

TAINÁ AGUIAR JUNQUILHO

**INTELIGÊNCIA
ARTIFICIAL NO
DIREITO:
LIMITES ÉTICOS**

2022

O objetivo desse capítulo é descrever como nasceu a concepção de IA que fundamenta a pesquisa e compreender por que seu uso ético deve ser encarado como direito fundamental.

1.1. COMPREENDENDO A HISTÓRIA DA IA

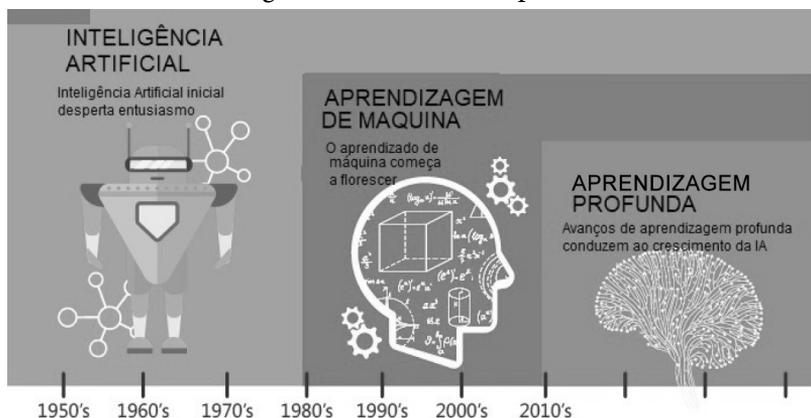
Historicamente, várias áreas como a matemática, a filosofia, a economia, a engenharia da computação, a neurociência, a psicologia e a linguística contribuíram, ao longo dos anos, com pesquisas fundamentais para o desenvolvimento da IA. Para citar exemplos, as teorias da decisão e dos jogos na economia e as descobertas da neurociência sobre o sistema nervoso e neuronal, foram e são responsáveis por dar importantes aportes na evolução da IA (RUSSELL; NORVIG, 2013).

Entretanto, embora a ficção científica e a filosofia refletissem sobre o tema há muitos anos, foi somente há sete décadas que o seminário “*Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*”, ocorrido no verão de 1955, organizado por John McCarthy, apresentou grandes nomes da pesquisa em IA, como Marvin Minsky, Nathaniel Rochester e Claude Shannon. A proposta exibida no seminário trazia pela primeira vez o termo IA e, também, a primeira definição desse novo ramo:

“[...] para a presente proposta o problema da inteligência artificial é fazer uma máquina se comportar de forma a ser chamada de inteligente como se a ação estivesse sendo praticada por um humano” (MCCARTHY et al., 2006, tradução nossa).

Mesmo após mais de setenta anos de pesquisas em IA, quem pretende conceituá-la com precisão ainda encontra dificuldades. Afinal, a história da IA não é de ascensão retilínea, como se pode observar na figura abaixo:

Figura 1- Linha do tempo IA



Fonte: Disponível em: <https://flammarion.wordpress.com/tag/inovacao/>.

Acesso em: 04 jun. 2021. (Tradução nossa).

A IA passou por períodos de “inverno”, nos quais eram poucas as pesquisas, produções acadêmicas e descobertas na área. Este é um campo que ainda não tem uma teoria firme, mas vem avançando em seus modelos empíricos principalmente a partir de 2010/2011, quando os “algoritmos de *machine learning* (ml, subárea da ciência da computação) e os de *deep learning* (dl, subárea da ml) receberam impulso de novos avanços, em *hardware* e *software*” (ARBIX, 2020, p. 402). Vive atualmente, portanto, um de seus “verões”, período de muitos investimentos e de importantes resultados.

Dentre os fatores que propiciaram a ascensão dos estudos sobre aprendizagem profunda, estão (a) crescente disponibilidade de grande quantidade de dados (*big data*), (b) maior capacidade computacional e (c) evolução dos algoritmos (KAUFMAN, 2018, p.24).

Mas, “afinal o que é um algoritmo? No seu nível mais fundamental, um algoritmo nada mais é que uma precisa série de instruções para executarem uma tarefa concreta” (KEARNS; ROTH, 2020, p.4). Isso é, um conjunto de códigos que permitem a efetivação de alguma operação.

