e validação.

Principais benefícios do uso dos sistemas baseados em indução:

1. Pode ser mais competente do que os humanos em adquirir ou ajustar certos tipos de conhecimentos.
2. Pode reduzir significativamente o alto custo dos recursos humanos na construção do sistema. O conjunto de dados de treinamento para o aprendizado de máquina indutivo usado na árvore de decisões
3. Outros métodos de indução incluem as regras de classificação que são geradas pelo algoritmo de aprendizagem como dados de treinamento

CARACTERÍSTICAS DE SISTEMAS ESPECIALIZADOS E SUAS ÁREAS DE APLICABILIDADE

A arquitetura de sistemas especialistas pode ser considerada um caso especial de arquitetura genérica de sistemas de apoio à decisão (SAD), cujo arcabouço conceitual identificando os componentes essenciais e interdependências entre eles.



Fonte: Adaptado de Holsapple e Burstein, 2018

Na Figura anterior apresenta a modalidade em que os quatro subsistemas do SAD são vinculados entre eles e seus usuários onde os componentes determinam as capacidades o comportamento:

- **Sistema de linguagem:** composto por todas as mensagens aceitos pelo sistema.
- **Sistema de visualização:** composto por todas as mensagens emitidos pelo sistema.
- **Sistema de conhecimento:** composto por todos os conhecimentos armazenados e mantidos pelo sistema.

Os dados destes subsistemas são utilizados pelo sistema para solução dos problemas (PPS) como quarto elemento do DSS. O componente procura reconhecer e obter soluções para os problemas no processo de tomada de decisões.

Os usuários podem ser: um desenvolvedor, o administrador, um humano ou um dispositivo de entrada de dados. E pode solicitar uma seleção de conhecimentos, de esclarecimentos ou respostas para questões anteriores, a resolução de problemas, a detecção de problemas, etc. E o papel dos usuários podem ser fatores decisivos ou participantes no processo de tomada de decisões.

A interface do usuário é definida por seus subsistemas de linguagem e de visualização, as habilidades na aquisição e emissão de conhecimentos bem como o conteúdo do sistema de conhecimento utilizado pelo PPS para interpretar os elementos de linguagem do sistema para processar o conhecimento nos elementos de visualização dos sistemas. O sistema de conhecimento fornece *expertise* e conhecimento necessário para resolver problemas difíceis e complexos, não estruturados ou semiestruturados e oferece ao sistema um atributo inteligente.

Nesta arquitetura genérica do SAD, podemos observar os aspectos cruciais e fundamentais comuns para todos os sistemas de apoio à decisão, assim como para os sistemas especialistas.

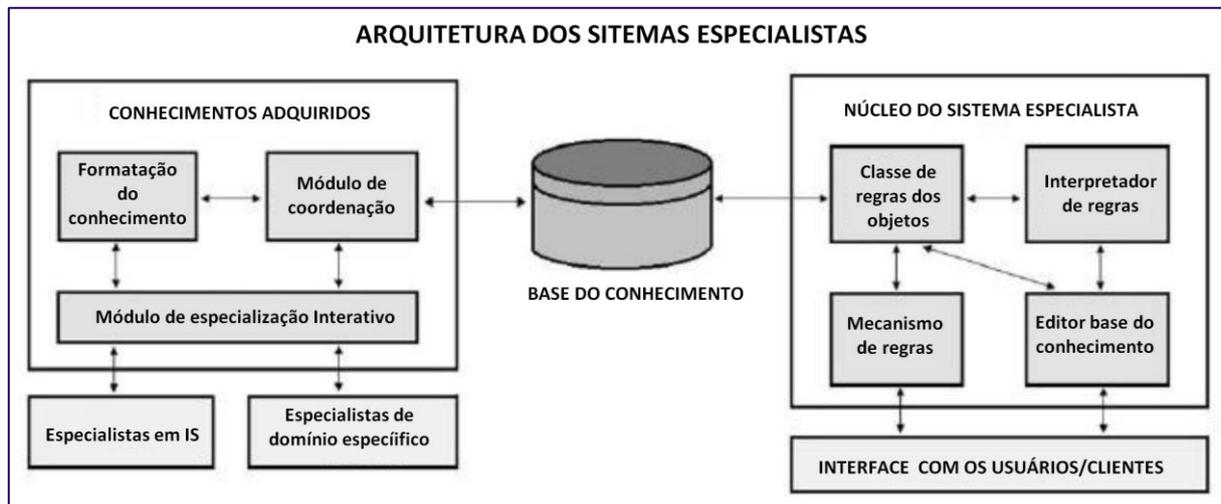
Para examinar a natureza específica do SAD, precisamos conhecer as solicitações especiais que compõem a linguagem do sistema, a resposta específica que compõe o sistema de visualização do sistema de conhecimento e as capacidades específicas de processamento do conhecimento pelo PPS.

Desta forma, a arquitetura do DSS com seus componentes, a interdependência entre eles com as bases teóricas necessárias para abordar as subclasses específicas do DSS que são os sistemas especialistas.

SISTEMAS ESPECIALISTAS

Nas abordagens dos sistemas especialistas podemos identificar três componentes principais que são a base da sua construção: base de conhecimento, motor de inferência e interface do usuário. À medida em que os sistemas especialistas evoluam com seus componentes, surgem novos elementos com o propósito de melhorar a capacidade e de ultrapassar seus limites que são o módulo interativo, módulo de coordenação, editor de regras e editor de conhecimento.

De acordo com (Atanasova & Krupka, 2013) a arquitetura de sistemas especialistas é baseada na separação clara do módulo de aquisição e o mecanismo de inferência composto por dois subsistemas, que são a base de conhecimento e o motor de inferência. Na arquitetura de sistemas especialistas representada na figura a seguir:



Fonte: Adaptado de Atanasova e Krupka, 2013

No modelo de refinamento da estrutura dos sistemas especialistas com componentes adicionais: o subsistema de aquisição de conhecimento, o espaço de trabalho, o módulo de aplicação e o módulo de refinamento de conhecimento.

Os usuários dos sistemas especialistas são:

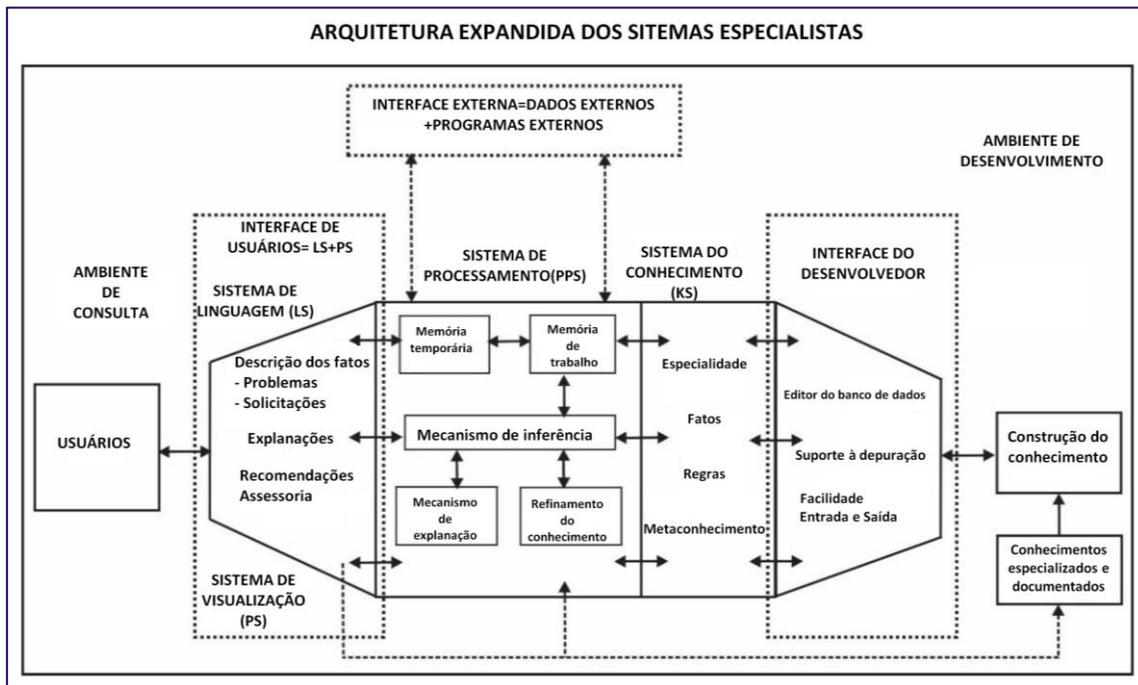
1. Que emitem solicitações ao sistema e recebem as respostas

Com a interface do usuário composta a partir do sistema de linguagem e do sistema de visualização. Podem emitir pedidos de conselhos, solicitação de explicação e introduzir dados que descrevem o problema atual (fatos). O sistema de visualização inclui mensagens que representam recomendações, conselhos e explicações.

2. O engenheiro do conhecimento.

Extrai os conhecimentos da área e os fornecidos pelos especialistas e tem à disposição a interface com os desenvolvedores de sistemas que contém: o editor da base de conhecimento, o suporte para depuração e as facilidades para entrada/saída. Com a ajuda do editor da base de conhecimento, podemos introduzir dados, modificar e monitorar as regras. O suporte para depuração permite detectar e remover os erros. As facilidades de entrada e saída durante a execução do sistema especialista solicitar as informações necessárias tantas vezes quanto estejam disponíveis. O sistema

especialista interage com os usuários do sistema com as duas interfaces, e a conexão com o exterior é assegurada pela interface externa que contém fontes de dados e programas externos.



Fonte: Adaptado de Avornicului, Diana-Aderina, 2015

O sistema de conhecimento contém fatos, regras de expertise e “meta-conhecimento” necessários na solução de problemas específicos. O conteúdo do KM é atualizado permanentemente pelo engenheiro do conhecimento e pelo componente do sistema para problemas de processamento.

O sistema de processamento contém o mecanismo de inferência, a área de trabalho, o módulo de explicação e o módulo de refinamento de conhecimento. As capacidades do PPS são dadas pelas características de seus componentes. O mecanismo de inferência confere-lhe capacidades de criação, revisão, eliminação, sendo de maior interesse a capacidade de fazer deduções lógicas com um conjunto de regras para produzir o aconselhamento solicitado a um utilizador. As soluções gerais (conselhos, recomendações) serão controladas pelo usuário, também pelo módulo de refinamento do conhecimento e pelo componente que contém o conhecimento da área. O PPS também tem a capacidade de explicar o seu comportamento pela presença do módulo explicativo e a capacidade de autoaprendizagem proporcionada pela presença do módulo de refinamento do conhecimento.

ÁREAS DE APLICAÇÃO DOS SISTEMAS ESPECIALISTAS

As principais áreas de aplicabilidade dos sistemas especialistas: produção, agricultura, educação, gestão ambiental, legislação e área da saúde sobretudo nas aplicações de: (Turban, Sharda, & Delen, 2011):

- **Finanças:** avaliações de seguros, análise de crédito, planejamento tributário, prevenção de fraudes, análise de relatórios financeiros, planejamento financeiro e avaliação de desempenho.
- **Processamento de dados:** Sistemas de planejamento, seleção e manutenção de equipamentos, avaliação de provedores e gerenciamento de rede.
- **Marketing:** Gestão do relacionamento com o cliente, a análise do mercado e o planejamento dos produtos.
- **Recursos humanos:** Planejamento de recursos humanos, avaliação de desempenho, agendamento de pessoal, gerenciamento de pensões e aconselhamento de funcionários.
- **Produção:** Planejamento e controle da produção, gerenciamento de qualidade, *design* de produto, reparo e manutenção de equipamentos.
- **Gestão de cuidados em saúde:** Bioinformação e outros problemas de gestão da saúde.

As mais recentes aplicações de sistemas especialistas podemos observar a solução das questões ligadas à gestão de riscos, automação das regras de negócio, rastreamento e acompanhamento automático do mercado, diagnóstico de doenças, sistemas robóticos, etc. Algumas das aplicações com resultados importantes:

- **Sistemas de análise de crédito:** Na avaliação do risco de crédito na análise financeira, em comparação com as abordagens estatísticas tradicionais.
- **Consultoria em sistemas de fundos de pensão dos funcionários:** Atualização contínua da base de dados para aconselhar os usuários sobre as mudanças de acordo com as novas normas e regulamentos.
- **Sistemas de diagnóstico de doenças e fungos em plantas e frutas:** Diagnóstico de doenças e pragas em frutas e plantas.
- **Sistemas robóticos:** para instruir robôs que transferem informações sobre o estado do tempo de uma região para outra. O primeiro robô recebe as informações do ambiente, o segundo recebe do primeiro, utilizando técnicas de sistemas especialistas, etc. O tempo de transmissão das informações é consideravelmente reduzido.

Analisando as abordagens dos sistemas especialistas (Nikolopoulos, 1997), (O'Brien & Marakas, 2007), podemos identificar três componentes principais que são: base de conhecimento, motor de inferência e interface do usuário. Esses componentes têm estado na base da construção dos sistemas especialistas tradicionais. À medida que os sistemas especialistas evoluíram em seus

componentes, surgiram novos elementos com o propósito de melhorar a capacidade dos sistemas especialistas e de ultrapassar seus limites.

De acordo com (Atanasova & Krupka, 2013) a arquitetura de sistemas especialistas é baseada na separação clara do módulo de aquisição e o mecanismo de inferência composto por dois subsistemas, que são a base de conhecimento e o motor de inferência. Na arquitetura de sistemas especialistas modernos representada em 5.

Diferentemente dos modelos arquitetônicos apresentados até agora (Turban, Sharda, & Delen, 2011) identificam na estrutura dos sistemas especialistas os próximos componentes adicionais: o subsistema de aquisição de conhecimento, o espaço de trabalho, o módulo de aplicação e o módulo de refinamento de conhecimento.

APLICAÇÕES ATUAIS E O FUTURO DOS SISTEMAS ESPECIALISTAS

Os métodos de implementação de sistema especialista podem ser classificados como sistema especialista clássico, sistemas especialistas neurais, sistemas especialistas difusos e os sistemas especialistas baseados em conjuntos aproximados, sistemas especialistas na tomada de decisão multicritério, sistemas especialistas neuro-difusos, sistemas especialistas baseados na web e sistemas especialistas multiagentes.

Hoje, os aplicativos de sistemas especialistas são frequentemente desenvolvidos em aplicações comerciais:

1. EXSYS Corvid (para Windows e Mac OS X) é um dos shells populares e um grande número de ferramentas de construção de ES específicas e de uso geral que estão disponíveis
2. Os aplicativos ES baseados na web ou na Internet para assessoria online como a Ferramenta SAP (CAST); Guide and Solvent Alternatives Guide da Pollution Prevention escrito em ColdFusion; metamodelo do processo de desenvolvimento para sistemas especialistas baseado na web do Consultor (LOMA); <http://www.MyMajors.com>, que fornece conselhos para alunos do ensino médio ou calouros da faculdade ajuda na seleção de um curso imitando um profissional orientador acadêmico.

O FUTURO

Os sistemas especialistas agora atingiram o estágio de maturidade em seu desenvolvimento. O as primeiras pesquisas estabeleceram a viabilidade dessa abordagem. Muitos comerciais sistemas foram desenvolvidos desde então para mostrar seu uso em ambientes reais. Hoje em dia, os encontramos como sistemas especializados híbridos ou incorporados na web páginas na Internet (29–33).

A limitação básica, atualmente, é construir sistemas especialistas com heurística e conhecimento empírico, em vez de conhecimento profundo, que inclui modelos de relações funcionais e causais subjacentes a um problema. No futuro, mais sistemas podem ser desenvolvidos usando modelos funcionais e causais usando uma variedade de representações.

Usando várias fontes de conhecimento (ou seja, especialistas no domínio) em uma cooperativa maneira ainda é um problema difícil esperando para ser resolvido.

CIÊNCIA DOS DADOS

O século XXI iniciou se com a era da economia dos dados, cujo DNA engloba percepções e conhecimentos. Na ciência aberta, observa se uma demanda crescente por dados com base nas pesquisas científicas, dos experimentos, observações, simulação para serem tratados na solução das questões globais e sociais como mudanças climáticas e degradação ambiental dos desafios políticos, econômicos, tecnológicas e sociais.



Fonte: Lee et al, 2018

A jornada evolutiva da análise de dados à ciência de dados começou com a comunidade de estatística e matemática em 1962 [Huber 2011]. Pois, ela é intrinsecamente uma ciência empírica e que o processamento de dados incluiu processamento de informações e a análise exploratória dos dados. Em 2001 expandiu as áreas técnicas da estatística para a Ciência dos dados.

Hoje, a análise descritiva constitui a base para a análise e relatórios padrão das tarefas e ferramentas em projetos e sistemas típicos e inteligência de negócios.

CIÊNCIA DE DADOS

O conceito de ciência de dados foi originalmente proposto pelos estatísticos e comunidade da matemática essencialmente voltada para a análise de dados.

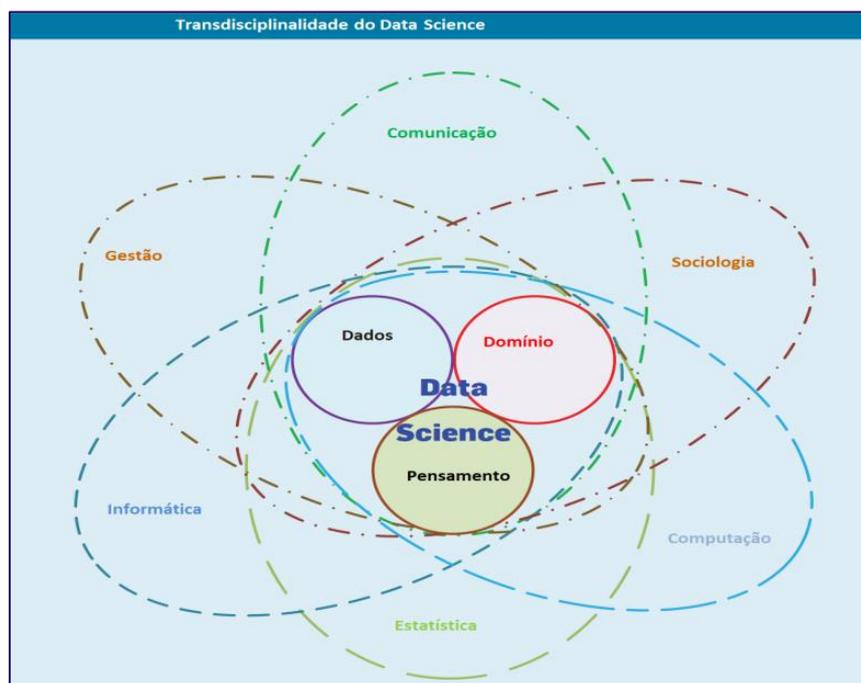
Trata da compreensão sistemática de dados complexos e nas questões relacionadas aos negócios, pois visa traduzir dados em *insights* e formulação de inteligências na tomada de decisão em sistemas complexos.

Definição. A ciência de dados é um campo transdisciplinar que se baseia e sintetiza série de disciplinas e campos do conhecimento como a estatística, informática, computação, comunicação, gestão, sociologia e no estudo do fluxo de dados sob a erige do pensamento da ciência de dados.

Ciência de dados é um termo dado à prática de analisar dados brutos para descobrir quaisquer padrões ocultos. Várias aplicações e ferramentas, como aprendizado de máquina e algoritmos sofisticados, são todos usados neste processo.

Diferentemente de outras formas de análise de dados, ele pode ser aplicado a dados estruturados e não estruturados.

Simplificando, os analistas de dados usam o histórico de processamento para explicar o comportamento dos dados. Esses são modelos que preveem a possibilidade de um determinado evento ocorrer no futuro. Também é útil para análises prescritivas, modelos inteligentes capazes de tomar suas próprias decisões e aprender dentro de parâmetros dinâmicos. A ciência de dados é uma abordagem focada no futuro na utilização de dados. Ao explorar dados atuais ou passados com ferramentas e algoritmos sofisticados, a ciência de dados é capaz de identificar padrões e fazer previsões precisas. Empregado corretamente, é capaz de responder a perguntas abertas, como o que, por que ou como certos eventos podem ocorrer.



Fonte: Cao, 2017

COMPLEXIDADES EM CIÊNCIA DE DADOS

Na ciência da computação, a estrutura de dados é uma maneira específica de organizar e armazenar dados num computador para que possa ser acessado e modificado com eficiência. Mais precisamente, uma estrutura de dados é uma coleção de valores de dados, os relacionamentos entre eles e as funções ou operações que podem ser aplicadas aos dados. O objetivo central da ciência de dados é a exploração das complexidades inerentes e na solução dos problemas na estrutura dados e nos sistemas de negócios.

As relações complexas dos dados originam dos diversos acoplamentos abrangentes que muitas vezes não podem ser descritíveis por meio das teorias e sistemas da associação, correlação, dependência e da causalidade. Tais acoplamentos podem incluir formas explícitas e implícitas, estrutural e não estrutural, semântica e sintática, hierárquica e vertical, local e global, relações tradicionais e não tradicionais, evolução e o efeito.

Desta forma, a complexidade dos dados refere-se ao grau de sofisticação das suas características, o significado, as abrangências, o comportamento, seus domínios, os fatores sociais associados, o contexto do meio ambientes e o grau de aprendizado associado. Reflete ainda o nível de sofisticações das circunstâncias incluindo a escala, suas dimensões, o grau de desequilíbrio extremo, nível de interatividade e o processamento *online* e em tempo real, as aplicações em mídia cruzada, as fontes mistas e híbridas, o grau de dinamicidade e e heterogeneidade, frequências associadas, grau de incerteza, os ruídos associados com os dados, nível de estruturas, a hierarquia e grau de distribuição pouco claras, a escassez e disponibilidade pouco clara às vezes críticas.

Outra questão importante para os cientistas de dados são as relações complexas e ocultas existentes na estrutura dos dados que são críticos para o correto entendimento na sua organização e análise.

A complexidade do comportamento refere-se à desafios envolvidos na compreensão do que realmente acontece nas atividades e nos negócios, conectando-se à semântica e aos processos, os assuntos comportamentais e objetos no mundo físico frequentemente ignorados ou simplificado no mundo dos dados gerado pela atividade física para aquisição e conversão de dados e o sistema de gestão.

As complexidades do comportamento são incorporadas e acopladas nos comportamentos individuais, coletivos e de grupo, nas redes, na divergência, convergência e recorrência de comportamentos, dinâmica do raciocínio em grupo, histórico dos fatos, descrição e previsão no mundo físico altamente heterogêneo, insights, impacto, utilidade e de efeito e de emergência e a gestão do comportamento e da inteligência.

Os dados das pesquisas sistemáticas e contextualizados estão disponíveis para quantificar, para representar, analisar, raciocinar e gerenciar os comportamentos complexos de forma compreensível.

Desta forma, os cientistas de dados reconhecem cada vez mais a complexidade de domínio dos dados como um fator crítico para descoberta das características intrínsecas dos dados, o seu valor e possíveis insights. Complexidades do domínio são refletidos em um problema domínio como fatores de domínio, domínio processos, normas, políticas, versões qualitativas conhecimento quantitativo do domínio, conhecimento especializado, hipóteses, meta-conhecimento, envolvimento e interação com especialistas e profissionais, múltiplos e entre domínios interações, aquisição de experiência, síntese homem-máquina e papéis e liderança no domínio.

No entanto, a análise de dados existente concentra-se principalmente no conhecimento de domínio.

A complexidade social está incorporada à atividade dos negócios e os dados relacionados e é uma parte essencial na compreensão dos dados e dos negócios. Pode ser incorporado nesses aspectos dos problemas de negócios como redes sociais, comunidade emergência, dinâmica social, impacto evolução, convenções sociais, contextos sociais, cognição social, inteligência social, mídias sociais, formação e evolução dos grupos, interação e colaboração em grupo, econômica e cultural fatores, normas sociais, emoção, processos de influência da opinião e dos sentimentos, questões sociais, incluindo segurança, privacidade, confiança, risco e responsabilidade em contextos sociais.

Oportunidades promissoras interdisciplinares surgem quando a ciência social encontra com a ciência dos dados.

A complexidade do ambiente é outro fator importante na compreensão dos dados complexos e problemas de negócios, como refletido em fatores ambientais (contextual), contextos de problemas e dados, de dinâmica, de envolvimento adaptativo, em contextos complexos de interações entre os negócios, sistemas ambientais e de dados, mudanças significativas no ambiente de negócios e seus efeitos nos sistemas de dados, variações e incerteza nas interações entre dados comerciais o ambiente de negócios.

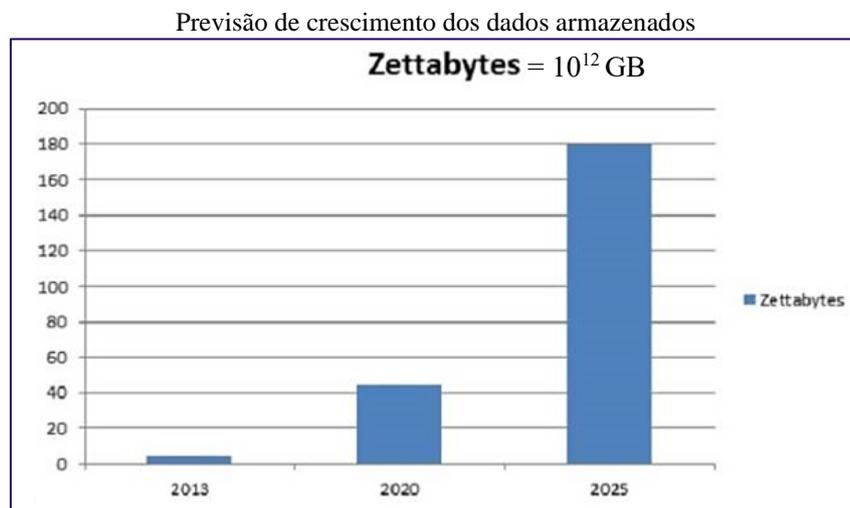
A complexidade de aprendizagem (processo e sistema) deve ser abordada para alcançar o objetivo da análise de dados. Pois os desafios na análise de dados incluem o desenvolvimento de metodologias, estruturas de tarefas comuns e aprendendo paradigmas para lidar com os dados, domínio, comportamental, social e complexidade ambiental.

Assim os cientistas dos dados devem ser capazes de aprender com as fontes e entradas heterogêneas, insumos paralelos e distribuídos e sua dinâmica infinita em tempo real; Apoio, suporte, a aprendizagem ativa e adaptativa on-the-fly dos grandes volumes de dados em ambientes de

computadores com poucos recursos (como sensores incorporados), bem como aprendizagem multisource considerando as relações e interações entre sensores; possibilitar aprendizado combinado com vários objetivos, fontes, conjuntos de recursos, métodos analíticos, estruturas e resultados; aprender relacionamentos de acoplamento de mistura de dados heterogêneos; garantir a transparência e certeza dos modelos de aprendizagem e seus resultados.

O universo digital é criado e definido por softwares aplicativos de diversas naturezas e origens que continuamente geram, coletam, analisam, estruturam e disponibilizam enorme quantidade de dados para a tomada de decisões dos humanos.

Os estudos da International Data Corporation (IDC-2019) estimam a produção aproximada de cinco trilhões de gigabytes (GB) de dados e que dobra a cada dois anos e projeta-se até o início da década de 2021 uma expansão acima de 44 trilhões de GB.



Fonte: IDC, 2019

Os dados são caracterizados como figuras ou fatos que podem ser armazenados ou processados por computadores. E são provenientes de variadas fontes incluindo os humanos conectados on-line, as organizações na gestão de seus negócios com a tecnologia da informação (TI) além da infinidade de sensores e de dispositivos eletrônicos inteligentes conectados à Internet.

Os humanos geram e consomem informações pelas inúmeras atividades como pesquisas na web, e-mails, *upload* e *download* de conteúdo e compartilhamento de arquivos de textos, desenhos, fotos, gráficos, áudio e vídeo nas redes sociais que contribui significativamente para o crescimento do universo digital.

Os termos "dados" e "informação" estão intimamente relacionados e é comum que ambos sejam usados de forma intercambiável. No entanto, é importante entender as suas diferenças.

Os dados por si só formam uma coleção de fatos que requerem ser processados para terem significados e, portanto, serem úteis. Por exemplo, um conjunto dos resultados de vendas anuais de uma organização é um dado. Quando os dados são processados e apresentados em um contexto específico como numa determinada região comparado com a previsão por produto, a lucratividade, fatia de mercado atingido ou por cliente, podem ser interpretados de uma maneira útil na tomada de decisão. Esses dados processados e organizados são chamados de informações.



DADOS NÃO ESTRUTURADOS

Possui a estrutura interna em forma de bits e bytes, mas não é predefinido com modelos ou esquemas, ou seja, não é organizado e rotulado para identificar relacionamentos significativos entre si. Pode ser textual ou não textual, gerado pelos humanos ou por máquinas e são armazenados em banco de dados não relacionais como o NoSQL¹.

¹ Not Only Structured Query Language, é um termo genérico que representa a classe de bancos de dados não relacionais que fornecem mecanismos para armazenamento e recuperação de dados, que não são modelados com relações tabulares usadas nos bancos de dados relacionais.

Dados não estruturados gerados por humanos

- **Arquivos de texto:** arquivos de processamento de texto, planilhas, apresentações, e-mails.
- **E-mail s:** formado em grande parte por textos, mas possui alguma estrutura interna pelos metadados ("para", "de", "data / hora", "assunto" inseridos num e-mail), que mescla dados não estruturados via corpo da mensagem. Desta forma, o e-mail também é conhecido como dados semiestruturados.
- **Mídias sociais:** geralmente são dados semiestruturados como e-mail, contém dados não estruturados (Tweet), dados estruturados (número de "Curtidas", "retweets", Sites: YouTube, Instagram etc. com dados não estruturados e estruturados).
- **Dados móveis:** mensagens de texto, locais.
- **Comunicações:** IMs (instant messages), gravações com ditafone.
- **Mídia:** MP3, fotos digitais, gravações de áudio e arquivos de vídeo.
- **Aplicativos de negócios:** documentos do MS Office, PDFs e similares.

Dados não estruturados gerados por máquinas

- **Imagens de satélite:** dados meteorológicos, formas geográficas, movimentos militares.
- **Dados científicos:** exploração de petróleo e gás, exploração espacial, imagens sísmicas e dados atmosféricos.
- **Vigilância digital:** CFTV.
- **Dados gerados pelos sensores em geral:** temperatura, câmeras, imagens em geral, batimentos cardíacos, sinais elétricos, áudios em geral, documentos eletrônicos em geral.

Os dados não estruturados são responsáveis pela maioria dos dados corporativos e está crescendo a uma taxa média de 55% a 65% ao ano devido a:

- redução de custos de armazenamento de dados e poder de processamento;
- aumento no uso da tecnologia na criação e gestão dos produtos e serviços (PCs, dispositivos móveis, sensores e IoT); e
- Internet e crescente interconexão de dispositivos e de dados.

DADOS ESTRUTURADOS

Residem em bancos de dados relacionais estabelecendo as relações entre os dados e são gerenciados por meio de um sistema de gerenciamento de banco de dados relacional ("RDBMS").

A estrutura do banco de dados é representada por uma tabela de linhas e colunas estabelecendo as relações.

ID	NOME	SOBRENOME	IDADE
01	José	Silva	32
02	Mário	Gomes	29
03	Sebastião	Pereira	40

A vantagem dos dados estruturados é a rotulagem para descrever os atributos e relacionamentos com outros dados. Essa estrutura permite a consulta humana gerada por um algoritmo.

SQL	RDBMS	DESCRIÇÃO
Linha	Tupla/registro	Conjunto de dados representando um único item, por exemplo Dados de José Silva descritos acima
Coluna	Atributo/campo	Elemento específico e marcado de uma coluna, por exemplo "Idade"
Tabela	Relação	Conjunto de linhas e colunas que compartilham os mesmos atributos, ou seja, organizando as mesmas informações sobre um conjunto de objetos de dados

BIG DATA

Big Data é um termo usado para denotar a coleção de conjuntos de dados grandes e complexos, dificultando o processamento usando aplicativos de processamento de dados herdados.

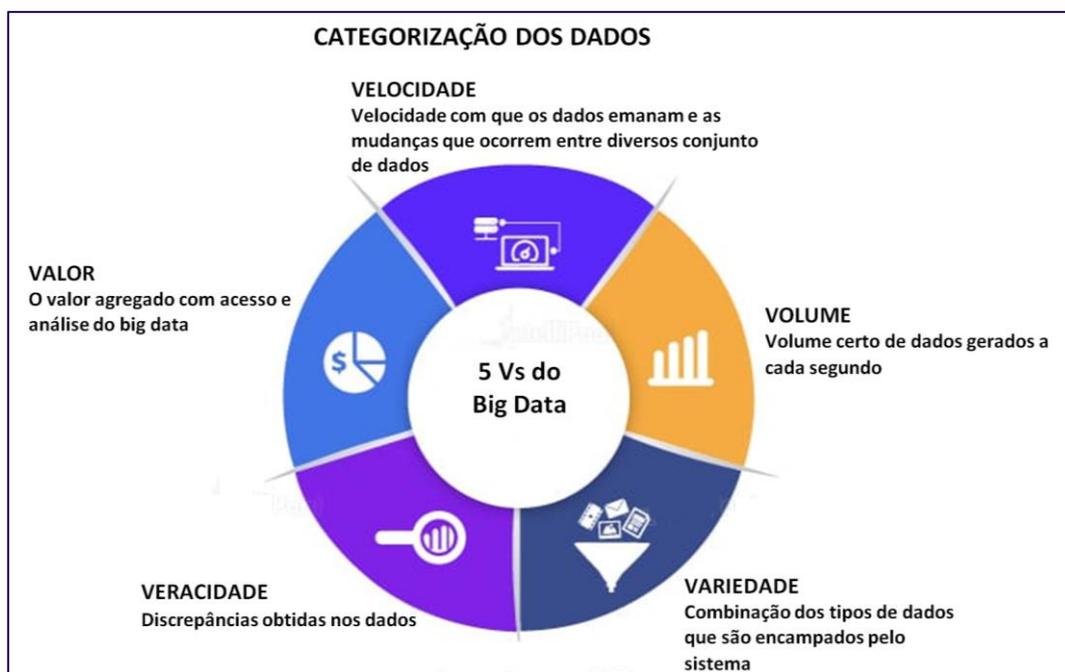
O Big Data refere-se a grandes conjuntos de dados coletados de forma privada ou pública por humanos, empresas e governos que são complexos e de grandes proporções que os métodos tradicionais de processamento de dados não são adequados para manipulação e estruturação para o seu bom entendimento. Esses conjuntos de dados são extremamente valiosos devido às vastas informações ocultas que abrigam nas suas estruturas. Requer desta forma serem analisados de forma precisa sobre seus padrões, tendências e associações ocultas, especialmente no contexto da tomada de decisão humana.

O termo Big Data foi cunhado por Doug Laney no início dos anos 2000 que inclui os seguintes conceitos:

- **Volume:** o tipo e o detalhe dos dados que estão sendo coletados para serem armazenados que continuam a crescer, e são armazenados nas nuvens públicas e/ou privadas.
- **Velocidade:** a velocidade na qual os dados são coletados em tempo real a taxas incrivelmente rápidas por meio de sensores e outros dispositivos de vídeo, áudio e imagens.

- **Variedade:** os tipos de dados que estão sendo coletados. Considerando que dados demográficos básicos, atitudes e opiniões e, possivelmente outras informações geográficas ou dados do comportamento do consumidor on-line que estão sendo monitorados e capturados entre outros.
- **Veracidade:** quantidades excessivamente grandes de dados podem dificultar a sua identificação no tocante à importância ou as distorções que descreve o nível de "ruído" existente.

Com a rápida evolução e a tecnologia de coleta dos dados foram categorizados como Big Data ampliando os conceitos iniciais como na figura a seguir:



Fonte: Gartner, 2017

1. **Volume:** refere-se ao volume de dados que são tremendamente grandes com o aumento exponencial a cada instante com os dispositivos de captura instalados em todos os ecossistemas.



Fonte: Gartner, 2017

2. Variedade: Uma razão para esse rápido crescimento do volume de dados é que os dados são provenientes de diferentes fontes em vários formatos.

Os dados são classificados da seguinte forma:



Fonte: Gartner, 2017

a) Dados estruturados: Os dados se apresentam num esquema organizado com todas as informações necessárias para fazer um sentido. Geralmente são armazenados num sistema de gerenciamento de banco de dados. Por exemplo, na tabela de produtos que estão num formato organizado conforme esquema a seguir:

Código do produto	Descrição	Quantidade	Preço	Fornecedor
3245685-32	Refrigerador 400 l, 220 V, branca	5	1200,00	GE
4568327-25	Liquidificador 220V, branca, 2l	8	300,00	Hamilton

- b) Dados semiestruturados:** os dados são apresentados tanto na forma semiestruturada quanto estruturada cuja estrutura não está claramente definida, por exemplo os dados no formato JSON, XML, CSV, TSV² e E-mail; dados do aplicativo da web não estruturados que contêm arquivos de histórico de transações, arquivos de log etc. Os sistemas OLTP (Online Transaction Processing)³ que são criados para trabalhar com dados estruturados e são armazenados em forma de relações ou tabelas.
- c) Dados não estruturados:** nesta categoria são todos os arquivos não estruturados como arquivos de vídeo, de log, de áudio e de imagem, estão incluídos. Quaisquer dados que possuam um modelo ou estrutura desconhecida são classificados como dados não estruturados. Como o seu tamanho é grande, os dados não estruturados possuem vários desafios em termos de processamento para extrair valor deles. Um exemplo disso é uma fonte de dados complexa que contém uma mistura de arquivos de texto, vídeos e imagens. Várias organizações têm muitos dados disponíveis, mas não sabem como extrair valor delas, pois os dados estão em sua forma bruta.
- d) Dados quase estruturados:** Neste formato incluem dados textuais com formatos inconsistentes que podem ser formatados com esforço e tempo, e com a ajuda de várias ferramentas. Por exemplo, os logs do servidor da web que é criado e mantido automaticamente por algum servidor que contém uma lista de atividades.

² São formatos principais de dados geralmente encontrados pelos analistas no mundo dos negócios:

JSON - JavaScript Object Notation: formato mais frequentemente encontrado em sistemas baseados na Web e foi criado para resolver a necessidade de representar um "estado" (isto é, o que aconteceu logo antes de um determinado evento) entre o servidor e o navegador.

