Inteligência Artificial: Avanços e Tendências

Organizadores:
Fabio G. Cozman
Guilherme Ary Plonski
Hugo Neri

Instituto de Estudos Avançados da Universidade de São Paulo

Center for Artificial Intelligence

DOI 10.11606/9786587773131

Inteligência Artificial: Avanços e Tendências

Este trabalho foi realizado pelo Instituto de Estudos Avançados (IEA) da Universidade de São Paulo (USP) em parceria com o Centro de Inteligência Artificial (C4AI-USP), com apoio da Fundação de Apoio à Pesquisa do Estado de São Paulo (processo FAPESP #2019/07665-4) e da IBM Corporation.

USP

ie] Instituto de Estudos Avançados da Universidade de São Paulo

Center for Artificial Intelligence

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

Inteligência artificial [livro eletrônico] : avanços e tendências / organizadores Fabio G. Cozman, Guilherme Ary Plonski, Hugo Neri. -- São Paulo : Instituto de Estudos Avançados, 2021.
PDF

Vários autores.
Bibliografia.
ISBN 978-65-87773-13-1
DOI 10.11606/9786587773131

1. Aspectos morais e éticos 2. Inteligência artificial 3. Inteligência artificial - Aspectos sociais 4. Inteligência artificial - Inovações tecnológicas I. Cozman, Fabio G. II. Plonski, Guilherme Ary. III. Neri, Hugo.

21-75618 CDD-006.3

Índices para catálogo sistemático:

1. Inteligência artificial 006.3

Eliete Marques da Silva - Bibliotecária - CRB-8/9380

Organizadores:
Fabio G. Cozman
GuilhermeAry Plonski
Hugo Neri

Inteligência Artificial: Avanços e Tendências

DOI 10.11606/9786587773131

Ficha técnica

Autores

Alexandre Moreli
Anna Helena Reali Costa
Bruno Moreschi
Guilherme Ary Plonski
Didiana Prata
Elizabeth Nantes Cavalcante
Fabio G. Cozman
Fábio Meletti de Oliveira Barros
Fernando Martins
Giselle Beiguelman
Glauco Arbix
Hugo Neri
Jaime Simão Sichman
José Afonso Mazzon
Jose F. Rodrigues-Jr.
Juliana Abrusio
Juliano Maranhão
Leliane Nunes de Barros
Liang Zhao
Lucas Antônio Moscato
Marcelo Finger
Marco Almada
Maria Cristina Ferreira de Oliveira
Murillo Guimarães Carneiro
Nuno Fouto
Osvaldo N. Oliveira Jr.
Solange Oliveira Rezende
Teixeira Coelho
Thiago Christiano Silva
Veridiana Domingos Cordeiro

Organização

Fabio G. Cozman, Guilherme Ary Plonski e Hugo Neri

Preparação e Revisão

Nelson Barboza

Projeto gráfico e diagramação

Vinicius Marciano

Produção

Fernanda Cunha Rezende

Sumário

O futuro da ciência e tecnologia com as máquinas inteligentes

Jose F. Rodrigues-Jr.[1]
Maria Cristina Ferreira de Oliveira[2]
Osvaldo N. Oliveira Jr.[3]

Nas discussões sobre o futuro e os rumos da humanidade, é comum ouvir que vivemos na era do conhecimento. Na verdade, pode-se afirmar que praticamente todas as épocas da trajetória do *"homo sapiens"* no planeta Terra foram eras do conhecimento. Mesmo quando o predomínio de nações ou povos aparentemente se deu por fatores como poderio bélico ou abundância de recursos naturais, o conhecimento subjacente sempre foi preponderante para conseguir os recursos que permitiram o estabelecimento de impérios. Talvez a característica marcante dos tempos atuais, que aparentemente induz a nossa crença de que esta seria uma era especial, é a velocidade observada no progresso da ciência e da tecnologia, muito mais intensa agora do que em qualquer outra época.

Essa diferença de velocidade poder ser apreciada analisando-se os diferentes paradigmas de geração e transmissão do conhecimento. Considerando o período a partir do qual há registros históricos escritos, definem-se quatro paradigmas do conhecimento (Hey et al., 2009). O Primeiro Paradigma, registrado na Grécia Antiga, baseava-se em observações empíricas e modelos abstratos sobre a matéria e o Universo. Cerca de dois mil anos depois, um salto qualitativo deu origem ao Segundo Paradigma,

1 Professor livre-docente do Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo. junio@icmc.usp.br.
2 Professora titular no Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo. cristina@icmc.usp.br.
3 Professor titular do Instituto de Física de São Carlos da Universidade de São Paulo. chu@ifsc.usp.br.

com o trabalho de cientistas como Galileu Galilei e Isaac Newton, quando o conhecimento passou a ser gerado a partir de modelos teóricos que explicam resultados experimentais. A combinação de teoria e experimento, incorporada de maneira arraigada no método científico, gerou enormes avanços que culminaram com o decifrar da estrutura da matéria no início do século XX. Dentre os muitos produtos desses avanços está o computador, essencial para estabelecer o Terceiro Paradigma, em que a geração de novos conhecimentos se dá com simulações computacionais que complementam teoria, experimento, e observações empíricas. Em sequência, a alta capacidade de geração de dados, em simultaneidade à acentuada interconectividade que emerge das redes de computadores, fez surgir uma nova maneira de gerar conhecimento, associada ao Quarto Paradigma. É o que se conhece por *Big Data*, ou Ciência dos Dados, em que a meta é gerar informação e conhecimento a partir do processamento de grandes quantidades de dados diversos e dispersos.

O que se observa é uma enorme redução da escala de tempo decorrida entre um paradigma e o subsequente. Enquanto foram necessários dois mil anos para que o Primeiro Paradigma evoluísse e amadurecesse, definindo as bases do Segundo Paradigma, apenas alguns séculos se seguiram até o advento do terceiro e, mais recentemente, algumas décadas bastaram para que o Quarto Paradigma se estabelecesse já no final do século XX. Neste novo milênio, o Quinto Paradigma se avizinha; nessa realidade, o conhecimento novo poderá ser gerado por máquinas, sem intervenção humana. Esse cenário em que será possível interagir com sistemas inteligentes deve trazer profundas consequências para o futuro da humanidade. Espera-se que esses sistemas atuem para o benefício da sociedade e contribuam para um futuro melhor para as pessoas. Entretanto, é preciso pensar esse futuro, e se preparar para ele.

O capítulo está organizado da seguinte maneira. Na segunda seção, apresentamos um cenário com destaque para um novo paradigma de desenvolvimento científico e tecnológico, amparado em evoluções significativas na capacidade de processamento automatizado de dados estruturados e não estruturados. Nas terceira e quarta seções, discorremos sobre os movimentos de Big Data e Processamento de Linguagem Natural (PLN), apresentando uma contextualização para não especialistas nessas áreas. Exemplos da convergência desses movimentos são apresentados na quinta seção. Na sexta seção, discutimos algumas questões éticas associadas ao uso destas novas técnicas, seguindo-se às conclusões, na sétima seção.

Dois movimentos convergindo ao Quinto Paradigma

A nossa perspectiva é a de que a capacidade de criar sistemas inteligentes aptos a gerar conhecimento sem intervenção humana dependerá da convergência de dois grandes movimentos, o que procuramos ilustrar na Figura 1. No diagrama da figura, à esquerda, tem-se o movimento denominado Big Data, sobre o qual discorremos na terceira seção. *Grosso modo*, ele pode ser definido como um conjunto de processos que investigam grandes quantidades de dados com potencial para produção de conhecimento. Técnicas de Inteligência Artificial (IA), e particularmente Aprendizado de Máquina (AM), são utilizadas para transformar dados dispersos e variados em informação e conhecimento. Um aspecto importante a ser observado no diagrama é que os dados devem ser passíveis de processamento por máquinas (em inglês, *machine-readable*), em virtude das limitações atuais das técnicas de IA. Não se pode, por exemplo, esperar que o sistema computacional (a máquina) leia textos em língua natural, que são dados "não estruturados". À direita no diagrama está refletido o movimento complementar, que aborda justamente o problema de

ensinar sistemas computacionais a processar textos, associado à área de pesquisa da IA denominada Processamento de Linguagem Natural (PLN), cujos desafios são abordados na quarta seção. Do ponto de vista conceitual, a transformação de dados em informação e conhecimento é semelhante ao que ocorre no movimento identificado como Big Data. A diferença é que, agora, os dados são textos, falados ou escritos, que formam os corpora, ou as bases a partir das quais os sistemas devem ser capazes de aprender.