Além disso, o “ABC das tecnologias de informação e comunicação (*analytics + big data + cloud computing*)” (MAGRANI, 2018, p. 25), caracteriza o contexto em que se desenvolveu a análise de grande

volume de dados, aliada ao grande potencial de armazenamento dos mesmos.

É a denominada Quarta Revolução Industrial, por meio da qual a tecnologia domina os meios de produção, provocando intensas transformações na sociedade e na economia. A Quarta Revolução Industrial caracteriza-se pela convergência de tecnologias digitais, dentre as quais a IA tem preponderância e destaque pela capacidade de lidar com grande conjunto de dados e com base neles prever como e quando ocorrerá determinado resultado com alto grau de assertividade (predição). O desenvolvimento da IA e a automatização de processos decisórios provocados pelas técnicas de aprendizado de máquina cumpre papel central nesse contexto (VERONESE; SILVEIRA; LEMOS, 2019; XAVIER; SPALER, 2019). O digital traz à tona a vivência do *homo digitalis*, assim chamado por Han (2019), para denominar a forma como o ser humano atual lida e interage com o mundo à sua volta. Nessa sociedade, “todo clique que eu faço é salvo. Todo passo que eu faço é rastreável. Deixamos rastros digitais em todo lugar. [...] A sociedade digital apresenta uma estrutura especial panóptica” (HAN, 2019, p.122).

Nesse cenário, os indivíduos tendem a estar sempre conectados à internet e, portanto, produzindo dados por meio dos mais variados objetos inteligentes do cotidiano que se comunicam (relógio, televisão, geladeira, carro, celular etc.), fenômeno chamado de internet das coisas (*IoT - Internet of Things*). “A hiperconectividade presente na era da *IoT*, gera circunstâncias que fazem o indivíduo estar permanentemente conectado (*always-on*); a possibilidade de estar prontamente acessível (*readily accessible*); a riqueza de informações; a interatividade; o armazenamento ininterrupto de dados (*always recording*)” (MAGRANI, 2018, p.21).

Na era da internet das coisas os objetos que servem às pessoas conectam-se permanentemente à internet. A máquina, conectada vinte e quatro horas por dia à internet, serve ao indivíduo e relaciona-se com o ambiente por meio de conexões permanentes, capturando, assim, dados que em sua maioria são pessoais, isso é, podem identificar de alguma forma os cidadãos que as utilizam.

A capacidade computacional de armazenamento desse grande volume de dados produzidos a todo instante pelos seres humanos,

possibilitou o *boom* das técnicas de aprendizado de máquina “para resolver problemas cada vez mais complexos” (LOURIDAS; EBERT, 2016, p. 110, tradução nossa)². A proeminência de produção, captação, conservação e guarda de dados em larga escala na era do *big data*, trouxe a “onda de novas demandas nos algoritmos de aprendizado de máquina” (JORDAN; MITCHELL, 2015, p. 257). O *big data* e o aumento da capacidade computacional foram pontos relevantes que possibilitaram a concretização do modelo de redes neurais, ao reduzir a velocidade do processamento e aumentar a oferta de dados quantitativa e qualitativamente.

Tais fatores podem ser apontados como concorrentes para o sucesso do desenvolvimento de modelos algorítmicos de IA, que dependem de muito poder de “processamento e grandes quantidades de dados para serem assertivos em tarefas de classificação de textos, de reconhecimento de faces, de reconhecimento de voz, de previsões e de agrupamentos de informações” (SILVA, 2019, p.43).

O movimento da IA, portanto, não é retilíneo, tendo sofrido invernos e períodos de relativo desprestígio. A partir de 2010/2011, contribuíram decisivamente para sua definitiva ascensão a era dourada dos dados (*Big Data*) e o aumento na capacidade de armazenamento desses, somado ao aperfeiçoamento dos algoritmos, sequência de códigos que automatizam decisões, fatores que proporcionaram o desenvolvimento da aprendizagem profunda. Por ser um campo de descobertas relativamente recentes, falta à IA ainda uma robusta teoria que a conceitue com precisão.