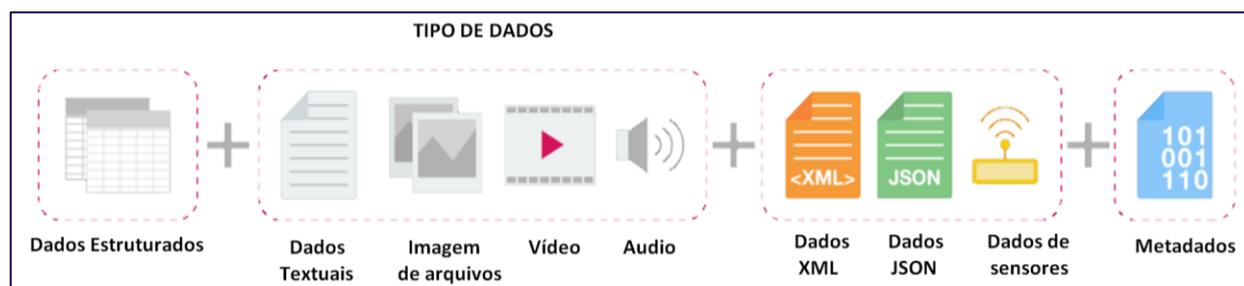
XML - Extensible Markup Language: formato extremamente detalhado tornando o tamanho do arquivo muito grande, o que o torna em um formato não ideal, onde o padrão e o tamanho do arquivo é considerado de grandes proporções.

CSV - Comma Separated Values: trata do formato de arquivo muito comum e é fácil entender pois minimiza os caracteres extras de formatação (para que haja apenas vírgulas e novas linhas) e mapeia 1 para 1 de como eles serão apresentados em um banco de dados ou planilha (ou seja, um ", " adicional significa uma célula). Se você estiver exportando de um banco de dados relacional que já existe em um espaço bidimensional, um CSV é a opção de formatação mais comum.

TSV - Tab Separated Values: TSV é separado usando tabs. Como a maioria dos editores de texto não distinguem facilmente o visual entre uma guia e os espaços.

³ OLTP é uma categoria de processamento de dados focada em tarefas orientadas a transações envolvendo a inserção, atualização e/ou exclusão de pequenas quantidades de dados em um banco de dados e lida principalmente com um grande número de transações por um grande número de usuários.

Desta forma podemos resumir os tipos de dados na seguinte estrutura:



Fonte: Gartner, 2017

Dados típicos não estruturados gerados por humanos incluem:

- **Arquivos de texto:** arquivos de processamento de texto, planilhas, apresentações em Power point, e-mails.
- **E-mail:** em grande parte são textos, mas possui alguma estrutura interna graças a seus metadados (por exemplo, incluindo os textos visíveis "para", "de", "data / hora", "assunto" inseridos para enviar um e-mail), mas também mescla de dados não estruturados via corpo da mensagem. Por esse motivo, o e-mail também é conhecido como dados semiestruturados.
- **Mídias sociais:** como e-mail, geralmente são dados semiestruturados, contendo dados não estruturados (por exemplo, um Tweet), mas também dados estruturados (por exemplo, o número de "Curtidas", "retweets", "data", "autor" etc.).
- **Sites:** YouTube, Instagram etc. contêm muitos dados não estruturados, mas também muitos dados estruturados, por exemplo como descrito acima para o Twitter
- **Dados móveis:** mensagens de texto, locais.
- **Comunicações:** IMs, gravações com ditafone.
- **Mídia:** MP3, fotos digitais, gravações de áudio e arquivos de vídeo.
- **Aplicações comerciais:** documentos do MS Office, PDFs e similares.

DADOS NÃO ESTRUTURADOS GERADOS POR MÁQUINAS

Tipos comuns de dados não estruturados gerados por máquina incluem:

- **Imagens de satélite:** dados meteorológicos, formas geográficas, movimentos militares.
- **Dados científicos: exploração** de petróleo e gás, exploração espacial, imagens sísmicas e dados atmosféricos.
- **Vigilância digital:** CFTV⁴.

⁴ circuito fechado de televisão: câmeras e sensores de vídeo instaladas em locais específicos possuem recursos de inteligência e tecnologia de alta performance, o que possibilita obter informações através das imagens e gerencia-las.

- **Dados legais não estruturados:** no contexto jurídico são comuns os dados nas seguintes áreas:
 - **Gerenciamento de documentos/e-mail:** embora a organização do DMS esteja estruturada (por exemplo, metadados básicos (dados sobre dados): nomes de arquivos, IDs de documentos, números de versão, datas de criação / edição / leitura, etc.), o conteúdo mais valioso não é estruturado, ou seja, o conteúdo dos documentos constituídos e e-mails. Assim, geralmente torna se dificultoso pesquisar e analisar esses dados de maneira específica como encontrar um texto de uma cláusula específica por não encontrarem estruturados sobre o conteúdo desse documento (ou seja, até o nível da cláusula ou da intra-cláusula), apenas os metadados básicos.
 - **Acesso eletrônico:** neste caso, a maior parte do conteúdo analisado é e-mail , anexos de e-mail (ou seja, documentos, imagens, PDFs e, por vezes, voz do MS Office) e, naturalmente, sofrem das mesmas limitações descritas para o gerenciamento de documentos e e-mails.
 - **Diligencia legal:** o conteúdo é quase exclusivamente documentos do MS Word e PDF, mas também às vezes planilhas e decks de slides - novamente, como antes, tudo é desestruturado além dos metadados básicos.
- 1. Velocidade:** a velocidade da acumulação de dados também desempenha um papel na determinação se os dados são categorizados como Big Data ou regulares.
- Como pode ser visto na imagem abaixo, os aplicativos da web compartilhados por uma infinidade de dispositivos muito fáceis de acessar, e portanto, muitos dados são gerados.



Fonte: Gartner, 2017

- 2. Valor:** Os benefícios com a extração de dados como mecanismo para extrair seus significados correto. Requer acima de tudo um processo de extração para transformar dados brutos em dados úteis. Em seguida, procede-se à análise dos dados com o descarte ou recuperação a partir dos dados brutos. Para depois garantir que qualquer análise que tenha feito beneficie os negócios para descobrir insights, atingir os resultados que não eram possíveis anteriormente.



Fonte: Gartner, 2017

Requer assim, a certeza de que, independentemente dos dados brutos obtidos houve a consistência e a seleção dos dados adequados para obter informações esperadas. Pois o grande desafio é que durante o processo de descarte do volume enorme de dados alguns pacotes podem ter sido perdidos.

- 3. Veracidade:** Significa confiabilidade e qualidade dos dados. Pois os pacotes de dados se perdem durante o processo de seleção e execução onde requer observar a importância do estágio de mineração de dados brutos na conversão em dados valiosos. Por exemplo, nas postagens do Facebook, com hashtags, abreviações, imagens, vídeos etc., que as tornam muitas vezes não confiáveis e dificultam a seleção qualitativa do seu conteúdo.

POR QUE O BIG DATA É TÃO IMPORTANTE?

O impacto do Big Data em seus negócios deve ser medido para facilitar a determinação do retorno do investimento. Portanto, o Big Data é definitivamente um problema que vale a pena investigar. Sempre que o navegador visita um site, pode ter notado que no painel direito ou no painel superior ou em algum lugar na tela encontrará um campo de recomendação que é basicamente um

anúncio relacionado às suas preferências. Como a empresa de propaganda sabe que você estaria interessado nela?

Bem, tudo o que navega na Internet é armazenado e todos esses dados são analisados corretamente para que se identifique pelo qual estiver interessado. Obviamente, o navegador estará interessado num anúncio em particular, mas estará navegando ainda mais em outras informações. Mas lembre-se que a quantidade de dados gerados a partir de um único usuário é tão grande que é considerada como Big Data

Dados e informações digitais

Em sua forma mais básica, os dados podem ser descritos como a representação não ordenada e não processada de quaisquer tipos de observações que são quantificadas e armazenadas em forma de símbolos (OECD, 2013).

Como parte da transformação digital, o uso de dados transformadas em informações se tornou cada vez mais valioso para as empresas e estimulou a criação de modelos de negócios totalmente novos. As informações contidas nos dados, frequentemente funcionam como uma entrada crítica nos processos de produção de bens e de serviços. Para algumas empresas, os dados são o ativo mais valioso que possuem. As empresas potencializam os dados como business inteligente, utilizam-nos com o propósito de otimização de processos, melhoria de produtos e serviços e em atividades de pesquisa e desenvolvimento (P&D) (Magalhães e Rosiera, 2017).

Para obter mais informações sobre como exatamente os dados podem gerar valor para as empresas a OECD (2015) apresentou o conceito da cadeia de valor dos dados em quatro estágios no processo transformação dos dados: i) coleta de dados, ii) agregação de dados, iii) análise de dados e iv) uso e monetização de dados que são sustentados por armazenamento de dados e fluxos de dados (além da fronteira) que ocorrem em toda a extensão conforme a figura seguinte:

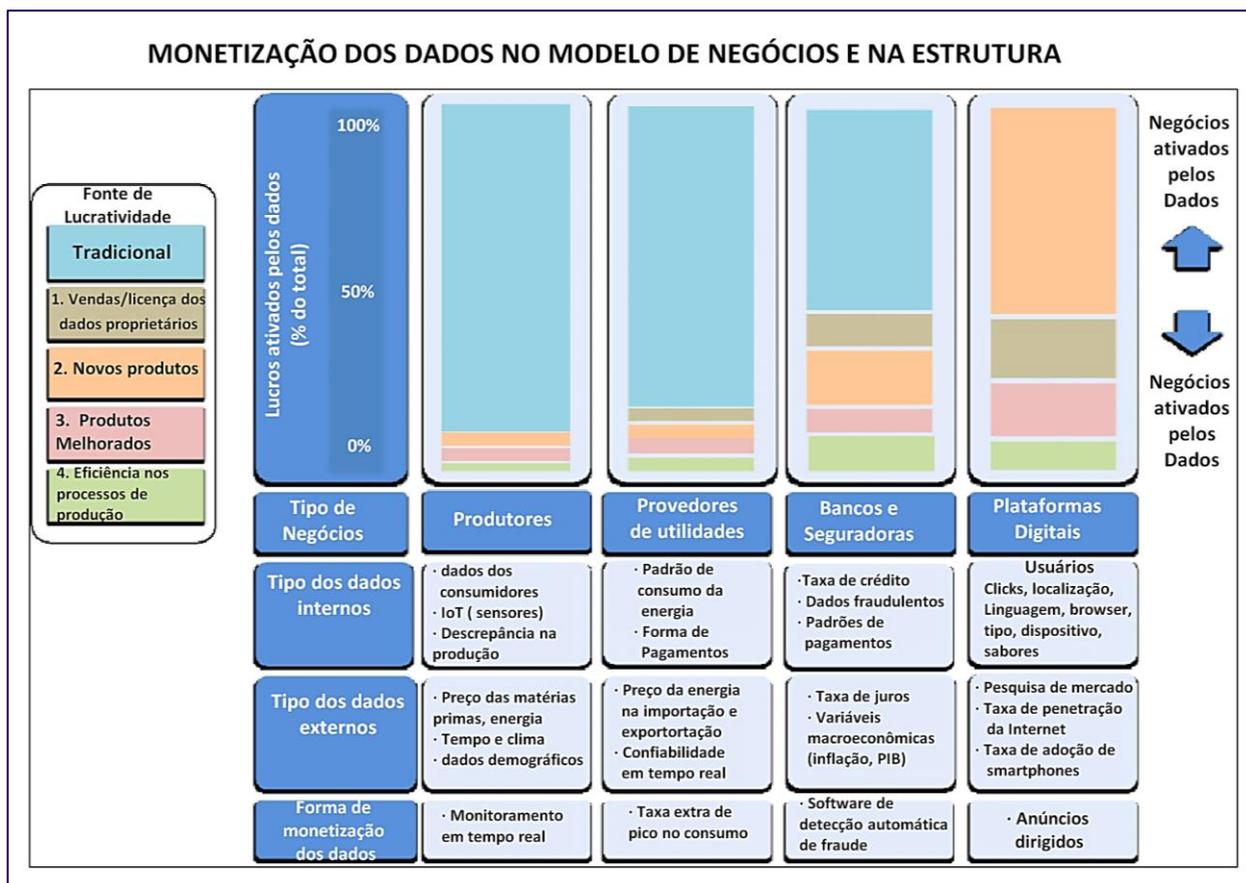


Estágio 1: A coleta de dados brutos pode ocorrer num único ponto ou em vários pontos em muitas localidades simultaneamente.

Estágio 2: Criação de valor adicional pode ser limitada quando as transações ocorrem com barreiras da distância física (transfronteiriças) no processo de transferência e na agregação de dados.

Estágio 3: Na etapa da análise de dados agregados ou múltiplos conjuntos de dados combinados pode ocorrer em simultaneamente em múltiplas localidades.

Estágio 4: Finalmente, o estágio de monetização gera dados adicionais que podem alimentar a cadeia de valor de dados para torná-la um “ciclo de valor de dados” global. Assim as empresas podem se beneficiar das economias de escala centralizando todos os seus dados num Data Center e/ou replicar cópias para proteção de desastres e para reduzir o tempo de acesso (latência) ou a garantia de continuidade do negócio. A computação em nuvem é uma tecnologia habilitadora chave desse processo.



Fonte: Adaptado de Nguyen e Paczos, 2020

Dados como ativo baseado em conhecimento

A literatura de ativos intangíveis contém série de *insights* quando se pensa em medir o valor dos dados, uma vez que bancos de dados e software formam uma categoria existente de ativos intangíveis. Corrado et al. (2005) descreve um método da “melhor prática” para medir o investimento intangível na estrutura das contas do ativo, distinguindo as informações digitalizadas, a propriedade inovadora e competências econômicas e do capital organizacional.

Esta abordagem pode ser útil quando se pensa em dados como um ativo intangível baseado em conhecimento que contribui para o processo de produção das empresas e de relatórios descritivos para medir o valor dos dados digitalizados (OECD, 2014).

O método baseado em custos usa dados do mercado de trabalho, para o número de desenvolvedores de software de computador (ou outras ocupações relacionadas a dados, como pesquisadores econômicos), seus custos médios por hora, combinados com suposições sobre o uso do tempo, custos de mão de obra não salariais e sobrecargas, a fim de estimar os valores das bases de dados.

Ciclo de vida dos dados

Os dados estruturados são armazenados no banco de dados relacional e são organizados em linhas e colunas na forma de tabelas nomeadas.

Os dados semiestruturados não possuem um modelo formal, mas possuem um padrão e estrutura aparentes e autoexplicativos, o que permite a sua análise. Nelas incluem planilhas com estrutura de linhas e colunas e definidos por um esquema de arquivos XML⁵.

Como um organismo biológico, os dados possuem um ciclo, da origem à vida ativa à mortalidade ou alguma forma de expiração. Os dados nunca existem no vácuo. Como um organismo vivo e inteligente ele sobrevive num ambiente que fornece suporte físico, contexto social e significado existencial. Desta forma é fundamental a sua compreensão, as oportunidades e os desafios que possa aproveitar ao máximo os dados digitais na obtenção da vantagem competitiva como na figura a seguir:



Fonte: Berman et al, 2018

Existe uma relação valiosa entre a ciência da computação, ciência da informação, domínio e comunidades de ciências físicas e da engenharia para reduzir lacunas e oportunidades entre o aprendizado de máquina, análise de dados e disciplinas relacionadas como estatísticas e pesquisa operacional.

⁵ XML (Extensible Markup Language) é uma recomendação da W3C (World Wide Web Consortium) para gerar linguagens de marcação para descrever dados. O padrão XML é uma maneira flexível de criar formatos de informações e compartilhar eletronicamente os dados estruturados através da Internet pública, bem como através de redes corporativas.

Os dados são rastreados, monitorados e identificados, mas existem desafios e questões como: podem ser identificados quem os criou? como são coletados e quem são os usuários? existem adequados controles de privacidade e segurança? os dados são bem cuidados? quem está comprometido em mantê-los, em que formatos, e por quanto tempo? quem são os financiadores e administradores de dados? como será armazenado e migrado para a próxima geração dos meios de comunicação?

Utilidade e forma de detecção dos dados: para quem são disponibilizados? quais serviços são necessários para melhorar o uso? quais metadados e outros formatos são necessários para sua reprodução e replicação?

A administração de dados é compatível com a política e boas práticas. Existem padrões e política apropriada em relação a relatórios, propriedade intelectual e outras questões e preocupações? Os direitos, licenças e outras propriedades que estão claras? E quais dados e os metadados devem ser mantidos, quem os possui e seus subprodutos, e quem tem acesso ao todo ou partes dele?

Os dados estão em toda parte, proporcionando crescentes desafios técnicos e sociais, de governança, ética, política e de privacidade e possuem uma série de características distintas, que contribuem para sua natureza única como ativo intangível.

Esses são um importante determinante da usabilidade em diferentes modelos de negócios. Em princípio, os dados podem ser explorados e reexplorados infinitamente com baixo custo marginal. Normalmente, é a infraestrutura e a análise de dados que determinam principalmente os custos de reutilização dos dados.

Laney (2011) discute uma série de fatores que afetam a utilidade dos dados sobretudo o valor econômico na perspectiva de negócios identificando as características objetivas e subjetivas dos dados.

Pois os dados têm maior probabilidade de fornecer uma base para monetização e criação de valor se forem:

- Mesclado com os outros conjuntos de dados;
- Acessível - facilmente recuperado e/ou integrado aos processos de negócios;
- Desagregado no nível desejado;
- Grau de atualização com frequência suficiente para atender aos requisitos de negócios (anualmente, diariamente ou em tempo real);
- Confiável pelos usuários; os dados são imparciais e não dependem do julgamento, interpretação ou avaliação de indivíduos (Open Data Watch, 2018);

- Representativo som campos ausentes, os dados e suficiente para atender aos requisitos do negócio;
- Raro ou secreto, proprietário e difícil de acessar.

Mais importante ainda, os dados são de natureza não rival (OECD, 2013a; Mandel, 2017) e não estão sujeitos ao desgaste normal dos ativos tangíveis. Isso significa que os dados podem ser usados várias vezes sem diminuir seu valor inerentemente.

Tipos de dados

A OCDE (2013a) classifica os tipos de dados com base na propriedade ou direito de uso, a fonte, métodos e financiamento de coleta, acesso e manutenção conforme a tabela seguinte:

VISÃO GERAL DO TIPO DE DADOS

Financiamento, coleta e manutenção dos dados	Setor privado	Financiado, criado, mantido e gerido pelas instituições privadas. Dados internos de processos de produção ou estoques no varejo mantidos em banco de dados
	Setor público	Dados sobre registros de pacientes em saúde ou registros fiscais individuais, ou dados originados dos satélites
Proprietário ou direito de uso	Proprietários	Dados com Propriedade claramente definida que são protegidos por Direitos de Propriedade Intelectual (LGPD) ou quaisquer outros direitos com efeito semelhante (OCDE, 2019a); isso pode incluir dados individuais, bem como dados organizacionais.
	Dados abertos (ou de domínio público)	Estão publicamente disponíveis (ao contrário de dados proprietários), de uso livre por qualquer pessoa para qualquer finalidade, sem quaisquer restrições legais (Habeas data). Não é protegido por direitos de propriedade intelectual, direitos autorais ou quaisquer outros direitos legais semelhantes.
Titularidade dos Dados	Dados pessoais	Dados individuais que permitem a identificação do titular. Pode cobrir dados do setor público e privado, por exemplo. conteúdo gerado pelo usuário (blogs, fotos, tweets) ou dados de geolocalização de celulares, bem como dados do setor público como RG, CNH, CPF ou números de previdência social).
	Dados organizacionais	Permitem a identificação e são geralmente controlados pelas próprias organizações, seja legalmente ou por motivos contratuais. Também pode ser mantido por órgãos públicos, como autoridades fiscais e geralmente são dados comercialmente confidenciais.
Geração de dados	Criados pelo usuário	disponibilizados por humanos (dados de rastreamento de telemetria, de comportamento do consumidor coletados por aplicativos móveis ou postagens da mídia social). Podem ser dados voluntários (ativos), dados observados (passivos ou implícitos) ou dados derivados de usuários.
	Gerados por máquina	Comunicação máquina a máquina (M2M); Internet das coisas (IoT), coletados de sensores.
Fonte de dados	Dados internos	Coletados e consolidados em diferentes áreas da organização. Pedidos de compra do departamento de vendas, transações contábeis ou qualquer fonte interna responsável por registrar informações sobre interações comerciais da empresa.
	Dados externos	Obtidos de uma fonte externa como acesso ao banco de dados proprietário sob licença ou outros dados adquiridos externamente.

O valor de um bem ou serviço é o preço de equilíbrio do mercado, ou seja, o ponto de encontro entre a demanda e a oferta. No entanto, existem vários desafios que tornam essa abordagem difícil de aplicar aos dados.

Primeiro: Identificar um mercado bem definido como preferências de marketing no conteúdo dos bancos de dados baseada em transações.

Segundo: o valor dos dados é altamente dependente do contexto e os dados podem ser avaliados de forma diferente entre diferentes fornecedores de dados, usuários e reguladores.

Em um estudo recente patrocinado pela IBM (IBM, 2019), o Ponemon Institute (2018) conduziu entrevistas com mais de 2.200 profissionais de TI, proteção de dados e conformidade de 477 empresas que experimentaram uma violação de dados entre julho de 2017 e julho de 2018. Eles descobriram que o custo médio de violações de dados neste período foi de US \$ 3,86 milhões. Embora as violações de dados envolvam todos os tipos de custos, isso pode pelo menos servir como uma indicação. De acordo com outra pesquisa com 1.800 tomadores de decisão de negócios globais conduzida pela NTT Security (NTT, 2018), as violações de dados acarretam custos significativos para as empresas, incluindo:

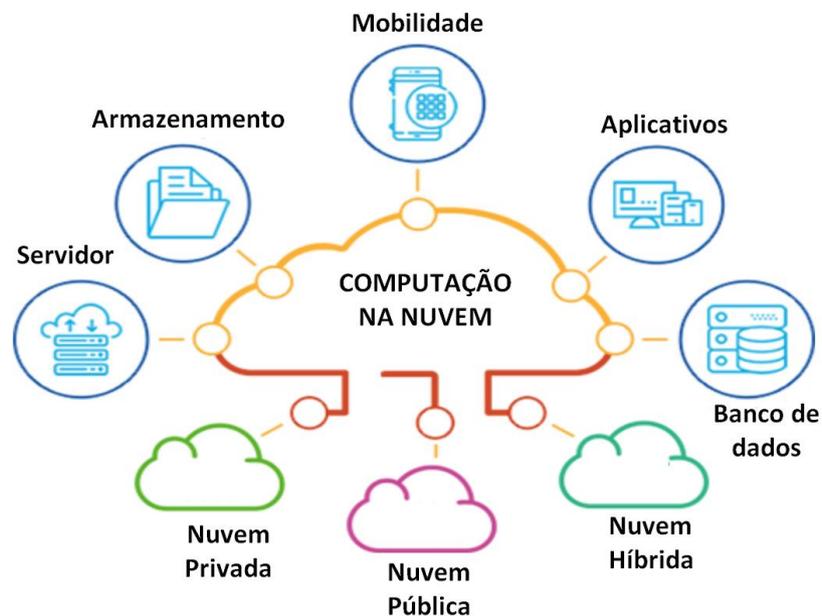
- Perda de confiança do cliente, danos à reputação e marca, perdas financeiras;
- Receita esperada cai acima de 10% em média;
- Custos de recuperação estimados em cerca de USD 1,5 milhões em média;
- Custos de seguro de segurança cibernética, que 38% das organizações possuem.

Esta seção discutiu uma série de abordagens que devem ser exploradas para medir o valor comercial dos dados a preços de mercado. Especificamente, propusemos examinar corretores de dados e seguros para violações de dados relacionadas aos próprios dados ou à infraestrutura de dados. A aplicação exata dessas fontes experimentais de informação para a avaliação dos dados no processo de produção teria necessariamente que ser mais desenvolvida e suas limitações precisariam ser abordadas (por exemplo, separação do 'valor dos dados' dos custos do processo legal ou desvinculação da contribuição de dados para emitir avaliação no caso de informações de seguro). No entanto, consideramos essas abordagens opções interessantes. As seções a seguir examinam outras abordagens que podem ser úteis ao considerar o valor dos dados e fluxos de dados.

COMPUTAÇÃO EM NUVEM

A computação em nuvem é a prestação de serviços incluindo servidores, sistemas de armazenamento, bancos de dados, redes, software, análises e a inteligência pela Internet (“a nuvem”)

para oferecer recursos da inovação flexíveis e economia de escala. Geralmente o custo é pelos serviços sob demanda reduzindo assim as despesas operacionais, uso eficiente da e escalar conforme as necessidades da dinâmica dos negócios.



Desta forma, existe a figura dos provedores de recursos na nuvem que alugam espaço de armazenamento e/ou ciclos de CPU nos computadores de outra empresa e/ou pessoa física sob demanda.

NUVEM PÚBLICA

Este é o modelo de implantação mais comum em que não possui hardware local para gerenciar ou manter atualizado - tudo é executado no hardware do provedor de nuvem. Em alguns casos, pode economizar custos adicionais compartilhando recursos de computação com outros usuários da nuvem.



Ícone de nuvem pública

Vantagens

- Alta escalabilidade/agilidade: não precisa comprar um novo servidor para escalar
- Pagamento sob demanda e não requer manutenção ou atualizações do hardware

- Conhecimento técnico mínimo para configurar e usar: uso das habilidades e os conhecimentos do provedor de nuvem para garantir que as cargas de trabalho sejam seguras, protegidas e altamente disponíveis

Desvantagens

- Nem todos os cenários se encaixam na nuvem pública.
- Requisitos de segurança específicos que não podem ser atendidos usando a nuvem pública
- Políticas governamentais, padrões do setor ou requisitos legais que as nuvens públicas não possam atender
- Hardware ou os serviços que não pode gerenciá-los. Ex. Requisitos comerciais exclusivos, como manter um aplicativo herdado

NUVEM PRIVADA

Criação do ambiente de nuvem no datacenter privado e fornece acesso de autoatendimento para calcular recursos para os usuários na própria organização. Oferece uma simulação de uma nuvem pública para os usuários, mas são responsáveis pela compra e manutenção dos serviços de hardware e software.



Ícone de nuvem privada

Vantagens

- A configuração possa suportar qualquer cenário ou aplicativo herdado
- Tem controle (e responsabilidade) sobre a segurança
- Podem atender a requisitos estritos de segurança, conformidade ou legais

Desvantagens

- Investimentos iniciais de adquirir o hardware para inicialização e manutenção
- Agilidade limitada para escalar pois requer comprar, instalar e configurar novo hardware
- Exigem habilidades e conhecimentos de TI na nuvem

O uso da nuvem privada seria quando a organização possui dados reservados e proprietárias que não podem ser colocados na nuvem pública, ou por razões legais.

NUVEM HÍBRIDA

Combina nuvens públicas e privadas permitindo a execução dos programas aplicativos no local mais apropriado. Por exemplo, pode se hospedar um site na nuvem pública e vinculá-lo a um banco de dados altamente seguro hospedado na sua nuvem privada (ou no data center local).



Ícone de nuvem híbrida

Isso é útil quando você tem alguns dados que não podem ser expostos publicamente (dados médicos, dados financeiros ou pessoais) que precisam ser mantidos em datacenter privado. Ou os aplicativos executados em hardware antigo que não podem ser atualizados. Nesse caso, pode se manter o sistema antigo em execução localmente e conectá-lo à nuvem pública para autorização ou armazenamento.

Vantagens

- Manter todos os sistemas em execução e acessíveis que usem hardware desatualizado ou um sistema operacional desatualizado
- Flexibilidade com o que executa localmente versus na nuvem
- Economias de escala dos provedores de nuvem pública para obter serviços e recursos mais baratos e, em seguida, complementar com o próprio equipamento privado
- Uso do equipamento próprio para atender a cenários de segurança, conformidade ou legados onde requer controlar completamente o ambiente.

Desvantagens

- Selecionar um modelo de implantação, pois envolve custos preliminares de configuração e de gerenciamento

O provedor de nuvem é responsável disponibilizar o hardware físico necessário para executar seu trabalho e por mantê-lo atualizado. Os serviços de computação oferecidos tendem a variar de acordo com o provedor da nuvem.

PODER DE COMPUTAÇÃO

E operações como interação em redes sociais, num e-mail, operação de consulta, compra, venda ou paga uma fatura online está interagindo com servidores baseados na nuvem que processam cada solicitação e retornando com uma resposta. Ou seja, todo consumidor, dependemos dos serviços de computação fornecidos pelos vários provedores de nuvem que compõem a Internet.

Ao criar soluções com base na computação em nuvem, existe a possibilidade de escolher como deseja que o trabalho seja realizado com base em seus recursos e necessidades.

A máquina virtual (VM) é uma emulação do computador - assim como um desktop/laptop/dispositivos móveis que permite ter mais controle e responsabilidade sobre a operação e manutenção dos recursos na nuvem. Pois, cada VM inclui um sistema operacional e hardware que aparece ao usuário como um computador físico executando o Windows, OS ou Linux. Você pode instalar o software necessário para executar as tarefas que deseja executar na nuvem. Assim, não requer ao usuário a aquisição de nenhum hardware ou instalar o sistema operacional. O provedor de nuvem executa sua máquina virtual em um servidor físico em um de seus datacenters - geralmente compartilhando esse servidor com outras VMs (isoladas e seguras) a um custo menor do que um computador físico.

As VMs não são a única opção de computação - existem duas outras opções populares: contêineres e computação sem servidor.

O que são contêineres?

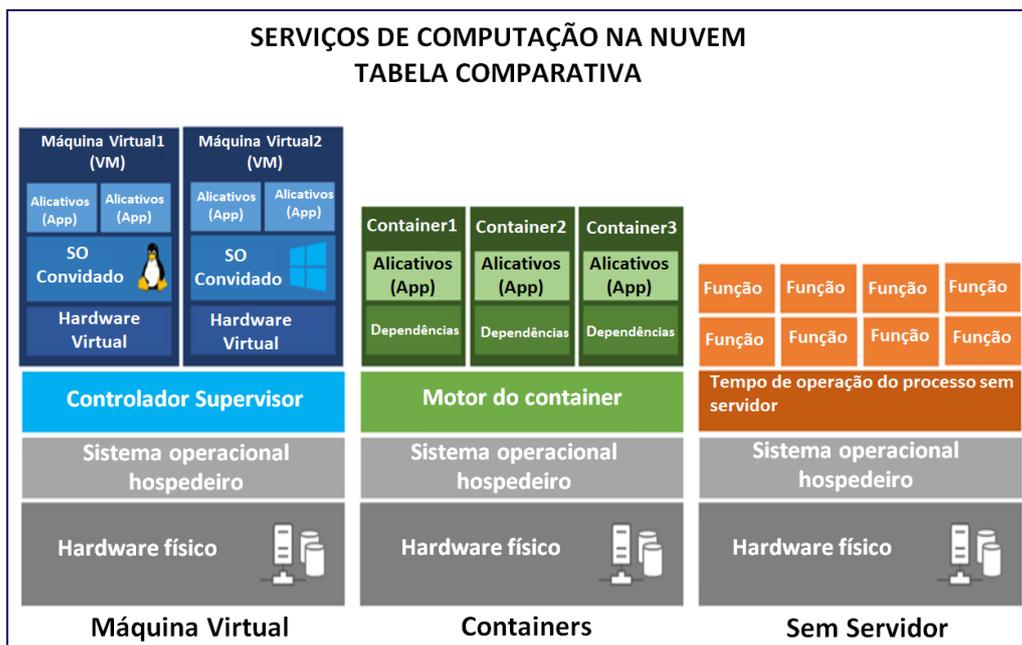
Os contêineres fornecem um ambiente de execução consistente e isolado para aplicativos. Eles são semelhantes às VMs, exceto que não requerem um sistema operacional convidado. Em vez disso, o aplicativo e todas as suas dependências são empacotados em um "contêiner" e, em seguida, um ambiente de tempo de execução padrão é usado para executar o aplicativo. Isso permite que o contêiner seja iniciado em apenas alguns segundos, porque não há SO para inicializar e inicializar, mas um aplicativo para inicialização.

O que é computação sem servidor?

A computação sem servidor permite executar o código do aplicativo sem criar, configurar ou manter um servidor. A ideia principal é que seu aplicativo seja dividido em funções separadas que são executadas quando acionadas por alguma ação. Isso é ideal para tarefas automatizadas - por exemplo, criar um processo sem servidor que envia automaticamente uma confirmação por e-mail depois que um cliente faz uma compra online.

O modelo sem servidor difere das VMs e contêineres, pois é sob demanda onde o custo é apenas pelo tempo de processamento usado por cada função que é executada. As VMs e contêineres são cobrados enquanto estão em execução - mesmo que os aplicativos neles estejam ociosos. Essa arquitetura não funciona para todos os aplicativos - mas quando a lógica do aplicativo pode ser separada em unidades independentes, pode se testá-los separadamente, atualizá-los separadamente e iniciá-los em microssegundos, tornando essa abordagem a opção mais rápida para implantação.

Aqui está um diagrama comparando as três abordagens de computação que abordamos:



Fonte: Adaptado de Microsoft, 2020

- **Armazenamento:** espaços e conteúdo como arquivos e bancos de dados

Os provedores de nuvem geralmente oferecem serviços que podem lidar com todos esses tipos de dados como armazenar um texto ou um clipe de filme, conjunto de relacionamentos, catálogo de endereços, imagens fixas ou dinâmicas, sinais de áudio, sinais elétricos ou uma abordagem mais estruturada, como usar um banco de dados.

A vantagem de usar o armazenamento de dados baseado em nuvem é a escalabilidade e a flexibilidade para atender às necessidades para expandir e/ou contrair automaticamente o espaço disponível.

- **Rede:** conexões seguras entre o provedor de nuvem e a empresa
- **Analytics:** visualização de dados de telemetria e desempenho

Principais benefícios da computação em nuvem

A computação em nuvem é uma grande mudança da maneira tradicional como as empresas pensam sobre os recursos de TI conforme as razões pelas quais as organizações estão recorrendo aos serviços:

Custo

A computação em nuvem elimina as despesas de capital na compra de hardware e software e na instalação e execução de datacenters e de servidores locais, o sistema de *no Break* no fornecimento da energia elétrica, sistemas de refrigeração especial e a estrutura de gerenciar a infraestrutura.

Rapidez

Os serviços de computação em nuvem são fornecida pela modalidade de autoatendimento e sob demanda; portanto, recursos de computação podem ser provisionadas em minutos, com apenas alguns cliques do mouse, dando às empresas muita flexibilidade e tirando a pressão do planejamento da capacidade.

Escala global

Os benefícios dos serviços de computação em nuvem incluem a capacidade de escalar elasticamente. Significa, portanto, fornecer a quantidade certa de recursos de TI como a capacidade de computação, armazenamento, largura de banda conforme a necessidade e na localização geográfica necessária.

Produtividade

Os datacenters locais normalmente exigem muita capacidade de "empilhamento" na configuração de hardware, aplicação de patches de software e outras tarefas demoradas de gerenciamento de TI. A computação em nuvem poupando as tarefas da equipe de TI e desta forma possam dedicar tempo à consecução de objetivos mais importantes.

Atuação

Os maiores serviços de computação em nuvem são executados em uma rede mundial de datacenters seguros, que são atualizados regularmente para a última geração de hardware de computação rápido e eficiente. Isso oferece vários benefícios em um único datacenter corporativo, incluindo latência de rede reduzida para aplicativos e maiores economias de escala.

Confiabilidade

A computação em nuvem torna o backup de dados, a recuperação de desastres e a continuidade dos negócios mais fácil e mais barato, porque os dados podem ser espelhados em vários sites redundantes na rede do provedor de nuvem.

Segurança

Muitos provedores de nuvem oferecem um amplo conjunto de políticas, tecnologias e controles que fortalecem sua postura geral de segurança, ajudando a proteger seus dados, aplicativos e infraestrutura contra ameaças em potencial.

Tipos de serviços em nuvem

Existem três categorias principais que são Infraestrutura como serviço (IaaS), Software como serviço (SaaS) e Plataforma como serviço (PaaS).

Infraestrutura como serviço (IaaS)

Trata-se da categoria mais flexível de serviços em nuvem. Seu objetivo é fornecer controle total sobre o hardware que executa seu aplicativo (servidores de infraestrutura de TI e máquinas virtuais (VMs), armazenamento, redes e sistemas operacionais). Nesta modalidade sob demanda, o hardware necessário pode ser alugado por ser uma infraestrutura de computação instantânea, provisionada e gerenciada pela Internet.

No entanto, a garantia de funcionamento do serviço é uma responsabilidade compartilhada onde, o provedor da nuvem garante a infraestrutura atualizada e disponível e os clientes com a configuração correta. Isso é chamado de modelo de responsabilidade compartilhada.

A opção IaaS é comumente usado nos seguintes cenários:

- **Migrando cargas de trabalho:** Normalmente, os recursos de IaaS são gerenciados de maneira semelhante à infraestrutura local e fornecem um caminho de migração fácil para mover aplicativos existentes para a nuvem.
- **Teste e desenvolvimento:** As equipes podem configurar e desmontar rapidamente os ambientes de teste e desenvolvimento, trazendo novos aplicativos ao mercado mais rapidamente. O IaaS torna os ambientes de desenvolvimento e teste de escala rápidos e econômicos.
- **Armazenamento, backup e recuperação:** As organizações evitam o desembolso de capital e a complexidade do gerenciamento de armazenamento, que normalmente exige pessoal qualificado para gerenciar dados e atender aos requisitos legais e de conformidade. O IaaS é

útil para gerenciar a demanda imprevisível e as crescentes necessidades de armazenamento. O IaaS também pode simplificar o planejamento e o gerenciamento de sistemas de backup e recuperação.

ÍCONE PAAS

PLATAFORMA COMO SERVIÇO (PAAS)

O PaaS é um ambiente completo de desenvolvimento e implantação na nuvem, com recursos desde aplicativos simples até os empresariais sofisticados habilitados para nuvem. Os recursos são adquiridos de um provedor de serviços em nuvem com base no pagamento conforme o uso e acessados por uma conexão segura à Internet.