Na Figura 1, é importante a distinção entre as elipses que representam os processos de Aprendizado de Máquina, desenhadas em linha tracejada ou em linha contínua. De maneira genérica, a aplicação dos algoritmos de aprendizado pode ser categorizada em dois grandes tipos de tarefas: as de classificação e as que requerem interpretação (Wallach, 2018). Avanços recentes mostraram que o AM em tarefas de classificação de dados é capaz de apresentar resultados com desempenho expressivamente superior ao de seres humanos. Além dos avanços em algoritmos, são requisitos para tal desempenho a disponibilidade de uma quantidade de dados suficientemente grande e capacidade de processamento. Reconhecimento facial (Parkhi et al., 2015), diagnóstico a partir de análise de imagens (Litjens et al., 2017), e classificação de textos em cenários controlados (Zhang; Zhao; Lecun, 2015) são alguns exemplos ilustrativos. Representamos esse tipo de tarefa na figura por meio das elipses com o traço contínuo para enfatizar que a geração de conhecimento por uma máquina já é possível, caso se considere tão somente o universo de tarefas de classificação de dados. Por outro lado, tarefas que demandam interpretação, ou seja, requerem respostas a perguntas do tipo “Como?” ou “Por quê?” ainda estão longe de poderem ser realizadas por algoritmos de aprendizado de máquina com desempenho semelhante ao de humanos. Por isso, as elipses referentes a esse tipo de tarefa estão representadas por linhas tracejadas: elas ainda remetem ao futuro. Os desafios para que esse futuro venha a se tornar realidade são discutidos na sétima seção.

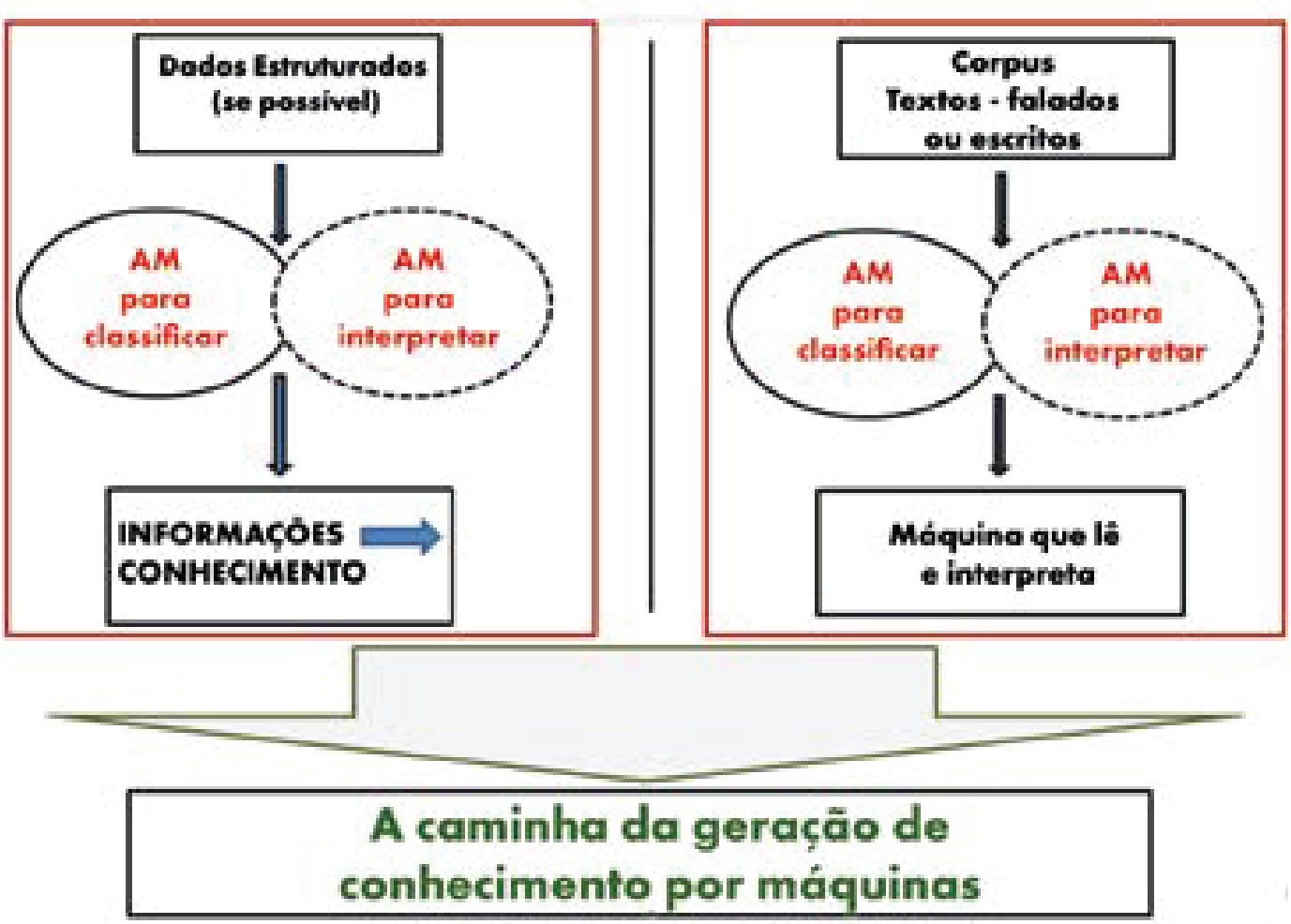

Figura 1 – Processos que conduzem ao Quinto Paradigma: à esquerda, o processamento massivo de informação por técnicas de Aprendizado de Máquina no contexto de Big Data; à direita, o processamento de informações não estruturadas por meio de Processamento de Linguagem Natural, o que amplia a quantidade de dados à disposição para a aquisição de novos conhecimentos.
Fonte: Elaborado pelos autores.

Algumas metáforas podem ilustrar a diferença entre os dois tipos de tarefa, de classificação de dados, ou de interpretação de dados. Podemos fazer uma analogia entre classificar e organizar agulhas de muitos tipos misturadas em um enorme palheiro (Wallach, 2018). Trata-se de uma tarefa que, ainda que não seja inerentemente complexa, é extremamente difícil para os recursos físicos e cognitivos de um ser humano. Entretanto, pode ser relativamente fácil para uma máquina, dado que o problema e a abordagem para resolvê-lo estão claramente especificados e a máquina não tem as limitações físicas ou cognitivas que dificultam a execução da tarefa pelo humano. Por outro lado, tentar explicar como a palha está organizada no palheiro, que é uma tarefa de interpretação, pode ser muito fácil para um humano, mas é difícil para um sistema computacional. Uma metáfora de natureza mais acadêmica diz respeito ao problema de fazer uma revisão sistemática da literatura em uma determinada área, o que, atualmente, pode ser realizada utilizando-se ferramentas computacionais. Já estão disponíveis ferramentas capazes de identificar os principais tópicos e suas conexões a um campo de pesquisa no qual exista uma vasta literatura científica de milhares de artigos (Silva et al., 2016), algo muito difícil para um pesquisador realizar, mesmo com o apoio de motores de busca sofisticados. Entretanto, ferramentas para revisão de literatura apenas conseguem classificar o conteúdo, sendo incapazes de capturar uma visão global e crítica da área a partir dessa organização, o que um pesquisador experiente na área consegue fazer. A ferramenta não consegue, por exemplo, interpretar os resultados e inferir os tópicos a serem abordados em um artigo de revisão da literatura, tarefa que faz parte da realidade de muitos pesquisadores. Observam-se, entretanto, avanços significativos nessa direção, um bom exemplo sendo a solução corporativa IBM Watson (IBM, 2019), bastante divulgada na mídia como capaz de processar milhões de arquivos de texto e produzir sumários, identificar atores, relacionamentos, palavras-chave, e papéis semânticos.