1.1.1. Pressuposto: o que é IA para essa pesquisa?

Inicialmente é preciso esclarecer que há críticas ao termo “Inteligência Artificial”, as quais citam principalmente o fato de que a expressão é imprecisa e dá a falsa expectativa de que a IA reproduz precisamente o ser humano, como nas ficções científicas. “O termo “inteligência”, para muitos observadores, sugere talvez que os sistemas mais recentes são, em certo sentido, realmente “conscientes”” (SUSS-

2. Essa capacidade tende a aumentar ainda mais com o desenvolvimento da chamada computação quântica, que aumentará a capacidade de processamento de dados e fará com que a IA desenvolva mais rapidamente (TAYLOR, 2020).

KIND, 2019, p. 265, tradução nossa)³. Entretanto, embora criticado, esse é o *topoi* (lugar comum) e termo discursivamente usado nas discussões acadêmicas, sendo fundamental precisá-lo como premissa para a pesquisa.

Para compreender o que é uma inteligência artificialmente construída é necessário antes entender o que é inteligência, característica que, em geral, é atribuída de forma exclusiva ao ser humano. Por isso, aprender sobre IA é também uma forma de autoconhecimento e de autocompreensão da raça humana (RUSSELL; NORVIG, 2013). Todavia, tampouco há consenso sobre o que é o pensamento inteligente. O que seria a inteligência? Capacidade de raciocínio lógico, de compreensão do mundo e de si, de conhecimento emocional, de criatividade, de aprendizado, de resolução de problemas etc.?

Em primeiro lugar, é necessário compreender a IA como substantivo não contável (é impreciso dizer, por exemplo, que foi desenvolvida “**uma** IA”), pois os estudos em IA estão inseridos na ciência da computação e tal como na matemática, na biologia, dentre outros campos, comportam diversos métodos de resolução de problemas específicos⁴. Não existe, tampouco, “uma” IA geral ou forte que imite completamente o comportamento cognitivo do humano em todas as suas nuances (ou que o supera – superinteligência/ultrainteligência), criação presente nos relatos de ficção científica. Ainda estamos distantes da “IA geral”, a tecnologia para todos os fins que pode fazer o mesmo que o humano é capaz” (LEE, 2019, p. 23).

O termo IA tem sido usado comumente para representar a “IA fraca ou estreita”, a qual abrange abordagens de resolução de problemas específicos. Nesse sentido, Garwood (2018, p. 10, tradução nossa), alerta que há inúmeras formas de conceituar IA, algumas publicadas em livros inteiros dedicados ao assunto que são extremamente técnicas, outras confusas. O manual de IA dos especialistas Russel e Norvig (2013) define IA como campo que estuda agentes inteligentes.

Segundo Cortiz (2020) “a IA pode ser entendida como um grande guarda-chuva que tem duas grandes abordagens: uma baseada em

3. Segundo Susskind (2019, p. 265, tradução nossa), “no jargão filosófico da IA, um sistema que é consciente seria uma exibição do que é conhecido como “IA forte”. A “IA forte”, entretanto, não é objeto dessa pesquisa pois, como se abordará mais a frente, ainda não é realidade no desenvolvimento da IA.

4. Disponível em: <https://course.elementsofai.com/1/1>. Acesso em: 30 abr. 2020.

conhecimento e o aprendizado estatístico”⁵. Os agentes baseados em conhecimento “são capazes de aceitar novas tarefas sob a forma de metas descritas de modo explícito, podem alcançar competência rapidamente ao serem informados ou adquirirem novos conhecimentos [...]” (RUSSELL; NORVIG, 2013, p. 349).

Essa abordagem compreende que “todo comportamento inteligente pode ser captado por um sistema que raciocine logicamente a partir de um conjunto de fatos e regras que descrevem o domínio”, isso é, que por meio de sistemas lógicos apreende a compreensão dos especialistas da área de determinado problema a ser solucionado. Atualmente essa abordagem é conhecida como GOFAI (*Good Old-Fashioned Artificial Intelligence*, ou a boa velha IA/a IA tradicional/à moda antiga) (RUSSELL; NORVIG, 2013, 827).