O objetivo básico do PaaS é no apoio para criar um aplicativo rapidamente, sem gerenciar a infraestrutura subjacente. Por exemplo, ao implantar um aplicativo Web usando PaaS, você não precisa instalar um sistema operacional, servidor Web ou mesmo atualizações do sistema.

O PaaS é comumente usado nos seguintes cenários:

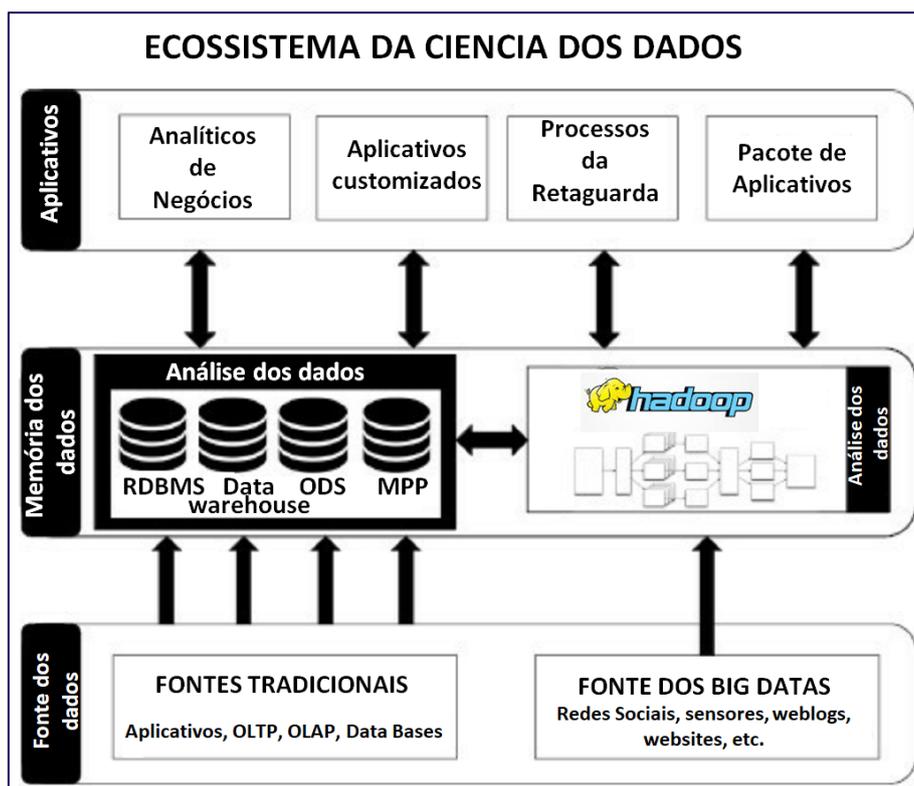
- **Quadro de desenvolvimento:** Fornece uma estrutura para desenvolver ou personalizar aplicativos baseados em nuvem. Assim como a macro do Microsoft Excel, o PaaS permite que os desenvolvedores criem aplicativos usando componentes de software internos. Recursos de nuvem como escalabilidade, alta disponibilidade e capacidade de vários locatários estão incluídos, reduzindo a quantidade de codificação que os desenvolvedores devem executar.
- **Análise ou inteligência de negócios:** As ferramentas fornecidas como um serviço com PaaS permitem que as organizações analisem e explorem seus dados. Eles podem encontrar idéias e padrões e prever resultados para melhorar as decisões de negócios, como previsão, design de produto e retorno do investimento.

Software como serviço (SaaS)

O SaaS é um software que é hospedado e gerenciado centralmente para o cliente final. Geralmente, ele é baseado numa arquitetura em que uma versão do aplicativo é usada para todos os clientes e licenciada por meio de uma assinatura mensal ou anual. O Office 365, Skype e Dynamics CRM Online são exemplos perfeitos de software SaaS.

Ecossistema da Ciência dos Dados

Outros requisitos para gerenciar e exploração de dados incluem desenho de experimentos e mecanismos. Aprendizagem inadequada pode resultar em informações enganosas ou prejudiciais resultados, como em um classificador que funciona para dados equilibrados, mas poderia por engano classificar casos tendenciosos e esparsos na detecção de anomalias.



Fonte: TERI_NORCE CRS,2019

Legenda: ***Hadoop** é uma implementação de código aberto, mantida pela Apache, do paradigma de programação Map-Reduce que foi introduzido pelo Google para processar e analisar grandes quantidades de dados. Uma tarefa complexa é dividida em várias tarefas menores. Elas são então executadas em máquinas diferentes e posteriormente combinadas para gerar a solução da tarefa mais complexa. Um exemplo comum de uso do Hadoop é a análise de padrões dos usuários em sites de e-commerce. Isso permite que novos produtos sejam sugeridos ao usuário.

RDBMS ("Relational Database Management System): é um subconjunto do DBMS (Database Management System) projetado especificamente para bancos de dados relacionais que armazena dados em um formato estruturado, usando linhas e colunas.

ODS (operational data store) é um banco de dados projetado para integrar dados de várias fontes para operações adicionais nos dados destinados para relatórios, controles e suporte a decisões operacionais e como fonte de dados para o Enterprise Data Warehouse (EDW).

MPP (massively parallel processing) é o processamento coordenado de um programa por vários processadores que funcionam em diferentes partes do programa com cada processador usando seu próprio sistema operacional e memória. Normalmente, os processadores MPP se comunicam usando alguma interface de mensagens. Em algumas implementações, até 200 ou mais processadores podem funcionar no mesmo aplicativo. Um arranjo de "interconexão" de caminhos de dados permite que as mensagens sejam enviadas entre processadores. Normalmente, a configuração do MPP é mais complicada, exigindo uma reflexão sobre como particionar um banco de dados comum entre processadores e como atribuir trabalho entre os processadores. Um sistema MPP também é conhecido como sistema "pouco acoplado" ou "nada compartilhado".

OLTP (Online Transaction Processing system) é um sistema de processamento de transações online. O foco principal é registrar a atualização, inserção e exclusão atuais durante a transação. As consultas OLTP são mais simples e curtas e, portanto, requerem menos tempo no processamento e também requerem menos espaço.

OLAP (Online Analytical Processing system), o banco de dados OLAP armazena dados históricos que foram inseridos pelo OLTP. Permite ao usuário visualizar diferentes resumos de dados multidimensionais. Usando o OLAP, você pode extrair informações de um grande banco de dados e analisá-las para tomada de decisão.

A complexidade de uma entrega de dados do produto ou "complexidade da entrega" torna-se uma obstrução quando acionável insight é o foco de uma ciência de dados inscrição. Tamanha complexidade requer identificação e avaliação dos resultados que satisfazem os requisitos técnicos significativo e ter altos negócios valor de um objetivo e uma perspectiva subjetiva. Os desafios

relacionados para os cientistas de dados também envolvem projetar a avaliação apropriada, apresentação, visualização, refinamento, e prescrição de aprendizado resultados e resultados para satisfazer as diversas necessidades de negócios, partes interessadas e apoio à decisão. Em geral, os dados entregues para usuários corporativos devem ser fáceis de entender e interpretáveis por não profissionais, revelando ideias que informam diretamente e permite tomada de decisão e possivelmente tendo um efeito transformador nos processos de negócios e resolução de problemas.

CICLO DE VIDA DA CIÊNCIA DOS DADOS

A ciência de dados está evoluindo rapidamente para ser um dos campos mais em evidência da indústria de tecnologia. Com os rápidos avanços no desempenho computacional que agora permitem a análise de conjuntos de dados massivos, podemos descobrir padrões e idéias sobre o comportamento do usuário e as tendências numa extensão sem precedentes.

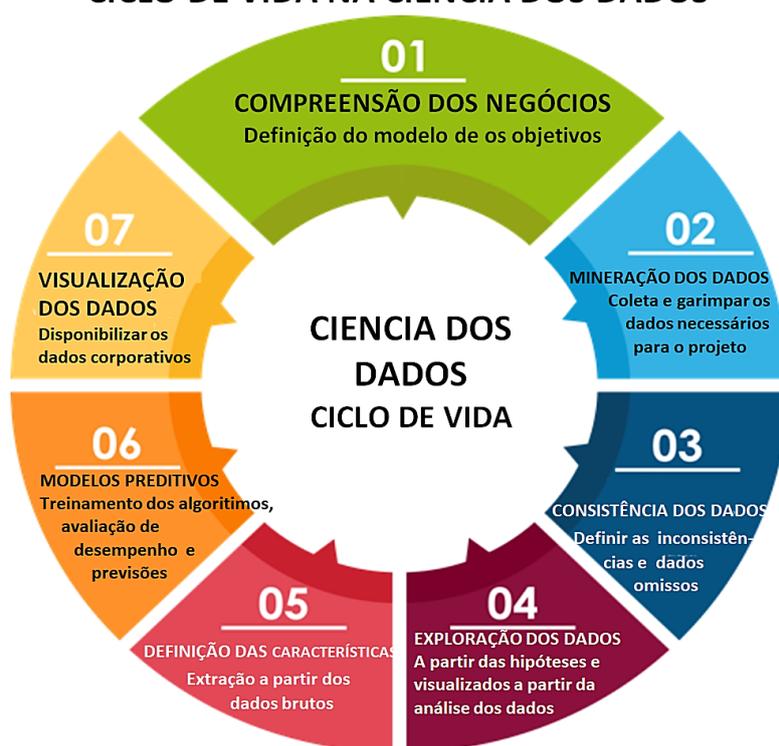
Com o influxo de inúmeras interpretações e no entendimento do campo da ciência de dados requer uma boa compreensão sobre o ciclo de vida no domínio dos negócios e das ações corporativas.

COMPREENSÃO DO NEGÓCIO

A ciência de dados procura garantir que todas as tomadas de decisões corporativas sejam apoiadas por dados concretos e precisos (alta probabilidade) para obter os resultados desejados. Ou seja, no projeto de ciência de dados é essencial o entendimento do problema que deseja solucionar para responder perguntas como:

- Quanto ou quantos? (regressão)
- Qual categoria? (classificação)
- Qual grupo? (agrupamento)
- Isso é estranho? (detecção de anomalia)
- Qual opção deve ser adotada? (recomendação)

CICLO DE VIDA NA CIÊNCIA DOS DADOS



Fonte: Adaptado de Ararwal, 2018

Nesse estágio, procura identificar os objetivos centrais do projeto e as variáveis que precisam ser previstas.

Caso seja uma regressão, pode ser algo como uma previsão de vendas ou programação de estoques. Se for um cluster, pode ser a definição do perfil do cliente. Compreender o poder dos dados e como se pode utilizá-los para obter resultados para os negócios, requer as perguntas certas que é mais uma arte do que uma ciência, e fazer isso bem vem com muita experiência. Para se adquirir essas experiências é consultar outras pessoas que têm o que dizer sobre o tópico, buscar informações em publicações e literaturas especializadas, simulações, participar de congressos e cursos especializados.

MINERAÇÃO DE DADOS

Após a definição dos objetivos do projeto, é a fase de coletar os dados. A mineração de dados é o processo de coleta de dados de diferentes fontes. Algumas perguntas como: quais dados são necessárias para o projeto? Onde estão? Como posso obtê-lo? Qual é a maneira mais eficiente de armazenar e acessar tudo isso?

Na maioria das vezes, encontrar os dados certos exige tempo e esforço. Se os dados residem em bancos de dados, o trabalho é relativamente simples pois ferramentas de consulta dos dados

relevantes como SQL ou manipulá-los usando uma ferramenta de quadro de dados como o Pandas. Caso contrário, Beautiful Soup é uma biblioteca popular usada via páginas da Web em busca de dados. Ou com ferramentas de aplicativo móvel e quiser acompanhar o engajamento e as interações do usuário para começar a obter dados valiosos dos clientes. O Google Analytics, por exemplo, permite definir eventos personalizados no aplicativo que podem ajudá-lo a entender como seus usuários se comportam e coletar os dados correspondentes.

CONSISTÊNCIA DOS DADOS

Após a coleta de todos os dados, passamos à etapa mais demorada de todas: efetuar a limpeza e preparar os dados pois nos projetos de Big Data, que geralmente envolvem terabytes de dados para trabalhar. Esse processo (também conhecido como "zeladoria de dados") geralmente pode levar de 50 a 80% do tempo. Pois existem muitos cenários possíveis para o processo de limpeza. Por exemplo, os dados também podem ter inconsistências na mesma coluna, o que significa que algumas linhas podem ser rotuladas como 0 ou 1 e outras podem ser rotuladas como não ou sim. Os tipos de dados também podem ser inconsistentes - alguns dos 0s podem ser inteiros, enquanto outros podem ser cadeias de caracteres. Se estivermos lidando com um tipo de dados categóricos com várias categorias, algumas delas podem ter erros de ortografia ou ter casos diferentes, como ter categorias para homens e mulheres. Este é apenas um subconjunto de exemplos em que você pode ver inconsistências e é importante capturá-las e corrigi-las nesta fase.

Outro fator importante nesta fase são os dados ausentes que podem gerar muitos erros na criação e no treinamento do modelo. Uma opção é ignorar as instâncias que possuem valores ausentes. Dependendo do seu conjunto de dados, isso pode não ser realista se houver muitos dados ausentes. Outra abordagem comum é usar algo chamado imputação média, que substitui os valores ausentes pela média de todas as outras instâncias. Isso nem sempre é recomendado porque pode reduzir a variabilidade dos seus dados, mas em alguns casos isso faz sentido.

EXPLORAÇÃO DE DADOS

Após a fase de consistência dos dados finalmente começa a etapa de análise. O estágio de exploração de dados é como o *brainstorming* da análise de dados para entender os seus padrões e significado. Pode acessar e analisar um subconjunto aleatório dos dados usando o Pandas, traçar um histograma ou curva de distribuição para ver a tendência geral ou até criar uma visualização interativa que permite mergulhar em cada ponto de dados e explorar a história por trás dos outliers.

Mediante essas informações, pode se formar hipóteses sobre os dados e a natureza do problema em questão. No exemplo de atribuição de notas de conceito dos alunos poderia tentar

visualizar a relação entre pontuações e a presença nas aulas. Na previsão dos preços dos imóveis no litoral, pode-se plotar os preços com o um mapa de calor em um gráfico espacial para ver se é possível detectar alguma tendência.

ENGENHARIA DE RECURSOS

No aprendizado de máquina, um recurso é uma propriedade ou atributo mensurável de um fenômeno que está sendo observado. Se estivéssemos prevendo as pontuações de um aluno, um recurso possível é a quantidade de presença ele possui. Em tarefas de previsão mais complexas, como reconhecimento de caracteres, os recursos podem ser histogramas contando o número de pixels pretos.

De acordo com Andrew Ng: "Criar recursos é difícil, demorado e exige conhecimento especializado". "O aprendizado de máquina aplicado" é basicamente engenharia de recursos ". A engenharia de recursos é o processo de usar o conhecimento do domínio para transformar seus dados brutos em recursos informativos que representam o problema de negócios que você está tentando resolver. Esse estágio influenciará diretamente a precisão do modelo preditivo que você construir no próximo estágio.

Normalmente, realizamos dois tipos de tarefas na engenharia de recursos - seleção e construção de recursos.

A seleção de recursos é o processo de reduzir os recursos que adicionam mais ruído do que informações. Isso geralmente é feito para evitar a maldição da dimensionalidade, que se refere ao aumento da complexidade resultante de espaços de alta dimensão (ou seja, muitos recursos como medidas estatísticas para atribuir pontuação a cada recurso), métodos de wrapper (enquadram a seleção de recursos como um problema de pesquisa e usam uma heurística para executar a pesquisa) ou métodos incorporados (use o aprendizado de máquina para descobrir quais recursos contribuem melhor para a precisão).

A construção de recursos envolve a criação de novos recursos a partir dos que já possui (e possivelmente abandonando os antigos). Um exemplo é a partir de uma variável contínua e para o conhecimento do seu domínio precisa apenas de uma variável de indicador com base num limite conhecido. No caso a partir do recurso da idade em que o modelo decide se a pessoa é adulta ou menor, pode colocar o limite de 18 e atribuir categorias diferentes a instâncias acima e abaixo desse limite. Pode ser também mesclar vários recursos para torná-los mais informativos, obtendo sua soma, diferença ou produto. No caso de se prever as pontuações dos alunos a partir do recurso o número de horas de sono em cada noite, convém criar um recurso que denotasse a média de presença que o aluno teve.

MODELAGEM PREDITIVA

A modelagem preditiva é onde o aprendizado de máquina finalmente entra em seu projeto de ciência de dados. Pois um bom projeto não é aquele que apenas treina um modelo e fica obcecado com a precisão, mas também usa métodos e testes estatísticos abrangentes para garantir que os resultados do modelo realmente façam sentido e sejam significativos.

Com base nas perguntas do estágio de entendimento de negócios, é aqui que se decide qual modelo escolher para o problema. Não é uma decisão fácil e não existe uma resposta certa. O modelo (ou modelos sempre deve testar vários) e dependerá do tamanho, tipo e qualidade dos dados, quanto tempo e recursos computacionais disponíveis ou deseja investir e o tipo de resultados que se pretende atingir. Existem algumas informações disponíveis online, com um fluxograma que ajuda a decidir o algoritmo certo com base no tipo de problema de classificação ou regressão em questão como as folhas de dicas do Microsoft Azure e as folhas de dicas do SAS.

Depois de treinar modelo é fundamental que se avalie o seu sucesso. Um processo chamado validação cruzada k-fold é comumente usado para medir a precisão de um modelo.

Envolve separar o conjunto de dados em k grupos de instâncias de tamanhos iguais, treinando em todos os grupos, exceto um, e repetindo o processo com diferentes grupos deixados de fora. Isso permite que o modelo seja treinado em todos os dados em vez de usar uma divisão típica de teste de trem.

Para modelos de classificação, geralmente testa se a precisão usando o PCC (classificação percentual correta), juntamente com uma matriz de confusão que divide os erros em falsos positivos e falsos negativos. Gráficos como curvas ROC, que são a verdadeira taxa positiva plotada contra a taxa de falsos positivos, também são usados para comparar o sucesso de um modelo. Para um modelo de regressão, as métricas comuns incluem o coeficiente de determinação (que fornece informações sobre a qualidade do ajuste de um modelo), erro quadrado médio (MSE) e erro absoluto médio.

DISPONIBILIZAÇÃO E VISUALIZAÇÃO DE DADOS

A visualização de dados é um campo relativamente complicado, principalmente porque parece simples, mas poderia ser uma das coisas mais difíceis de se fazer bem. Isso porque o conjunto de dados relaciona os campos da comunicação, psicologia, estatística e arte, com o objetivo final de comunicar os dados de uma maneira simples, porém eficaz e visualmente agradável. Depois de extrair as ideias pretendidas do seu modelo, você deve representá-las de maneira que os diferentes principais interessados no projeto possam entender.

Neste campo da visualização de dados existe o pipeline de análise e visualização em um bloco de anotações interativo em Python como o Jupyter, que possibilita obter o código e visualizações

comparativas permitindo uma rápida interação com bibliotecas como Seaborn e Bokeh. Ferramentas como o Tableau e o Plotly tornam muito fácil arrastar e soltar seus dados em uma visualização e manipulá-los para obter visualizações mais complexas. Se você está criando uma visualização interativa para a Web, não há melhor ponto de partida do que o D3.js.

Podemos observar de que o ciclo é portanto, um processo iterativo. Pois o sucesso do modelo está relacionado com o correto entendimento do negócio existente. Ele aborda os problemas identificados? A análise produz alguma solução tangível? Se foram encontrados novos *insights* na interação do ciclo de vida, agora pode se inserir esse conhecimento na próxima iteração para gerar insights ainda mais poderosos e liberar o poder dos dados para obter resultados ainda mais importantes para o negócio ou o projeto como no quadro a seguir:



Fonte: TERI-NORCE CRS, 2019

FLUXOS DE TRABALHO EM CIÊNCIA DE DADOS

A ciência de dados é um campo complexo que requer experiência, habilidade, paciência e tomada sistemática de decisões para ser bem-sucedido. Para se prosperar e agregar valor às pessoas e aos negócios é essencial o desenvolvimento do fluxo de trabalho em ciência de dados como seguinte roteiro:

1. Demarque cada projeto em fases

O projeto de ciência de dados de cima para baixo (top down) fará com que você se sinta sobrecarregado. Se nada mais, comprometerá sua capacidade de dar passos tangíveis. A melhor estratégia é demarcar cada projeto de ciência de dados em quatro fases distintas:

- **Fase 1: Análise Preliminar:** Esta é a etapa de preparação em que os dados são coletados, as metas são definidas e os objetivos são esclarecidos. Essa fase é a mais importante para que o restante do fluxo de trabalho seja eficiente e produtivo.
- **Fase 2: Dados Exploratórios:** Durante esta fase, os dados são limpos, analisados e avaliados. Esta etapa em que perguntas específicas são feitas e o mal entendido é esclarecido.
- **Fase 3: Visualização de Dados:** Com os dados analisados e armazenados em planilhas, é hora de visualizar os dados para que possam ser apresentados de maneira eficaz e persuasiva.
- **Fase 4: Descoberta de conhecimento:** Finalmente, modelos são desenvolvidos para explicar os dados. Os algoritmos também podem ser testados para obter resultados e possibilidades ideais.

2. Use a combinação adequada de hardware e software

Quando se trata de fluxos de trabalho de ciência de dados, os fatores como a velocidade e a eficiência são de extrema importância para garantir o sucesso do projeto inteiro. Requer desta forma obter o equilíbrio e alinhamento entre os recursos do computador e o software para garantir o processo em harmonia para minimizar os atrasos e assegurar o menos risco de corrupção de dados.

3. Assegurar a transparência do fluxo de trabalho

Independentemente do porte do projeto, se é pequeno e isolado ou envolvido numa tarefa muito maior com uma variedade de pessoas e grupos, é necessário garantir que seu fluxo de trabalho seja claro, óbvio e aparente para todos.

4. Envolver o número adequado de pessoas

Manter a equipe inicial pequena para limitar o ruído externo e garantir a continuidade e imune a opiniões excessivas e estratégias diversas e manter o foco.

5. Selecione os KPIs (Key Performance Indicators) apropriados

As métricas e indicadores devem apontar para a estratégia e os objetivos gerais do negócio no longo prazo, além de criar itens acionáveis no curto prazo.

Com a geração de Big Data, o KPIs como "o número de usuários que comprem um determinado produto por mês", "número de usuários ativos que abrem o aplicativo em cinco segundos após a ação X".

Nenhum KPI é perfeito, mas é essencial ter alinhamento com os objetivos do negócio.

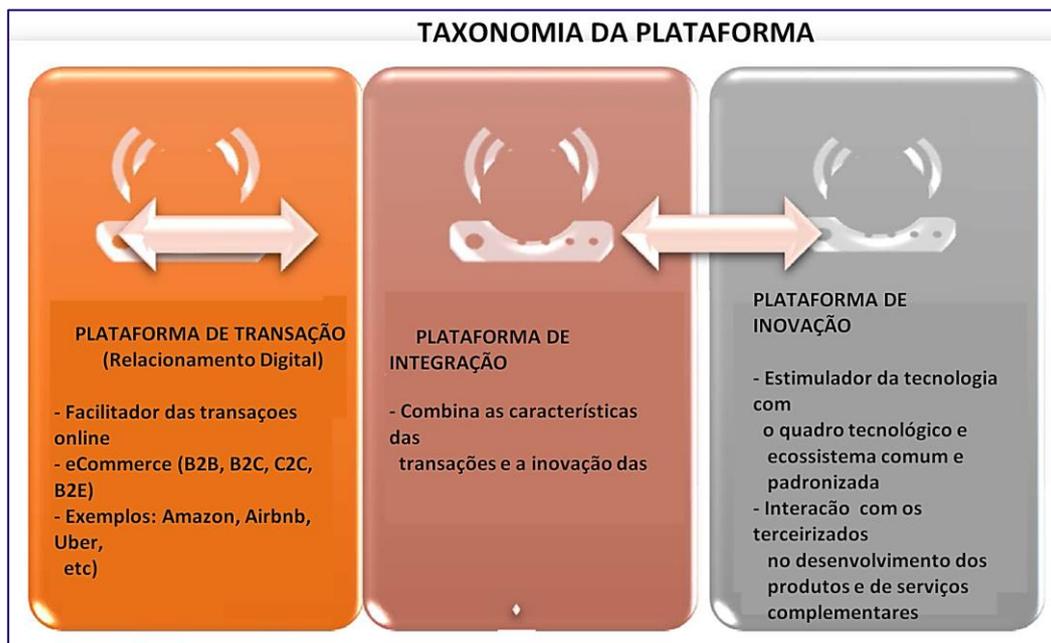
As metas e objetivos no front-end articuladas com o back-end para garantir que os resultados sejam analisados de maneira objetiva.

Quanto mais cedo estabelecer os KPIs e começar a analisar seus resultados, mais eficaz será o seu fluxo de trabalho.

Pois, há casos em que as medidas de sucesso devem ser extremamente técnicas, mas para a maioria das situações, métricas mais simples são mais eficazes e deve evoluir com o tempo.

PLATAFORMAS

As plataformas são modelos de negócio plug-and-play que permite aos múltiplos participantes (produtores e consumidores) a conexão e interação entre si e para criação e troca de valores.



Fonte: Moro, 2019

A história da plataforma está intimamente relacionada com a transformação digital dos serviços e, mais amplamente, os processos de fabricação. As informações baseadas em regras e tecnologia de comunicação, as tecnologias digitais proporcionam a homogeneização no formato dos dados, na capacidade de edição, reprogramabilidade, distribuição e auto-referência em que as transações são mediadas por players complementares que compartilham um ecossistema de rede.

Desta forma, as atividades de serviço são alteradas quando podem ser formalizadas, os processos codificáveis e computáveis com regras claramente definidas para sua execução.

Os dados são (não necessariamente informações) são convertidos em formato legível por computador pelo processo de digitalização. Onde Kallinikos, Aaltonen e Marton (2013) identificam os artefatos digitais como qualquer elemento do projeto de desenvolvimento de software que inclui documentação, planos de teste, imagens, arquivos de dados e módulos executáveis. E destacam as dimensões da editabilidade, interatividade, abertura, reprogramabilidade e distribuição e os atributos de modularidade e granularidade. Ainda, segundo os autores, os artefatos digitais são editáveis como elementos e podem ser reorganizados excluindo ou adicionando elementos ou modificando funções

de elementos (atualização do conteúdo ou dados). Possuem ainda as propriedades interativas, pois oferecem maneiras alternativas pelas quais os usuários podem ativar as funções dos elementos ou explorar informações (funcionalidade de pesquisa). Além disso, eles observam que os artefatos digitais são abertos e reprogramáveis na medida em que sua estrutura lógica pode ser alterada e modificada por outros objetos digitais (funcionalidade existente é alterada por novos aplicativos). Afirmam ainda os recursos de interoperabilidade e abertura, e em seguida, distribui os artefatos digitais à medida que são montados as funções, componentes, ou itens de informação de diferentes fontes ou instituições. Portanto, os artefatos digitais são modulares como eles são montados de blocos distintos e auto-suficientes que são interdependentes e granulares como os blocos podem ser rastreados até unidades elementares (por exemplo, bits e bytes) (Kallinikos et al., 2013).

Yoo et al. (2010) concluem que isso permite a homogeneização de dados como conteúdo digital multimídia (vídeo, gráfico, texto, áudio ou sinais elétricos) que podem ser consumidos e manipulados usando qualquer dispositivo ou rede digital. No entanto, eles argumentam que isso torna os artefatos digitais autorreferenciais, o que significa que dependem de outros os objetos digitais (tecnologia digital como dispositivos móveis). Assim, artefatos digitais estão constantemente evoluindo e mudando ao longo do tempo (Kallinikos et al., 2013). A literatura sobre a plataforma digital identifica quatro artefatos digitais: (a) infraestruturas digitais, (b) plataformas, (c) complementos digitais e (d) ecossistemas digitais que interagem entre si.

Esses quatro artefatos digitais são definidos a seguir:

INFRAESTRUTURA DIGITAL

São os recursos de TI predominantemente tangíveis (tecnologias de hardware, software e rede e telecomunicações), os aplicativos de negócios ou base sobre a qual o portfólio de TI (recursos técnicos e humanos, sistemas e processos) podem ser construídos tanto dentro como fora das empresas (fornecedores, parceiros ou terceirizados).

As infraestruturas digitais são consideradas como o primeiro artefato digital de fundo invisível por parte dos usuários, pois está embutida em outras estruturas sociais e tecnológicas, transparente de usar, tem escopo além de eventos únicos, pode ser aprendido ao se tornar um membro, é moldado por convenções, pode ser conectado a outras infraestruturas usando padrões, baseia-se uma base instalada e só se torna visível após ocorrência de problemas técnicos.

Hanseth e Lyytinen (2010, p. 4) definem as infraestruturas de informação como um “compartilhamento compartilhado, aberto (e sistema sócio-técnico heterogêneo e em evolução ... constituído por um conjunto de recursos e suas comunidades de usuários, operações e projetos. Eles descrevem esses sistemas (chamados de base instalada) como a especificação existente da

infraestrutura de informação. Elas são compartilhadas porque os componentes podem ser usados por vários usuários para executar ações específicas. Além disso estão abertas na medida em que novos componentes podem ser integrados à base instalada existente desde que sejam compatíveis entre si e ajudam várias partes interessadas a orquestrar suas necessidades. Como novos elementos técnicos e sociais são adicionados e integrados, a diversidade técnica e social aumenta, resultando em heterogeneidade que causa evolução ilimitada (Hanseth & Lyytinen, 2010).

As infraestruturas digitais diferenciam outras formas de infraestrutura por sua capacidade de coletar, armazenar e distribuir dados digitais em vários sistemas e dispositivos (Constantinides, Henfridsson & Parker, 2018).

Além dessa conceituação, Henfridsson e Bygstad (2013) identificam quatro fluxos de pesquisa entre infraestruturas digitais:

- a) **Modelos de ativos estratégicos:** são infra-estruturas dentro de empresas onde os gerentes implementam mudanças no portfólio de TI para aprimorar o alinhamento entre TI e estratégia da empresa;
- b) **Perspectiva relacional:** considera as infra-estruturas como relações sociotécnicas em uma comunidade que é mediada pela TI;
- c) **Visualização dos modelos de rede:** A infraestrutura é vista como uma rede de atores humanos que compartilham seus interesses em tecnologia; e
- d) **Estudos orientados:** Os atores humanos ou organizacionais que usam a TI para se adaptar se com o meio ambiente.



Fonte: Gartner, 2020

Os componentes da infraestrutura digital podem ser vistos como “amplos” que incluem os recursos da Internet, redes de intercâmbio eletrônico de dados ou padrões abertos como o Protocolo de controle de transmissão (TCP / IP) que são mais frequentemente associados com a necessidade dos profissionais e departamentos de TI. Ou “limitados” que são componentes mais baratos, implantados pelos usuários finais que incluem os dispositivos smartphones ou tablets), aplicativos ou plataformas existentes (Facebook)

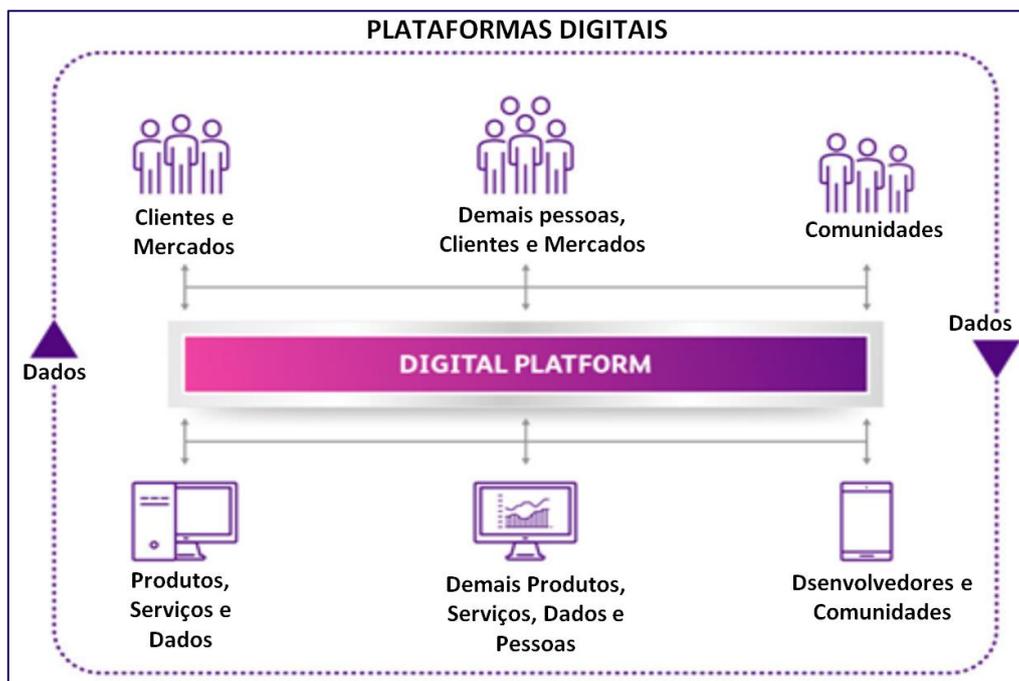
Constantinides et al., (2018) destacam que os componentes de ambas as dimensões estão se tornando cada vez mais complementares à medida que os smartphones e a Internet os recursos de geolocalização e permitem modelos de negócios como o Uber ou AirBnb.

Ao resumir esta seção, as infraestruturas digitais podem ser vistas como compartilhadas, sistemas sócio-técnicos abertos e heterogêneos e em evolução (ou seja, a base instalada com componentes técnicos e humanos, juntamente com suas relações) que formam a base para coletar, armazenar, e distribuir dados digitais em vários sistemas e dispositivos.

PLATAFORMA DIGITAL

Como você definiria uma plataforma digital?

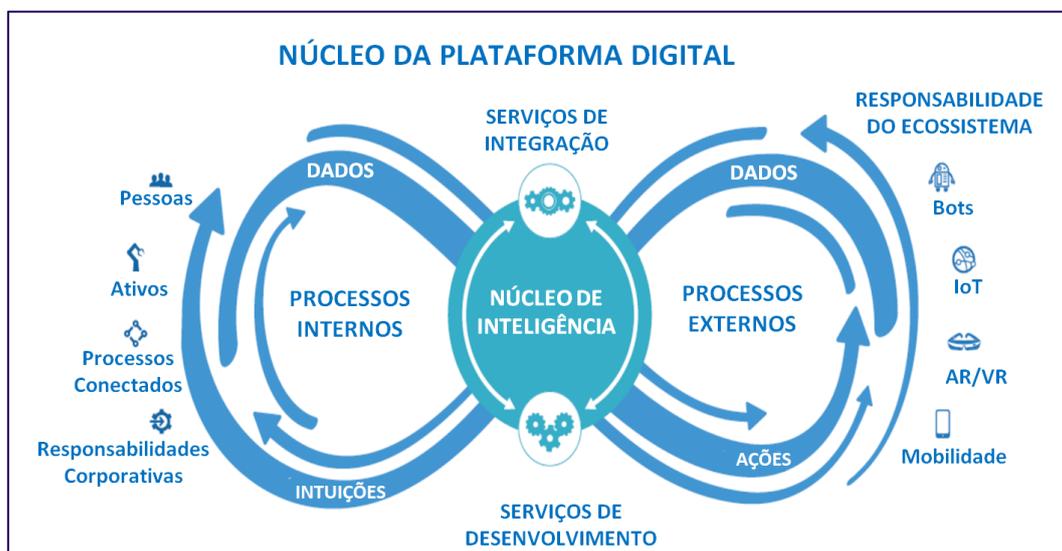
Existem muitas definições na Web, complexas ou simplistas e vamos definir uma plataforma digital com seus principais atributos.



Fonte: Ward, 2020

1. É um modelo de negócios habilitado pela tecnologia.
2. Facilita trocas entre vários grupos - por exemplo, usuários finais e produtores - que não necessariamente se conhecem.
3. Oferece um valor proporcional ao tamanho da comunidade. Existem efeitos de rede. Uma plataforma digital não vale nada sem a sua comunidade.
4. É um facilitador de confiança: deve gerar confiança com termos e condições gerais claros sobre a propriedade intelectual e a propriedade dos dados. Também ajuda consumidores e fornecedores a confiarem uns nos outros dentro da rede, graças aos mecanismos de pontuação.
5. Possui conectividade aberta: compartilha dados com desenvolvedores de terceiros para criar novos serviços e ampliar o ecossistema. Isso é feito com APIs e participa da economia da API.
6. Ele pode ser dimensionado massivamente para atender a milhões de consumidores sem degradar o desempenho.
7. Ele oferece uma experiência atraente para o usuário: fácil de usar, sem necessidade de treinamento, de autoatendimento.
8. Possui modelos de negócios inovadores baseados no valor imediato.

Quais são os benefícios de uma plataforma digital?



Fonte: Adaptado de IDC, 2020



Fonte: Gartner, 2020

- Gerar receita
- Reduzir custos
- Promover a colaboração e a inovação para novos produtos e serviços
- Ganhe velocidade para colocar produtos nos mercados-alvo.

As Plataformas em geral fornecem várias categorizações sob vários pontos de vista.

Sob o ponto de vista econômico, Evans e Schmalensee (2007) distinguem quatro tipos de plataformas:

- Plataformas de trocas:** mediam entre compradores e vendedores para auxiliam na procura de assuntos correspondentes (eBay, Mercado livre);
- Mídia suportada por publicidades:** cria conteúdo ou pela compra de conteúdo de outras pessoas para atrair usuários, que atraem anunciantes (Twitter ou Facebook);
- Sistemas de transações:** facilitam as transações, fornecendo métodos de pagamento (Bitcoin ou PayPal); e
- plataformas de software:** nas quais os desenvolvedores de aplicativos podem acessar serviços para implementar e vender seus aplicativos na plataforma.
- Plataformas sociais:** Facebook, Twitter, Instagram, LinkedIn - modelo de negócios de “publicidade”.
- Plataformas internas:** aumenta a flexibilidade para projetar novos produtos dentro das empresas;

g) Lojas de aplicativos: Apple / Google Play - modelo de negócios “digital good”.