A revolução do Aprendizado de Máquina com Big Data

Quatro fatores têm se mostrado determinantes no avanço do tratamento computacional de problemas antes considerados inviáveis: (i) o salto na escala de capacidade de processamento numérico decorrente da popularização de unidades de processamento dedicadas, denominadas Graphics Processing Units (o termo Graphics deve-se a razões históricas); (ii) o aperfeiçoamento de técnicas computacionais relacionadas a Aprendizado de Máquina, incluindo técnicas de otimização e Redes Neurais Artificiais; (iii) a alta disponibilidade de arcabouços e de linguagens de programação capazes de alavancar a pesquisa e o desenvolvimento em Ciência da Computação; e, sobretudo, (iv) a explosão na produção e armazenamento de dados, no fenômeno identificado genericamente como Big Data. Impulsionado pelos avanços tecnológicos em sensores, aquisição, transmissão, e armazenamento de dados, o Big Data é um fenômeno com múltiplas denominações e que pode ser considerado segundo diversos enfoques. Do ponto de vista da indústria, diz respeito à produção de dados suficientemente intensa para inviabilizar a gestão e uso da informação por meio de recursos centralizados em uma única empresa ou instituição. Outro enfoque se refere à produção de dados de maneira contínua e em altíssima escala, não raramente em ordem planetária. Comum a essas concepções está o fato de que o tratamento de dados em escala demanda técnicas inovadoras de hardware e de software com o intuito de viabilizar a identificação de padrões, tendências, associações, e extração de modelos e conhecimento relativos à atividade humana em suas diversas modalidades.

A evolução do Big Data

Um exemplo prático de um cenário Big Data são os mapas digitais planetários. Mesmo em obras de ficção científica, jamais se imaginou que seria possível catalogar, fotografar e gerar

modelos tridimensionais de (quase) todas as localidades urbanas do planeta. Entretanto, já existem algumas instâncias comerciais bastante populares deste tipo de informação que vêm sendo continuamente aprimoradas como o Google Maps[4] e o Apple Maps.[5] Outro exemplo ilustrativo é a disponibilização de Prontuários Médicos Eletrônicos; diversas organizações e governos já atuam para armazenar o histórico clínico de milhões de pacientes em bilhões de consultas e procedimentos médicos. Esse conteúdo é objeto de pesquisa que, combinado a técnicas de Inteligência Artificial, pode propiciar avanços em medicina preventiva, prognóstico automatizado, previsão e interpretação de fenômenos epidêmicos, e diagnóstico auxiliado por computador. Alguns trabalhos relatam capacidade de detecção de câncer de mama com precisão na ordem de 99% (Liu et al., 2018), contra estimados 62% por parte de especialistas humanos. Não está no horizonte substituir médicos por algoritmos; no entanto, é bem possível que enfermidades mais frequentes venham a ser pré-diagnosticadas por um computador em um futuro não muito distante.

De modo mais geral, a Inteligência Artificial alimentada por dados em larga escala tem dado aos computadores a capacidade de executar tarefas antes exclusivas de especialistas humanos, como descrever o conteúdo de uma imagem, ou compor um texto. Especula-se que as máquinas não cheguem ao ponto de substituir os humanos em atividades de natureza mais complexa, como em gerência administrativa ou de ensino em sala de aula. Entretanto, é esperado que os humanos que fazem uso efetivo do auxílio computacional substituam aqueles que não o fazem (Brynjolfsson; Mcafee, 2017).

A Figura 2 ilustra que a crescente produção de dados tende a continuar à medida que mais empresas investem em tecnologias baseadas em Inteligência Artificial.

4 Disponível em: <https://www.google.com/maps>.

5 Disponível em: <https://www.apple.com/ios/maps/>.

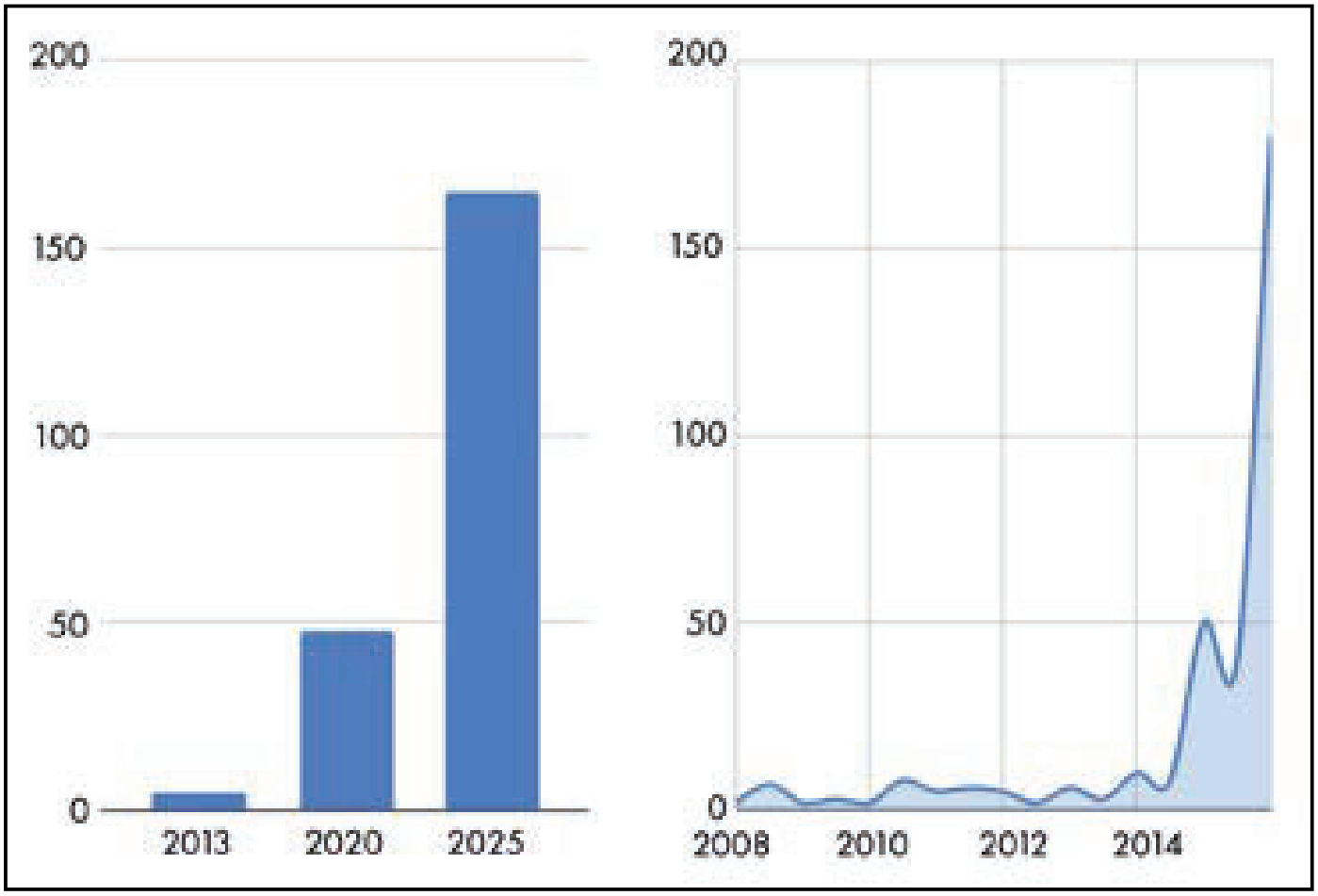

Figura 2 – À esquerda, projeção sobre a quantidade de dados que será produzida pela atividade humana, com dados IDC. À direita, número de empresas de tecnologia norte-americanas que anunciaram ganhos advindos de técnicas de Inteligência Artificial, dados Bloomberg. Fonte: Gráficos elaborados pelos autores.

A evolução do Aprendizado de Máquina

A capacidade de processamento da tecnologia computacional, hoje, está longe de ser comparável à do cérebro humano, como ilustrado na Figura 3. Um dos maiores experimentos em simulação cerebral (Furber et al., 2014), realizado na Universidade de Manchester, no Reino Unido, alcançou menos de 1% da capacidade do cérebro humano, apesar de demandar uma infraestrutura comparável à dos grandes *mainframes* da década de 1960.