Ensinar computadores a pensar por meio da codificação de regras lógicas (se A, então B), era a tarefa dos desenvolvedores do campo baseado em regras ou sistemas simbólicos/especialistas (LEE, 2019). Já o aprendizado estatístico ou aprendizado de máquina (AM) (*machine learning*, termo cunhado por Arthur Lee Samuel em 1959), compreende a investigação e o estudo e concepção de algoritmos capazes de assimilar e prever dados.

As pesquisas em AM despertaram o surgimento de cinco escolas de pensamento: a) os simbolistas, para os quais o manuseio de símbolos é o que importa para reproduzir a inteligência, que trabalham com a dedução inversa (encontrar algoritmo capaz de generalizar uma dedução); b) os conexionistas, que compreendem o aprendizado inteligente como função do cérebro e utilizam *backpropagation* para reproduzir as conexões neuronais; c) os evolucionários, que entendem que o aprendizado vem da seleção natural e, por isso, usam a programação genética como norte para criação algorítmica; d) os bayesianos, para os quais todo aprendizado a ser reproduzido é incerto e por isso usam algoritmos de inferência probalística e adotam a teoria de Bayes e seus consequentes e os e) analogistas, que acreditam que o aprendizado se dá por identificação de semelhança entre os acontecimentos e sua reprodução, que empregam máquina de vetores de suporte (DOMINGOS, 2017).

5. O professor Diogo Cortiz cita essas abordagens em seu curso sobre IA. Disponível em: <<https://www.youtube.com/watch?v=Ze-Q6ZNWpco&t=1136s>>. Acesso em: 02 mai. 2020.

O AM comporta diferentes tipos de abordagens, dentre eles o supervisionado, o não supervisionado e o por reforço⁶. Cada uma delas tem um conjunto de técnicas e é adequada para resolver determinados tipos de situações.

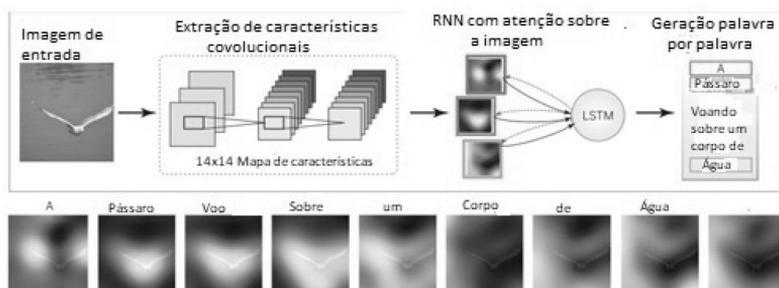
Pelo aprendizado supervisionado, sabendo-se do resultado desejado, o conjunto grande de dados inseridos para criação do modelo são rotulados previamente com a resposta correta necessária para a resolução do problema. Isso é, “o conjunto de treinamento contém dados e a saída correta da tarefa com esses dados” (LOURIDAS; EBERT, 2016, p.110, tradução nossa). Por exemplo, quando o modelo a ser construído objetiva criar algoritmo que separe os *e-mails* que são *spam* daqueles que não são e o conjunto de dados (*dataset*) apresenta previamente exemplos classificados do que é e o que não é *spam*. O aprendizado supervisionado se vale das técnicas de regressão (que inclui predição, árvores de decisão, redes bayesianas, regressão linear, dentre outras) e classificação (que abrange regressão lógica, redes neurais artificiais, dentre outras).

Uma das áreas que mais evoluiu pelo aprendizado supervisionado foi o da aprendizagem profunda (*Deep Learning* (DL)) que faz “uso de algoritmos de otimização baseados em gradiente para adaptar os parâmetros por meio de uma rede de múltiplas camadas baseadas em erro da saída (*output*)” (JORDAN; MITCHELL, 2015, p. 257, tradução nossa). Na Figura 2 demonstra-se o funcionamento de uma rede neural convolucional treinada para interpretar imagens, nela o *output* (saída) é utilizado:

[...] por uma rede neural recorrente treinada para gerar uma legenda (em cima). A sequência na parte inferior mostra o foco palavra por palavra da rede em diferentes partes da imagem de entrada enquanto gera a legenda palavra por palavra. (JORDAN; MITCHELL, 2015, p. 257, tradução nossa)

6. Existe também um aprendizado “intermediário” denominado semi-supervisionado por meio do qual a máquina “aprende com uma pequena quantidade de dados rotulados e com uma grande quantidade de dados não rotulados” (HAO; ZHANG, 2016, p. 426).