Já Gawer (2009, 2014) classificam plataformas como:

- a) plataformas da cadeia de suprimentos:** para aumentar a eficiência ao longo da cadeia de suprimentos;
- b) plataformas da indústria:** estabelecer relações de negócios com as empresas incluindo inovação complementar externa;
- c) plataformas de mercados multi-facetados:** mediação das transações entre dois ou mais lados do mercado;
- d) Plataformas StackOverflow:** modelo de negócios de "publicidade" como Motor de busca do Google.
 - **Mercados:** Amazon marketplace, mecanismos de compras: modelo de negócios de "comércio eletrônico".
 - **Plataformas de mídia:** Spotify, Deezer - modelo de negócios de “assinatura”.
 - **Plataformas de afiliados:** Commision Junction... - modelo de negócios “ganhe conforme você executa”
 - **Plataformas de crowdsourcing:** Uber, BlablaCar, AirBnB - modelo de negócios "pague conforme o uso"
 - **Plataformas de repositório:** GitHub - modelo de negócios "freemium": você paga apenas se não quiser compartilhar seu código com a comunidade.
 - **Plataformas de infraestrutura (IaaS):** AWS, Azure - modelo de negócios "pague conforme o uso"
 - **Plataformas de classificados:** imóveis... - modelo de negócios de “publicidade”

Assim, as plataformas internas e plataformas da cadeia de suprimentos podem ser vistas como plataformas internas ao passo que as plataformas industriais e os mercados multifacetados podem ser vistos como plataformas externas.

Do ponto de vista econômico, as plataformas digitais são predominantemente vistas como voltados aos mercados que envolvem grupos de agentes que se interagem via "plataformas", onde o benefício de um grupo de ingressar numa plataforma depende do tamanho do outro grupo que se juntam à determinada plataforma. Assim, aa perspectiva econômica em plataformas limita sua visão a um relacionamento comprador-vendedor que trata os usuários finais ou desenvolvedores

simplesmente como consumidores e não fornece uma resposta sobre como as plataformas devem ser projetadas (Gawer, 2014).

A literatura de design de engenharia começou com a visualização de plataformas à luz das "hierarquias de design" levando a plataformas de produtos e finalmente integrando-se às famílias de produtos (Gawer, 2014). Cedo definições incluem a “reutilização sistemática de componentes em diferentes produtos dentro de uma família de produtos, que permite economias de escopo na produção” (Gawer, 2014, p. 1242).

Os primeiros pesquisadores veem a plataforma como um conjunto de componentes ou funções padrão que formam uma tecnologia de uso geral com arquitetura e padrões integrados como base para terceiros desenvolvedores a desenvolver ativos complementares (Bresnahan & Greenstein, 1999; Fichman, 2004; Taudes, Feurstein e Mild, 2000; West, 2003). Assim, uma plataforma não está mais totalmente controle do proprietário da plataforma, mas renuncia ao controle de desenvolvedores de terceiros (Gawer & Cusumano, 2008). O que essas conceituações iniciais de plataforma têm em comum é o foco em arquitetura e, mais especificamente, na arquitetura tecnológica modular (Gawer, 2014). Na linha com isso, Baldwin e Woodard (2009, p. 19) definem plataformas como “um conjunto de componentes estáveis que suporta variedade e capacidade de evolução em um sistema, restringindo os vínculos entre os outros componentes”. Eles distinguem explicitamente entre componentes “principais” com baixa variedade que representam a plataforma e os componentes “periféricos” complementares com alta variedade. Desenhando nessa definição, Tiwana, Konsynski e Bush (2010, p. 675) especificam com mais precisão: “O base de código extensível de um sistema baseado em software que fornece a funcionalidade principal compartilhada pelos módulos que interoperam com ele e as interfaces pelas quais eles interoperam”. Boudreau (2010) afirma ainda que especialmente as normas técnicas apoiam a interoperabilidade entre os módulos. Para resumir a perspectiva da engenharia, as plataformas fornecem uma solução modular e integrada a arquitetura com um conjunto de componentes estáveis e funcionalidade principal (ou seja, o núcleo da plataforma) compartilhada pelos módulos / complementos que interoperam com ele usando interfaces ou regras padronizadas.

No entanto, a perspectiva da engenharia não leva em conta como as plataformas evoluem ao longo do tempo e como eles competem entre si (Gawer, 2014). A perspectiva organizacional também conceitua plataformas como consistindo em uma arquitetura com recursos tangíveis e intangíveis. No entanto, a interação de criação de valor entre atores (externos e internos) e recursos são mais pronunciados nessa perspectiva (Constantinides et al., 2018; Lusch & Nambisan, 2015). Ciborra (1996) enfatiza que isso é ativada por um núcleo de plataforma estável com estruturas organizacionais relativamente estáveis e temporárias acordos com estruturas, produtos, rotinas ou rotinas organizacionais em constante mudança de capacidades. Essas organizações ou meta-organizações em

evolução podem ser caracterizadas por três aspectos: (a) coordenam os atores (que podem ser indivíduos ou empresas e podem levar diferentes papéis), que podem competir e inovar (que é frequentemente chamado de coopetição [Brandenburger e Nalebuff, 1996]); (b) fazer uso de economias de escopo tanto na oferta quanto na exigência; e (c) consistir em uma arquitetura tecnológica modular com um núcleo e periféricos componentes (Gawer, 2014). Esses elementos técnicos, juntamente com os processos organizacionais e padrões (de Reuver et al., 2018) podem ser realizados e implementados por recursos dinâmicos (Thomas et al., 2014). Para resumir essa visão, as plataformas são vistas como organizações ambídestras que alavancam recursos, fornecendo uma arquitetura tecnológica modular com um nível relativamente estável núcleo da plataforma e uma periferia em mudança dinâmica. Portanto, é necessário para a plataforma organização para coordenar os atores da organização para permitir a cooperação e concorrência por inovação e economias de escopo.

O que significa que não há um único proprietário que possua o núcleo da plataforma capaz de ditar as hierarquias de *design* (Henfridsson et al., 2014). As plataformas são facilitadoras da troca (de bens, serviços e informações) entre diferentes tipos de partes interessadas que não poderiam interagir de outra forma.

Convencionalmente usado, o termo “plataformas” refere-se a múltiplas estruturas digitais que moldam e intermediam as regras para interação dos participantes.

No campo da ciência da computação, as plataformas forneçam um conjunto de técnicas compartilhadas, tecnologias e interfaces para um amplo conjunto de usuários para construção de artefatos e soluções personalizadas.

Assim, as plataformas são “cyberplaces” habilitados por algoritmos onde os protagonistas podem agir, interagir e negociar.

A maioria das definições observa o papel em que uma entidade desempenha atividades intermediárias de outras entidades ou de pessoas. Como “uma empresa que controla uma rede, uma instalação ou um insumo essencial onde os fornecedores de um bem ou prestadores de serviços complementares” devem “confiar”. Ou enfoca o papel da infraestrutura em que as empresas estruturam o acesso aos mercados ou facilitar as transações. Desta forma, são recursos facilitadores da troca (de bens, serviços e informações) entre diferentes tipos de partes interessadas que não poderiam interagir de outra forma. As transações são mediadas por players complementares que compartilham um ecossistema de rede (Rochet e Tirole, 2003; Armstrong, 2006). Devido às suas características têm um alcance que lhes dá o potencial de escalar globalmente.

Dentro dessa perspectiva, uma plataforma pode ser categorizada em termos do escopo do processo de produção como:

1. plataformas internas, permitindo a recombinação de subunidades dentro da empresa;
2. plataformas da cadeia de suprimentos coordenando fornecedores externos ao redor de uma montadora; e
3. plataformas da indústria em que um líder de plataforma agrupa recursos externos de complementadores (Gawer, 2014).

Nos dois últimos tipos, as plataformas não apenas fornecem um núcleo estável, mas também mediam entre diferentes grupos de usuários como compradores e vendedores, normalmente é indicado como uma plataforma de vários lados (Boudreau e Hagiu, 2009).

Como as plataformas reúnem vários grupos de usuários, cria os chamados efeitos de rede ou externalidades de rede que implicam no aumento da utilidade da tecnologia à medida que aumenta a base instalada de usuários (Katz e Shapiro, 1985; Shapiro e Varian, 1998). Pois, o valor da plataforma depende do número de usuários no mesmo grupo de usuários conectando ou usando os serviços fornecidos.

Observa-se que a força dos efeitos da rede pode variar drasticamente e podem moldar tanto a criação quanto a captura de valor. Quando os efeitos de rede são fortes, o valor fornecido pela plataforma continua a aumentar acentuadamente conforme o número de participantes.

Por exemplo, como o número de usuários numa determinada plataforma aumentam, assim como a quantidade e variedade de conteúdo interessante e relevante também aumentam.

Em muitas situações, a melhor estratégia de crescimento da plataforma é se conectar diferentes redes de plataformas umas às outras. Pois qualquer negócio de plataforma, o sucesso depende da aquisição/manutenção do grande número de usuários e acumulação de dados sobre suas interações. Esses ativos são sempre muito valiosos em vários cenários e mercados. Ao alavancá-los, as empresas conseguem diversificar-se em diferentes linhas de negócios e melhorar sua economia de escala podem criar sinergias importantes.

PLATAFORMAS DIGITAIS

As principais empresas de plataformas digitais combinam com governança eficaz para obter efeitos positivos na rede, criação de valor e escalabilidade nas suas atividades.

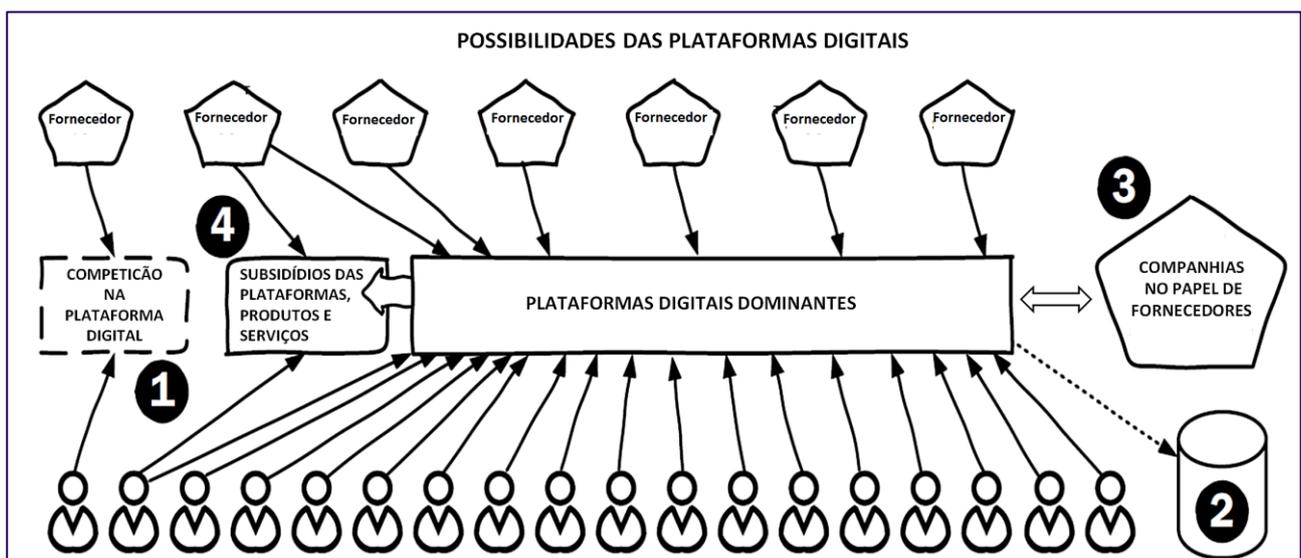
Digitalização é o processo de conversão de dados (não necessariamente informações) num código cujo formato é legível pelos computadores. E as tecnologias digitais implicam homogeneização de dados, capacidade de edição, reprogramabilidade, distribuição e auto-referência (Yoo et al., 2010; Kallinikos et al., 2013). Tais recursos podem levar a várias heranças em distribuídos

configurações, o que significa que não há um único proprietário que possua o núcleo da plataforma e dite sua hierarquia de design (Henfridsson et al., 2014).

Essas ações são altamente diversas e são categorizadas por mercado, função social ou de caráter técnico. Cada plataforma envolve sua própria diversidade seja questões computacionais e/ou questões de mercado. Isto é, há plataformas de B2C (business to consumer), as plataformas de compartilhamento que geralmente são C2C (consumer to consumer) ou B2B, e as plataformas para Industrie 4.0 com recursos do IoT (Internet das coisas) em que os objetos são ligados por conexões cibernéticas em que as regras e propriedade dos dados são gerados por máquinas.

A figura a seguir destaca os tipos de plataformas e suas características. Em primeiro identifica a concorrência por haver um líder de mercado cujo domínio pode sufocar a inovação. Em segundo, a plataforma pode coletar muitos dados confidenciais, dados pessoais com milhões de perfis de pessoas que não é tão claro onde estão os limites éticos e legais. A terceira questão é um possível viés nos rankings criados pela plataforma. Isso pode variar de filtrar certas opiniões para classificar certas questões de maneira injusta. Nesses casos, o algoritmo é manipulado para fornecer um resultado diferente, dando vantagem a terceiros (possivelmente sendo a própria plataforma).

O quarto problema é que grandes plataformas digitais tendem a integrar o mercado horizontalmente, às vezes até verticalmente, usando sua base de usuários como um ativo. Além disso, a receita gerada por uma plataforma dominante pode ser usada para subsidiar a expansão em outros setores.



Fonte: Aalst et. Al, 2019

Legenda: (1) nenhuma concorrência real é possível devido à necessidade de escala; (2) dados confidenciais estão sendo coletados e usados para obter vantagens competitivas ou receita adicional; (3) a plataforma pode ser tendenciosa (porque o papel do fornecedor e o papel da plataforma) são combinados) e (4) a receita e a base de usuários gerados pela plataforma dominante são usados para expandir horizontal ou verticalmente.

As plataformas digitais diferem dos sistemas físicos como produtos, porque os limites do subsistema são vagamente definidos, o que torna menos onerosos a recombinação de elementos e a informação não é conflitante e compartilham uma visão colaborativa. E são baseadas em algoritmos digitais e estruturas de software que processam dados armazenados na nuvem. A história da plataforma está intimamente relacionada com a transformação digital dos serviços e, mais amplamente, os processos de fabricação e também as informações baseadas em regras e tecnologia de comunicação.

PLATAFORMAS DIGITAIS COM TECNOLOGIAS DA INTELIGÊNCIA ARTIFICIAL (IA)

As tecnologias da inteligência artificial (IA) estabelecem relacionamento amigável entre humanos e máquinas. Pois a IA proporciona a capacidade da máquina executar funções cognitivas em que associamos às mentes humanas como percepção, raciocínio, aprendizado, interação com o meio ambiente, resolução de problemas, tomada de decisão e até demonstração de criatividade como propor sugestões, avaliarem a capacidade física ou traçar perfil dos humanos.

Ou seja, os aspectos das tarefas nas plataformas digitais são alinhados com os recursos dos agentes de IA (por exemplo, velocidade, precisão, confiabilidade, escalabilidade) e outras competências de agentes humanos (criatividade, empatia, julgamento) cujo resultado criam oportunidades para tarefas nas plataformas a serem executadas por novos híbridos humano-IA que variam desde a substituição de tarefas (substitutos dos humanos pela IA) ao aumento de tarefas (IA e humanos se complementam) à montagem de tarefas (IA e humanos são dinamicamente reunidos para funcionar como uma unidade integrada).

Temos assim, plataformas com base no software (iOS e Android), redes sociais (Facebook) e plataformas de negociação (Amazon Marketplace) cunhou inicialmente o termo plataforma digital, que hoje temos muitas outras categorias como plataformas de trabalho on-line (Twago, Freelancer.com).



Fonte: Gartner, 2020

Existem várias conceituações de plataformas digitais

São definidas como artefatos puramente técnicos onde plataforma é uma base de código extensível e o ecossistema compreende os módulos de terceiros que complementam essa base de código. Pode também ser caracterizado como uma assembléia sociotécnica abrangendo os elementos técnicos (de software e hardware) e processos e padrões organizacionais associados (Tilson et al., 2012). Ou Ghazawneh e Henfridsson (2015) constroem bases em Tiwana et al. (2010), definindo como plataformas externas baseadas em software que formam a base de código extensível que fornece o núcleo de funcionalidade compartilhada pelos módulos que interoperam com as interfaces com as APIs¹.

São também plataformas externas baseadas em software que consistem na base de código extensível de um sistema baseado em software que fornece a funcionalidade principal compartilhada pelos módulos que interoperam com ele e as interfaces pelas quais eles interoperar” (Tiwana et al., 2010). As plataformas de software representam um ponto de encontro tecnológico em que aplicativos desenvolvedores e usuários finais convergem (Evans et al., 2006). Já Spagnoletti et al. (2015, p. 364) definem uma plataforma digital como “um bloco de construção que fornece uma função essencial a um sistema tecnológico e serve como base sobre a qual produtos complementares, tecnologias ou serviços podem ser desenvolvidos”.

¹ A interface de programa de aplicativo (API) é um conjunto de rotinas, protocolos e ferramentas para criar aplicativos de software. Basicamente, uma API específica como os componentes de software devem interagir com o núcleo da plataforma e cabe ao programador juntar os blocos que facilita o desenvolvimento do programa fornecendo todos os componentes. Além disso, as APIs são usadas ao programar componentes da interface gráfica do usuário (GUI).

A plataforma digital incorpora vários módulos como "subsistemas de software complementares" que permite estender a funcionalidade do produto de software na forma de programas aplicativos projetados e desenvolvidos por terceiros. Os programas aplicativos são porções executáveis de software, serviços ou sistemas disponíveis para usuários finais (Ghazawneh e Henfridsson, 2013, p. 175).

Assim, as plataformas digitais são a base de modelos de negócios habilitados pela tecnologia que facilitam as trocas entre vários grupos - como usuários finais e produtores que não necessariamente se conhecem. O valor gerado é proporcional ao tamanho da comunidade, com efeitos de rede escalonáveis graças à conectividade e a interação digital com a Internet. As plataformas seguem paradigmas inovadores, onde as partes interessadas criam e compartilham valor e incorporam curvas de aprendizado (economias de experiência) que são alimentadas por big data em tempo real.

As tecnologias digitais implicam homogeneização de dados, na capacidade de edição, na reprogramabilidade e distribuição e auto-referenciabilidade (Yoo et al., 2010; Kallinikos et al., 2013). Tais características da digitalidade podem levar a múltiplas configurações, o que significa que não há um proprietário que possua o núcleo da plataforma e que seja capaz de ditar a sua hierarquia de design (Henfridsson et al., 2014). Além disso, ao combinar a modularidade de bens físicos com a arquitetura em camadas do software, tem como resultado o acoplamento por meio das interfaces padronizadas, levando a produtos abertos para novas aplicações e segmentos após a fabricação (Yoo et al., 2010, p. 729).

A seguir a visão geral do conceito:

Definições dos principais conceitos em plataformas digitais

CONCEITO	DEFINIÇÃO
Plataformas Multifacetadas	Mediação de diferentes grupos de usuários, como compradores e vendedores
Plataformas Multimercados	Reúne (e/ou combina) grupos distintos que o valor de um grupo aumenta à medida que o número de participantes do outro grupo aumenta
Redes externas diretas	O valor da plataforma depende do número de usuários no mesmo grupo de usuários
Redes externas indiretas	O valor da plataforma depende do número de usuários externos diferentes dos grupos
Visão técnica das plataformas Digitais	Base de código extensível à qual podem ser adicionados módulos complementares de terceiros
Visão sociotécnica das plataformas Digitais	Elementos técnicos (de software e hardware) e processos organizacionais padrões associados
Ecosistema (visão sociotécnica)	Uma coleção de complementos (aplicativos) para a plataforma técnica principal, fornecida principalmente por terceiros
Ecosistema (visão técnica)	Coleção de empresas interagindo para contribuição complementar.
Aplicações	Software executável oferecido como aplicativos, serviços ou sistemas para usuários finais
Recursos da fronteira	Ferramentas e regulamentos de software que facilitam o relacionamento de longo prazo entre os envolvidos
Grau de Abertura da plataforma	O limite de extensão em que a plataforma suporta os recursos complementares

Fonte: Reuver et al, 2017

Por exemplo, permitindo a separação de forma e função, o teclado do smartphone pode ser definido por software e se adaptar dinamicamente à necessidade específica de entrada para cada tipo de aplicação. Além disso, os usuários podem adaptar a configuração padrão do teclado conforme a língua ou aplicação. Ou seja, os desenvolvedores de aplicativos combinam os recursos modulares existentes em camadas dos sistemas operacionais a vários elementos de hardware, os kits de desenvolvimento de software e uma variedade de interfaces de programação de aplicativos públicos (APIs) com aplicativos novos não considerados quando os smartphones e software básicos foram concebidos.

Ao contrário de plataformas não digitais, plataformas digitais contêm componentes em diferentes níveis, como dispositivos, o sistema operacional e os aplicativos e o processo de inovação são dinâmicos e depende de suas dependências com plataformas em diferentes níveis técnicos da arquitetura.

Tipos de plataformas Digitais

As plataformas digitais redesenham as relações entre clientes e prestadores de serviços e bens, funcionários e empregadores, estado e sociedade, à medida que as TIC (Tecnologia da Informação e Comunicação) penetram em quase todas as esferas de nossa vida: desde a compra de bens casuais, pagamento e o recebimento de serviços públicos até a realização complexos tarefas de profissionais.

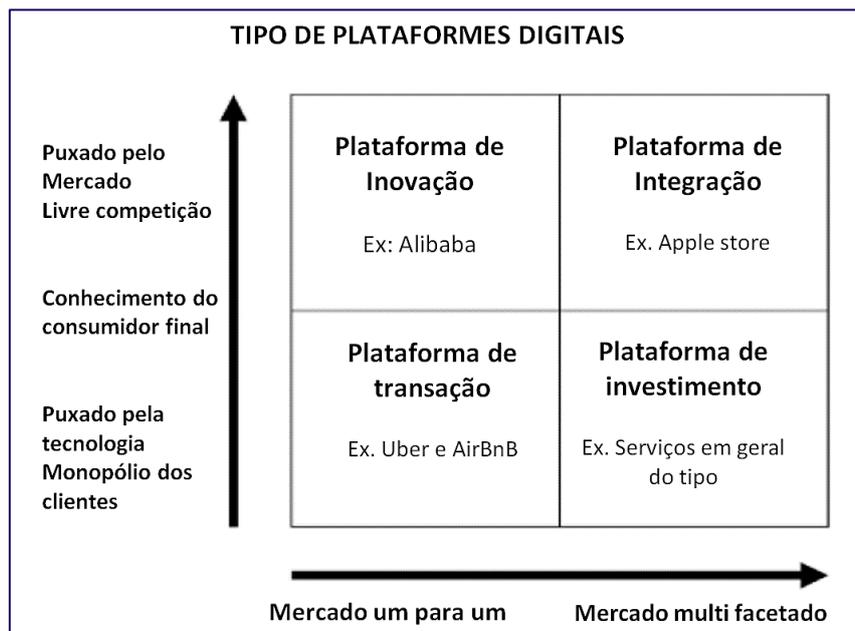
À medida que o poder computacional cresce, e mais pessoas no mundo participam da economia digital, é necessário analisar mais profundamente os recursos, princípios e vantagens das transformações digitais para minimizar os prazos de integração de novas ferramentas de interação. Apesar de as plataformas digitais terem uma essência comum, as empresas as integram e as utilizam no mercado de maneira diferente.

As principais plataformas estão fornecendo serviços de IA, com estas características:

- Estruturas de código aberto e bibliotecas de modelos de aprendizado de máquina para desenvolvedores e pesquisadores (Google TensorFlow, Azure), pacotes de aprendizado de máquina (PyTorch do Facebook, Microsoft Computational Network Tool Kit)
- Interfaces de programa de aplicativo (APIs) para serviços como visão, fala, idioma, conhecimento e pesquisa (nuvem do Google), API de conversão, API de interpretação de idioma natural da nuvem e API de conversão de voz em texto da nuvem)
- Ferramentas de arrastar e soltar para criar, testar e implantar modelos de IA personalizados (por exemplo, Azure Machine Learning Studio), Modelos de IA pré-treinados para os usuários implantarem em seus respectivos contextos

- Plataformas para atividades de governança. (Google Play para revisão e monitoramento de aplicativos de detecção de fraudes, conteúdo inapropriado, aplicativos prejudiciais e violações de propriedade intelectual e políticas de segurança).

Como resultado, Evans e Gawer (2016) apontaram 4 tipos de plataformas digitais, de acordo com suas principais funções:



Fonte: Ablyazov e Rappgof, 2019

- **plataformas de transações:** agindo como um meio, facilita as trocas ou transações entre vários usuários, compradores ou fornecedores;
- **plataformas de inovação:** usadas como base para outras empresas (não da melhor maneira integrada no sistema ecológico da inovação) para desenvolver tecnologias, produtos ou serviços complementares;
- **plataformas integradas:** sendo simultaneamente uma plataforma de transações e uma plataforma de inovação. Esta categoria inclui empresas como a Apple, que possui a plataforma App Store, e um grande sistema ecológico de desenvolvedores com recursos que suporta a criação de um conteúdo nessa plataforma;
- **plataformas de investimento:** constituídas por empresas que desenvolveram a estratégia de portfólio de plataformas e atuam como holding, investidor ativo de plataforma ou ambos.

A principal característica das plataformas digitais como ferramenta de interação socioeconômica é o efeito de rede: aumento do valor do produto ou serviço da empresa, aumento do efeito positivo (ou negativo) no aumento do número de usuários.

O fenômeno em si aparece não apenas no ambiente digital, mas no cotidiano da vida também. Por exemplo, com a ampliação da empresa em um território específico e, conseqüentemente, aumentar. Nos locais de trabalho, o número de habitantes do bairro também aumenta. No entanto, no uso de plataformas digitais, a taxa de propagação do efeito de rede aumenta pela progressão geométrica.

Plataforma Digital de Trabalho

São ambientes on-line onde os serviços digitais são adquiridos e entregues em troca de remuneração, com componentes de tarefas para determinados serviços, executados e coordenados por agentes humanos e de IA. O trabalho realizado nessas plataformas é, por definição digital e, portanto, pode ser modularizado em tarefas que exigem uma gama de habilidades cognitivas para execução e coordenação, fornecendo um rico contexto híbrido de humanos e da IA.

São identificamos um três características de digitalização que estão transformando as plataformas de trabalho digital:

- 1. Modularização de tarefas:** as arquiteturas modulares em camadas de plataformas digitais que permitem a especificação de entradas e saídas de tarefas digitais em níveis granulares, encapsulamento de atividades em módulos para executar tarefas e definição de interfaces padronizadas entre módulos que permite experimentação e flexibilidade na alocação e coordenação de tarefas.
- 2. Geração automatizada de dados:** as entradas, processos e resultados associados às tarefas digitais podem ser rastreados de forma granular juntamente com parâmetros contextuais. Infraestruturas de nuvem e tecnologias de fluxo de trabalho geram dados relacionados a tarefas para alimentar algoritmos de aprendizado de IA.
- 3. Modalidades de aprendizagem de IA:** inovações na ciência e na tecnologia aos métodos de IA (redes neurais convulsivas) acopladas às arquiteturas computacionais permitem a expansão das modalidades de aprendizado de IA (supervisionado ou autônomo com reforço, autodidático (máquina única) e máquina a máquina) que pode ser aplicado a tarefas em plataformas de trabalho digital.

Assim a modularização de tarefas, geração de dados relacionados a tarefas para aprendizado de IA e a expansão das modalidades de aprendizado de IA permite alavancar uma variedade de agentes humanos e de IA na execução de tarefas em plataformas de trabalho digitais.

PLATAFORMAS DIGITAIS HÍBRIDAS	
TEMA	DESCRIÇÃO
Digitalização das plataformas	Modularização de tarefas, rastreamento automatizado de dados granulares relacionados a tarefas e modalidades de aprendizado de IA.
Tipos de plataformas híbridas IA-humano	Os híbridos podem variar conforme o nível de interdependência humano-IA: - Substituição: IA substitui os humanos - Colaborativo: Humanos e IA de complementam - Integração: IA e humanos são reunidos para formar uma unidade integrada. A execução das tarefas com eficiência dependerá dos papéis dos seres humanos e dos sistemas de IA em gerar as entradas, execução, comunicação, implementação e saídas de tarefas (Tomada de decisões e soluções).
Desenho das Plataformas sem viés	Identificação das limitações dos dados ou algoritmos codificados de treinamento no projeto. Mitigação de vieses nos processos dinâmicos de aprendizado envolvendo seres humanos e sistemas de IA. Competências necessárias para novas categorias de tarefas e implicações na interface humano-IA sem viés nos processos de tomada de decisão.
Tutela dos dados	Questões relacionadas à responsabilidade e ética na tomada de decisões, segurança e privacidade de dados na implantação de plataformas.
Domínio da plataforma versus fortalecimento	A democratização das ferramentas de IA por meio de plataformas de código aberto, combinadas com recursos de fronteiras abertas como APIs, permitem freelancers e novos players de plataforma para inovar e competir. A dinâmica dos novos recursos possibilita acionar, catalisar e dimensionar movimentos para capacitar indivíduos e transformar a colaboração para explorar novas necessidades coletivas e heterogêneas.

Com base na interdependência entre os papéis da IA e dos agentes humanos na execução de tarefas formando um espectro de híbridos humano-IA que variam da substituição à potencialização da capacidade humana como segue:

- 1. Substituição de tarefas:** onde plataformas de trabalho digital implantam sistemas de IA para substituir o papel de agentes humanos na execução de tarefas. Ex. Assistentes virtuais robotizados para tarefas de atendimento ao cliente.
- 2. Potencialização das tarefas:** permitem que agentes humanos e de IA complementem um ao outro para executar uma tarefa. Ex. Designers e engenheiros humanos em manufatura, arquitetura e engenharia usam ferramentas de IA para aumentar a exploração das soluções. Ex. Conforme as metas de projeto (juntamente com parâmetros relacionados a materiais, métodos de fabricação e restrições de custo), o software gera rapidamente alternativas de design, explorando um grande número de soluções por meio do processo evolutivo de aprendizagem.

Além disso, os seres humanos também podem aumentar os agentes de IA como:

- Treinar os agentes de IA para identificar melhor as nuances da linguagem e o conhecimento contextualizado, corrigir erros e reforçar o índice de sucessos.
- Explicar decisões tomadas por algoritmos de aprendizado de máquina de caixa preta para indivíduos sem conhecimento técnico, melhorar a transparência deste processo de tomada de decisão e a responsabilidade de suas ações.
- Aplicar julgamento e valores morais para garantir que os agentes de IA tenham um bom desempenho ao longo do tempo em áreas como éticas e conformidade.

3. Montagem de tarefas: onde se reúnem dinamicamente agentes humanos e de IA para executar uma tarefa emergente. Na prática, a IA e os agentes humanos são reunidos de forma contextual e temporal e funcionam como uma unidade integrada para executar as tarefas. Como tal, os agentes de IA tornam-se extensões incorporadas das capacidades cognitivas dos agentes humanos na execução das tarefas digitalmente.

Por exemplo, os cirurgiões humanos e robôs movidos a IA podem funcionar como uma unidade integrada para realizar cirurgias minimamente invasivas. Da mesma forma, empresas de manufatura estão usando “cobots” (robôs colaborativos) sensíveis ao contexto dos “exoesqueletos” que são dispositivos robóticos vestíveis acoplados que se adaptam e integram aos trabalhadores como colaboração dinâmica entre os agentes humanos e de IA que geram inteligência coletiva sensível ao contexto.

Escalabilidade das plataformas Digitais

As redes se tornam mais valiosas e importantes com o aumento do tamanho e interconexões conforme lei de Metcalfe (Odlyzko e Tilly, 2005), segundo a qual o efeito de uma rede de telecomunicações é proporcional ao quadrado do número de usuários conectados do sistema (n^2). A lei d expressa matematicamente o número de conexões únicas possíveis num nó de uma rede.

Assim, temos o seguinte postulado: “Se uma rede é composta por n pessoas e cada uma delas acessa a rede, o valor proporcional ao número de outros participantes, o valor em que todas as n pessoas acessam a rede é: $n * (n-1) = n^2 - n$

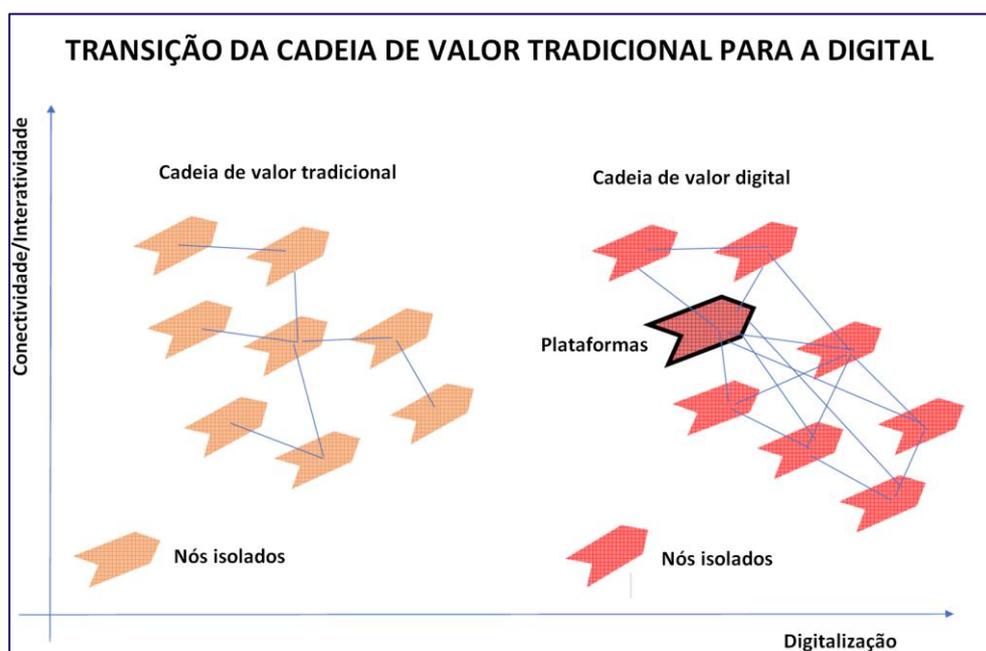
Law expresses mathematically the number of unique possible connections in a network of nodes.

If a network is composed of n people and each of them assigns to the network a value proportional to the number of other participants, then the value that all n people assign to the network is:

As plataformas digitais aumentam a lucratividade escalável com possibilidade de crescimento exponencial do ecossistema. A escalabilidade é um elemento essencial de qualquer negócio. Indica a capacidade de um processo, rede ou sistema de lidar com uma quantidade crescente de trabalho ou seu potencial ampliado para acomodar o processo de crescimento. Observa-se o crescimento da receita acompanhado de aumento proporcional dos custos variáveis. Escalabilidade é a capacidade de um dispositivo se adaptar a um ambiente em mudança conforme a mudança das necessidades do cliente. Em termos mais amplos, escalabilidade significa flexibilidade (incorporando opções reais para expandir, adiar, abandonar os negócios), o que permite atender melhor e atender às necessidades específicas dos clientes com uma certa dose de personalização. Os interesses e gostos das pessoas, bem como as condições ambientais, evoluem e se ajustam continuamente. A escalabilidade é portanto, vital, pois contribui para a competitividade, eficiência e qualidade (Moro Visconti, 2020). A escalabilidade também aponta e ajuda a resposta sem degradação evitando o desperdício fazendo uso eficiente dos recursos disponíveis (Gupta et. al., 2017).

A teoria de redes está principalmente relacionada a plataformas digitais, que por sua vez representam um catalisador de intangíveis escaláveis.

As interações entre a empresa em rede, a plataforma digital e outras partes interessadas (Stakeholders) externas podem ser examinadas com uma análise da cadeia de valor que descreve seus recursos digitais e em rede. A cadeia de valor é digitalizada pelos dispositivos/tecnologias a ela conectadas conforme a figura a seguir.



Fonte: Moro, 2019

As cadeias de valor digitais tendem a ser mais planas do que as cadeias de valor tradicionais e a plataforma de ponte atua como uma coordenação cubo. A rede digital é intrinsecamente mais valiosa devido à sua arquitetura altamente interconectada (maior número de links); O valor também depende do tráfego crescente de dados ou transações entre os nós vinculados.

O fornecimento digital e as cadeias de valor podem ser representados por dois ecossistemas de rede inicialmente separados que eventualmente interagir, dentro de uma rede multicamada (Bianconi, 2018). Essa interpretação é consistente com a fabricação em nuvem paradigma, uma forma avançada de manufatura em rede. Esse processo é baseado em uma combinação de sistemas de manufatura e tecnologias emergentes, como computação em nuvem, manufatura virtual, agilidade manufatura, grade de manufatura, IoT e tecnologias orientadas a serviços (Akbaripour et al., 2015). Fornecimento global cadeias (e cadeias de valor relacionadas) estão se tornando cada vez mais conectadas devido ao aumento da globalização em termos de tamanho da rede, força e conectividade, mostrando mudanças intertemporais significativas e maior cluster (Tsekeris, 2017).

REDES MULTICAMADAS

Muitos sistemas complexos do mundo real são, portanto, melhor modelados por redes multiplex (multidimensionais) de camadas em interação (Lee et al., 2015). Estes sistemas interconectados são muito sofisticados e podem dar uma explicação melhor do funcionamento no campo da análise de rede, economia, gerenciamento de operações, finanças, etc., sendo consistente com a governança corporativa preocupações.