Poderemos observar um salto enorme nessa capacidade de processamento se a computação quântica se tornar realidade, pois computadores quânticos resolveriam problemas em tempo logarítmico em comparação aos tradicionais. Entretanto, seu princípio computacional é radicalmente diferente do clássico computador eletrônico. Os fundamentos algorítmicos também mudam, de modo que até os problemas computacionais básicos precisarão ser reformulados (Ghosh et al., 2018). Pode-se prever que o uso de computadores quânticos também impulsione o Aprendizado de Máquina (Bahdanau; Cho; Bengio, 2014), ainda que estejamos longe de conhecer o universo de possibilidades de modo mais amplo (Biamonte et al., 2017). O ritmo dos avanços em Inteligência Artificial deve continuar acentuado na próxima década, em ritmo semelhante ao que já vem sendo observado no desenvolvimento tecnológico há mais de um século.

Um dos exemplos mais expressivos em Aprendizado de Máquina é o desafio ImageNet Large Scale Visual Recognition Challenge (ILSVRC).[6] Realizado anualmente, entre 2010 e 2017, o ILSVRC pede aos competidores que apresentem sistemas de classificação para classificar imagens de um conjunto contendo um milhão delas, separadas em mil classes. A partir de 2012, quando os desafiantes adotaram técnicas baseadas em Redes Neurais Artificiais profundas, a taxa de erro dos sistemas passou a cair de

6 Disponível em: <http://www.image-net.org/>.

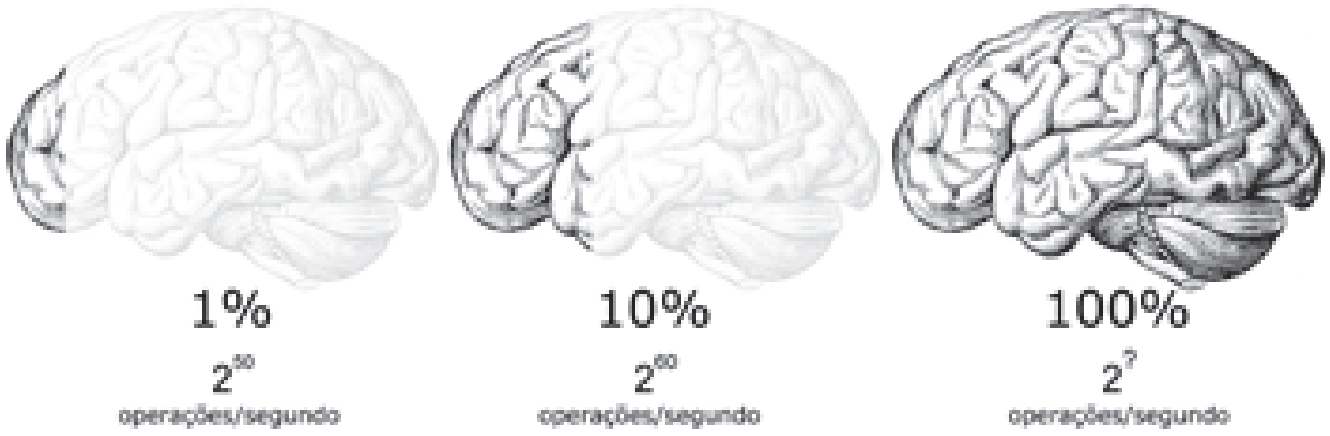

Figura 3 – Capacidade de simulação do cérebro humano considerando as tecnologias atuais. Um supercomputador atual com capacidade computacional em petaescala (2^{50} operações/s) consegue simular apenas 1% das conexões e processamento do cérebro humano; a próxima geração de supercomputadores com exoescala computacional (2^{60} operações/s) conseguirá 10%; ao passo que ainda não se sabe o quanto é necessário para reproduzir completamente um cérebro humano. Fonte: Jordan et al. (2018).

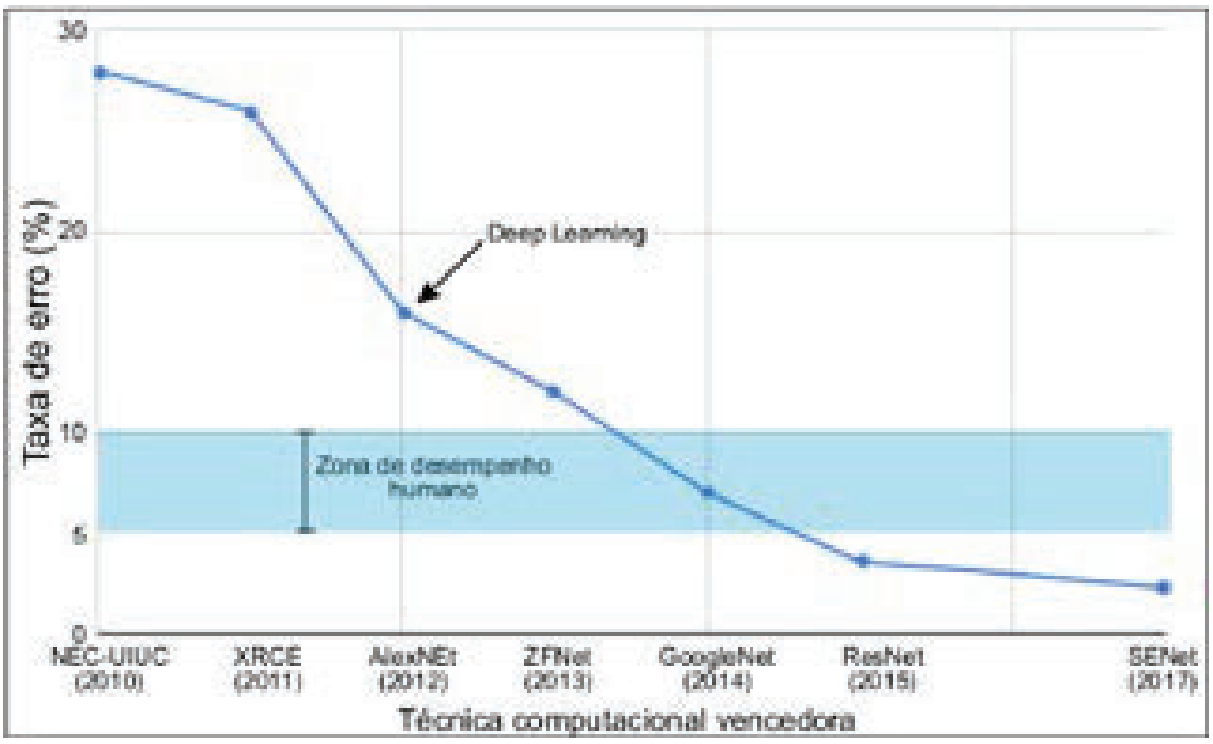

Figura 4 – Resultados do ILSVRC, realizado entre 2010 e 2017. A partir de 2012, os competidores começaram a usar técnicas de *Deep Learning*. Em 2015, o desempenho foi superior ao que é alcançado por seres humanos. Fonte: Elaborado pelos autores.

modo marcante (Figura 4). Em 2015 os resultados ultrapassaram a capacidade humana de classificação, com taxa de erro inferior a 5%. Esses progressos provocaram uma revolução nas áreas de Visão Computacional e Aprendizado de Máquina. Alguns dos trabalhos vencedores do desafio (Krizhevsky; Sutskever; Hinton, 2012; Szegedy et al., 2015; He et al., 2016) tornaram-se marcos científicos que impulsionaram o emprego de Redes Neurais Artificiais profundas, estratégia conhecida como *Deep Learning* (Goodfellow; Bengio; Courville, 2016).

Os pioneiros da área de *Deep Learning* Yoshua Bengio, Geoffrey Hinton, e Yann LeCun, alvo de crescente interesse e impacto como ilustrado na Figura 5, foram agraciados com o Prêmio Turing em 2019, considerado a mais alta distinção na área de Ciência da Computação (Lecun; Bengio; Hinton, 2015). Com efeito, os avanços iniciados em problemas de análise de imagens se estenderam para uma ampla gama de aplicações, como tradução de textos (Bahdanau; Cho; Bengio, 2014), reconhecimento de voz (Hannun et al., 2014), processamento de vídeo (Bertinetto et al., 2016), entre muitas outras (Deng et al., 2014).