Figura 2 - Geração automática de legendas de texto para imagens com redes profundas



Fonte: JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, v. 349, n. 6245, 2015 (Tradução nossa)

O AM inclui, portanto, a aprendizagem profunda, que se trata de pilhas de multicamadas de modelos sujeitos a aprendizado, o qual também se desenvolveu no tipo de aprendizado não supervisionado (LECUN; BENGIO; HINTON, 2015).

[...] Embora grande parte do sucesso prático da aprendizagem profunda venha de métodos de aprendizagem supervisionada para descobrir essas representações, também foram feitos esforços para desenvolver algoritmos de aprendizagem profunda que descobrem representações úteis da entrada sem a necessidade de dados de treinamento rotulados. (JORDAN; MITCHELL, 2015, p. 258, tradução nossa)

Por aprendizado não supervisionado compreende-se aquele que envolve a análise de conjuntos de dados (*dataset*) não rotulados previamente. Esse pode valer-se de técnicas para segmentação (que compreende técnicas como algoritmos genéticos, por exemplo) e para redução de dimensão (que utiliza técnicas como o aprendizado múltiplo, análise fatorial, projeções aleatórias, dentre outras).

A Figura 3 resume as duas abordagens de AM e técnicas que podem ser usadas.

Figura 3 - Abordagens de Aprendizado de Máquina e Técnicas.



Fonte: LOURIDAS, P.; EBERT, C. Machine Learning. **IEEE SOFTWARE**, 2016
(Tradução nossa)

Ambos os paradigmas são usados quando se está diante de uma grande massa de dados, o que não é necessário na aprendizagem por reforço. A aprendizagem por reforço, inspirada principalmente na psicologia comportamental e neurociência, trabalha com sistema de recompensas para acertos da aprendizagem da máquina. Esse tipo de aprendizado é feito em ambientes controlados que permitam que o sistema aprenda por tentativa e erro (JORDAN; MITCHELL, 2015). Em resumo, na aprendizagem por reforço o algoritmo “aprende como fazer ações através de observações. Cada ação tem um impacto no meio ambiente. O aprendizado por reforço faz julgamentos com base no *feedback* do ambiente observado” (HAO; ZHANG, 2016, p. 426, tradução nossa).

A IA geral, conforme foi afirmado, não existe. As abordagens de IA fraca ou estreita resolvem problemas específicos. Conclui-se que no AM, “cada método de aprendizado, usado separadamente, tem aplicações limitadas, pois requer tipos especiais de *input* e conhecimento e aprende um tipo específico de informação” (TECUCI, 2012, p. 175).

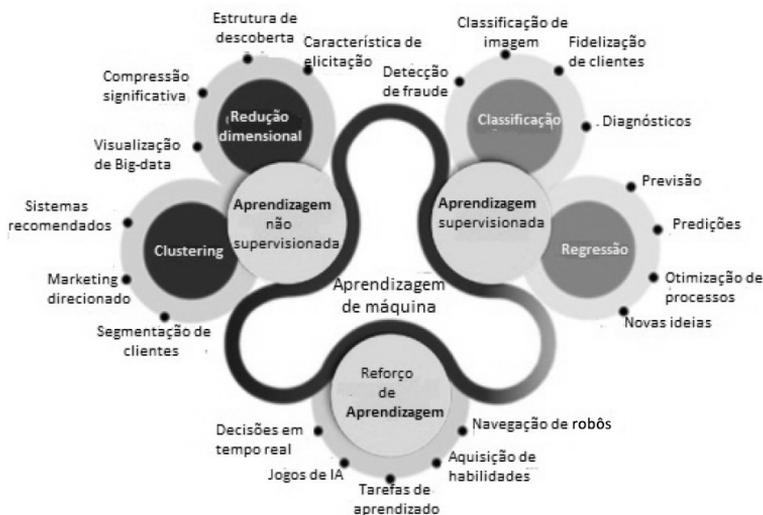
Cada técnica de abordagem se adéqua melhor para desenvolvimento de uma tarefa particular. O aprendizado supervisionado