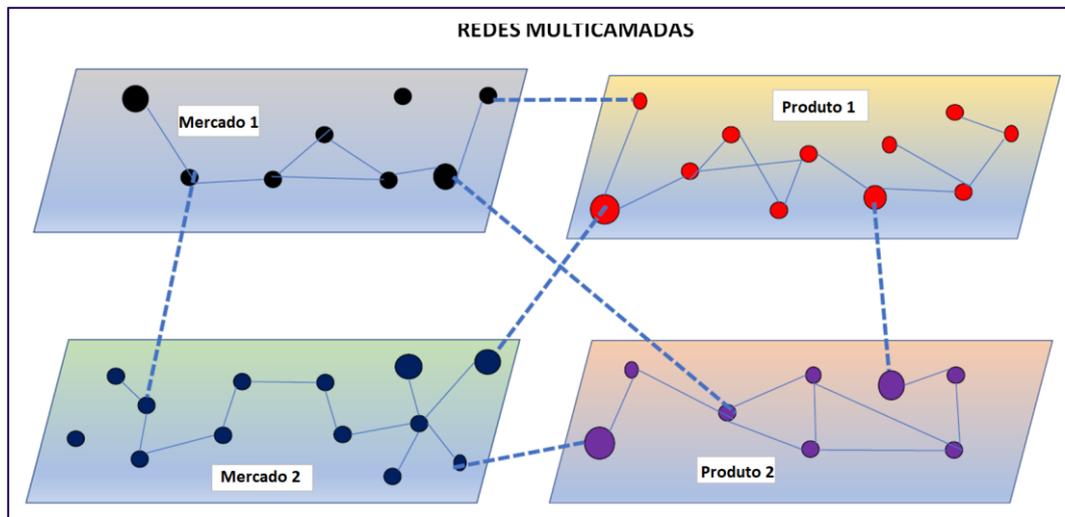
Redes multidimensionais complexas hospedam vários tipos de relações (multiplex, multicamada, multinível, multirelacional, interconectados, interdependentes etc.) e podem gerar informações valiosas em muitos campos interdisciplinares. Estas redes de redes podem dizer respeito a redes sociais que envolvem diferentes tipos de conexões, redes de aeroportos conectados por diferentes transportadoras aéreas, múltiplas infra-estruturas de um país que estão mutuamente conectadas etc.

Nós pertencentes simultaneamente a diferentes camadas (redes) podem ser representados matematicamente por adjacência tensores com bordas entre camadas que conectam uma rede à outra. Esses links aumentam o valor geral da rede de redes, impulsionando a formulação da Metcalfe.

Enquanto a matemática sofisticada que explica essas relações (ver Bianconi, 2018) vai muito além disso o estudo preliminar, algumas implicações econômicas podem valer a pena considerar.

As arestas entre as camadas (links) entre os diferentes nós vão além de cada camada e conectam duas (ou mais) camadas adjacentes, representando uma rede de redes com vários subsistemas e propriedades de conectividade. Se os links entre os nós aumentam (ambos na mesma

camada e graças a uma conexão entre camadas), existe uma correspondente valorizar o crescimento da rede sistêmica de redes que pode ser estimada com a lei da Metcalfe.



Fonte: Moro, 2019

A Figura anterior mostra, à primeira vista, que as bordas de ponte entre redes (que vinculam nós no mercado 1 ao mercado 2, produto 1 e produto 2) agregam valor a todo o ecossistema da rede. Esse valor incremental pode ser estimado provisoriamente (com um diferencial sem / com abordagem) comparando redes não relacionadas com redes vinculadas.

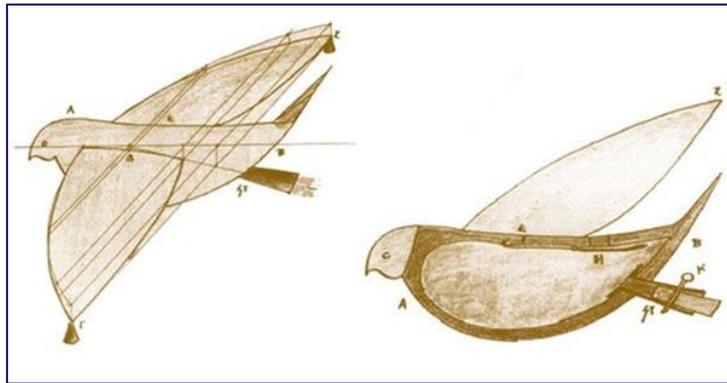
Uma interpretação econômica das redes multiplex é - pelo melhor conhecimento do autor - ainda pouco explorada e pode generalizada (incluindo outras camadas de interação em um ecossistema dinâmico), dando uma explicação inovadora de interações entre o fornecimento eletrônico e as cadeias de valor eletrônico. As plataformas digitais podem mais uma vez representar a borda de ligação virtual entre as redes.

ROBÓTICA

Os inventores e engenheiros nos tempos antigos tentaram construir artefatos que se assemelhavam a animais, engenhos e seres humanos.

O Deus Hefesto da mitologia grega¹ era responsável pela criação de seres mecanizados e projetou para uso pessoal o gigante de bronze Talos que guardava a ilha de Creta.

A possibilidade de imitar os Deuses fizeram dos gregos, os primeiros experimentadores em engenharia mecânica. No século V a.C, Archytas de Tarentum - o fundador da mecânica matemática - criou junto com a sua equipe um pombo de madeira a vapor, (figura a seguir) frequentemente considerado como o primeiro robô da história.



Três séculos mais tarde, engenheiros desconhecidos criaram uma calculadora astronômica - o Mecanismo Antikythera - cujo mecanismo provavelmente ficava alojado em uma caixa de madeira que era operado por uma manivela. O complexo aparelho semelhante a um relógio, com mostradores e engrenagens em bronze e previa com precisão eclipses lunares e solares e indicava as posições do Sol, da Lua e de alguns planetas, além de determinar as datas dos Jogos Olímpicos. Também foi descoberto num barco mecanizado naufragado nos arredores de Creta em 1901. Os cientistas que examinaram formato das letras gregas nas inscrições estimaram que o mecanismo foi montado entre 150 e 100 a.C. que era tão preciso que nada de compatível surgiria nos próximos mil anos. Embora não fosse programável no sentido moderno, há quem diga que foi o primeiro computador analógico.

¹ A mitologia grega surgiu por volta do ano 700 a.C. – a palavra mito tem origem grega e quer dizer narrar ou contar. Na época, os gregos eram politeístas (acreditavam em vários deuses) e antropomórficos (os deuses se assemelhavam aos homens). Assim, a mitologia grega é a narração de histórias que tentam explicar a origem dos fenômenos naturais e do mundo – com direito a divindades, semideuses, heróis e criaturas fantásticas!



No mesmo século V a.C. o engenheiro chinês King-Shu Tse projetou um pássaro mecânico e um cavalo mecânico que poderia pular. Já na Idade Média, Al-Jazari, de Diyarbakir (atual Turquia), foi provavelmente o maior gênio científico do século XIII. Em 1206, ele publicou o livro “*The Book of Knowledge of Ingenious Mechanical Devices*” que continha instruções para a construção de dezenas de aparelhos mecânicos. Projetou os mais famosos dos seus autômatos - um bando de músicos mecânicos - para o governador turco Artuklu, Nasreddin Muhammad.

Trezentos anos depois, Leonardo da Vinci fez uma façanha semelhante para o rei francês François I, construindo um autômato-leão capaz de se levantar sobre as patas traseiras. No século XVIII, os autômatos estavam em toda parte, desde um pato defecador feito por Jacques de Vaucanson até o "Tigre de Tipu" projetado quase em tamanho real para um sultão na Índia agredir o soldado europeu.

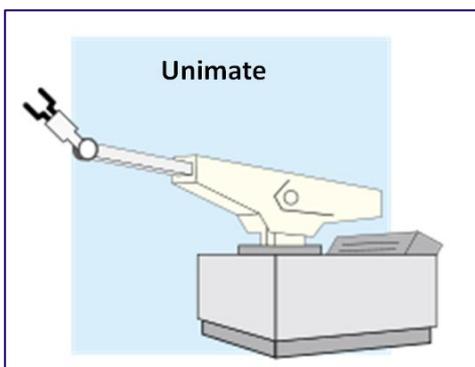


Em 1893, o escritor americano Ambrose Biercrystalizou descreveu os horrores em "Moxon's Master", uma pequena história sobre um autômato que mata seu mestre. Três décadas depois, o dramaturgo tcheco Karel Capek escreveu "Os Robôs Universais de Rossum", onde cunhou a palavra "robô" derivado da palavra tcheca “robotá”, que significa “trabalho forçado”.

Em seu filme “Metrópolis”, de 1927, o diretor Fritz Lang levou a ideia adiante com a Maria, a malvada “*machinenmensch*” (máquina humana), cuja aparência realista a torna indetectável.

A maioria dos robôs são projetados para agirem como humanos, e com sua própria inteligência executam tarefas perigosas, insalubres e arriscadas como desativar bombas e acesso a locais nocivos, contaminados ou inseguros.

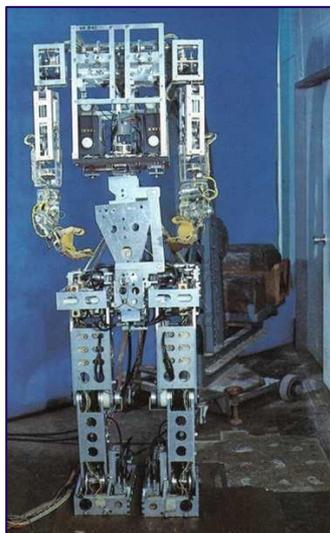
O Unimate, primeiro robô programável industrial inventado por George Devol na década de 1950, cuja patente original registrada em 1954 e concedida em 1961 (U.S. 2.988.237). Foi comprado pela General Motors em 1961 para uso em linhas de montagem de automóveis.



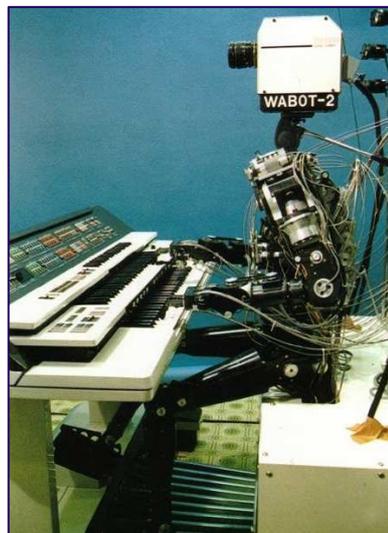
George Charles Devol Jr. (1912-2011) inventor americano ganhou o título de "Avô da Robótica". O primeiro braço robótico programável operado digitalmente representa a base da moderna indústria de robótica, e fundou a primeira empresa de fabricação de robôs do mundo com Joseph Engelberger, seu parceiro de negócios. O Unimate original tinha uma grande caixa com computador conectada a outra caixa com o braço, e os programas armazenados numa memória a tambor para execução de tarefas sistemáticas. A máquina assumiu o trabalho de transportar peças fundidas na linha de montagem para soldar nas carrocerias de automóveis. São tarefas perigosas onde os trabalhadores correm riscos de envenenamentos por vapores tóxicos ou sofrer acidentes no ambiente insalubre e perigosos.

Melhorando significativamente o Unimate, em 1972, os pesquisadores da Universidade de Waseda (Tokyo-Japão) construíram o primeiro robô humanoide inteligente do mundo, o WABOT-1.

WABOT-1

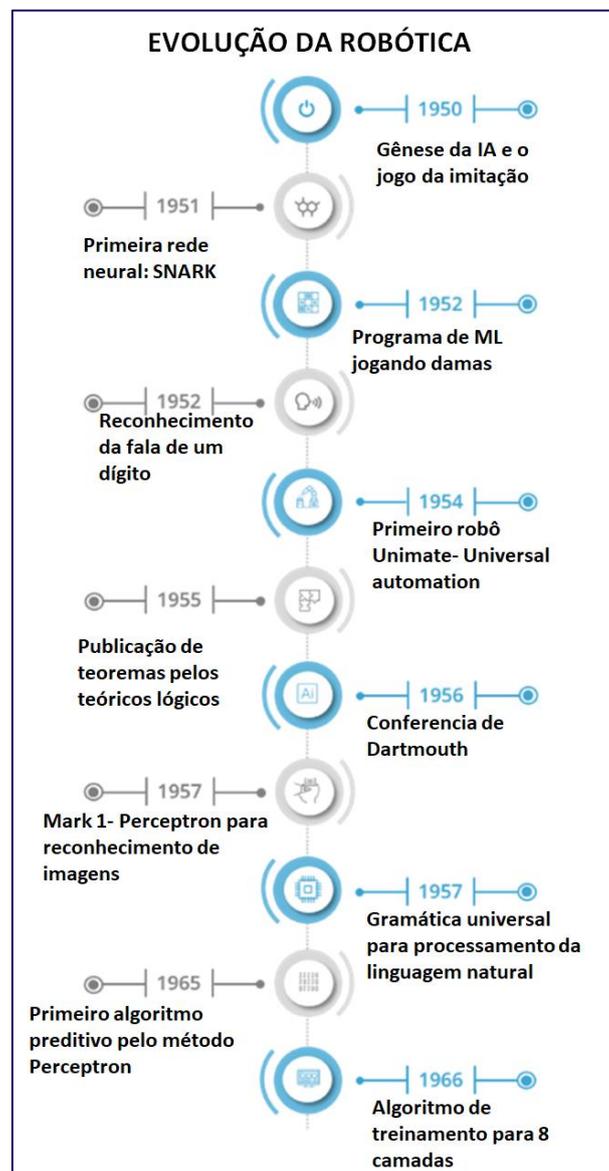


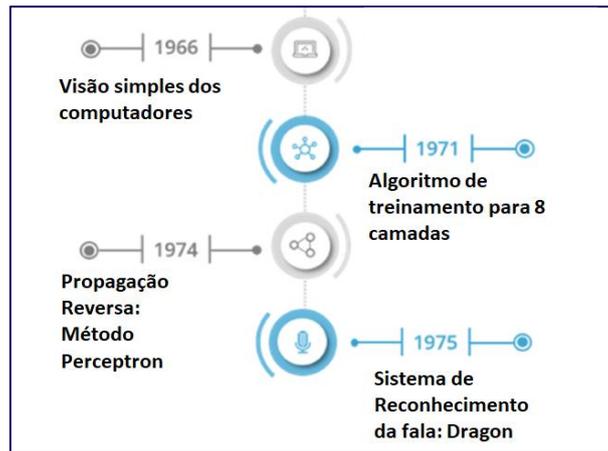
WABOT-2



Embora fosse semelhante a um brinquedo, o sistema de membros lhe permitia andar, agarrar e transportar objetos com as mãos; o sistema de visão e audição (olhos e ouvidos artificiais) permitiu medir distâncias e visualizar direções dos objetos; com a boca artificial permitiu diálogos em japonês. Esta base permitiu inovações no campo da visão de máquina e na criação de robôs para operações de empilhar blocos de objetos.

Robert Williams foi a primeira vítima causada por um robô. O acidente aconteceu na fábrica da Ford em Flat Rock, Michigan, no dia 25 de janeiro de 1979. Foi atingido por um braço do Robô no exato momento em que o operário subiu numa prateleira para pegar alguns moldes e vindo a óbito. Ironicamente, era o robô que tinha a função de pegar esses moldes. A justiça determinou que o acidente foi causado por medidas de segurança inadequadas. Sua família recebeu uma indenização de 10 milhões de dólares.





Fonte: Aggarwal, 2018

Os robôs são normalmente definidos como agentes físicos que executam uma variedade de tarefas manipulando o mundo físico.

“Robô” é uma palavra cunhada por Josef Capek (1887-1945), pintor, escritor e poeta checo. “Robota” significa “trabalho forçado” em sérvio e sua raiz é “rab”, escravo.

O escritor russo-americano Isaac Asimov (1920-1992) instituiu as três leis da robótica: um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano sofra algum mal; um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que entrem em conflito com a Primeira Lei; um robô deve proteger a própria existência, desde que tal proteção não entre em conflito com a Primeira ou Segunda Leis.

Conforme IRF (The International Federation of Robotics), o robô industrial (definido pela ISO 8373:2012) é formado por um atuador controlado automaticamente, reprogramável, de multiuso e programável em três ou mais eixos que pode ser fixo no local ou móvel para uso em aplicações de automação industrial.

Os termos usados na definição significam:

- **Reprogramável:** projetado para que os movimentos programados ou funções auxiliares possam ser alterados sem a alteração física;
- **Multiuso:** capaz de ser adaptado para diferentes aplicações com alteração física;
- **Ajustes físicos:** alteração do sistema mecânico (o sistema mecânico não inclui mídia de armazenamento como ROMs e SSDs)
- **Eixo:** usada para especificar a direção do movimento do robô em um modo linear ou rotativo.

A figura a seguir mostra tipos de robôs cobertos por essa definição e, portanto, incluídas nas pesquisas. Um robô que possui seu próprio sistema de controle e não é controlado pela máquina deve ser incluído nas estatísticas, embora possa ser dedicado a uma máquina especial. Outros robôs industriais dedicados não devem ser incluídos nas estatísticas.

Os manipuladores de wafer têm seu próprio sistema de controle e devem ser incluídos nas estatísticas de robôs industriais. Os manipuladores de bolachas podem ser robôs articulados, cartesianos, cilíndricos ou SCARA.



Outros robôs industriais dedicados são:

- Equipamentos de carga/descarga de bens, máquinas e ferramentas
- Equipamento de montagem dedicado, por exemplo para montagem em placas de circuito impresso
- Manipuladores de circuitos integrados (seleção e inserção)
- Sistemas automatizados de armazenamento e recuperação
- Veículos guiados automatizados (AGVs)

TIPOS DE ROBÔS

Classificação por estrutura mecânica:

- Robôs lineares (incluindo robôs cartesianos e de pórtico)
- Robôs SCARA
- Robôs articulados
- Paralelos / robôs delta

- Robôs cilíndricos
- Outros não classificados

A Figura a seguir ilustra a configuração mecânica desses tipos de robôs. O número de eixos deve ser entendido como a característica básica fornecida pelo produtor e não como eixos adicionados posteriormente pelo usuário.

Os robôs decompostos por estrutura mecânica baseiam-se nas seguintes definições:

- **Robô cartesiano:** robô cujo braço possui três articulações prismáticas e cujos eixos estão correlacionados com um robô SCARA do sistema de coordenadas cartesianas: um robô com dois paralelos juntas rotativas para garantir a conformidade em um avião Robô articulado: um robô cujo braço possui pelo menos três juntas rotativas Robô paralelo / delta: um robô cujos braços têm juntas prismáticas ou rotativas simultâneas Robô cilíndrico: um robô cujos eixos formam um sistema de coordenadas cilíndricas.

Os robôs estão equipados com atuadores como pernas, rodas, braços e pinças. O único objetivo dos atuadores é exercer força física no meio ambiente. Atualmente, os robôs também usam toda uma gama de sensores, incluindo câmeras, sonares, LiDARs, radares e lasers para avaliar e medir o seu ambiente de interação, bem como giroscópios e acelerômetros para medir seu próprio movimento.

Os robôs de hoje se enquadram em uma das três categorias principais que são os manipuladores, robôs móveis e os manipuladores móveis.

- **Manipuladores:** geralmente chamados de braços de robô e estão fisicamente ancorados em seu local de trabalho ou atuação como linhas de montagem de fábricas e estações espaciais. O movimento do manipulador geralmente envolve uma cadeia de juntas controláveis permitindo que eles posicionem seus atuadores em qualquer posição dentro do posto de trabalho. A maioria dos robôs industriais pertencem a esta categoria e também podem ser encontrados em hospitais auxiliando cirurgias, centros de distribuição nos processos de movimentação e armazenamento, às vezes em sistemas embarcados em dispositivos móveis.
- **Robôs móveis:** podem mover-se usando rodas ou pernas robóticas. Os veículos terrestres não tripulados (UGVs) podem dirigir autonomamente em ruas, rodovias e também fora de estrada. Veículos Aéreos Não Tripulados (UAVs, também conhecidos como drones) são comumente usados para vigilância, operações militares, pulverização de colheitas e até entregas de mercadorias. Os Veículos Subaquáticos Autônomos (AUVs) são usados em explorações em alto mar e buscas subaquáticas. Também podem ser encontrados robôs móveis entregando

comida em hospitais, movendo contêineres nas docas de carregamento, entregando pacotes aos clientes e aspirando pisos em escritórios e residências.

- **Manipuladores móveis:** Às vezes chamados de robôs humanoides, normalmente imitam os movimentos humanos. Os manipuladores móveis podem aplicar seus atuadores numa área muito maior do que os manipuladores típicos que são ancorados. No entanto, sua tarefa é dificultada, pois não têm a rigidez que as âncoras fornecem.

Robôs reais devem lidar com ambientes parcialmente observáveis, estocásticos, dinâmicos e contínuos. Pois os sistemas robóticos práticos precisam desenvolver conhecimentos prévios sobre o robô, seu ambiente físico e as tarefas que podem executar para que possam aprender rapidamente e executar as tarefas com segurança.

Para tal, a robótica reúne vários conceitos de IA e Machine Learning em questões como estimativa de estado probabilístico, percepção, aprendizado não supervisionado e aprendizado por reforço, entre outros. O sucesso dos robôs depende bastante do *design* de sensores e atuadores adequados à tarefa para a qual o robô está sendo implantado.

Vamos examinar brevemente alguns dos domínios de aplicativos mais populares e comuns da tecnologia robótica.

- **Indústria e Agricultura:** Máquinas pesadas usadas para colheita, mineração ou escavação de terra são cada vez mais robotizadas. Os robôs têm sido usados para remover a tinta dos navios pelo menos 50 vezes mais rápido que os humanos, com impactos ambientais resultantes muito reduzidos. Os robôs foram usados para gerar mapas de alta precisão de minas e sistemas de esgoto abandonados. Eles também foram encontrados para ser muito mais rápido do que os seres humanos no transporte de minério de minas subterrâneas.
- **Assistência médica:** Os robôs são frequentemente usados para auxiliar os cirurgiões na colocação de instrumentos na cirurgia de operar em órgãos intrincados, como cérebro, coração e olhos. Devido à sua alta precisão, eles se tornaram auxílios indispensáveis em procedimentos cirúrgicos, como a substituição da anca. Fora da sala de operações, eles podem ser utilizados como auxiliares de pessoas idosas e deficientes, como andadores robóticos inteligentes e dispositivos inteligentes que lembram as pessoas a tomar seus medicamentos a tempo e proporcionam conforto a eles.
- **Serviços pessoais:** Os robôs de serviço ajudam as pessoas a realizar tarefas diárias, como aspirar escritórios ou residências, cortar a grama ou atuar como caddies de golfistas. Os robôs de serviço em locais públicos atuam como quiosques de informações em shopping centers,

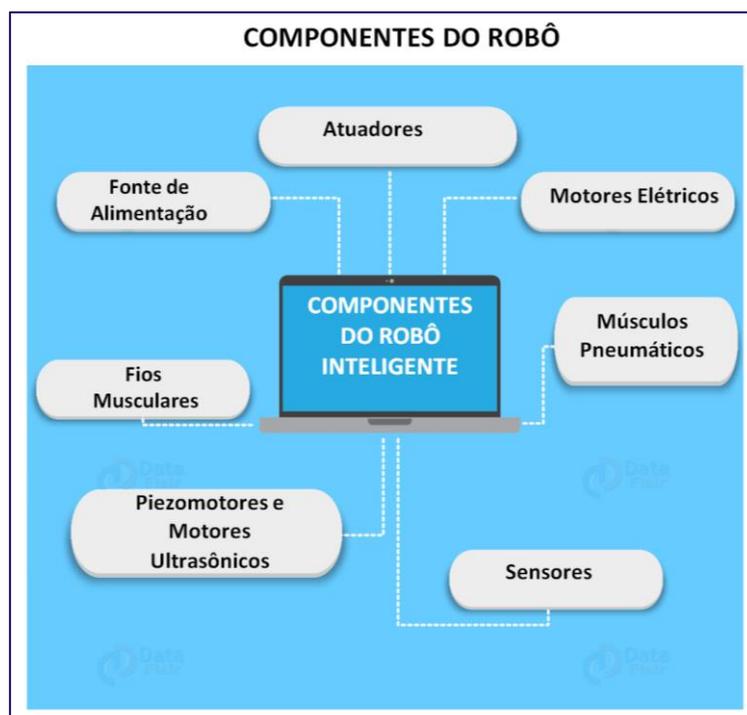
feiras e museus. Eles não apenas interagem com os seres humanos, mas possuem a capacidade de lidar com ambientes imprevisíveis e dinâmicos.

- **Ações humanas aumentadas:** os dispositivos robóticos facilitam as pessoas a andar ou mover os braços fornecendo forças adicionais nos locais apropriados. A telecirurgia robótica envolve a realização de tarefas a longas distâncias com o auxílio de dispositivos robóticos. Uma configuração regular é a de mestre-escravo, onde o atuador do robô emula o movimento da cirurgia remota medido através de uma interface háptica. Os veículos subaquáticos telecomandados são enviados para profundidades consideradas perigosas para os seres humanos.

SISTEMAS DE PERCEPÇÃO ROBÓTICA INTELIGENTE

A percepção robótica está relacionada a muitas aplicações em robótica, onde dados sensoriais e técnicas de inteligência artificial/aprendizado de máquina (AI/ML) estão envolvidos. Exemplos de tais aplicações são detecção de objetos, representação do ambiente, entendimento de cenas, detecção de pedestres/humanos, reconhecimento de atividades, classificação de locais semânticos, modelagem de objetos, entre outros.

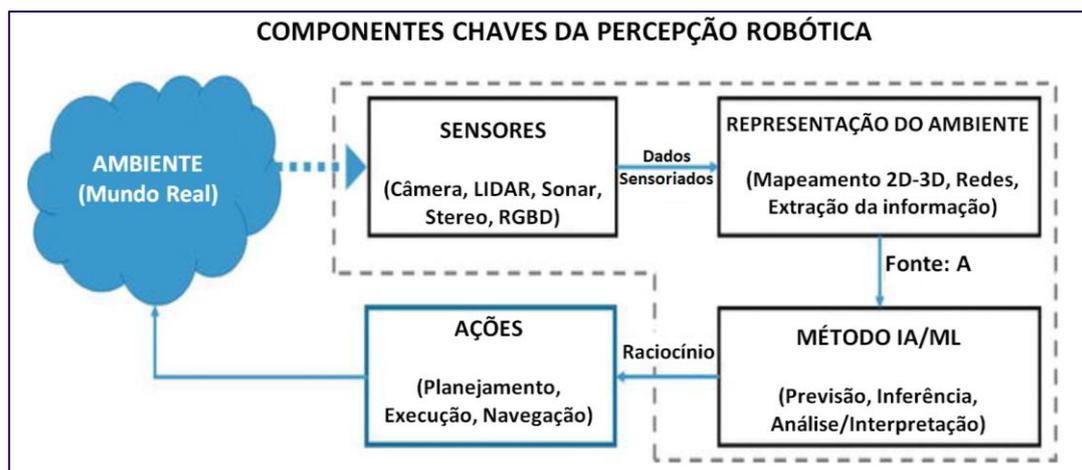
A percepção robótica, no escopo deste tópico, abrange os algoritmos e as técnicas de ML que capacitam os robôs a aprender com dados sensoriais e, com base em modelos aprendidos, a reagir e tomar decisões em conformidade.



Fonte: DataFlyer, 2020

Os recentes desenvolvimentos no aprendizado de máquina, ou seja, abordagens de aprendizado profundo (DL) são evidentes e, conseqüentemente, os sistemas de percepção robótica estão evoluindo de uma maneira que novas aplicações e tarefas estão se tornando realidade. Os recentes avanços na interação homem-robô, tarefas robóticas complexas, raciocínio inteligente e tomada de decisão são, em certa medida, os resultados da notória evolução e sucesso dos algoritmos de ML.

A capacidade de percepção é entendida como um sistema que confere ao robô a capacidade de perceber, compreender e raciocinar sobre o ambiente circundante. Os componentes principais de um sistema de percepção robótica são essencialmente processamento de dados sensoriais (enfocando aqui a percepção visual e de alcance); representações de dados específicas para as tarefas em questão; algoritmos para análise e interpretação de dados (usando métodos IA/ML); e planejamento e execução de ações para interação robô-ambiente ilustrado na figura a seguir:



Fonte: Premebida et al, 2018

Como a IA forte ainda está longe de ser alcançada em aplicativos de robótica do mundo real, este capítulo trata-se de IA fraca, isto é, abordagens padrão de aprendizado de máquina.

A percepção robótica é crucial para um robô tomar decisões, planejar e operar em ambientes do mundo real por meio de inúmeras funcionalidades e operações, desde o mapeamento da grade de ocupação até a detecção de objetos.

Alguns exemplos de subáreas de percepção robótica incluindo veículos-robô autônomos são detecção de obstáculos, reconhecimento de objetos, classificação semântica de locais, representação do ambiente 3D, reconhecimento de gestos e vozes, classificação das atividades e tarefas repetitivas, classificação do terreno, detecção e conhecimento de estradas e vias, detecção de veículos, detecção de pedestres e outros objetos e animais, rastreamento de objetos, detecção de imagens nos humanos e detecção de alterações ambientais entre outros.

Atualmente, a maioria dos sistemas de percepção robótica usa técnicas de aprendizado de máquina (ML), variando de abordagens clássicas a profundas. O aprendizado de máquina para percepção robótica pode estar na forma de aprendizado não supervisionado ou classificadores supervisionados usando recursos artesanais ou redes neurais de aprendizado profundo (por exemplo, rede neural convolucional (CNN)), ou mesmo uma combinação de vários métodos.

Independentemente da abordagem de ML considerada, os dados do(s) sensor(es) são o ingrediente chave na percepção robótica. Os dados podem vir de um único ou vários sensores, geralmente montados a bordo do robô, mas também podem vir da infraestrutura ou de outro robô (por exemplo, câmeras montadas em UAVs voando nas proximidades). Na percepção de múltiplos sensores, na mesma modalidade ou multimodal, geralmente é necessária uma abordagem eficiente para combinar e processar dados dos sensores antes que um método de ML possa ser empregado. As etapas de alinhamento e calibração de dados são necessárias, dependendo da natureza do problema e do tipo de sensores utilizados.

A representação/mapeamento do ambiente baseado em sensor é uma parte muito importante de um sistema de percepção robótica. O mapeamento aqui abrange tanto a aquisição de um modelo métrico quanto sua interpretação semântica, sendo, portanto, sinônimo de representação do ambiente/cena.

Este processo de mapeamento semântico usa ML em vários níveis, por exemplo, raciocínio sobre ocupação volumétrica e oclusões, ou identificação, descrição e correspondência ideal das regiões/locais de diferentes variáveis/modelos de tempo, ou seja, não apenas interpretações de nível superior. No entanto, na maioria das aplicações, o papel principal do mapeamento do ambiente é de modelar os dados captados de sensores montados a bordo do robô, a fim de permitir o raciocínio e inferência sobre o ambiente do mundo real onde o robô opera.

As funções de percepção do robô, como localização e navegação dependem do ambiente em que o robô opera. Essencialmente, um robô foi projetado para operar em duas categorias de ambientes: interior ou exterior.

As diferenças e os desafios entre um robô móvel que navega em um ambiente interno versus um externo é o solo, ou terreno.

A maioria dos robôs internos supõe que o solo seja regular e plano, o que, de alguma maneira, facilita os modelos de representação do ambiente; por outro lado, para robôs de campo (ao ar livre), o terreno geralmente está longe de ser regular e, como consequência, a modelagem do ambiente é um desafio e, sem uma representação adequada, as tarefas de percepção subsequentes são afetadas negativamente. Além disso, no exterior, a percepção robótica deve lidar com as condições climáticas e as variações nas intensidades e espectros de luz.

Inteligência artificial e aprendizado de máquina aplicados à percepção da robótica

Uma vez que um robô é (auto) localizado, pode prosseguir com a execução de sua tarefa mediante a localização dos objetos de interesse no ambiente operacional e a captura deles. Numa configuração típica, o robô navega para a região de interesse para criar um mapa 3D para planejar uma rota sem colisões para localizar objetos de destino. O alvo pode ser uma mesa ou contêiner em que algo precisa ser enviado ou um objeto a ser retirado conforme os 6 graus de liberdade de um objeto.

ROBÔS COLABORATIVOS

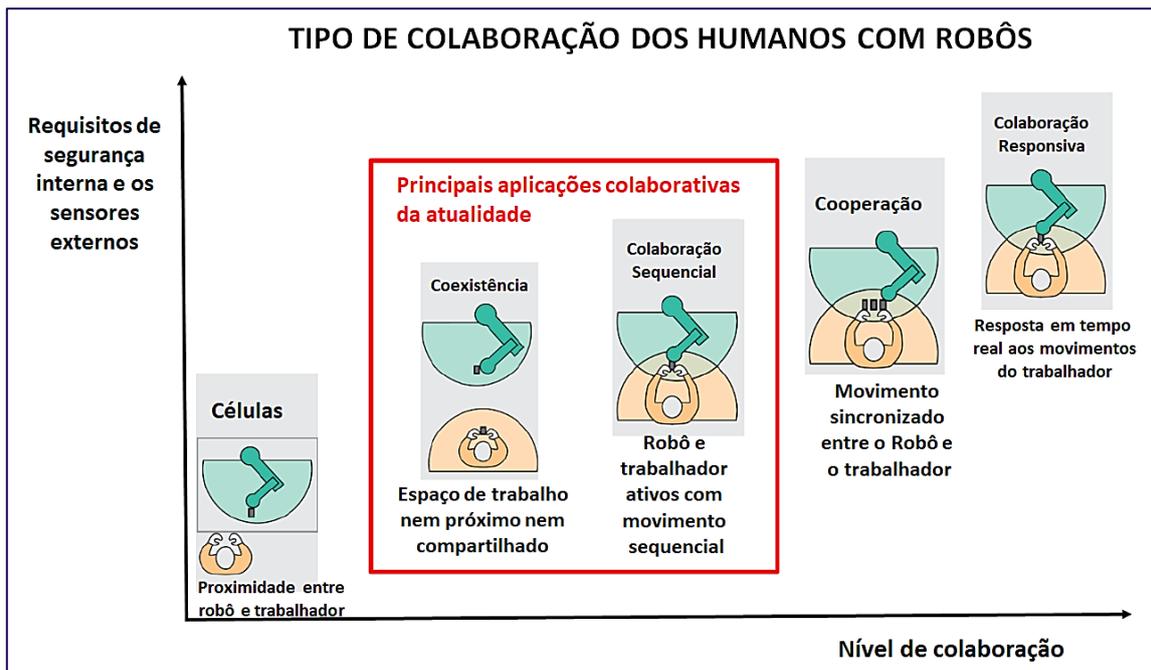
A International Federation of Robotics define dois tipos de robôs projetados para uso colaborativo para realizar tarefas em colaboração com trabalhadores em setores industriais e na prestação de serviços.

- Grupo que atendem à norma ISO 10218-1² da Organização Internacional de Padrões, que especifica os requisitos e diretrizes para um projeto seguro inerente, medidas de proteção e informações para o uso de robôs industriais.
- Grupo que não atendem aos requisitos da ISO 10218-1. Isso não significa que esses robôs sejam inseguros pois seguem diferentes padrões de segurança, por exemplo, padrões nacionais ou internos.

Há uma variação considerável nos tipos de robôs colaborativos que atendem às especificações acima e no nível de contato entre o robô e o trabalhador em aplicativos colaborativos.

No âmbito do espectro técnico se encontra o espaço de trabalho entre o robô e o trabalhador geralmente equipado com sensores que detectam o movimento humano e garantem que o robô ajuste a velocidade conforme o seu ritmo.

² International Organization for Standardization (ISO) define robôs de acordo com o uso pretendido. Para a ISO 8373: 2012, o robô industrial é um manipulador multiuso, controlado automaticamente e programável em três ou mais eixos, que podem ser fixos no local ou móvel para uso em aplicações de automação industrial.



Fonte: Bauer et al., 2016

Os Cobots possuem uma variedade de recursos técnicos que asseguram segurança ao trabalhador em contato direto, deliberadamente ou por acidente. Tais recursos incluem materiais leves, contornos arredondados, sensores que medem e controlam a força e a velocidade que não excedam os limites definidos em caso de contato físico.

Eles podem ser usados para automatizar a linha de produção de peças de pequeno a médio porte com mudanças mínimas no resto da linha com melhorias de produtividade e qualidade oferecidos por robôs.

As *interfaces* de programação cada vez mais intuitivas sem o redesenho de todo o processo de produção e podem reimplantar o robô para uma nova tarefa.

Os Cobots são geralmente leves que além de ser facilmente transportados pela fábrica ocupam menos espaço na fábrica, sendo um fator de custo significativo para os fabricantes.

Os robôs industriais geralmente operam em uma montagem fixa, mas há demanda por robôs industriais móveis que combinem uma base móvel e um robô (colaborativo). Esses robôs podem, por exemplo, transportar materiais de uma estação de trabalho e descarregá-los ou alimentar uma máquina em uma segunda estação de trabalho.

A escolha certa do robô - tradicional ou colaborativo - é determinada pela aplicação pretendida. Quando a velocidade e a precisão absoluta são os principais critérios de automação, é improvável que qualquer forma de aplicação colaborativa seja economicamente viável. Nesse caso, um robô industrial tradicional com cerca é - e continuará sendo - a escolha preferida. Se a peça que está sendo manipulada puder ser perigosa quando em movimento, por exemplo, devido a arestas

vivas, será necessária alguma forma de cerca. Isso se aplica até mesmo para cobots que param no contato. Outro fator que influencia a viabilidade econômica é até que ponto o robô deve ser integrado a outras máquinas em um processo. Quanto mais integração necessária, maior será o custo de instalação do robô.

Os robôs industriais colaborativos são ferramentas para apoiar os funcionários em seu trabalho, livrando-os de muitas tarefas pesadas, não ergonômicas e tediosas, como segurar uma peça pesada firme na posição necessária para que o trabalhador coloque os parafusos. No entanto, ainda existem muitas tarefas que são fáceis para humanos, mas difíceis de automatizar, por exemplo, lidar com peças não classificadas e formas irregulares ou flexíveis. As aplicações de acabamento, como polimento e retificação, que exigem um ajuste fino contínuo da pressão aplicada à superfície, também são difíceis de automatizar de maneira econômica. A robótica colaborativa permite que os fabricantes aumentem a produtividade usando robôs para complementar as habilidades humanas (consulte também Documento de Posicionamento IFR: Robôs e o Local de Trabalho do Futuro para obter mais informações sobre como os robôs auxiliam os trabalhadores em manufatura, logística e saúde).

ROBOTIC PROCESS AUTOMATION (RPA)

O RPA é um software que simula as etapas que o humano desempenha em tarefas repetitivas com base em regras e mantendo a velocidade e a precisão constantes na execução das atividades.