A evolução do processamento computacional de linguagem natural

Já mencionamos que a Ciência de Dados é essencial para transformar dados em informação e conhecimento, e que as técnicas clássicas de Aprendizado de Máquina requerem dados de entrada estruturados. Esse requisito, que pode ser bastante restritivo, decorre do fato de que os algoritmos pressupõem que os dados de entrada satisfaçam a uma determinada estrutura. Em outras palavras, as máquinas não conseguem ler, pelo menos não no sentido estrito desse verbo, com a implicação que ainda não é possível para um sistema computacional interpretar texto apresentado em língua natural. Essa limitação é crucial, pois uma

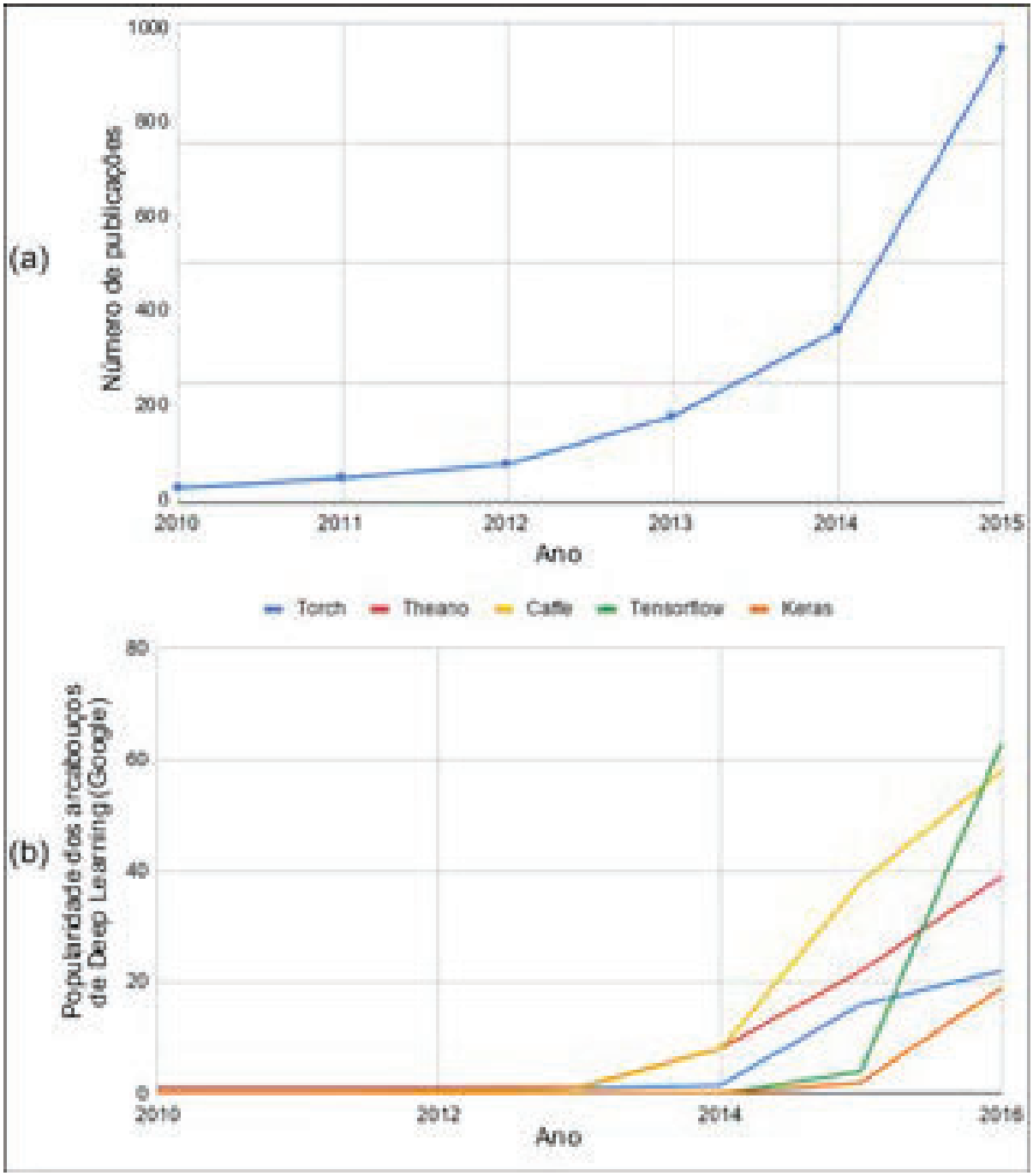

Figura 5 – Fatos sobre *Deep Learning* (Goh; Hodas; Vishnu, 2017). (a) O crescente número de publicações indexadas pelo International Scientific Indexing (ISI). (b) A popularidade (Google Trends Score) dos principais arcabouços de software para tarefas de *Deep Learning* atualmente: Torch, Theano, Caffe, TensorFlow e Keras. Fonte: Elaborado pelos autores.

grande parcela do Big Data é de dados embutidos em documentos textuais – como na literatura técnica, e em registros de patentes. O desafio da comunicação com máquinas em língua natural é enfrentado por pesquisadores em Processamento de Linguagem Natural (PLN). A área é tão antiga quanto o advento de computadores, pois uma das primeiras tarefas que se imaginou para um computador era justamente a de traduzir textos (Bar-Hillel, 1960). Até há pouco tempo, os resultados de tradução automática, assim como com outras tarefas de PLN, eram bastante limitados, o que por vezes gerou a impressão de que tradução automática de qualidade seria uma tarefa impossível. Houve uma transformação radical nesse cenário quando técnicas de Deep *Learning* passaram a ser empregadas para tradução (Wu et al., 2016). Para compreender a evolução da área de PLN e os potenciais impactos dessa evolução, cabe um breve histórico, apresentado a seguir.

As tarefas realizadas com PLN são de diversas naturezas, sendo as mais comuns: tradução automática, revisão ortográfica e gramatical, recuperação de informação, classificação de textos, sumarização automática, geração automática de texto, sistemas de perguntas e respostas, e busca "inteligente" em bases de dados (Shaalan; Hassanien; Tolba, 2017). Há também aplicações específicas para texto falado, como o reconhecimento de fala, tradução automática e assistentes virtuais (robôs) (Deng; Liu, 2018). Por décadas, duas estratégias predominaram na solução desses problemas, identificadas como abordagem simbólica e abordagem estatística. A abordagem simbólica, que foi utilizada em muitos revisores gramaticais (Martins et al., 1998), é baseada em regras explícitas; a abordagem estatística também é conhecida por outras nomenclaturas, como abordagem conexionista (Reilly; Sharkey, 2016); e, mais recentemente, abordagem baseada em *corpus* (Caseli; Nunes, 2004). Ela difere da abordagem simbólica por não se calcar na explicitação de regras, e sim na identificação de padrões frequentes que ocorrem nos textos escritos ou falados.

O sucesso dessa abordagem sempre se mostrou bastante dependente do domínio de aplicação, e por muito tempo acreditou-se que o ideal seria adotar abordagens híbridas, associando regras à análise de padrões estatísticos ou outros modelos.

O cenário mudou drasticamente a partir do fenômeno de Big Data, quando algoritmos de aprendizado de máquina puderam ser aplicados a grandes *corpora*, ou seja, a volumes massivos de textos (falados ou escritos). Hoje, as aplicações mais bem-sucedidas de PLN são todas apoiadas por Aprendizado de Máquina. Os exemplos mais marcantes, e com maior impacto na sociedade, são os tradutores automáticos e os assistentes pessoais com reconhecimento de fala. Praticamente não se aventa mais o uso de uma abordagem simbólica, embora a incorporação de algumas regras possa se mostrar benéfica para acelerar o aprendizado em tarefas específicas de PLN. A abordagem simbólica tem sido gradualmente substituída pela abordagem probabilística subjacente às Redes Neurais Artificiais, mais especificamente, as redes neurais recorrentes, capazes de aprender a partir de sequências de palavras (Klein et al., 2017).

A despeito do enorme progresso observado recentemente em PLN, ainda não é possível antever quando as máquinas serão capazes de efetivamente ler e escrever de maneira totalmente independente. Todavia, exemplos específicos ilustram a viabilidade de processamento sofisticado, ainda que não se possa considerar tal processamento como "interpretação" do texto. Mencionamos aqui dois exemplos, a título de ilustração.