Na automação tradicional, todas as ações são baseadas principalmente na programação/ script, APIs³ ou outras formas de métodos de integração aos sistemas de back-end ou aplicativos internos. Por distinção, o RPA automatiza o software que pode migrar o trabalho do ser humano para o computador, o que pode parar de pagar o trabalho humano pela automação, processamento mais rápido de transações de front e back office, próximo à integração "Instant On" com o menor custo, otimização de Interface do usuário para reduzir os longos tempos de chamada / transação, acelerar os objetivos de transformação digital, eliminar erros e melhorar a produtividade, tornando os trabalhadores mais inteligentes.

O Instituto Europeu de Patentes (EPO) reconhece Cyrille Bataller e Adrien Jacquot como inventores do termo RPA que definem como sendo “uma tecnologia que permite automatizar a execução de atividades repetitivas e manualmente intensivas.

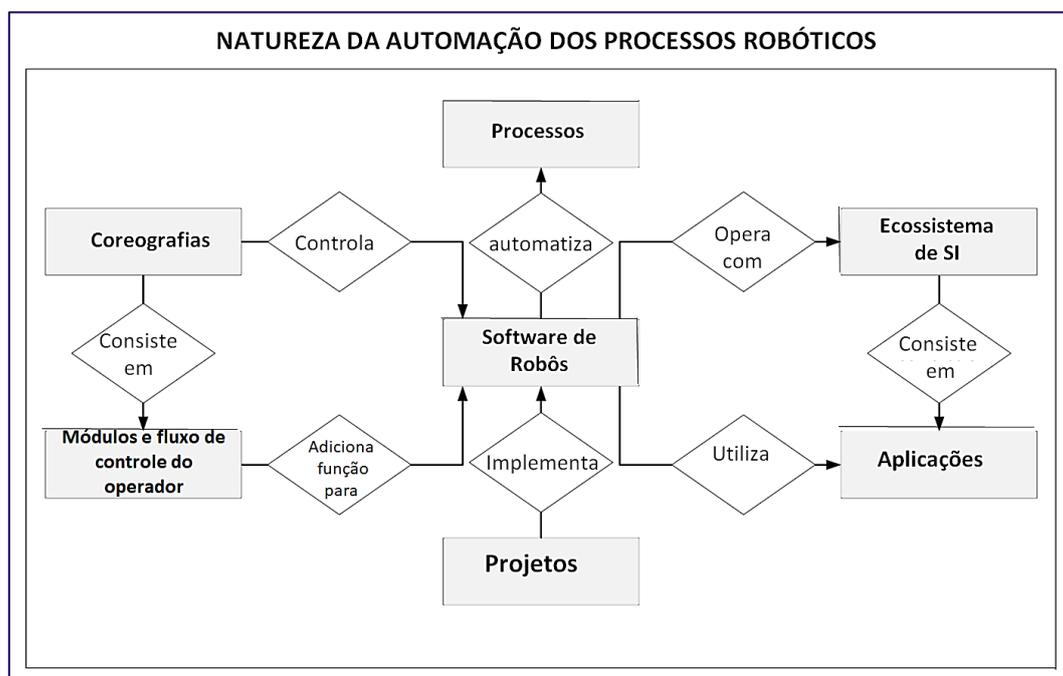
O Instituto Gartner propõe outra definição do RPA: “Trata-se de uma ferramenta de produtividade (comercializada como software licenciado) que permite ao usuário configurar um ou

³ API provém do Inglês *Application Programming Interface*, é um conjunto de rotinas e padrões estabelecidos por um software para a utilização das suas funcionalidades por aplicativos que não pretendem envolver-se em detalhes da implementação do software, mas apenas usar seus serviços.

mais scripts (alguns fornecedores referem se como "bots") para ativar o acionamento das teclas de forma automatizada.

Sutherland (2019) considera que a automação robótica é a “Aplicação de tecnologias e metodologias específicas para a utilização do computador ou < ETI ou robô virtualizado>, ou seja a manipulação de um software aplicativo (ERPs, bancos de dados, sistemas de gestão) da mesma maneira que os humanos processam uma transação ou conclui um processo. ”

Assim, os RPAs seguem uma coreografia de módulos tecnológicos e operadores de controle de fluxo enquanto opera nos ecossistemas de TI usando aplicativos estabelecidos. Facilidade de uso e a adaptabilidade permitem que as empresas concebam e implementem o RPA com projetos (ágeis).



Fonte: Adaptado de Hofmann et al., 2019

Ambiente do RPA

A automação de TI é ativada por software, incluindo APIs, scripts, trabalhos, agendas, programas, eventos e uma ampla gama de ferramentas de automação. Com o advento da computação em nuvem, os recursos são expostos por APIs e acessíveis em tempo real. Quando combinadas com conjuntos de ferramentas de automação, criação e liberação automatizadas de aplicativos, carga de trabalho, recursos de servidor e infraestrutura, as atualizações de gerenciamento de configuração agora são viáveis. Os ambientes locais são pressionados a alcançar a velocidade e flexibilidade fornecidas pelos ambientes em nuvem. Os conjuntos de ferramentas de automação estão sendo Estratégia organizacional e de TI, estruturas de governança e de gerenciamento usado para aprimorar ambientes locais.

A automação heurística que aproveita o Machine Learning é uma classe emergente de automação de TI. No entanto, atualmente existem poucos fornecedores com ofertas de nicho e um aumento da maturidade está previsto nos próximos 3-5 anos.

Para entender mais sobre automação por TI, a própria automação refere-se a uma forma não manual de concluir uma tarefa. O termo assume um significado diferente com base no contexto e na situação.

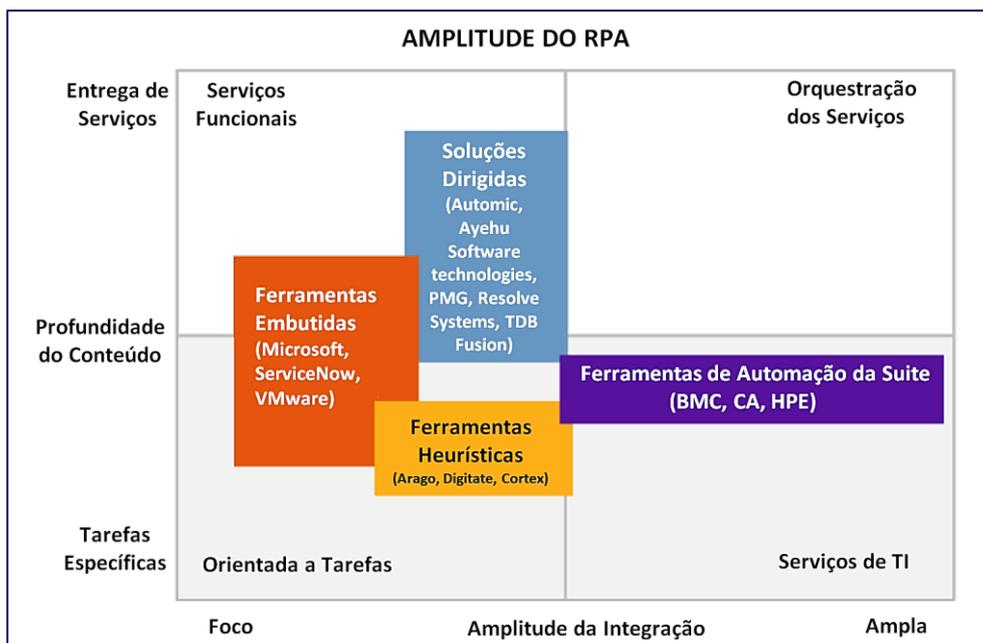
Dentro da TI, existem cinco tipos amplos de componentes de automação e tarefas associadas, como automação de processos, automação de serviços, automação de carga de trabalho, automação de infraestrutura e automação de criação e lançamento de aplicativos. A Automação de Processo Robótico (RPA) é um subconjunto da Automação de Processo.

Pois, o RPA é uma aplicação de tecnologia programada como um "robô" para capturar e interagir com aplicativos existentes para processar uma transação, manipular dados, acionar respostas e se comunicar com outros sistemas digitais. A RPA pretende ajudar ou remover uma atividade humana, usando o software para realizar tarefas. A ferramenta RPA pode ser acionada manual ou automaticamente, mover ou preencher dados entre locais prescritos, documentar trilhas de auditoria, realizar cálculos, executar ações e acionar atividades a jusante.

AUTOMAÇÃO ROBÓTICA

A automação robótica é uma soma de duas partes:

- **Robotic Desktop Automation (RDA)** - usada como robôs pessoais pelos funcionários, front office (central de atendimento, varejo, filiais) e back office, barras de ferramentas, assistentes, aprimoramentos de interface do usuário e automação de tarefas.
- **Automação de Processo Robótico (RPA)** - Robôs autônomos que replicam 100% do trabalho algorítmico, operações de back-office, de natureza repetitiva, para aliviar o trabalho de front office, melhoria em subgrupos menores



Fonte: Adaptado de Gartner, 2016

O RPA permite que as organizações implementem a automação de maneira não invasiva, o que torna único entre outros tipos de automação. Também pode reduzir a necessidade de envolvimento constante da TI e no desenvolvimento de aplicativos.

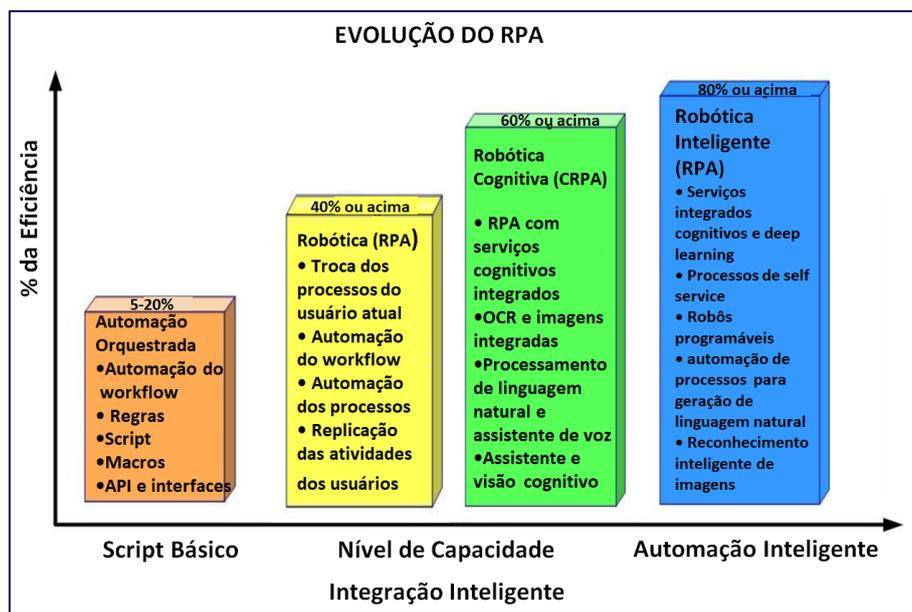
- **Redução de custo geral** - automatiza os processos eliminando as lacunas de integração de dados sem alterações nos investimentos em tecnologia subjacentes. Reduz o custo das operações e aumenta a satisfação do cliente. O custo médio da implementação e execução de um robô é muito menor que os custos equivalentes dos funcionários em tempo integral e diminui a implantação em larga escala. Ao automatizar, padronizar tarefas que exigem muita mão-de-obra e propensas a erros, erros de configuração e processos demorados e dispendiosos são reduzidos.
- **Velocidade e produtividade** - o RPA é tipicamente 2X-3X mais rápido que os humanos, mesmo que os robôs trabalhem no mesmo ritmo que os humanos, porém eles podem trabalhar o tempo todo, diferentemente dos humanos.
- **Escalabilidade e flexibilidade** - os robôs podem ser facilmente escalados para cima e para baixo para lidar com flutuações de demanda e variações sazonais.
- **Precisão e conformidade** - os robôs trabalham com níveis de precisão de 100% e permitem a conformidade, evitando erros humanos e economizando custos. Redução de vulnerabilidades de segurança com o aprimoramento de software e aumento da padronização da configuração.

EVOLUÇÃO DO RPA

As ferramentas RPA evoluem para incluir funcionalidades como inteligência artificial e recursos de aprendizado de máquina e observa se a gradativa queda nos preços. Nos segmentos industriais, serviços financeiros e no varejo assumiram a liderança com a RPA seguido pela manufatura e a saúde.

Aqui estão algumas das projeções mais interessantes sobre o futuro da robótica:

Uma porcentagem significativa de aplicativos robóticos comerciais será na forma de "robô como serviço" (RaaS). Isso ajudará a reduzir significativamente o custo da implantação do robô.



Fonte: Adaptado de Gartner, 2016

Como os robôs são usados?

Os robôs de hoje, sejam robôs industriais, robôs colaborativos ou robôs de serviço profissional, são usados em uma ampla gama de indústrias. Estes podem incluir:

- Automotivo
- Montagem
- Soldagem
- Pintura
- Eletrônicos
- Aparelho médico
- Embalagem
- Manuseio de materiais
- Nuclear

- Farmacêutico
- Mineração
- Agricultura
- Serviço
- Comida
- Segurança
- Semicondutor

Essas são algumas das principais indústrias e aplicações que utilizam os robôs modernos, mas essa lista está longe de ser exaustiva.

Uma porcentagem significativa de todas as implantações robóticas serão robôs colaborativos inteligentes que operarão três vezes mais rápido do que muitos robôs sendo usados hoje e serão seguros para trabalhar com seres humanos.

Mais de 50% dos robôs dependerão de software baseado em nuvem para definir novas habilidades, habilidades cognitivas e programas de aplicativos, levando à formação de um mercado de nuvem robótica.

Em um futuro próximo, 50% das 200 principais empresas globais de comércio eletrônico e omnichannel implementarão sistemas robóticos em suas operações de atendimento de pedidos, armazenamento e entrega.

Até 2020, mais de 40% dos robôs comerciais estarão conectados a uma malha de inteligência compartilhada, resultando em uma melhoria de 200% na eficiência operacional robótica geral. - Raja Mitra, 2018

Desafios na adoção tradicional de RPA

- **Ritmo de mudança** - A maioria das empresas está passando por uma jornada típica de dois anos para fazer algum progresso real na automação;
- **Pessoas qualificadas** - Disponibilidade limitada de recursos qualificados - domínio + tecnologia em relação à demanda do mercado;
- **Governança** - Modelo operacional de governança ruim / ineficiente na gestão e mitigação de riscos;
- **Entrada de dados** - muitas empresas são inundadas pelos desafios de começar, pois ignoram a arquitetura, a entrada e a codificação de dados;

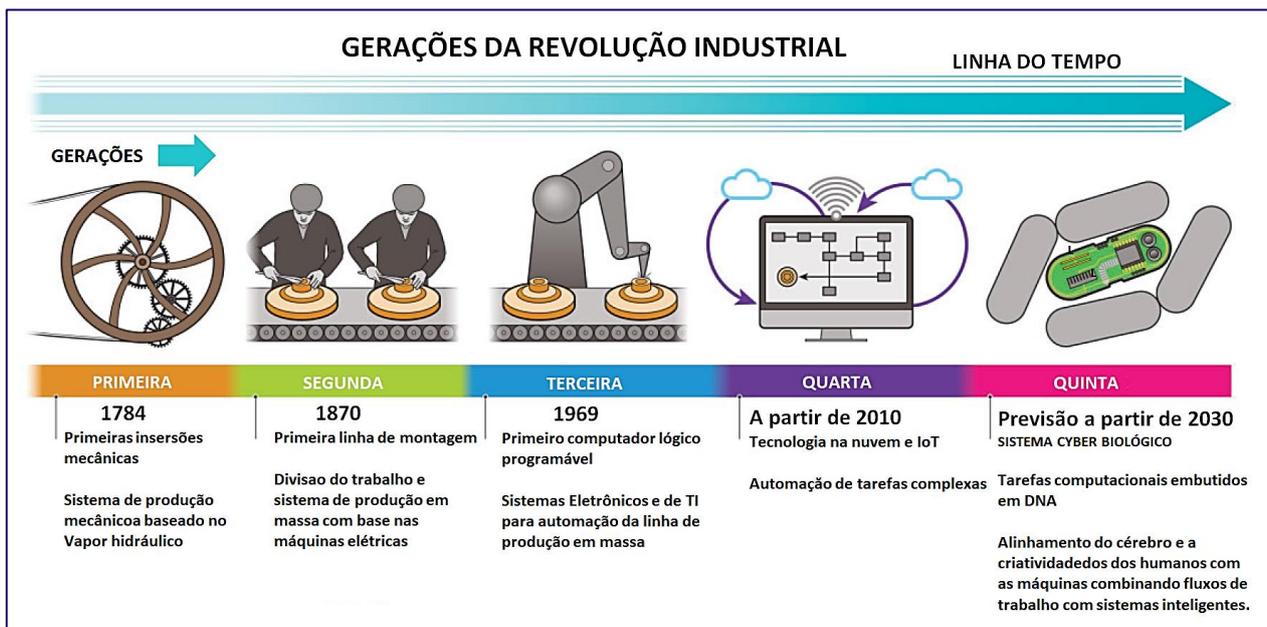
- **Experiência no setor** - À medida que o programa de automação evolui, o caminho para o que deve ser automatizado se torna menos claro e mais específico para o setor;
- **Força de trabalho digital** - Muitas empresas não planejam a manutenção dos recursos necessários quando a automação quebra ou os sistemas subjacentes são alterados;
- **Sobrecarga de tecnologia** - um obstáculo especialmente grande é o número de fornecedores de automação e a maior do que a vida alega alguma entrada 93% das organizações dizem que não estão totalmente preparadas para lidar com os problemas decorrentes de sua jornada de automação. – Forrester.

REVOLUÇÕES INDUSTRIAIS

Remontando aos anos 1780, a Primeira Revolução Industrial nasceu com a geração de energia da água, vapor e combustíveis fósseis. Na segunda revolução, a energia elétrica foi a base dos fabricantes com linhas de montagem e produção em massa na década de 1870.

Empregando tecnologias eletrônicas e da informação (TI), a Terceira Revolução Industrial teve como suporte ao conceito da automação na década de 1970.

A quarta fase da revolução utiliza o IoT (Internet of Things) e a computação em nuvem com a interface em tempo real entre o ambiente virtual e o mundo físico, os chamados sistemas cyberfísicos (CPS) com a automação dos processos e a introdução da computação distribuída e inteligente conforme a representação visual a seguir.



Fonte: Adaptado de Peccoud, 2016

Com o crescimento agressivo e a disseminação da Internet, prevê-se que mais 3 bilhões de pessoas acessarão a Internet até 2025.

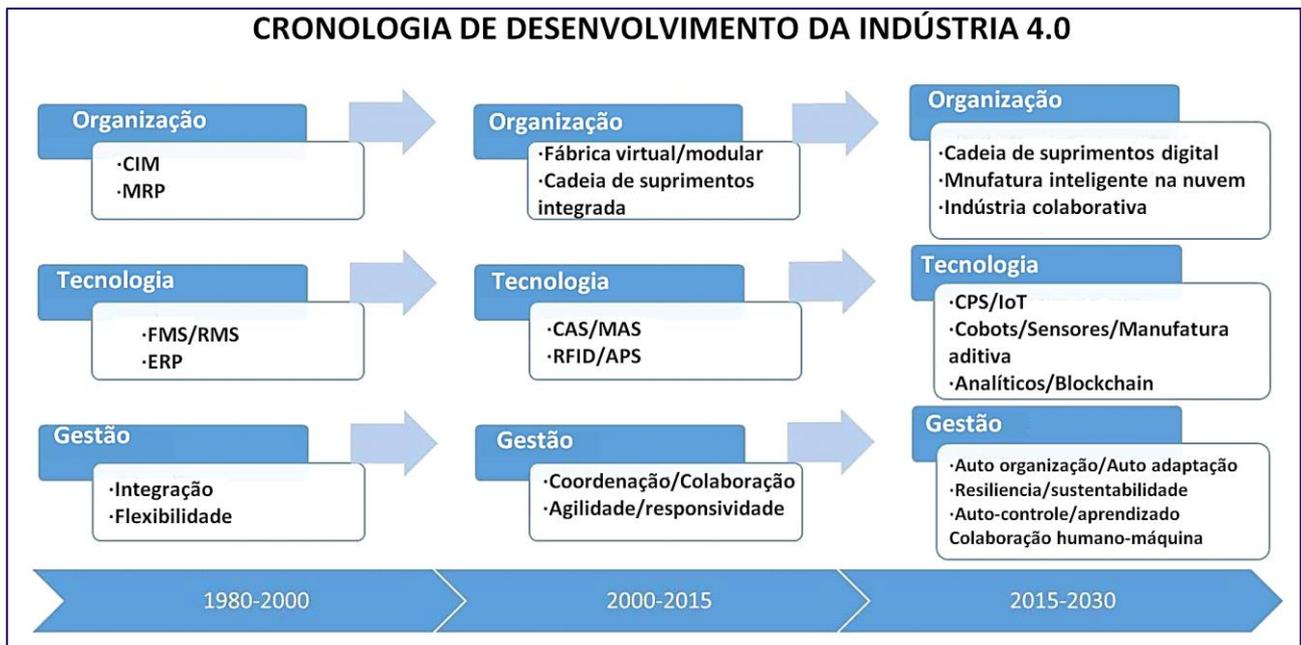
Além dos profissionais e corporações, a definição de 14.0 varia entre as disciplinas acadêmicas.

No campo das disciplinas como a engenharia, gestão, controle e ciência de dados é caracterizado como avanços tecnológicos, re-designs organizacionais, desenvolvimentos de gestão de operações (OM) e transformações de mercado (Yin et al. 2018, Tang e Veelenturf 2019, Oztemel e Gursev 2020).

Em operações, para Piccarosi et al. (2018) a "I4.0 refere-se à integração de tecnologias de IoT para criação do valor industrial onde os fabricantes na cadeia de valor totalmente digitalizadas, conectadas, inteligentes e descentralizadas são capazes de fornecer maior flexibilidade e robustez para a competitividade da empresa na construção de estruturas de negócios flexíveis e capacidades de produção adaptáveis conforme a evolução, mudança e desenvolvimento do ambiente de negócios como resultado da estratégia formulada e implementada ao longo do tempo".

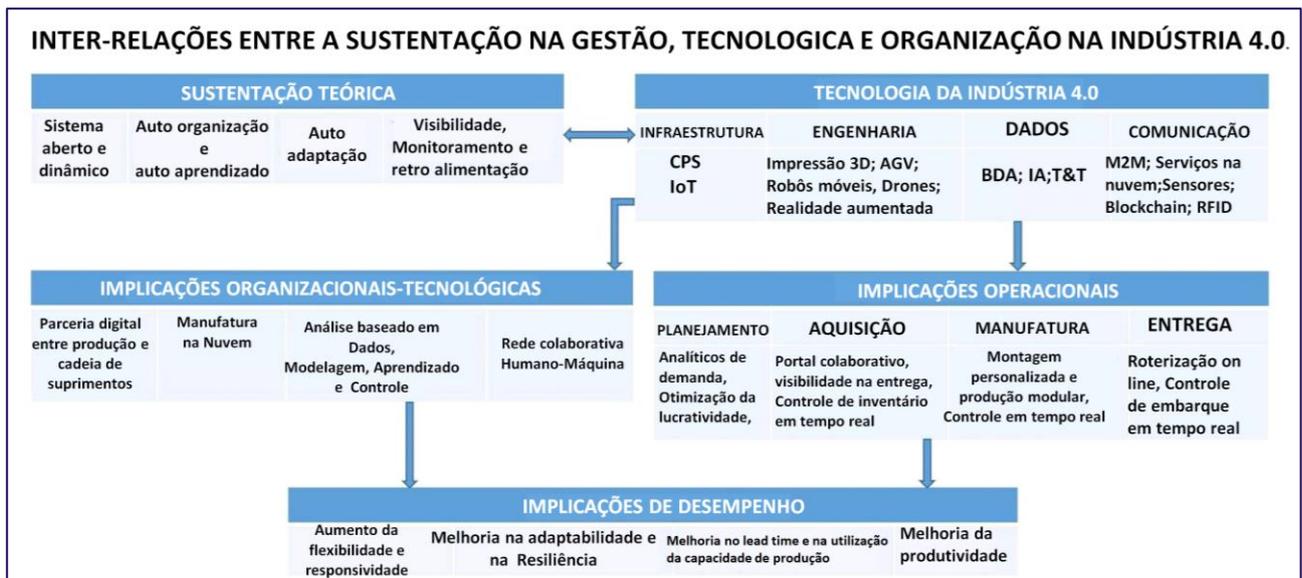
Oztemel e Gursev (2020) definem a 14.0 como a "metodologia para gerar transformação da manufatura dominante das máquinas para a manufatura digital". Nesse contexto, tem se concentrado predominantemente nas aplicações da manufatura aditiva, IoT, Blockchain, robótica avançada e inteligência artificial.

Ou seja, I4.0 trata da integração de tecnologias, conceitos organizacionais e princípios de gerenciamento subjacentes a uma rede econômica, responsiva, resiliente e sustentável, orientada pela dinâmica dos dados estruturalmente adaptável conforme as mudanças no ambiente da oferta e demanda com rearranjo rápido e realocação dos componentes e recursos.



Fonte: Ivanov et al., 2020

Observa-se que a Indústria 4.0 baseia-se em princípios fundamentais da teoria de sistemas na formação de sistemas abertos e dinâmicos com auto-organização, auto-adaptação e autoaprendizagem, bem como visibilidade, monitoramento e controle de feedback. A Figura a seguir classifica a estrutura I4.0 em quatro áreas, ou seja, infraestrutura, tecnologia de engenharia, tecnologia de dados e comunicação. A Indústria 4.0 abre novos *designs* tecnológicos de organização, como manufatura em nuvem, parcerias digitais, modelagem orientada a dados e indústria colaborativa. No nível operacional, são classificadas as implicações de I4.0 de acordo com os processos SCOR (planejar-fonte-fazer-entregar). Finalmente identifica as principais implicações de desempenho da utilização do I4.0 e também adiciona novos elementos ao OM para ajudar a projetar cadeias de suprimentos e operações, bem como métodos e modelos de tomada de decisão em nível operacional para planejamento, fornecimento, produção e logística.



Fonte: Ivanov et al., 2020

Primeiro, a fábrica inteligente da Audi em Baden-Württemberg na Alemanha, cujo projeto de processo altamente flexível com sequenciamento das ordens de produção para atingir um alto grau de personalização a eficiência (Audi 2019). Em segundo lugar, o sistema operacional aberto baseado em IoT na plataforma de manufatura na nuvem da Siemens MindSphere gera informações para analíticos avançados para gerenciar digitalmente diferentes sistemas e máquinas interconectadas nas diversas plantas fabris da empresa (Siemens 2018).

Além disso, o mundo tem visto um grande aumento na poluição ambiental a partir da Segunda Revolução Industrial. No entanto, ao contrário das últimas décadas, a indústria está cada vez mais focada em controlar os diferentes aspectos da geração e gestão de resíduos e na redução dos impactos adversos ao meio ambiente decorrentes de sua operação.

A conscientização ambiental torna-se um fator de vantagem competitiva devido à grande quantidade de apoio do governo; organizações internacionais como ONU, OMS, e até mesmo uma base importante de clientes de nicho apoiam cada vez mais as empresas ambientalmente corretas.

O foco da I4.0 não tem um forte direcionamento na proteção ambiental bem como as tecnologias para melhorar a sustentabilidade do ambiente da Terra, embora muitos algoritmos de IA diferentes tenham sido usados para investigar a perspectiva da sustentabilidade na última década.

Embora a I4.0 ainda não esteja na maturidade plena, os líderes de tecnologia observam o futuro com a Quinta Revolução Industrial: fabricação autônoma com inteligência humana dentro e no circuito conforme a representação visual a seguir.

Trazendo de volta trabalhadores humanos para o chão da fábrica, a Quinta Revolução Industrial formará a parceria humano e máquina para utilizar ainda mais a capacidade do cérebro

humano e a criatividade para aumentar a eficiência do processo combinando os fluxos de trabalho com sistemas inteligentes.

Embora a principal preocupação na Indústria 4.0 seja a automação, a Indústria 5.0 buscará obter maior sinergia entre humanos e máquinas autônomas.

Desta forma, a força de trabalho autônoma será perceptiva e informativa voltada para as intenções e os desejos humanos. Haverá aproximação da forma em que os humanos trabalharão ao lado de robôs, não apenas sem medo, mas também com paz de espírito como colegas de trabalho robóticos que os compreendem adequadamente e têm a capacidade de colaboração efetiva.

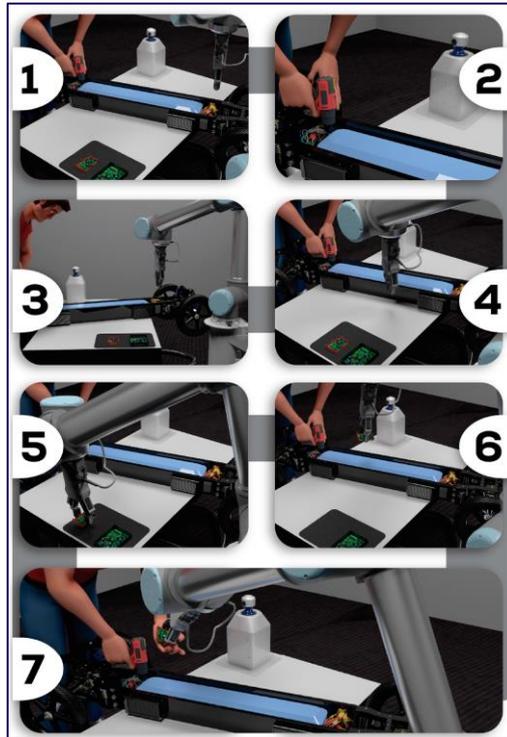
Isso resultará em processo de produção excepcionalmente eficiente de valor agregado, florescente, com autonomia confiável e redução do desperdício e equilíbrio nos custos associados.

Na Indústria 5.0 mudará a definição da palavra “robô”. Os robôs não serão mais apenas uma máquina programável para executar tarefas repetitivas, mas também se transformará em um companheiro humano ideal para alguns cenários.

A próxima geração de robô, comumente denominado de Cobots⁴ (Collaborative Robots) terá um aprendizado acelerado com critérios de segurança e risco compartilhando objetivos e expectativas com o operador humano.

O conceito da Indústria 5.0 pode ser visualizado usando um exemplo de linha de produção mostrado na figura a seguir onde os robôs trabalham com humanos para aumentar a eficiência da produção, não para substituir os trabalhadores.

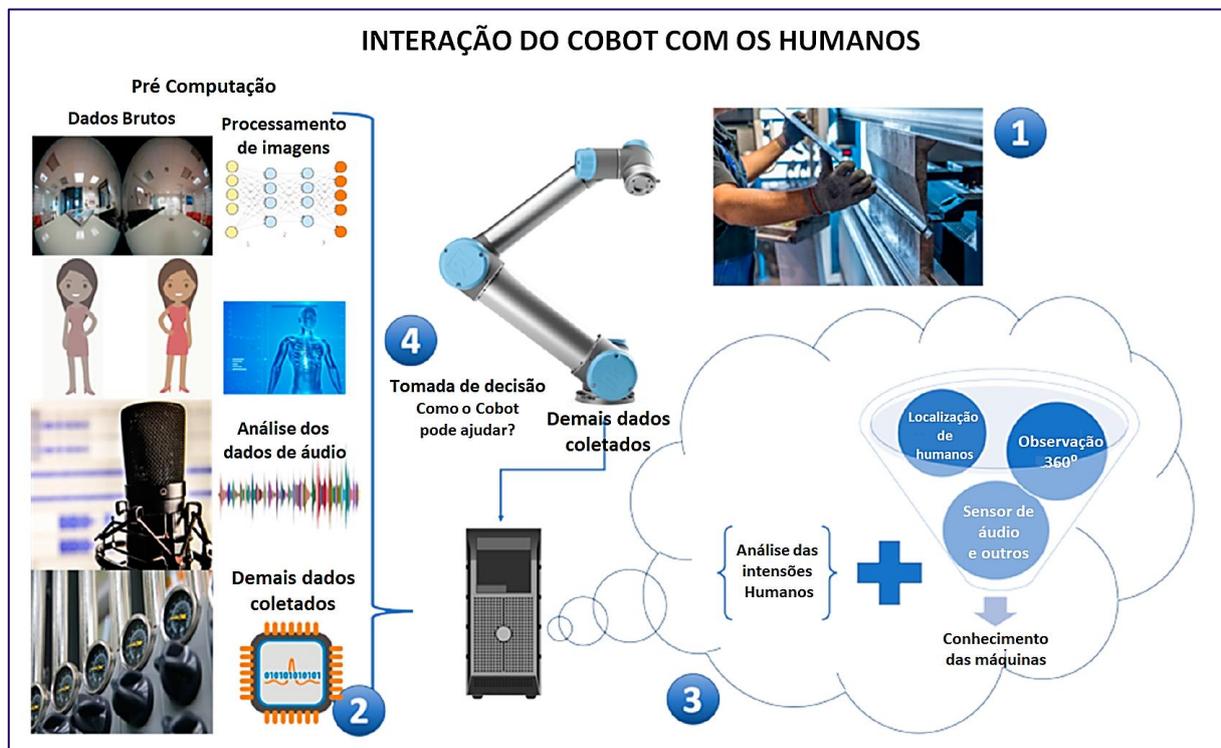
⁴ Robôs colaborativos destinados à interação direta com humanos em espaços compartilhados.



Mostra um trabalhador humano na montagem de uma máquina eletromecânica.

1. O humano inicia uma tarefa e o robô observa o processo com uma câmera que funciona como o seu olho que está conectado ao computador que captura, armazena e processa a imagem e aprende os padrões usando os recursos do aprendizado de máquina.
2. O robô observa os movimentos do humano, monitora o ambiente e infere a próxima ação do operador fará a seguir usando os recursos da análise intenção humana do Deep Learning. O sensor de espectroscopia ativa fNIRS(functional Near-Infrared Spectroscopy) para entender a intenção humana no canal de comunicação sem fio para recuperar sinais do cérebro humano. O fNIRS tem forma de um fone de ouvido e não requer procedimentos demorados para configuração e calibração.
3. Será como se o outro humano estivesse ajudando o trabalhador humano. Neste exemplo, o robô prevê que o operador humano usará uma determinada peça na próxima etapa do a tarefa.
4. Em seguida, ele vai buscar a peça com antecedência e a entrega ao ser humano quando necessário. O processo ocorre perfeitamente para que o operador humano não precise fazer nenhum ajuste em seu processo de trabalho.
5. O robô escolhe um objeto de interesse do trabalhador humano.
6. O robô trazendo o objeto para o trabalhador.
7. O robô entrega o objeto conforme a necessidade e aceite pelo trabalhador humano.

A CRO (Chief Robotics Officer) é uma entidade responsável por tomar decisões sobre máquinas ou robôs e suas interações com humanos a serem adicionados ou removidos do ambiente/chão de fábrica para atingir desempenho e eficiência ideais. CROs irão têm experiência em robótica, inteligência artificial, modelagem de fatores humanos e homem-máquina interação. Os CROs estão melhor equipados com tecnologias robóticas colaborativas e, aproveitando poder de avanços na computação, será colocado adequadamente para causar um impacto positivo no meio ambiente gestão também. Isso acabará aumentando a sustentabilidade da civilização humana, reduzindo poluição e geração de resíduos e preservação da Terra. Prevemos que a Indústria 5.0 criará uma nova função de manufatura: Chief Robotics Officer (CRO). Um CRO é um indivíduo com experiência em compreender robôs e suas interações com humanos.



Fonte: Nahavandi, 2019

Como a inteligência artificial é aplicada em robôs? Quais são os benefícios?

A inteligência artificial em robôs proporciona oportunidades de aumentar a produtividade, tornar o trabalho mais seguro e economizar tempo valioso para as pessoas.

Permite expandir a funcionalidade do robô sobretudo na detecção e na resposta ao seu ambiente ampliando a gama de execução das suas funções. A integração com os sistemas de informação móveis e interativos em vários ambientes, nos espaços públicos, hospitais e lojas de varejo, e economizando tempo aos públicos em geral.

- Mobilidade:** Embora os robôs sejam móveis por mais de 60 anos (o primeiro veículo automatizado guiado foi introduzido em 1953), a IA permite a mobilidade do robô em ambientes imprevisíveis. Os robôs móveis têm sido tradicionalmente programados para executar um conjunto específico de manobras de forma linear, guiados por sinais (magnéticos, laser, radiofrequência, IP) de dispositivos instalados para esta finalidade em seu ambiente. O robô móvel habilitado para IA vai de A para B construindo um mapa em tempo real (ou atualizando um mapa pré-programado em tempo real) de seu ambiente e de sua localização dentro desse ambiente, planejando um caminho para a meta programada, detectando obstáculos e replanejando um caminho in situ. Estão em uso no comércio, nas indústrias, prestação de serviços como: Coleta e transporte de mercadorias em fábricas, armazéns, hospitais, gestão de inventário (robôs móveis utilizando scanners RFID ou tecnologias de visão); serviços de limpeza - desde escritórios a grandes peças de equipamento, como cascos de navios; exploração de ambientes perigosos como exploração no mar profundo, espaços, ambientes contaminados entre outros.
- Detecção e resposta:** Tradicionalmente, os robôs são capazes de pegar objetos dentro de uma trajetória pré-programada na qual o objeto deve ser conhecido e no lugar esperado. Robôs equipados com sensores são programados usando algoritmos de inteligência artificial para identificar objetos específicos, independentemente de sua localização espacial. O software de visão 3D permite que o robô detecte objetos que estão ocultos por outros objetos. Por meio do aprendizado de máquina, o robô pode aprender em muito pouco tempo como pegar um objeto que não encontrou antes, aplicando o nível de força adequado e melhorando com a experiência. A tecnologia de coleta permite objetos que não sejam rígidos como mercadorias em embalagens plásticas ou materiais flexíveis, ou com formas irregulares e variáveis como frutas e vegetais incluindo a inspeção para detectar falhas em frutas e vegetais a vazamentos em dutos subaquáticos e subterrâneos.
- Otimização de Processo IA:** Para otimizar a precisão e a confiabilidade do robô analisando dados em tempo real para prever manutenção preventiva para evitem tempo de inatividade com a manutenção corretiva. O desempenho remoto com rastreamento de dados dos sensores como movimento e consumo de energia. O programa do robô pode ser ajustado automaticamente com base na saída do algoritmo de IA. Pois na tecnologia automação de manufatura em grande escala, os robôs são normalmente conectados a outras máquinas - incluindo outros robôs - e a IA é usada para otimizar todo o processo, analisando dados de todas as máquinas melhorando a velocidade e a precisão de todo o processo resultando em economia de custos. Em projetos de

- **Robôs móveis para prestação de informações:** Estão sendo usados como cabines de informações para auxiliar os clientes em ambientes como hotéis, hospitais, aeroportos, lojas e outros locais públicos como museus, zoológicos, parques. Podem responder a perguntas, conduzir os clientes aos produtos ou locais solicitados e podem conectar o cliente por vídeo a um agente de serviços humanos.

A Indústria 5.0 trará desafios sem precedentes no campo da interação homem-máquina (HMI) pois vai colocar as máquinas muito próximas da vida cotidiana de qualquer ser humano. Inovações dos dispositivos auxiliares programáveis e carros programáveis antes não considerados como uma versão de Cobots principalmente por causa da sua forma. Os Cobots serão muito diferentes, pois sua organização e introdução contêm funcionalidades semelhantes às humanas, como agarrar, beliscar e interagir com base na intenção e demais fatores ambientais.

Conforme relato da Australian Research Council e iTrust Lab. do Instituto para Pesquisa e Inovação de Sistemas Inteligentes (IISRI) na Deakin University, Austrália - a Quinta Revolução Industrial se viabiliza com a convergência dos dispositivos inteligentes, os sistemas inteligentes e a automação inteligente com o mundo físico em cooperação com a inteligência humana.

Ou seja, a abrangência do termo "automação" se amplia com os robôs autônomos como agentes inteligentes colaborando simultaneamente com os humanos no mesmo espaço de trabalho. A confiança e confiabilidade entre ambos alcançarão a eficiência promissora, a produção sem falhas com mínimo de desperdício e fabricação personalizada.

Desta forma espera-se que os humanos adicionem tarefas de alto valor à ao processo da manufatura com as políticas, padronização e a legalização para alinhar questões da tecnologia, dos negócios e a sociedade.

Surge desta forma novos desafios da fabricação rápida, processos transparentes e altamente eficiente que pode resultar em fenômeno da superprodução; princípios éticos a serem consideradas com a incorporação dos sistemas autônomos; surgimento de lacunas de habilidades essenciais que devem ser objeto de verificação e validação e as futuras gestões e as funções executivas que devem ser abordadas.

Isso será possível com a introdução de robôs em custos menores e altamente capazes, compostos por materiais avançados, como fibra de carbono e materiais leves, mas fortes, alimentados por baterias altamente otimizadas, proteção reforçados contra ataques cibernéticos e processos de manipulação de dados (big data e inteligência artificial) potentes e rede de sensores inteligentes. Indústria 5.0 aumentará a produtividade e eficiência operacional, ser ecologicamente correto, reduzir acidentes de trabalho e redução do ciclo e encurtar o tempo de produção. No entanto, ao contrário da

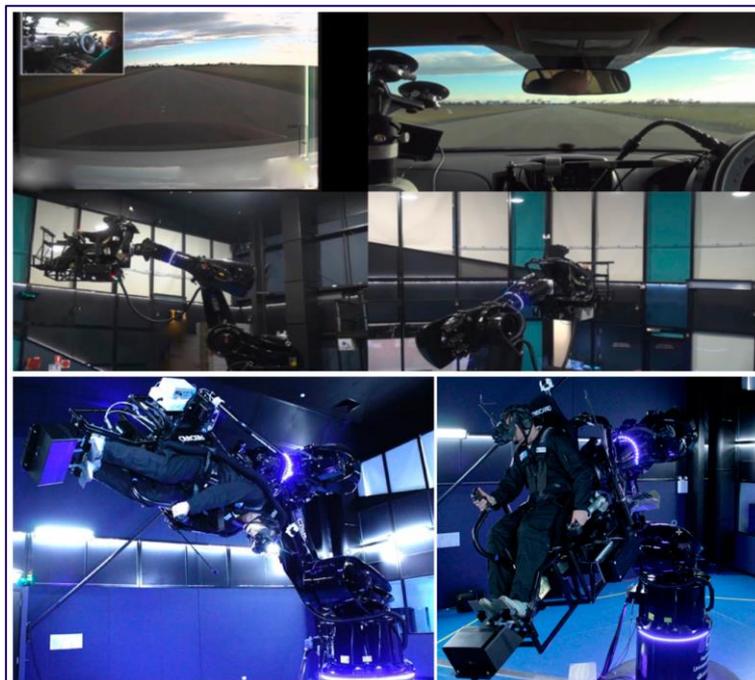
intuição imediata, a Indústria 5.0 criará mais empregos na área de sistemas inteligentes, IA e programação robótica, manutenção, treinamento, invenção e reaproveitamento (*retrofitting*) e da nova geração de robôs de fabricação. Além disso, a redução das tarefas repetitivas permite ao trabalhador humano envolver com criatividade e inovação no processo de trabalho e o uso de robôs de diferentes formas no local de trabalho.

Como um impacto direto da Indústria 5.0, um grande número de empresas iniciantes construirá um novo ecossistema de fornecimento de soluções robóticas personalizadas, em termos de hardware e software em todo o mundo. Isso impulsionará ainda mais a economia global e aumentará o fluxo de caixa em todo o mundo.

Treinamento Virtual

Trata-se do aprendizado simulado em ambientes virtuais em que o instrutor e o treinando em diferentes localidades aprende uma tarefa ou habilidade específica com redução significativa dos custos e o tempo para ambas as partes. Permite flexibilidade para ser atualizado e reconfigurado para novos cursos e conteúdo de treinamento para treinar motoristas, pilotos, bombeiros, médicos, e outros profissionais longe do perigo e dos riscos dos locais reais ou sem impor risco para os outros.

Simulador de movimento universal no Instituto de Pesquisa e Inovação de Sistemas Inteligentes-Deakin University.



ANALÍTICOS

A análise dos dados com base nos fatos históricos e de estrutura simples está se transformando com uso de testes de hipóteses em conjunto de dados complexos e heterogêneos e apoiados nos recursos dos algoritmos inteligentes com o propósito de elevar o nível de conhecimento e descoberta de *insights*.

Desta forma, os dados e os analíticos formam um espectro abrangente que cobre todo o ciclo de vida dos dados do passado, ao presente e o futuro abordando a análise explícita (conhecida), da compreensão reativa para implícita (desconhecidas) e a previsão com intervenção antecipada e proativa com a exploração dos dados (análises descritivas e preditivas), a entrega de ideias e decisões acionáveis (análises prescritivas) com o propósito da entrega do conhecimento (CAO et al. 2010).

O CICLO DOS DADOS PARA *INSIGHTS* E TOMADA DE DECISÕES

O conjunto de dados são transferidos para *insights* e processo de tomada de decisões em diferentes períodos e dos estágios analíticos são incorporados ao longo da análise do ciclo de vida. Esta operação pode ser mais representativa mediante a identificação em termos da variedade dos objetivos analíticos (O) e as abordagens (A) para atingir a meta do conjunto dos dados para a tomada de decisão.

- **Dados passados (históricos):** o foco principal é explorar "o que aconteceu?" com os dados e os negócios, e obter *insights* sobre "como e por que isso aconteceu?" por meio da modelagem computacional e design experimental. Neste estágio, "nós sabemos os fatos!" para obter a compreensão reativa do que aconteceu.
- **Dados presentes:** concentra-se em explorar "o que está acontecendo?" para gerar ideias sobre "como e por que isso acontece?". Nesta fase procura "entender o que não sabemos" com alertas gerados sobre eventos suspeitos ou grupos ou padrões apresentados nos dados e negócios. Os *insights* são extraídos para a tomada de decisão sobretudo no gerenciamento de risco e intervenção em tempo real para identificar "quais são os principais fatores impulsionadores?"
- **Dados futuros:** análise preditiva é para investigar "o que vai acontecer?" no futuro e obter *insights* sobre "como e por que isso pode acontecer?", estimando desta forma a ocorrência de eventos futuros com agrupamento e padrões dos dados. O objetivo desta etapa é "estimar

o que não sabemos” buscando a compreensão proativa e efetuando a previsão, predição e prevenção de forma precoce.

- **Decisão acionável:** análises prescritivas e entrega de conhecimento acionável são recursos para investigar "qual a melhor ação a ser tomada?" com a interpretação dos dados passados, presentes ou futuros. Isso proporciona *insights* sobre "qual é a próxima melhor ação?" e permite que as recomendações ideais sejam realizadas com base nas alternativas. O objetivo desta etapa é “como gerenciar de forma ativa e otimizada os problemas identificados?” mediante recomendações e intervenções acionáveis.



Fonte: Adaptado de Cao, 2017

EVOLUÇÃO DOS ANALÍTICOS EXPLÍCITOS PARA OS IMPLÍCITOS

Nas últimas décadas do século XX observou-se a transferência da análise de dados simples e limitados com testes de hipótese para análise dos dados complexos e de grande volume com hipóteses livres com base em conhecimentos e descobertas de *insights*.

A figura a seguir apresenta uma visão conceitual do espectro e evolução dos componentes analíticos e das tarefas organizadas em duas dimensões.



Fonte: Adaptado de Cao, 2017

- **Dimensão 1:** Nível de visibilidade, capacidade de automação e estado da arte: determina o nível de complexidade de dados cujas análises automatizadas são visíveis para os usuários.
- **Dimensão 2:** Grau de complexidades X, inteligência X e valor: abordagens e componentes analíticos típicos como relatórios com base na análise estatística, previsão e alertas, modelagem preditiva, otimização, análise prescritiva e entrega de conhecimento acionável (ações baseadas em *insights* para tomada de decisão nas operações de negócios)

Nos componentes e tarefas da evolução analítica podem ser categorizadas em duas eras sob a perspectiva do desenvolvimento disciplinar:

- **Era 1:** Análise explícita: concentra-se na análise descritiva com abordagens em relatórios, análises estatísticas, alertas e previsões.
- **Era 2:** Análise implícita: foco nas análises profundas com abordagens em modelagem preditiva, otimização, análise prescritiva e entrega de conhecimento acionável.

Observa-se que a figura indica o caminho unidirecional e linear da evolução da família analítica. Numa organização isso não indica o caminho da análise dentro de uma organização específica que conduz e utiliza análises. Também não indica um caminho linear de evolução analítica.

Frequentemente, um backend-quarta abordagem iterativa é feita em uma equipe de análise e análises múltiplas componentes podem estar envolvidos em paralelo para explorar observações multifacetadas e entendimentos.

A era da análise explícita: Análise descritiva

Elementos e tarefas típicas têm no passado, focava na análise descritiva explícita e tinha os seguintes recursos:

- **Goal:** sabemos o que sabemos e, portanto, visamos identificar e descrever a distribuição, geração e tendências de dados e problemas de negócios;
- **Natureza do problema:** semelhante a pessoas com visão reconhecem um elefante, sabemos o que é a ser analisado por abordagens baseadas em hipóteses, e para que fins;
- **Abordagem: análise** dirigida por domínio para a qual as hipóteses estão disponíveis em domínios específicos conhecimento e especialistas; a análise de dados testa tais hipóteses, e os dados verifica e explica as hipóteses;
- **O resultado:** métodos focados estão disponíveis em matemática e estatística, bem como da computação. Esses métodos descrevem e apresentam o que aconteceu, está acontecendo, ou acontecerá em dados geralmente pequenos ou altamente manipulados.

A era da análise implícita: Análise profunda

A limitação da análise explícita tem recentemente foi mais amplamente reconhecido na comunidade analítica, como no tratamento dados latentes, incertos e não IID [Cao 2014, 2015a]. Como resultado, o foco tem mudou recentemente para análises implícitas e para análises profundas. Ganhos analíticos profundos uma compreensão aprofundada de por que e como as coisas aconteceram, estão acontecendo ou acontecerá. Esses porquês e comos não podem ser tratados por análises descritivas e pode determinar a próxima melhor ou pior situação, bem como planejar a intervenção ideal estratégias.

- **Goal:** não sabemos o que não sabemos e, portanto, almejamos ganhar um, mas latente compreensão genuína de dados e problemas de negócios visíveis e invisíveis fontes;
- **Natureza do problema:** semelhante a pessoas cegas reconhecem um elefante, não sabemos o que deve ser analisado, ou mesmo por que e o que podemos obter;
- **Abordagem:** descoberta baseada em dados pela qual percepções interessantes, mas ocultas, são aprendidas de dados; dados criam uma visão invisível para nós e explica as razões

invisíveis ou indicadores, para complementar as hipóteses e observações baseadas no domínio;

- **Resultado:** o foco está em obter um entendimento profundo, intrínseco e completo de percepções invisíveis, conhecimento e sabedoria de dados, comportamentos e ambiente sobre o que aconteceu, está acontecendo ou acontecerá nos dados e nos negócios.

Mudança de paradigma de análise descritiva para preditiva para prescritiva

A mudança de paradigma da análise de dados para a ciência de dados constitui o chamado "novo paradigma" [Nelson 2009; Hey et al. 2009], ou seja, descoberta baseada em dados. A história de análise da perspectiva do espectro e dinâmica abrange duas eras principais da análise, conforme mostrado na Figura 3. As práticas analíticas viram uma mudança de paradigma significativa em três estágios principais: (1) Estágio 1: análise descritiva e relatórios, (2) Estágio 2:

No passado, o armazenamento e o processamento de dados limitaram o processo de análise limitada. Com a evolução da inteligência artificial e dos algoritmos de aprendizado de máquina (ML) e do aprendizado profundo (Deep Learning) permitem manipular grandes quantidades de dados em formatos multimídia.

Isso significa que deixamos de perguntar o que aconteceu (descritivos) e o que deve acontecer (preditivos) ao solicitar que as máquinas automatizem e aprendam por conta própria a partir dos dados - e até nos digam quais perguntas fazer.

De forma crescente, a maioria das organizações trata a análise como um ativo estratégico, e é central para muitas funções e habilidades funcionais com os aplicativos de processamento de linguagem natural. Os computadores usam a PNL para interpretar fala e texto. Os chatbots usam a PNL para responder a perguntas de atendimento ao cliente ou oferecer conselhos de investimento nas janelas de bate-papo online. Eles também podem oferecer sugestões por script para os funcionários do call center ao vivo.

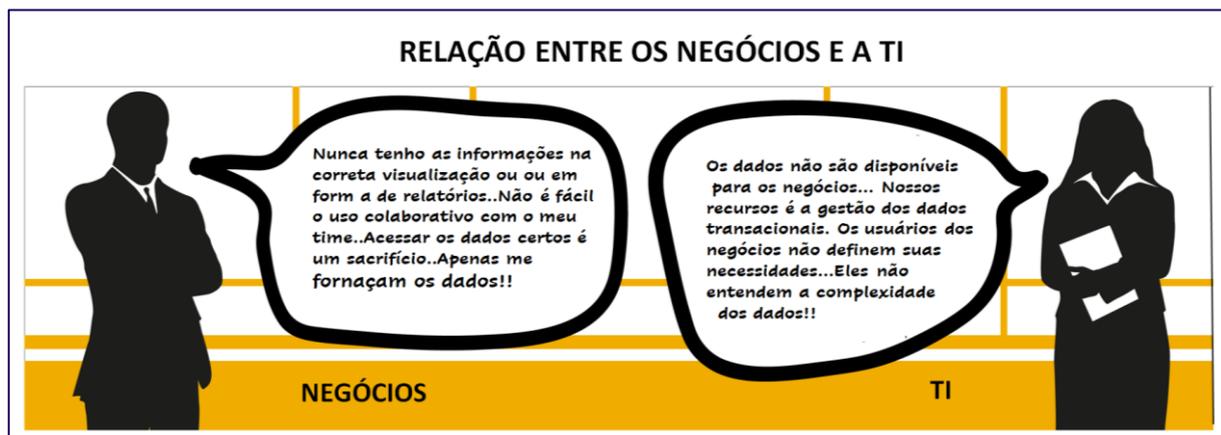
O aprendizado de máquina e a inteligência artificial também nos trouxeram aplicativos úteis, como carros autônomos e motores de recomendação, que prometem nos taxiar enquanto assistimos às próximas séries de TV recomendadas com base em nossos gostos.

Obviamente, a análise molda mais do que o nosso tempo de lazer. Com computadores mais rápidos e mais poderosos, há muitas oportunidades para o uso de análises e inteligência artificial no contexto empresarial. Seja determinando o risco de crédito, desenvolvendo novos medicamentos, encontrando maneiras mais eficientes de fornecer produtos e serviços, prevenindo fraudes, descobrindo ameaças cibernéticas ou retendo os clientes mais valiosos, a análise pode ajudá-lo a

entender o que impulsiona o sucesso da sua organização - e como isso é importante para o mundo em torno dele.

Como os Analíticos funcionam?

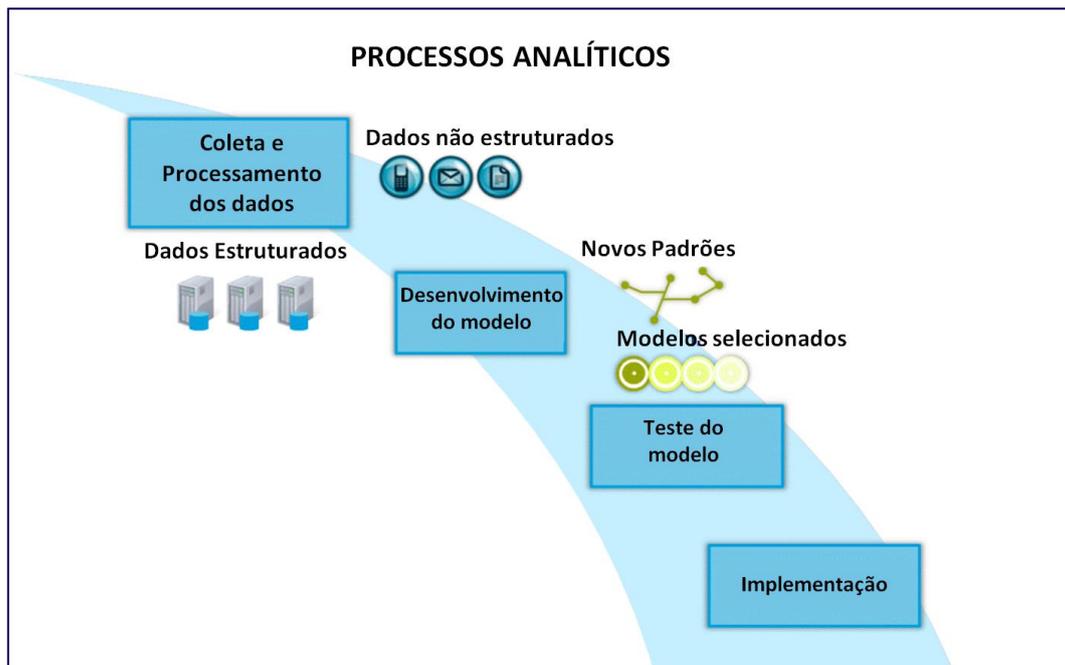
Toda a organização é uma empresa de análise.



Todo processo é um processo de análise pronto para melhorias. E todo funcionário pode ser um usuário de análise de alguma forma. Não importa o que se planeja realizar com a análise, mas o primeiro requisito para qualquer projeto de análise é o de dados. Pois, a partir dos dados é que se procedem as análises para impulsionar a tomada de decisão. Quanto mais rápidas as organizações puderem passar pelo ciclo de vida analítico, mais rapidamente obterão um valor tangível de seus investimentos em análise.

Ciclo de vida dos analíticos

Como etapa interativa do ciclo de vida dos analíticos existem três categorias – os dados, a descoberta e sua implantação, independentemente do escopo ou escala do seu projeto.



Fonte: Sark7, 2019

Dados

Os dados e as transações são onipresentes e quantificados via sistemas de informações e disponibilizados em aplicativos móveis e sociais. A quantificação dos dados ocorre a qualquer hora e em qualquer lugar por qualquer pessoa, de qualquer forma, de maneira não tradicional, sem limite de extensão, profundidade, variedade e velocidade.

O reconhecimento dos valores e potencial da ciência e análise de dados e seu rápido crescimento impulsionaram a promoção e a evolução da nova economia baseada em dados estimulando a transformação do modelo de negócios na indústria e serviços com inteligência artificial,

O surgimento de tecnologias avançadas na análise de negócios permite descobertas de correlações e padrões anteriormente ocultas com maior clareza e certeza em muitos negócios e gerenciar a complexidade crescente com melhor controle sobre as informações e prever as consequências pela redução da incerteza e na formulação de respostas rápidas e assertivas.

Tratamento de dados

Trata-se da etapa da exploração, visualização e construção de modelos com o algoritmo certo, pois geralmente é processo de tentativa e erro onde a identificação do algoritmo que terá o melhor desempenho requer experimentar várias abordagens. A escolha do algoritmo certo depende de vários fatores, incluindo o tamanho dos dados, as necessidades de negócios, o tempo de treinamento, os parâmetros, os pontos de dados entre outros.

De fato, é comum na fase de descoberta comparar diferentes modelos escritos em uma linguagem de programação diferente com diferentes recursos de dados.

Por exemplo, um projeto de análise recente que usou a detecção de objetos para identificar tumores em exames médicos de fígados começou com uma exploração de várias redes neurais e algumas semanas comparando e documentando os resultados de diferentes modelos.

Esse processo colaborativo funciona melhor quando cientistas de dados com diferentes conjuntos de habilidades podem escrever código no idioma de sua escolha e não programadores podem usar uma interface visual de apontar e clicar para explorar os resultados de diferentes abordagens analíticas.

Desdobramento, desenvolvimento e Implantação

Estima-se que a preparação de dados leve até 80% do tempo gasto em um projeto de analíticos. A plataforma de analíticos inteligente otimiza a preparação de dados com os mecanismos de acesso nativos, qualidade integrada e ferramentas de preparação de dados que automatizam com a IA sobretudo das tarefas demoradas.

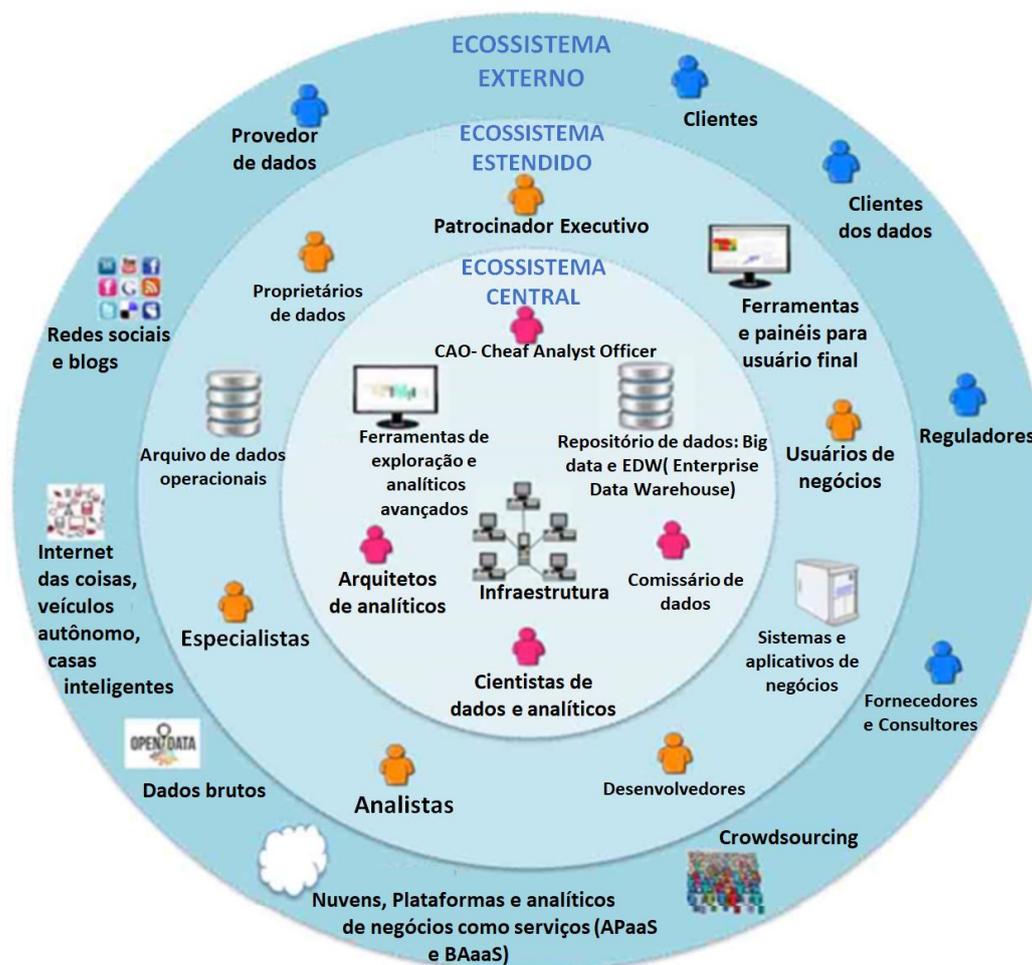
O gerenciamento de modelos fornece o controle de versão e ajuda a registrar, validar e gerenciar centralmente seus modelos. Ele ajuda a desenvolver procedimentos e regras para implantação e monitoramento de modelos com transparência sobre o uso de dados e modelos.

Por fim, a governança de dados garante que seus dados sejam confiáveis, pois, a fonte e o conteúdo são conhecidos e pode se assim monitorar a qualidade dos dados. A governança de dados também facilita a proteção de dados quando apropriado.

ECOSSISTEMA DOS ANALÍTICOS

Um ecossistema é uma rede de entidades interconectadas e interdependentes (veja a figura a seguir). O conhecimento de biologia, da ecologia e práticas de negócios fornecem muitos subsídios sobre como introduzir, sustentar e evoluir com a prática do ecossistema. Pois, o ecossistema de analíticos da big data contém indivíduos e grupos - equipes técnicas e de negócios com várias habilidades, os parceiros de negócios e clientes, dados internos e externos, ferramentas, software e infraestrutura de recursos. Além disso, uma organização pode ser visualizada num amplo ecossistema de dados que conectam com outras organizações e entidades que compartilham e trocam dados para gerar um valor econômico robusto.

ECOSSISTEMA DOS ANALÍTICOS



Adaptado de Fattah, 2014

Desta forma, a interação e a colaboração entre as esferas do ecossistema de análise da big data formada pelo ecossistema central, o estendido e externo definem o ecossistema de análise.

- **Ecossistema central:** indivíduos e tecnologias reúnem os dados necessários, analisam os dados para gerar *insights*¹ e determinam as ações para alcançar resultados nos negócios. Concentra as principais funções e tecnologias necessárias para introduzir e manter os recursos dos analíticos.
- **Ecossistema estendido:** indivíduos, grupos e sistemas que direcionam os projetos de análise, colaboram com a equipe central, fornecem os dados brutos, consomem os resultados e agem com base nas ideias.

¹ Capacidade de obter uma compreensão intuitiva, precisa e profunda de uma pessoa ou uma coisa.

- **Ecosistema externo:** clientes, parceiros de negócios, fornecedores, provedores de dados e consumidores que interagem com a organização para ajudar a fornecer todo o potencial das metas do *big data*.

O ecossistema central não implica necessariamente numa unidade organizacional única ou separada. Pode ser organizado em várias configurações, dependendo das necessidades da empresa: centralizada, descentralizada ou outras variações híbridas. O critério principal é a visibilidade do valor da análise e a capacidade de responder às necessidades em toda a organização.

O arranjo mais usual é o modelo Center of Excellence (CoE) composto por áreas comuns como treinamento, introdução de novas ferramentas, processos de inovação e comunicação entre as partes interessadas nos analíticos. Pode-se também visualizar resultados de negócios para atrair diferentes partes da organização a adotarem análises e incentivar a replicação de padrões comuns.

As principais funções do ecossistema analítico central e outras funções de apoio - como administrador e gerente de projeto podem ser compartilhados ou desempenhados por um único indivíduo de forma multifuncional.

Diretor de análise (CAO²)

O seu objetivo principal é mover a organização para novas descobertas e experimentos com os analíticos para tornar-se um aplicativo sistemático e abrangente em diferentes áreas e grupos com valor mensurável e visível para os negócios. O CAO pode personificar a mudança em direção a ações orientada aos dados em que os usuários de negócios procuram para obter *insights*. Assim o CAO lidera o ecossistema dos Analíticos e coordena os recursos para garantir que os analíticos sejam utilizados para fornecer os resultados comerciais desejáveis.

Cientista de dados e/ou analistas

O cientista de dados com conhecimentos matemáticos e estatísticos tem o papel de obter informações a partir dos dados e liberar o seu valor econômico. Utiliza as ferramentas e técnicas avançadas de análise e mineração de dados para descobrir padrões ocultos que podem ser explorados para alcançar os resultados desejados. No entanto, essas habilidades também podem ser adquiridas com a combinação de funcionários e aprimoradas por treinamentos e aumentada por consultores externos e modernas ferramentas de análise.

² CAO- Chef Analytics Officer

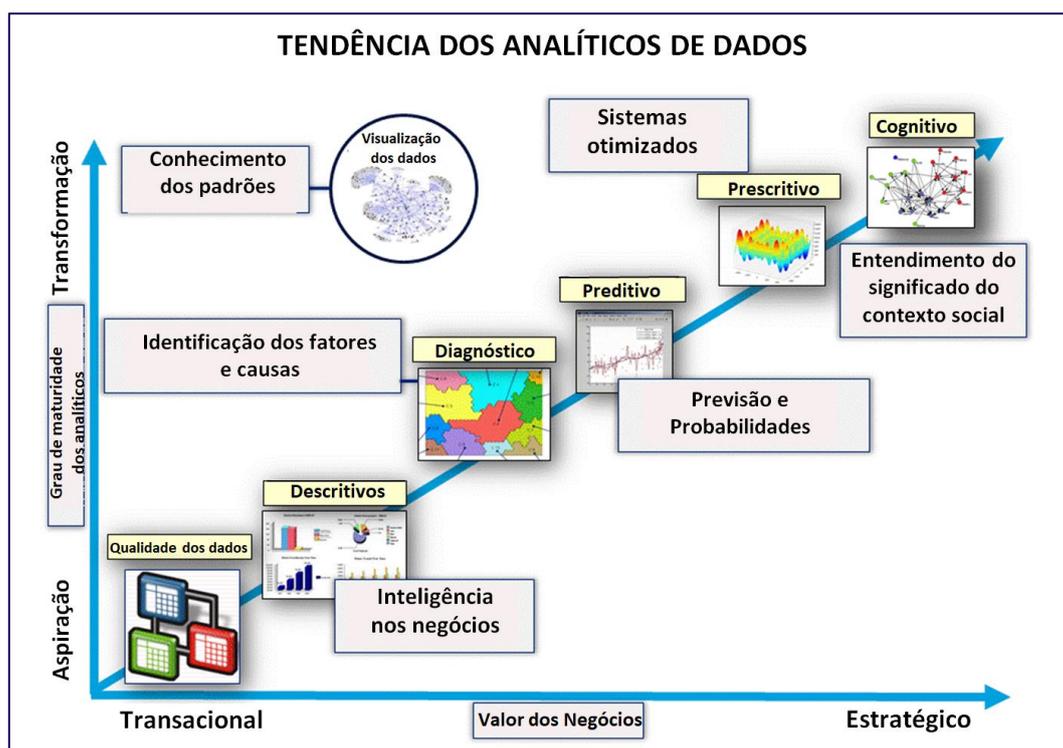
Ferramentas como a plataforma de análise preditiva da IBM® SPSS Modeler que podem executar automaticamente algumas das tarefas repetitivas e tediosas, mas necessárias, como preparação de dados e geração de modelos associativos.

Gestão das Informações - quais são as necessidades dos analíticos?

O mundo hoje requer os dados. Diante do enorme fluxo de dados sempre há novas inclusões/atualizações na organização por meio dos processos automatizados para sua adequada gestão.

As empresas cada vez mais baseadas em valor há necessidade analítica das informações como suporte para o processo de tomada de decisão.

Quanto mais complicada a estrutura dos dados e informações, mais difícil para entender e agir sobre os mesmos. Daí surge a importância dos analíticos que são classificados com base na sua complexidade e no valor agregado, pois os mais valiosos são também os mais complexos.



Fonte: Sark7, 2019

1. Analíticos Descritivos

Os analíticos descritivos respondem à pergunta: o que está acontecendo agora? Como foi o histórico?

Procura assim descrever um fato e são considerados de primeira geração

Tem o foco na descoberta: estatística clássica típica (geralmente inspirada na abordagem das ciências sociais) e na visualização de dados para reconhecer padrões e dar uma resposta ao que aconteceu no passado até os dias de hoje.

Proporciona muito conhecimento com base nos dados de referência e séries temporais que são sempre relevantes e muitas vezes disponíveis em tempo real que podem trazer informações valiosas sobre os negócios.

Os dados históricos são conhecidos, disponíveis e geralmente documentados, permitem a manipulação e visualização com mais facilidade.

No entanto, limita a visão futura para a tomada de decisões críticas dos negócios e sobre as ações preventivas/corretivas de forma oportuna e capazes de direcionar ao resultado desejado.

2. Analíticos Diagnósticos

São analíticos de segunda geração. Monitora os dados históricos e transacionais em tempo real para comparar com os resultados planejados e deduzir automaticamente o que está certo ou o que deu errado, e por quê? Grandes quantidades de dados são usadas para encontrar dependências, relacionamentos e identificar padrões para fornecer uma visão profunda de um problema específico.

Trata-se também de olhar para o futuro, fazer previsões e cenários: o que provavelmente acontecerá? Todos nós queremos saber!

Normalmente é utilizada para o diagnóstico da causa raiz para analisar e entender por que as coisas deram errado.

3. Analíticos Preditivos - Analíticos de terceira geração

Estes analíticos de terceira geração podem ser descritos como a área dos analíticos prescritivos, que fornecem o aconselhamento orientado a dados para a pergunta: o que devemos fazer agora? a decisão é certa?

Informa assim o que provavelmente acontecerá no futuro com base em todos os dados históricos disponíveis e acessáveis para identificar padrões definidos de eventos e prever os acontecimentos ou os resultados. Tem como base nos resultados dos analíticos descritivos e diagnósticos para detectar tendências, correlações, estimar os resultados e identificar as exceções para prever tendências futuras.

As ferramentas são os algoritmos avançados como modelos de previsão para visualizar as tendências futuras. É essencial entender que a previsão é apenas uma estimativa, cuja precisão depende muito da qualidade dos dados e da estabilidade da situação, por isso requer um tratamento cuidadoso e otimização contínua, examinadas nos capítulos anteriores.

Aumenta assim a responsabilidade em cada etapa, desde a estruturação dos dados relevantes até a discussão sobre os aconselhamentos orientados por dados, na tomada de decisão e as ações a serem direcionadas.

4. Analíticos Prescritivos

Os analíticos prescritivos de quarta geração são utilizados para prescrever literalmente a ação a ser tomada diante de uma questão.

Ele usa um vasto conjunto de dados e a inteligência para propor um conjunto de cenários e uma análise dos possíveis resultados para cada ação resultante de cada tomada de decisão.

São analíticos inteligentes que requerem além dos dados históricos, os dados operacionais, mas também de informações externas de especialistas humanos (especialistas), fontes de dados especializados e os algoritmos adequados para escolher a possível tomada decisão.

Os analíticos prescritivos utilizam ferramentas e tecnologias como aprendizado de máquina, regras de negócios e algoritmos, o que a torna mais sofisticada na implementação e gerenciamentos.

Por exemplo, as ferramentas IBM Runbook Automation ajudam as equipes de operações de TI a simplificar e automatizar tarefas repetitivas. Eles incluem procedimentos para todos os cenários previstos e geralmente usam árvores de decisão passo a passo para determinar a resposta efetiva para um cenário específico.

5. Analíticos Cognitivos

Na quinta geração, encontramos os analíticos cognitivos; os modelos de automatizados de autoaprendizagem funcionam como seres humanos.

Basicamente é inspirado no processamento das informações do cérebro humano que tira conclusões, codifica instintos e experiências para aprender, a entender não apenas as palavras de um texto, mas também o contexto completo do que está sendo escrito ou falado.

Combinam várias tecnologias inteligentes como inteligência artificial, algoritmos de aprendizado de máquina, aprendizado profundo para aplicar a inteligência semelhante ao cérebro humano na execução de tarefas e nos processos de tomada de decisão.

Desta forma:

Os analíticos ajudam a formar hipóteses. Melhora a qualidade das perguntas.

A estatística ajuda a testar hipóteses. Melhora a qualidade das respostas.



Fonte: Gartner, 2017

Como podemos observar, a Inteligência Artificial é um campo vasto, heterogêneo e dinâmico.

Assim como a Revolução Industrial ampliou a capacidade física do ser humano e a tornou possível construir produtos mais capazes e replicá-los em um ritmo mais rápido, a IA pode fazer o mesmo com as habilidades mentais das pessoas.

A partir do que fazemos as coisas hoje, tomaremos decisões baseado em todas as informações do mundo, ao invés de apenas a informação que por acaso tenho na minha cabeça.

A IA continuará a mudar os tipos de empregos que existem, mas não vai mudá-los tão drasticamente, porque em muitas dessas situações, a necessidade da intuição humana, empatia e o sentimento.

O conceito de aprendizado contínuo precisa ser incorporado à forma como as pessoas trabalham para que possam evoluir à medida que a tecnologia evolui.

Estamos aprendendo enquanto construímos a tecnologia, por isso é importante definir e ajustar os princípios éticos da IA desde o início que são inquestionavelmente boas ou inquestionavelmente ruins.

Pois, o aumento nas tecnologias de compartilhamento social para incluir conveniências para pessoas fora dos núcleos urbanos: não apenas bicicletas e carros compartilhados, mas recursos agrícolas, alimentos e até mesmo conhecimentos e saberes compartilhados inseridos no recurso do modo de vida que está longe de ser o único modelo ou mesmo o modelo economicamente mais eficiente do ponto de vista do cidadão.

Olhando para os próximos anos do que compõe a infraestrutura das cidades inteligentes como estradas, sistema de transporte e muitas construções com sensores em todos os lugares, do ruído à poluição, aplicações de segurança além da iluminação inteligente com a interação humana.

Então, à medida que abrimos essa visão, haverá evolução em todos os lugares e em todos os setores da economia e da sociedade.