O supercomputador Watson (Chen; Argentinis; Weber, 2016) foi projetado pela IBM para atuar no jogo de perguntas e respostas *Jeopardy* transmitido por um canal de televisão norte-americano. Nesse jogo, pessoas competem respondendo perguntas sobre conhecimentos gerais, em qualquer área do conhecimento. As perguntas são formuladas de maneira relativamente elaborada e as respostas são, geralmente, curtas. Essa característica favorece o

tratamento por um sistema computacional, pois o texto da pergunta, em geral, já fornece pistas para a resposta. O Watson conseguiu derrotar campeões do *Jeopardy*, um feito considerável tanto do ponto de vista de "interpretação" das perguntas como da habilidade de encontrar as respostas corretas. Certas características do Watson são impressionantes (Rennie, 2011): ocupava um espaço equivalente ao de 10 refrigeradores, sendo composto por 90 servidores com 256 GB de memória RAM, cada qual com 3.290 unidades de processamento, podendo armazenar cerca de 200 milhões de páginas. Ele consegue processar ("ler") 500 GB/s, o que seria equivalente a um milhão de livros em um segundo. Esse número é sugestivo da grande diferença entre máquina e ser humano: apesar da dificuldade em desenvolver sistemas computacionais com capacidade de interpretação de texto, a capacidade de processar rapidamente um volume massivo de material permite antever aplicações futuras que incorporem algum nível de interpretação.

A tecnologia do Watson foi, posteriormente, empregada em outras aplicações que demandam processamento intensivo de texto, como direito (Mills, 2016) e medicina (Ahmed et al., 2017). Apesar de evidentes limitações, os resultados do Watson demonstram a viabilidade de se realizar tarefas sofisticadas usando-se PLN. Também é digna de nota a sua versatilidade, pois a mesma tecnologia pode ser explorada em outros domínios e aplicações. O custo ainda é o principal limitante para a difusão desta tecnologia, pois não se justifica o uso de um supercomputador para viabilizar a execução de tarefas consideradas corriqueiras.

Outro exemplo de sistema inteligente que realiza atividades sofisticadas é o robô Todai, desenvolvido pela equipe da Profa. Noriko Arai, do Instituto Nacional de Informática do Japão (Arai; Matsuzaki, 2014). O Todai foi preparado para fazer os testes de múltipla escolha e as provas dissertativas dos exames de ingresso nas universidades japonesas, a partir de 2013. Alguns dos feitos são marcantes. O robô conseguiu ser classificado no estrato de 1%

dos estudantes mais bem-sucedidos nos exames de Matemática. Em outras disciplinas, especialmente as que requerem interpretação de texto, o desempenho cai consideravelmente, e fica significativamente abaixo dos ingressantes de universidades mais seletivas, como a Tokyo University. Ainda assim, o Todai seria aprovado em dois terços das universidades japonesas. Esse é mais um indicativo de que sistemas computacionais já conseguem executar tarefas sofisticadas, às vezes com desempenho superior ao de muitos humanos. A limitação, assim como no caso do Watson, é o custo de desenvolver uma ferramenta para executar uma tarefa bastante específica.

A convergência dos movimentos: dois exemplos

Seja na Medicina – para a automação do conhecimento médico – ou na Ciência de Materiais – na descoberta de novos materiais –, o aproveitamento de conjuntos massivos de dados, inclusive não estruturados (textuais), para inferir conhecimento, fornece as primeiras evidências de que o Quinto Paradigma se aproxima. Nessa nova realidade, torna-se possível prever propriedades e comportamentos sem que seja necessário construir a priori um modelo abstrato descrevendo as sutilezas de um dado fenômeno para que ele possa ser tratado computacionalmente. De outra maneira, no Quinto Paradigma, recorre-se à observação empírica automatizada de quantidades massivas de instâncias representativas do fenômeno de interesse, fazendo-se uso da alta capacidade algorítmica, e de memória e processamento dos computadores para obter controle sobre processos complexos. Trata-se de uma forma ainda rudimentar de conhecimento gerado pela máquina, mas que já representa um avanço significativo.

A seguir discutimos, a título de exemplo, dois cenários em que já é possível observar resultados que refletem o movimento convergente na direção ao Quinto Paradigma.

Caso 1: Medicina

A onipresença de exames clínicos de naturezas diversas, com sensores e biossensores na prática clínica, os recursos de monitoramento de pacientes, e a pesquisa farmacêutica sinalizam que o diagnóstico clínico no futuro deverá ocorrer no contexto de grandes quantidades de dados. Equipamentos de diagnóstico baseados em sensores permitem construir bancos de dados contendo anos de registros digitalizados de pacientes, registros de biomarcadores, relatórios de pesquisa em texto não estruturado, e farta literatura médica que demanda PLN. Ao possibilitar diagnósticos precoces, esses bancos de dados trazem o potencial de tratamento médico personalizado, induzindo a redução dos custos de assistência médica (Obermeyer; Emanuel, 2016). De fato, há um forte movimento global para o compartilhamento de dados hospitalares, como a iniciativa Observational Health Data Sciences and Informatics (OHDSI, or Odyssey, Hripcsak et al., 2015), que visa estabelecer uma comunidade de desenvolvedores, pesquisadores e, sobretudo, repositórios de dados médicos para fins analíticos.

Já os sensores e biossensores não se resumem à prática clínica profissional, seu uso já é uma realidade no autodiagnóstico de condições fisiológicas que requerem atenção. Auxiliados por aplicativos funcionando em telefones móveis, os sensores detectam condições das mais variadas, como alteração de pressão sanguínea, arritmia cardíaca, e crises de hiperglicemia, entre outras possibilidades (JR et al., 2016). Inserindo nesse cenário a alta disponibilidade de redes de transferência de dados, vislumbra-se a disseminação da telemedicina, antecipada há anos (Grigsby; Sanders, 1998), mas que ainda enfrenta obstáculos (Kahn et al., 2015). Um impacto mais óbvio na medicina já pode ser observado. Serviços globais de busca de informação usam ferramentas de indexação de dados munidas de técnicas de PLN capazes de traduzir, compilar, e relacionar dados textuais. A consequência

é uma farta disponibilidade de informações relativas à pesquisa, diagnósticos, tratamentos, e ação de medicamentos, fazendo com que muitos pacientes se apresentem ao médico com uma perspectiva inicial sobre sua condição e sobre como deverá ser seu tratamento. Embora esta prática receba críticas de profissionais da saúde, é uma tendência irreversível. E que tem um efeito colateral benéfico de desafiar médicos e profissionais da saúde a proverem informações mais valiosas e precisas do que as que os pacientes já são capazes de obter por si próprios.

Caso 2: Ciência dos Materiais: descoberta de novos materiais

As técnicas mais recentes de Aprendizado de Máquina têm sido aplicadas aos mais variados domínios, e a área de ciência e engenharia de materiais é ilustrativa do grande potencial a ser explorado. As Redes Neurais Artificiais, em particular, têm permitido identificar materiais ainda desconhecidos que apresentem propriedades desejadas para aplicações específicas. A metodologia considera o uso associado de conjuntos de dados que descrevem a composição e as informações físico-químicas de um material, bem como de suas propriedades elétricas, térmicas, oxidativas, reativas, tóxicas, e de ligação química, entre outras. Esses dados só estão disponíveis em quantidade suficiente para viabilizar esse tipo de abordagem em razão dos avanços em técnicas de PLN, hoje capazes de extrair informações diretamente de textos da literatura científica, bancos de patentes, e relatórios técnicos (Banville, 2006; Krallinger et al., 2017). Ao serem apresentadas a várias instâncias dessa informação relativas a um conjunto relativamente grande de compostos conhecidos, as técnicas de aprendizado conseguem extrair padrões físicos subjacentes à matéria. A partir daí, pode-se prever as propriedades de um novo composto, sem que seja necessário sintetizá-lo e experimentá-lo, acelerando enormemente o processo de descoberta (JR et al., 2019).

Outra modalidade de Aprendizado de Máquina, os denominados algoritmos genéticos (Whitley; Sutton, 2012), inspi-

rados nas ideias de Charles Darwin sobre a Teoria da Evolução, imitam o princípio "sobrevivência do mais apto" para estabelecer um procedimento de otimização. Nos algoritmos genéticos, cada característica composicional ou estrutural de uma molécula é interpretada como um gene. O genoma refere-se ao conjunto de todos os genes em um composto, enquanto as propriedades resultantes de um genoma são denominadas de fenótipo. Exemplos de genes químicos incluem fração dos componentes constituintes em um determinado material, tamanhos de blocos de polímeros, composições de monômeros, e temperatura de processamento. A tarefa de um algoritmo genético é varrer o espaço de busca definido pelos domínios dos genes para identificar os fenótipos mais adequados, medidos por alguma função de adequação (fitness function). Nesses algoritmos, compostos funcionais conhecidos são cruzados junto com um fator de mutação para produzir novos compostos; o fator de mutação introduz novas propriedades (genes) nas mutações. Novos compostos sem propriedades úteis são desconsiderados, enquanto aqueles que exibem maior aptidão são selecionados para produzir novas combinações. Após um certo número de gerações (ou iterações), novos compostos funcionais emergem com propriedades herdadas de seus ancestrais, complementadas com outras propriedades adquiridas ao longo de sua via de mutação. Esta é uma descrição simplificada do processo, que depende de modelagem precisa dos compostos, da definição adequada do procedimento de mutação, e de uma avaliação robusta da propriedade desejada. A avaliação de uma dada propriedade pode ser feita por meio de cálculos, como no caso de condutividade ou rigidez, reduzindo-se a necessidade de experimentos. Para uma revisão abrangente, com foco na ciência dos materiais, remetemos o leitor ao trabalho de Paszkowicz (2009).

Implicações: aspectos éticos

Sistemas computacionais dotados de algum tipo de "inteligência" têm sido empregados em diversas atividades. Em muitas situações eles já determinam, de maneira direta ou indireta, quem será contratado ou demitido, quem conseguirá um empréstimo, quanto custará o plano de saúde de um indivíduo, ou mesmo se um indivíduo representa um perigo para a sociedade. Na China, um gigantesco sistema de vigilância denominado Xue Liang ("Olhos Afiados") utiliza reconhecimento facial para monitorar a atividade cotidiana de pessoas em aeroportos, bancos, hotéis, e até em banheiros públicos. O objetivo é rastrear suspeitos e comportamentos perigosos com o intuito de coibir crimes. Todavia, o sistema já levanta desconfianças de seletividade étnica e coerção relacionada à maneira como os cidadãos pensam (Leibold, 2019), e um projeto de expansão do monitoramento poderá avançar para o monitoramento de chamadas telefônicas, hábitos de crédito, e até mesmo coleta não autorizada de DNA (Qiang, 2019). Em princípio, o sistema chinês não chega a causar temor na sociedade em geral, haja vista que o país limita, oficialmente, as liberdades individuais. Ainda assim, ele ilustra tecnologias e práticas passíveis de se difundirem por todo o mundo, caso condições políticas favoreçam esse tipo de atividade governamental.

Um problema mais universal, tema de investigação acadêmica, são os chamados "Algoritmos Tendenciosos" (tradução livre do termo em inglês "Algorithmic Bias"). O termo refere-se a algoritmos que, não necessariamente de maneira intencional, resultam em práticas discriminatórias ou de exclusão de determinados indivíduos. Não é difícil entender a origem do problema. Uma parcela significativa dos algoritmos de Aprendizado de Máquina, como já discutido, aprende a partir de conjuntos de dados com exemplos ilustrativos do conceito a ser aprendido. Caso não sejam "alimentados" com uma diversidade adequada de dados,

os exemplos ausentes simplesmente não farão parte do aprendizado. Por exemplo, treinar um algoritmo de reconhecimento facial requer utilizar exemplos de faces representativas de todas as etnias; no entanto, algumas etnias são minoritárias, tornando escassa a disponibilidade de exemplos correspondentes. Com poucos exemplos, talvez nenhum, representativos desses indivíduos, o algoritmo não aprende o suficiente sobre eles, tratando-os como exceções (Buolamwini, 2017). Para combater problemas deste tipo já existem mecanismos para auditar algoritmos, como a ferramenta Gender Shades (Raji; Buolamwini, 2019), cujo intuito é aumentar a imparcialidade e a transparência de tais sistemas, e, de modo geral, permitir a identificação de viés em algoritmos, avaliar a capacidade de inclusão dos sistemas, e promover a conscientização dos seus desenvolvedores.

Outros problemas são menos óbvios do que o cenário envolvendo reconhecimento de faces. No livro intitulado *Weapons of math destruction*, a autora Cathy O'Neil (2016) discute como algoritmos usados em áreas como marketing, análise de risco (crédito, seguros), educação (seleção, financiamento), e policiamento, podem agravar desigualdades. Um dos exemplos mais citados de seu trabalho hipotetiza sobre um aluno que tem seu financiamento estudantil negado por um banco em razão de seu endereço – possivelmente porque o histórico dos dados reflete alta taxa de inadimplência de indivíduos da mesma região. Tal decisão o excluiria da oportunidade de obter uma educação capaz de tirá-lo da pobreza, estabelecendo, assim, um círculo vicioso. Segundo a autora, decisões baseadas em recomendações fornecidas por ferramentas matemático-computacionais são obscuras e de difícil contestação. Ademais, tais ferramentas não são regulamentadas, e como agravante atuam em larga escala, afetando camadas cada vez maiores da sociedade. Paradoxalmente, a expectativa seria de que o uso de algoritmos que adotam critérios objetivos deveria promover a igualdade, pois todas as pessoas seriam avaliadas segundo o

mesmo conjunto de critérios. Mas, ao contrário, ao desconsiderar questões socioeconômicas relevantes na definição de critérios e tomada de decisões, esses algoritmos tendem a agravar o problema da desigualdade.

Conclusão e perspectivas

Neste capítulo, e na ampla literatura, muito se discutiu sobre o Aprendizado de Máquina guiado por exemplos capazes de ensinar ao computador o que é certo e o que é errado com relação a uma tarefa. Essa modalidade de aprendizado é conhecida como "supervisionada". Apesar de estudado de modo abundante e apresentar resultados reconhecidos, o aprendizado supervisionado depende de grandes bases de dados organizados e rotulados, e é inflexível com relação aos parâmetros iniciais do problema. Em razão dessas limitações, Yoshua Bengio (2019) afirma que o futuro da Inteligência Artificial deve se voltar para outra modalidade de aprendizado, denominado "não supervisionado". Mais desafiadora, essa modalidade prevê que o computador aprenda sem que precise inspecionar exemplos rotulados, tendo como critério para identificar padrões tão somente o objetivo alvo. Os principais usos dessa modalidade ainda são bastante rudimentares, com aplicações em problemas clássicos de Aprendizado de Máquina, como detecção de agrupamentos, redução de dimensionalidade, seleção de características, entre outros (Celebi; Aydin, 2016). No que diz respeito às Redes Neurais Artificiais, a despeito do progresso em ritmo acelerado, pode-se afirmar que ainda há muito a ser feito para viabilizar um uso amplo e bem-sucedido em tarefas de aprendizado não supervisionado. Nessa modalidade, as pesquisas em Inteligência Artificial avançam, com a introdução de técnicas de aprendizado por reforço, transferência de conhecimento, modelos adversariais, e modelos autorregressivos (Schmidhuber, 2015; Mesnil et al., 2011).

Esse movimento orientado ao aprendizado não supervisionado é essencial quando se discute o advento do Quinto Paradigma. O aproveitamento das informações geradas pela humanidade e a conversão dessas informações em conhecimento dependem de uma abordagem mais autônoma e exploratória, que demande menos supervisão humana, e não exija uma pré-concepção explícita sobre o que se busca. Ademais, são necessárias técnicas de representação do conhecimento gerado (Sowa, 2014) que possam ser utilizadas na execução de tarefas e atividades que demandem o uso de raciocínio analítico, de modo análogo ao ser humano.

Agradecimentos

Este trabalho teve apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (2013/14262-7, 2017/05838-3, 2018/17620-5), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (406550/2018-2), e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Código de Financiamento 001).

Referências

AHMED, M. N. et al. Cognitive computing and the future of health care cognitive computing and the future of healthcare: the cognitive power of ibm watson has the potential to transform global personalized medicine. *IEEE pulse, IEEE*, v.8, n.3, p.4-9, 2017.

ARAI, N. H.; MATSUZAKI, T. The impact of ai on education – can a robot get into the university of tokyo. In: Proc. ICCE. [s.l.: s.n.], p.1034-42. 2014.