Não temos capacidade para combinar tudo isso num sistema abrangente de cidade inteligente ainda.

A quinta geração da Internet é um ponto de virada muito interessante para as operadoras se conectarem praticamente tudo à rede inteligente.

A confluência de hardware e sistema de comunicações, mas potente, sensores melhores, avanços na aprendizagem das máquinas, e uma melhor conectividade nos permitirá fazer coisas que não foram possíveis até então. Vai permitir uma saúde muito melhor sobretudo com o rápido envelhecimento da população e como gerenciá-lo.

Muitos desses casos de uso são orientados a dados e computacionalmente intensivos com base em dispositivo que envia todos os dados para a nuvem com preocupações de privacidade e de alguma forma, requer controlar e a necessidade de regulamentação e de proteção e da transparência.

Outro fator está na IA pessoal no diálogo com um agente ou robô social que o conhece muito bem. Ao longo dos anos, pode agir como companheiro de aprendizado e torna-lo mais produtivo, um agente motivador quando você está entediado ou para baixo, ajuda organizar sua vida, na saúde e bem-estar, e proativamente defende e age em seu nome.

Com a IA pessoal, teremos nosso próprio perfil que armazena nossas preferências e conhece nossos estados emocionais e cognitivos longitudinais. Pode tornar o documento de identidade e de passaporte tecnológico conosco em cada interação tecnológica que temos em casa, no trabalho, em veículos autônomos, em jogo ou em qualquer outro espaço no mundo digital ou físico que habitamos de forma perceptiva e sensível.

A interface homem-máquina onipresente irá se manifestar e a interagir da mesma forma em que os humanos interagem entre si tornando as nossas experiências e a vida melhores e mais fácil. Esses sistemas e tecnologias serão perfeitamente integrados a nossa vida cotidiana, os dispositivos que usamos, os carros que dirigimos, as casas em que vivemos.

As tecnologias vestíveis com sensores em nossas roupas, sapatos e mochilas vai sentir o ambiente e medir a qualidade do ar que respiramos, para monitorar a nossa resposta.

Quando enchemos um copo com água, os sensores no vidro vão monitorar a qualidade da água, o nosso metabolismo, temperatura corporal, níveis de estresse e outros fatores de saúde indicadores serão constantemente medidos.

Pois, viveremos em uma era de saúde personalizada, onde uma combinação de nossa composição genética, exposições ambientais e estilo de vida serão usados para prever nossos resultados de saúde e nos guiam em direção a escolhas ideais. Se alguém é particularmente sensível à má qualidade do ar externo, você será orientado da forma que minimize exposição ambiental. Se alguém é sensível a certos minerais dissolvidos em sua água, será guiado para filtrá-lo.

A tecnologia nos ajudará a ser mais saudáveis e eficientes e reduzir os impactos negativos que temos no ambiente, mas nós, como seres humanos, também teremos que nos adaptar à natureza.

Por fim agradeço a Deus na conclusão desta obra.

AI

Loukianoff, G. 1936: “Une statue parlante ou oracle du dieu Ré-Harmakhis,” *Annales du Service des Antiquités de l’Égypte* 36, 187-93.

Hans-Dieter Wehle, Artificial Intelligence, ResearchGate, January 2019

Jason Griffey, Artificial Intelligence and Machine Learning in Libraries; Library Technology Reports, American Library Association, 2019

Michael Haenlein¹ and Andreas Kaplan, A Brief History of Artificial Intelligence: On the Past, Present and Future of Artificial Intelligence, California Management Review 2019

LEE, Kai-Fu, The Four Waves of AI, Fortune.com, nov. 2018

Joseph Byrum, Taking advantage of the AI revolution: An overview of what artificial intelligence can bring to society and businesses, ISE Magazine, June 2018

Sunil Kumar Srivastava, Artificial Intelligence: Way Forward for India, Journal of Information Systems and Technology Management, Vol. 15, 2018

Samuel Greengard, What is Artificial Intelligence?, Datamation, May 24, 2019

VARGHESE S. Jacob, JAMHS C. Moore, ANDREW B. Whinston Artificial Intelligence and the Management Science Practitioner: Rational Choice and Artificial Intelligence, Interfaces 18: 4 July-August 1988 (pp. 24-35)

Barr, A. and Feigenbaum, E., eds. 1981, The Handbook of Artificial Intelligence, Vol. 1, William Kaufman, Los Altos, California.

Bobrow, D. G. and Hayes, P J., eds. 1985, "Artificial intelligence — Where are we?" Artificial Intelligence, Vol. 25, No. 3, pp. 375-415.

Tom Culver, Lee Green, and Jim Redden, Peering into the Future of Intelligent Systems-Lessons from the SPRING Program- The SPRING strategic foresighting exercise provides a vision of the future of intelligent systems, Research-Technology Management • May—June 2019

Bishop, P., and Hines, A. 2012. Teaching about the Future. New York: Palgrave Macmillan.

Bishop, P., and Hines, A. 2007. Thinking about the Future: Guidelines for Strategic Foresight. Washington, DC: Social Technologies.

Gayle Sheppard , Intelligent Systems: The Big Picture, Research-Technology Management, July—August 2019

Groopman, J. 2007. *How Doctors Think*. New York: Houghton Mifflin.

V. Kumar, Bharath Rajan, Rajkumar Venkatesan, Jim Lecinski, Understanding the Role of Artificial Intelligence in Personalized Engagement Marketing, California Management Review, 2019, Vol. 61(4) 135–155

Kaplan, Andreas, and Michael Haenlein (2019), “Siri, Siri, in My Hand: Who’s the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence,” *Business Horizons*, 62 (1), 15–25.

Gianvito Lanzolla (Cass Business School, London), Annika Lorenz (U. of Utrecht, the Netherlands), Ella Miron-Spektor (Technion, Israel and Cass Business School, London), Melissa Schilling (NYU Stern School of Business, USA), Giulia Solinas (Ludwig-

Maximilian Universit’at, DE), Christopher Tucci (EPFL, Switzerland), Academy of Management Discoveries (AMD), “Digital Transformation: What Is New If Anything?”, 2018, Vol. 4, No. 3, 378–387.

RUGGIERO JR., MURRAY A., Artificial intelligence: A comeback story, *FUTURES*, October 2009

OECD DIGITAL ECONOMY PAPERS, VECTORS OF DIGITAL TRANSFORMATION, January 2019 No. 273, visit www.oecd.org/going-digital.

Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM(2018) 237 final.

Artificial Intelligence: A Modern Approach”, S. Russell and P. Norvig, Prentice Hall, 3rd edition, 2009.

THOMAS H. DAVENPORT AND RAJEEV RONANKI, Artificial Intelligence for the Real World, *HARVARD BUSINESS REVIEW* JANUARY–FEBRUARY 2018

Howard Gardner, *Frames of Mind: The Theory of Multiple Intelligences*, Basic Books, NY, 2011

MACHINE LEARNING

JAMES B. AIMONE, Neural Algorithms and Computing Beyond Moore’s Law, *Communications of the ACM*, APRIL 2019

Marvin Minsky and Seymour A. Papert, *Perceptrons: An Introduction to Computational Geometry* (Cambridge, MA: MIT Press, 1969).

Ted G. Lewis and Peter J. Denning, Learning Machine Learning, *Communications of the ACM*, December, 2018

Gabe Dickey, Sandra Blanke e Lloyd Seaton, Machine Learning in Auditing, *The CPA Journal*, June 2019

Nikki Castle, “Supervised vs. Unsupervised Machine Learning, *Datascience.com Blog*, July 13, 2017: <http://bit.ly/2VqkT8s>.

Tom M. Mitchell, *Machine Learning and the Study of Computer Algorithms*, McGraw-Hill Science/Engineering/Math; (March 1, 1997

WONG, WILLIAM, What’s the Difference Between Machine Learning Techniques? Machine learning is a hot topic, but what does this subset of artificial intelligence really mean?, *Electronic Design*, Jun2017

Domingos, P. 2015. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books.

Andriy Burkov, *The Hundred-Page Machine Learning Book (English Edition) [Print Replica]* eBook Kindle

John D. Kelleher, Brian Mac Namee, *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*, The MIT Press, 2018

John D. Kelleher, Brian Mac Namee, eBook, 2018

John D. Kelleher, Brian Mac Namee, *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*, The MIT Press, 2018

Mitchell, T. (1997). *Machine Learning*. McGraw Hill.

ALGORITHMS AND DATA SCIENCE

Longbing Cao, Data Science: Challenges and directions, *Communications of the ACM*, August 2017

ARTHUR S. JAGO, Algorithms and authenticity, *Academy of Management Discoveries*, 2019, Vol. 5, No. 1, 38–56.

FRANCINE BERMAN, ROB RUTENBAR, BRENT HAILPERN, HENRIK CHRISTENSEN, SUSAN DAVIDSON, DEBORAH ESTRIN, MICHAEL FRANKLIN, MARGARET MARTONOSI, PADMA RAGHAVAN, VICTORIA STODDEN, AND ALEXANDER S. SZALAY, Realizing the Potential of Data Science, *Communications of the ACM*, April 2018

Bahman MEHRI, From Al-Khwarizmi to Algorithm, *Olympiads in Informatics*, 2017, Vol. 11, Special Issue, 71–74

Algorithms Taiwo Oladipupo Ayodele, Types of Machine Learning Algorithms, University of Portsmouth United Kingdom, 2014

DEEP LEARNING

Roger Parloff, The Deep-Learning Revolution, *fortune.com*, October, 2016

BRENT KITCHENS, DAVID DOBOLYI, JINGJING LI, AHMED ABBASI, Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data, *Journal of Management Information Systems*, 2018,

BIG DATA

John E. Grable, Angela C. Lyons, An Introduction to Big Data, *Journal of Financial Service Professionals*, 2018

ANN

JEAN-CHRISTOPHE B., Loiseau, Rosenblatt's perceptron, the first modern neural network: A quick introduction to deep learning for beginners, *Towards Data Science*, 2019 [acesso em https://towardsdatascience.com/rosenblatts-perceptron-the-very-first-neural-network-37a3ec09038a](https://towardsdatascience.com/rosenblatts-perceptron-the-very-first-neural-network-37a3ec09038a)

Oleksii Kharkovyna, A Comprehensive Guide to Neural Networks for Beginners, *Towards Data Science*, 2019, <https://towardsdatascience.com/a-comprehensive-guide-on-neural-networks-for-beginners-a4ca07cee1b7>, acesso em out/2019

Lina M. KHAN, The Separation of Platforms and Commerce, *Columbia Law Review*, 2019

PLATFORMS

FENG ZHU , MARCO IANSITI , Why Some Platforms Thrive and Others Don't, Harvard Business Publishing, 2019

Arun Rai, Saonee Sarker, Next-Generation Digital Platforms: Toward Human–AI Hybrids, *MIS Quarterly Vol. 43 No. , March 2019*

Michael Blaschke, Kazem Haki, Stephan Aier, Robert Winter , Taxonomy of Digital Platforms: A Platform Architecture Perspective Conference Paper · February 2019

Geoffrey Parker, Marshall Van Alstyne, Xiaoyue Jiang, PLATFORM ECOSYSTEMS: HOW DEVELOPERS INVERT THE FIRM, *MIS Quarterly Vol. 41 No. 1*, pp. 255-266/March 2017

Wil Van der Aalst , Christof Weinhardt, Oliver Hinz , Big Digital Platforms: Growth, Impact, and Challenges, September 2019

Gawer, A. (2014). Bridging Differing Perspectives on Technological Platforms: Toward an Integrative Framework, *Research Policy* 43(7): 1239–1249.

Boudreau, K.J., and Hagiu, A. (2009). Platform Rules: Multi-sided Platforms as Regulators, In A. Gawer (Ed.), *Platforms, Markets and Innovation* (pp. 163–191), Cheltenham, UK: Edward Elgar Publishing Limited.

Mark de Reuver, Carsten Sørensen, Rahul C. Basole, The digital platform: a research agenda, *Journal of Information Technology*, 2017

Tiwana, A., and Konsynski, B. (2010). Complementarities Between Organizational

IT Architecture and Governance Structure, *Information Systems Research* 21(2): 288–304.

Tiwana A., Konsynski B., Bush A.A. (2010), Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental

Dynamics, *Information Systems Research*, 21, 4, pp. 675–687.

Evans D.S., Hagiu A., Schmalensee R. (2006), *Invisible Engines. How Software Platforms Drive Innovation and Transform Industries*, Cambridge, MIT University Press.

Boudreau, K.J. (2012). Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation, *Organization Science* 23(5): 1409–1427.

Tilson, D., Sørensen, C. and Lyytinen, K. (2012). Change and Control Paradoxes in Mobile Infrastructure Innovation: The Android and iOS Mobile Operating Systems Cases, In: 45th Hawaii International Conference on System Science (HICSS 45), Maui, HI.

Tiwana, A., Konsynski, B., and Bush, A.A. (2010). Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics, *Information Systems Research* 21(4): 675–687.

Ghazawneh, A., and Henfridsson, O. (2013). Balancing Platform Control and External Contribution in Third-Party Development: The Boundary Resources Model, *Information Systems Journal* 23(2): 173–192.

Kallinikos, J., Aaltonen, A., and Marton, A. (2013). The Ambivalent Ontology of Digital Artifacts, *MIS Quarterly* 37(2): 357–370.

Visconti, Roberto Moro, Digital Platforms Network Theory, Researchgate, November 2019

Yoo, Y., Henfridsson, O., and Lyytinen, K. (2010). The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research, *Information Systems Research* 21(4): 724–735.

Henfridsson, O., Mathiassen, L., and Svahn, F. (2014). Managing Technological Change in the Digital Age: The Role of Architectural Frames, *Journal of Information Technology* 29: 27–43.

Roberto Moro Visconti, Corporate Governance, Digital Platforms and Network Theory: Information and Risk-Return Sharing of Connected Stakeholders, Università Cattolica del Sacro Cuore, Milan, Italy

Odlyzko A., Tilly B. (2005), *A refutation of Metcalfe's law and a better estimate for the value of networks and network interconnections*, Minneapolis, University of Minnesota.

Moro Visconti R. (2020), *The Valuation of Digital Intangibles*, Palgrave-Macmillan.

Gupta A., Christie R., Manjula R. (2017), Scalability in internet of things: features, techniques and research challenges, *International journal of computational intelligence research*, 13, 7, pp. 1617-1627.

Bianconi G. (2018), *Multilayer networks*, Oxford, Oxford University Press.

Akbaripour H., Houshmand M., Valilai O.F. (2015), Cloud-Based Global Supply Chain: A Conceptual Model and Multilayer Architecture, *Journal of Manufacturing Science Engineering*, 137, 4, p. 040913.

Tsekeris T. (2017), Global value chains: Building blocks and network dynamics, *Physica A: Statistical Mechanics and its Applications*, 488, pp. 187-204.

ROBOTICA

Peter Hofmann, Caroline Samp, Nils Urbach, Robotic Process Automation, Finance and Information Management Research Center-The International Journal on Networked Business, 2019

Sutherland, C. (2013). *Framing a Constitution for Robotistan*. Vinay Kommera, Robotic process Automation, ResearchGate, · October 2019

Cristina-Claudia OSMAN, Robotic Process Automation: Lessons Learned from Case Studies, *Informatica Economică* vol. 23, no. 4/2019

IEEE Corporate Advisory Group. (2017). *IEEE Guide for Terms and Concepts in Intelligent Process Automation*.

IEEE Corporate Advisory Group. (2017). *IEEE Guide for Terms and Concepts in Intelligent Process Automation*.

Gartner, Robotic Process Automation (RPA), *IT Glossary*, <https://www.gartner.com/it-glossary/robotic-process-automation-rpa>, [Accessed: July 12, 2019].

Cristiano Premebida, Rares Ambrus and Zoltan-Csaba Marton, Intelligent Robotic Perception, Systems, IntechOpen, 2018

INTERNET

Artificial Neural Networks

Lao Fei, Wang Xiu yan, Study on the E-commerce Platform Based on Artificial Intelligence, Iberian Journal of Information Systems and Technologies, 2015

Nguyen G, Dang T T, Hluchy L. (2007). Agent platform evaluation and comparison.

Rapport technique, *Institute of Informatics, Bratislava, Slovakia*.43–49.

Jeng B, Lee W P., Chen D N, (2008). An agent-based model for consumer-to-business electronic commerce. *Expert Systems with Applications*, 34(1), 469–481.

Layth Abdulbari Al-Jaberi, Neural Networks-ARTIFICIAL NEURAL NETWORKS, researchgate, 2018

Mehdi Akhoondzadeh, Kamal Azizi , Artificial Neural Network (ANN), Mehdi Akhoondzadeh, University of Tehran, 2019
<https://www.researchgate.net/search.Search.html?type=publication&query=artificial%20NEURAL%20NET%20WORK%20concepts>

ANN

<https://data-flair.training/blogs/artificial-neural-networks-for-machine-learning/>

<https://medium.com/coinmonks/the-artificial-neural-networks-handbook-part-1-f9ceb0e376b4>

<http://www.easytipsntricks.com/machine-learning-aiartificial-intelligence-2019/>

<https://towardsdatascience.com/why-deep-learning-is-needed-over-traditional-machine-learning-1b6a99177063>

<https://www.edureka.co/blog/ai-vs-machine-learning-vs-deep-learning/>

<https://data-flair.training/blogs/machine-learning-tutorial/>

<https://www.bayesserver.com/docs/introduction/bayesian-networks>

<https://www.omicsonline.org/an-overview-of-health-analytics-2157-7420.1000132.php?aid=18916> health analytics

<https://decidesoluciones.es/en/lets-talk-about-advanced-analytics/> five types analytics

Algoritmos genéticos

Vijini Mallawaarachchi, Introduction to Genetic Algorithms — Including Example Code, Loves Bioinformatics, Data Science, 2017

Data visualization

https://www.sas.com/en_us/insights/analytics/what-is-analytics.html

https://www.sas.com/en_us/insights/big-data/data-visualization.html
<https://www.klipfolio.com/resources/articles/what-is-data-visualization>
<https://www.tableau.com/learn/articles/books-about-data-visualization>
<https://www.finereport.com/en/data-visualization/data-visualization-2.html>
<https://towardsdatascience.com/february-edition-data-visualization-18e1b663edc4>
<https://towardsdatascience.com/the-power-of-visualization-in-data-science-1995d56e4208>
<https://lawtomated.com/a-i-technical-machine-vs-deep-learning/>

ALGORITMO do caixeiro viajante

<https://medium.com/@aronmartins/k-nearest-neighbors-e-o-caixeiro-viajante-ed46f4926272>

ALGORITIMOS

<file:///C:/Users/Getulio/OneDrive/Downloads/azure-machine-learning-algorithm-cheat-sheet-nov2019.pdf>
<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-select-algorithms>
<https://www.researchgate.net/search.Search.html?type=publication&query=machine%20learning%20algorithms%20types>
<https://www.researchgate.net/search.Search.html?type=publication&query=machine%20learning%20algorithms%20classification>
<https://www.researchgate.net/search.Search.html?type=publication&query=machine%20learning%20algorithms>
<https://www.researchgate.net/search.Search.html?type=publication&query=machine%20learning%20algorithms>
<https://www.topbots.com/algorithmic-solutions-to-ai-bias-technical-guide/>
<https://www.geeksforgeeks.org/introduction-to-algorithms/>
<https://www.educba.com/what-is-an-algorithm/>
<https://becominghuman.ai/top-machine-learning-algorithms-you-should-know-to-become-a-data-scientist-17b16bc85077>
<https://www.bayesserver.com/docs/introduction/bayesian-networks>
<https://www.thebalancecareers.com/alison-doyle-2058389> Alison Doyle Analytical Skills
<https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>
<https://towardsdatascience.com/rosenblatts-perceptron-the-very-first-neural-network-37a3ec09038a>

<https://www.mathworks.com/help/stats/machine-learning-in-matlab.html#bvni0g6-1> Machine Learning in MATLAB

<https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861><https://www.xenonstack.com/blog/log-analytics-deep-machine-learning/> Deep Learning and AI

https://ogrisel.github.io/scikit-learn.org/sklearn-tutorial/tutorial/text_analytics/general_concepts.html Machine Learning 101: General Concepts

*<https://towardsdatascience.com/a-tour-of-the-top-10-algorithms-for-machine-learning-newbies-dde4edffae11>

<https://machinelearningmastery.com/start-here/#algorithms>

<https://data-flair.training/blogs/data-science-algorithms/>

*<https://intellipaat.com/blog/tutorial/data-science-tutorial/data-science-algorithms/>

<https://medium.com/ml-research-lab/machine-learning-algorithm-overview-5816a2e6303>

<https://medium.com/ml-research-lab/ensemble-learning-the-heart-of-machine-learning-b4f59a5f9777>

http://share.opsy.st/56e7090e92b6c-MathWorks_Figure+1_Machine+Learning+Types.jpg

<https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>

<https://www.educba.com/types-of-machine-learning-algorithms/>

Algoritmos Genéticos

<https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3>

Algoritmos Inteligentes

http://www.dainf.ct.utfpr.edu.br/~fabro/IA_I/busca/IA_Estrategias_Busca_Inf.pdf

Research Gate

<https://www.researchgate.net/search/Search.html?type=publication&query=>

Big Data

<https://intellipaat.com/blog/tutorial/hadoop-tutorial/big-data-overview/>

<https://www.bigdataframework.org/data-types-structured-vs-unstructured-data/>

<https://www.educba.com/what-is-big-data-analytics/> Kasheeka Goel

AI

<https://www.hisour.com/pt/automaton-43233/>

<https://www.analyticsinsight.net/understanding-three-types-of-artificial-intelligence/>

https://www.tutorialspoint.com/artificial_intelligence_with_python/artificial_intelligence_with_python_primer_concepts.htm

J. McCarthy, "What is Artificial Intelligence?," 2001. [Online]. Available: <http://lidecc.cs.uns.edu.ar/~grs/InteligenciaArtificial/whatisai.pdf>. [Accessed: 26-Apr-2016].

<http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>

Cetic.br- Regional Center of Studies on the Development of the Information Society <https://www.cetic.br>

Beijing AI principles

<https://www.baai.ac.cn/blog/beijing-ai-principles>

European Commission AI principles

<https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines>

US Department of Defense AI principles https://media.defense.gov/2019/oct/31/2002204458/-1/-1/0DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF

Google AI principles

<https://www.blog.google/technology/ai/ai-principles/>

Microsoft AI principles

<https://www.microsoft.com/en-us/ai/our-approach-to-ai>

OCDE AI principles

<https://www.oecd.org/going-digital/ai/principles/>

OECD Libraries

https://www.oecd-ilibrary.org/science-and-technology/oecd-digital-economy-papers_20716826

OECD, Broadband Portal,

www.oecd.org/sti/broadband/oecdbroadbandportal.htm

https://s3.amazonaws.com/MLMastery/MachineLearningAlgorithms.png?__s=6wwdud6ih59ewgmjamfz

<https://news.gotchamobi.com/category/artificial-intelligence/3232926/what-is-artificial-intelligence>

<https://www.researchgate.net/search.Search.html?type=publication&query=Artificial%20intelligence%20concepts>

<https://www.scnsoft.com/blog/artificial-intelligence-types>

<https://www.cisin.com/coffee-break/technology/7-types-of-artificial-intelligence-ai.html>

<https://data-flair.training/blogs/artificial-intelligence-ai-tutorial/>

Computer time line

https://www.atariarchives.org/deli/Time_Line.php

Data science

<https://www.smartdatacollective.com/unlocking-the-power-of-better-data-science-workflows/>

<http://sudeep.co/data-science/Understanding-the-Data-Science-Lifecycle/> <https://intellipaat.com/blog/what-is-data-science/>

https://algorithmxlab.com/blog/data-science-8-powerful-applications/#What_is_Data_Science_used_for

<https://www.topbots.com/top-ml-research-papers-2019/>

<https://medium.com/ml-research-lab/ensemble-learning-the-heart-of-machine-learning-b4f59a5f9777>

<https://azure.microsoft.com/en-us/overview/what-is-cloud-computing/#benefits>

DADOS

<https://lawtomated.com/structured-data-vs-unstructured-data-what-are-they-and-why-care/>

ML

<https://steelkiwi.com/blog/what-is-machine-learning/>

https://sebastianraschka.com/Articles/2014_intro_supervised_learning.html Predictive modeling, supervised machine learning, and pattern classification

<https://lawtomated.com/a-i-technical-machine-vs-deep-learning/>

https://americas.signiflow.com/features?gclid=EAIaIQobChMI9-rZhMCA5wIVAgYRCh1foA3UEAAYBCAAEgLUg_D_BwE

Data Analytics

<https://www.smartdatacollective.com/big-data-20-free-big-data-sources-everyone-should-know/>

<https://www.ibmbigdatahub.com/blog/going-beyond-data-science-toward-analytics-ecosystem-part-1>

<https://www.ibmbigdatahub.com/blog/going-beyond-data-science-toward-analytics-ecosystem-part-2>

<https://www.researchgate.net/search.Search.html?type=publication&query=Data%20Analytics%20type>

<https://www.ulster.ac.uk/cognitive-analytics-research/cognitive-analytics>

<https://www.thebalancecareers.com/analytical-skills-list-2063729>
Scott Allen Mongeau, Emerging Trends in Data Analytics, SAS Institute, 2014

Arun Kottolli, 4 Types of Data Analytics, Texas A&M University, 2018

<https://www.slideshare.net/ibmsverige/gene-villeneuve-moving-from-descriptive-to-cognitive-analytics>

<https://blogs.sap.com/2018/04/19/predictive-is-the-next-step-in-analytics-maturity-its-more-complicated-than-that/>

<https://www.architectureandgovernance.com/digital-transformation/examining-analytics-transformation-and-what-the-journey-looks-like/>

<https://www.linkedin.com/pulse/analytics-management-information-jacob-jensen>

<https://www.google.com.br/url?sa=i&url=https%3A%2F%2Fwww.linkedin.com%2Fpulse%2Fanalytics-management-information-jacob-jensen&psig=AOvVaw1k7TC5ytmODFCsfbPYvpwn&ust=1585689610768000&source=images&cd=vfe&ved=2ahUKEwi16oehkMPoAhVrMbkGHbcECiYQr4kDegQIARBxData Science>

<https://www.researchgate.net/search.Search.html?type=publication&query=Data%20Science>

Deep Learning

<https://www.researchgate.net/search.Search.html?type=publication&query=Deep%20Learning>

Expert Systems

<https://www.researchgate.net/search.Search.html?type=publication&query=concept%20of%20%20expert%20systems>

Machine Learning

<https://www.quora.com/What-is-the-difference-between-deep-learning-and-usual-machine-learning>

<https://www.researchgate.net/search.Search.html?type=publication&query=machine%20learning%20concepts>

Platform

<https://azure.microsoft.com/en-us/overview/what-is-machine-learning-platform/>

https://www.researchgate.net/publication/332114370_Digital_Platforms_Porosity_and_Panorama

<https://www.researchgate.net/search/publication>

<https://www.researchgate.net/search.Search.html?type=publication&query=Digital%20platforms>

<https://lawtomated.com/google-document-understanding-a-i-features-screenshots-and-use-cases/>

<https://www.innovationtactics.com/platform-business-model-complete-guide/>

https://www.researchgate.net/figure/Digital-platform-types-13_fig1_332164974
<file:///C:/Users/Getulio/Desktop/LIVROS%20PRELO/INTELIGENCIA%20ARTIFICIAL/AI%2028%2009/Platforms/Digital%20Platforms%20Network%20Theory.pdf>

<https://www.researchgate.net/search/publication?q=digital+platforms>

https://www.researchgate.net/publication/337368318_The_Evolution_of_Digital_Platforms

https://www.researchgate.net/publication/337186627_Digital_platform_ecosystems?_sg=QoH5HcH1Dy72LIRAs7fe5P3hl1dWAzn-J8B1rjJvqeRmRJCKmiCHfIEEi89rnPk5gyKKM_yttNzjG_w

https://www.researchgate.net/publication/335628797_Globalization_in_the_Age_of_Digital_Platforms

https://www.researchgate.net/publication/335690578_Data_Governance_in_Digital_Platforms

<http://beyondplm.com/2015/01/06/how-plm-can-build-itself-using-artificial-intelligence-technologies/>

RPA and ROBOTICS

<https://data-flair.training/blogs/ai-robot/http://cs.brown.edu/people/tdean/courses/cs148/02/architectures.html>

<https://riaricis1795.com/>

<https://www.intechopen.com/books/applications-of-mobile-robots/intelligent-robotic-perception-systems>

[https://related.forbes.com/search.cfm?&cq=pN5WQWJW1adKB5faaWsfBJfdLf3N1adLf5ivNU9&vsid=2272217016723565&vi=1584323302342941490&kp=3&kbc=&kt=65&ki=2282119&ktd=824650498560&kbc2=rps%3D0.01%7C%7Cr_c%3D0.03%7C%7Cp_r%3D0.26%7C%7Cc%3D-33%7C%7Crpc%3D0.26%7C%7Clvl%3D1.17&kmb=MC4wMTY%3D&fdkt=375&fp=G_IRe64-7i5EeY9qABvjqwjiWbTAeTCQkBki3pvAHRyQ7qB_QyHorowWR_1NQ_lpgOJRuWVRwtBFRlmOV5f2mCtxD_BRgt-6irXREHDUP07wc5K-0VL3g%3D%3D&c=pFfCC5TCCfjmRJrtOFAKQ&cme=OUWa_3mbJ_ZRyQ-K_ASS-JMZijl_dH3W1-JPplefTIKBjqT5S23Gir_FoPQecQF6Ib6zwnvuUAH9QUndIKO1VcPy6XydAdj3cxcUPsE7TSGOWGlrQdk38jkwZ3wvm8FXwDBWN3t8uMzsiKNDuQR7bOul94r7PbNbkMpj-mwt-bOE3VZUpfhyh9GOZ6Kep07h%7C%7CNDHRnZ9Gz3KXII-i9OnZqQ%3D%3D%7C5gDUJdTGiJzedmq9hanWYg%3D%3D%7CN7fu2vKt8_s%3D%7CtL2CLYCgZXcKYqcffJCzBPLAyaH041T5-z2wsoX9yfomiHsFJWDFUSEmKi1OEhC_aewOxnlteTgJAn9Xtla3F0FyLrATShV5CmjITPHOxvfuuhqM6YDaQ%3D%3D%7Cy2SqoJcE0s9nfXn920_qJvMroXhLPTa3-yxAj0rmay8%3D%7CzqOloebru98qS9QAC3XGFKb6Gwmjg8yHXObhBR68Pb4uljO2u7FKWN8_6rcxcYSDXvamq1BbA05yuhG4ZJw-AAhRIYZPp3hrYIMMQAoW8ZmTWHK7hQHhcYEUXU_yyA0IBAqX-XoQI36GIDNOU52WZvp9kUqN8DtL0gNdN6N8449mHHbAW4DyT1GHb9en36w9-1Xz6z2HhM9ODqhYkT4x9FtS7oZQ8KEi%7C&ib=0&cid=8CU2T3HV4&crid=147984832&size=800x200&lpid=&tsid=15032&ksu=224&chid=&https=0&kwdsMaxTm=400&ugd=4&pid=8PO7OCV8B&rms=1584323307&dym=1584323307793&rtbsd=6&vgd_isiols=1&vgd_hbReqId=T1584322610C8S2U727&kbbq=%26sde%3D1%26adepth%3D1%26ddepth%3D1&tdAdd\[\]=%7C%40%7Csde%3D1%7C%40%7Cadept%3D1%7C%40%7Cdepth%3D1&verid=121199&hvsid=00001584323306822031339672352997&upk=1584323305.21655&sttm=1584323306822&vgd_ifrmode=00&tdAdd\[\]=%7C%40%7Ccabp%3A1%3A2&tdAdd\[\]=%7C%40%7Ccabp%3A1%3A2&&sc=SP&oksu=224&asn=28573&ecref=w77EQ%3ASSBBB.yummyjJ.NmY.GLS&bd=-3%23640%231138&lgpl=eQ7L8O%3Affhffuh9uFhfAXFX%7CUQU%3Aazxjj%7CNkxO%3Af99ffF%7C8QM1YE%3A9%7CLJzQ8IJ%3AF9XMfh9%7CJN7%3AHy&kct=36675&abpl=2](https://related.forbes.com/search.cfm?&cq=pN5WQWJW1adKB5faaWsfBJfdLf3N1adLf5ivNU9&vsid=2272217016723565&vi=1584323302342941490&kp=3&kbc=&kt=65&ki=2282119&ktd=824650498560&kbc2=rps%3D0.01%7C%7Cr_c%3D0.03%7C%7Cp_r%3D0.26%7C%7Cc%3D-33%7C%7Crpc%3D0.26%7C%7Clvl%3D1.17&kmb=MC4wMTY%3D&fdkt=375&fp=G_IRe64-7i5EeY9qABvjqwjiWbTAeTCQkBki3pvAHRyQ7qB_QyHorowWR_1NQ_lpgOJRuWVRwtBFRlmOV5f2mCtxD_BRgt-6irXREHDUP07wc5K-0VL3g%3D%3D&c=pFfCC5TCCfjmRJrtOFAKQ&cme=OUWa_3mbJ_ZRyQ-K_ASS-JMZijl_dH3W1-JPplefTIKBjqT5S23Gir_FoPQecQF6Ib6zwnvuUAH9QUndIKO1VcPy6XydAdj3cxcUPsE7TSGOWGlrQdk38jkwZ3wvm8FXwDBWN3t8uMzsiKNDuQR7bOul94r7PbNbkMpj-mwt-bOE3VZUpfhyh9GOZ6Kep07h%7C%7CNDHRnZ9Gz3KXII-i9OnZqQ%3D%3D%7C5gDUJdTGiJzedmq9hanWYg%3D%3D%7CN7fu2vKt8_s%3D%7CtL2CLYCgZXcKYqcffJCzBPLAyaH041T5-z2wsoX9yfomiHsFJWDFUSEmKi1OEhC_aewOxnlteTgJAn9Xtla3F0FyLrATShV5CmjITPHOxvfuuhqM6YDaQ%3D%3D%7Cy2SqoJcE0s9nfXn920_qJvMroXhLPTa3-yxAj0rmay8%3D%7CzqOloebru98qS9QAC3XGFKb6Gwmjg8yHXObhBR68Pb4uljO2u7FKWN8_6rcxcYSDXvamq1BbA05yuhG4ZJw-AAhRIYZPp3hrYIMMQAoW8ZmTWHK7hQHhcYEUXU_yyA0IBAqX-XoQI36GIDNOU52WZvp9kUqN8DtL0gNdN6N8449mHHbAW4DyT1GHb9en36w9-1Xz6z2HhM9ODqhYkT4x9FtS7oZQ8KEi%7C&ib=0&cid=8CU2T3HV4&crid=147984832&size=800x200&lpid=&tsid=15032&ksu=224&chid=&https=0&kwdsMaxTm=400&ugd=4&pid=8PO7OCV8B&rms=1584323307&dym=1584323307793&rtbsd=6&vgd_isiols=1&vgd_hbReqId=T1584322610C8S2U727&kbbq=%26sde%3D1%26adepth%3D1%26ddepth%3D1&tdAdd[]=%7C%40%7Csde%3D1%7C%40%7Cadept%3D1%7C%40%7Cdepth%3D1&verid=121199&hvsid=00001584323306822031339672352997&upk=1584323305.21655&sttm=1584323306822&vgd_ifrmode=00&tdAdd[]=%7C%40%7Ccabp%3A1%3A2&tdAdd[]=%7C%40%7Ccabp%3A1%3A2&&sc=SP&oksu=224&asn=28573&ecref=w77EQ%3ASSBBB.yummyjJ.NmY.GLS&bd=-3%23640%231138&lgpl=eQ7L8O%3Affhffuh9uFhfAXFX%7CUQU%3Aazxjj%7CNkxO%3Af99ffF%7C8QM1YE%3A9%7CLJzQ8IJ%3AF9XMfh9%7CJN7%3AHy&kct=36675&abpl=2)

<https://medium.com/@montouche/robotics-in-the-era-of-artificial-intelligence-6747f0d9ede2>

<https://www.researchgate.net/search.Search.html?type=publication&query=concept%20of%20%20Robotic%20Process%20Automation>

<https://www.researchgate.net/search.Search.html?type=publication&query=concept%20of%20%20RPA>

<https://www.researchgate.net/search.Search.html?type=publication&query=concept%20of%20Robotic%20%20and%20RPA>

<https://www.researchgate.net/search.Search.html?type=publication&query=Robotic%20%20and%20RPA>

<https://www.researchgate.net/search.Search.html?type=publication&query=Robotic%20conceptualization%20and%20RPA>

<https://www.bloomberg.com/professional/blog/new-era-artificial-intelligence-nowbiggest-tech-disrupter/>

<https://www.ft.com/content/f04128de-c4a5-11e7-b2bb-322b2cb39656>

<https://www.danieljhatton.com/fun-facts-about-robotics/>

[https://www.economicsofai.com/nber-conference-toronto-2017/.](https://www.economicsofai.com/nber-conference-toronto-2017/)

<https://ifr.org/standardisation>

https://ifr.org/downloads/press2018/WR%20Industrial%20Robots%202019_Chapter_1.pdf

<https://www.dramandaforeman.com/the-wall-street-journal-7-march-2015/>

<https://www.iso.org/standards.html>

Python

https://www.tutorialspoint.com/artificial_intelligence_with_python/artificial_intelligence_with_python_getting_started.htm

<https://data-flair.training/blogs/python-tutorial/>

5G

<https://support.chinavasion.com/index.php?/Knowledgebase/Article/View/284/42/1g-2g-3g-4g---the-evolution-of-wireless-generations>

Charu, Rajeev Gupta, A Comparative Study of Various Generations in Mobile Technology, Researchgate, 2015

https://www.gta.ufrj.br/ensino/eel879/trabalhos_vf_2008_2/ricardo/2_1.html

Ilustração/ICONS

<https://www.vecteezy.com/vector-art/518305-ai-artificial-intelligence-technology-for-auto-and-transportation-icon-and-design-element>

https://www.freepik.com/free-vector/abstract-artificial-intelligence-background_1533310.htm

REALIZAÇÃO:

SEVEN
publicações acadêmicas

ACESSE NOSSO CATÁLOGO!



WWW.SEVENPUBLI.COM

CONECTANDO O **PESQUISADOR** E A **CIÊNCIA** EM UM SÓ CLIQUE.