BAHDANAU, D.; CHO, K.; BENGIO, Y. Neural *machine* translation by jointly *learning* to align and translate". *arXiv preprint arXiv:*1409.0473, 2014.

BANVILLE, D. L. Mining chemical structural information from the drug literature. Drug discovery today. *Elsevier*, v.11, n.1-2, p.35-42, 2006.

BAR-HILLEL, Y. The present status of automatic translation of languages. *Advances in computers,* v.1, n.1, p.91-163, 1960.

BENGIO, Y. What's next for AI. 2019. Disponível em: <https://www.ibm.com/watson/advantage-reports/ future-of-artificial--intelligence/yoshua-bengio.html>.

BERTINETTO, L. et al. Fully-convolutional siamese networks for object tracking". In: SPRINGER. In: European Conference On Computer Vision. [s.l.], p.850-65, 2016.

BIAMONTE, J. et al. Quantum machine learning. *Nature*, v.549, n.7671, p.195, 2017.

BRYNJOLFSSON, E.; MCAFEE, A. The business of artificial intelligence. *Harvard Business Review*, v.6, 2017.

BUOLAMWINI, J. A. Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers. Massachusetts, 2017. Thesis (Ph.D) — Massachusetts Institute of Technology.

CASELI, H. M.; NUNES, M. G. V. Alinhamento sentencial e lexical de *corpus* paralelos: recursos para a tradução automática. In: 52º SEMINÁRIO DO GEL - Simpósio de Perspectivas com *Corpus* para Tradução e Terminologia: Projetos de Pesquisa e Ferramentas. *Caderno de resumos*, [s.l.: s.n.], 2004. p.369-70.

CELEBI, M. E.; AYDIN, K. *Unsupervised learning algorithms*. [s.l.]: Springer, 2016.

CHEN, Y.; ARGENTINIS, J. E.; WEBER, G. Ibm watson: how cognitive computing can be applied to *big data* challenges in life sciences research. Clinical therapeutics. *Elsevier,* v.38, n.4, p.688-701, 2016.

DENG, L.; LIU, Y. *Deep Learning in Natural Language Processing*. [s.l.]: Springer, 2018.

DENG, L. et al. Deep learning: methods and applications. Foundations and Trends•R in Signal Processing. *Now Publishers*, Inc., v.7, n.3-4, p.197-387, 2014.

FURBER, S. B. et al. The spinnaker project. *Proceedings of the IEEE*, IEEE, v.102, n.5, p.652-65, 2014.

GHOSH, D. et al. Automated error correction in ibm quantum computer and explicit generalization. *Quantum Information Processing*, Springer, v.17, n.6, p.153, 2018.

GOH, G. B.; HODAS, N. O.; VISHNU, A. Deep *learning* for computational chemistry. Journal of computational chemistry, *Wiley Online Library*, v.38, n.16, p.1291-307, 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, Ma.: MIT Press, 2016. Disponível em: <http://www.deeplearningbook.org>.

GRIGSBY, J.; SANDERS, J. H. Telemedicine: where it is and where it's going. Annals of internal medicine. *American College of Physicians*, v.129, n.2, p.123-27, 1998.

HANNUN, A. et al. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:*1412.5567, 2014.

HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.l.: s.n.], 2016. p.770-8.

HEY, A. J. et al. *The fourth paradigm:* data-intensive scientific discovery. [s.l.]: Microsoft research Redmond, WA, 2009. v.1.

HRIPCSAK, G. et al. Observational health data sciences and informatics (ohdsi): opportunities for observational researchers. Studies in health technology and informatics. *NIH Public Access*, v.216, p.574, 2015.

IBM. Natural Language Understanding. 2019. Disponível em: <https://www.ibm.com/watson/services/natural-language-understanding/>. Acesso em: 9 out. 2019.

JORDAN, J. et al. Extremely scalable spiking neuronal network simulation code: from laptops to exascale computers. *Frontiers in Neuroinformatics, Frontiers*, v.12, p.2, 2018.

JR, J. F. R. et al. On the convergence of nanotechnology and *big data* analysis for computer-aided diagnosis. *Nanomedicine, Future Medicine*, v.11, n.8, p.959-82, 2016.

JR, J. F. R. et al. A survey on big data and machine learning for chemistry. *arXiv preprint arXiv:*1904.10370, 2019.

KAHN, J. M. et al. Virtual visits — confronting the challenges of telemedicine. *N Engl J Med*, v.372, n.18, p.1684-5, 2015.

KLEIN, G. et al. Opennmt: Open-source toolkit for neural *machine* translation. *arXiv preprint arXiv:*1701.02810, 2017.

KRALLINGER, M. et al. Information retrieval and text mining technologies for chemistry. *Chemical reviews, ACS Publications*, v.117, n.12, p.7673-761, 2017.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems.* [s.l.: s.n.], 2012. p.1097-1105.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep *learning*. *Nature, Nature Publishing Group*, v.521, n.7553, p.436, 2015.

LEIBOLD, J. Surveillance in china's xinjiang region: Ethnic sorting, coercion, and inducement. *Journal of Contemporary China*, Taylor & Francis, p.1-15, 2019.

LITJENS, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, v.42, p.60-88, 2017.

LIU, Y. et al. Artificial intelligence–based breast cancer nodal metastasis detection: Insights into the black box for pathologists. *Archives of Pathology & Laboratory Medicine*, the College of American Pathologists, 2018.

MARTINS, R. T. et al. Linguistic issues in the development of regra: A grammar checker for brazilian portuguese. *Natural Language Engineering*, v.4, n.4, p.287-307, 1998.

MESNIL, G. et al. Unsupervised and transfer learning challenge: a deep *learning* approach. In: JMLR. ORG. *Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning workshop*, v.27, p.97-111, 2011.

MILLS, M. *Artificial intelligence in law:* The state of play 2016. [s. l.]: Thomson Reuters Legal executive Institute, 2016.

OBERMEYER, Z.; EMANUEL, E. J. Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, v.375, n.13, p.1216, 2016.

O'NEIL, C. *Weapons of math destruction*: How big data increases inequality and threatens democracy. [s.l.]: Broadway Books, 2016.

PARKHI, O. M. et al. Deep face recognition. *bmvc*., v.1, n.3, p.6, 2015.

PASZKOWICZ, W. Genetic algorithms, a nature-inspired tool: Survey of applications in materials science and related fields. *Materials and Manufacturing Processes*, Taylor & Francis, v.24, n.2, p.174-97, 2009.

QIANG, X. The road to digital unfreedom: President xi's surveillance state. *Journal of Democracy*, Johns Hopkins University Press, v.30, n.1, p.53-67, 2019.

RAJI, I. D.; BUOLAMWINI, J. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In: *AAAI/ACM Conf. on AI Ethics and Society*. [S.l.: s.n.], 2019. v.1.

REILLY, R. G.; SHARKEY, N. *Connectionist approaches to natural language processing*. [s.l.]: Routledge, 2016.

RENNIE, J. How ibm's watson computer excels at jeopardy. *Aralık,* v.14, p.2014, 2011.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, v.61, p.85-117, 2015.

SHAALAN, K.; HASSANIEN, A. E.; TOLBA, F. *Intelligent Natural Language Processing*: Trends and Applications. [s.l.]: Springer, 2017. v.740.

SILVA, F. N. et al. Using network science and text analytics to produce surveys in a scientific topic. J*ournal of Informetrics*, v.10, n.2, p.487-502, 2016.

SOWA, J. F. *Principles of semantic networks*: Explorations in the representation of knowledge. [s.l.]: Morgan Kaufmann, 2014.

SZEGEDY, C. et al. Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.l.: s.n.], 2015. p.1-9.

WALLACH, H. Computational social science= computer science+ social data. *Communications of the ACM*, v.61, n.3, p.42-4, 2018.

WHITLEY, D.; SUTTON, A. M. Genetic algorithms — a survey of models and methods. *Handbook of natural computing, Springer*, p.637-71, 2012.

WU, Y. et al. Google's neural machine translation system: Bridging the gap between human and *machine* translation. *arXiv preprint arXiv:*1609.08144, 2016.

ZHANG, X.; ZHAO, J.; LECUN, Y. Character-level convolutional networks for text classification. In: *Advances in neural information processing systems*. [s.l.: s.n.], 2015. p. 649-57.