pode ser usado, por exemplo, para desenvolver sistemas antifraude, reconhecimento de imagem e de fala, tradução de idiomas, no desenvolvimento de filtros de *spam* e de sistemas de recomendação (predição). O aprendizado não supervisionado para pontuação de crédito (*credit score*), em detecção de fraudes, segmentação de clientes, dentre outros. E o aprendizado por reforço pode ser usado para criação de jogos e robótica (DOMINGOS, 2012).

LeCun, Bengio e Hinton (2015, p. 100, tradução nossa) apresentam exemplos de aplicação do aprendizado profundo que possibilitaram o *hype* atual da IA: para manuseio de *big data* em marketing de vendas e promoções; reconhecimento de objetos e padrões para encontro de “pontos fracos”, previsão de fluxos e vídeo para direção autônoma, avaliação e análise de imagens para identificação de formas distintas (como em imagens médicas ou reconhecimento facial e impressão digital) e “heurísticas de segurança que soltam ataques padrão para proteger, por exemplo, portas ou redes”.

Na Figura 4 pode-se verificar um resumo das abordagens de AM e usos que podem ser feitos, por técnica.

Figura 4 – Aprendizado de Máquina e utilizações por técnica.



Fonte: Disponível em: <<https://pt.the-ai-academy.com/machine-learning-boot-camp>>. Acesso em: 08 mai. 2020 (Tradução nossa)

De forma resumida, é possível observar que a moderna compreensão sobre IA envolve entendê-la como algo que pode resolver problemas específicos. Um problema de aprendizagem pode ser definido como o problema de melhorar alguma medida de desempenho ao executar alguma tarefa, através de alguns tipos de experiência de treinamento (JORDAN; MITCHELL, 2015, p. 255).

Note-se que, longe da compreensão que a princípio o termo pode nos levar, a IA não é a reprodução exata do comportamento humano, mas apenas algumas funções específicas do cérebro humano. Defini-la passa por compreender como funcionam os modelos de aprendizagem da máquina e saber identificar qual a melhor técnica se ajusta ao problema a ser resolvido. A pesquisa dá enfoque ao modelo de aprendizado de máquina e à técnica de redes neurais artificiais, que é atualmente a mais utilizada.

1.2. POR QUE É (DIREITO) FUNDAMENTAL EXIGIR APLICAÇÃO ÉTICA DA IA?

A compreensão de IA para essa pesquisa foca, portanto, na aprendizagem profunda, modelo que vem sendo mais utilizado e que obteve maiores avanços recentes, pois permite o desenvolvimento de soluções tecnológicas que proporcionam progressos nas mais variadas esferas da vida humana⁷. Existem numerosos exemplos de aplicações de sucesso de IA, que colaboram tanto no desenvolvimento de políticas públicas, quanto na otimização de processos e no crescimento exponencial do setor privado. Nesse sentido, inúmeras experiências podem ser apontadas em alguns campos que experimentam profundas transformações⁸.

No setor de transportes os maiores exemplos são os veículos autônomos, as tecnologias de IA incorporadas aos automóveis⁹ e as plataformas que traçam melhores rotas em tempo real (GORDON-MURNANE, 2018). As mudanças trazidas podem refletir na diminuição

7. Disponível em: <<https://blogs.scientificamerican.com/observations/ethics-in-the-age-of-artificial-intelligence/>>. Acesso em: 13 maio 2020.

8. No site da Comissão Norte-Americana de proteção aos consumidores, é possível encontrar vários exemplos de aplicações atuais bem sucedidas de AM. Disponível em: <<https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms>>. Acesso em: 20 mai. 2020.

9. Algumas descobertas interessantes se encontram na seguinte reportagem: <https://www.nytimes.com/2018/01/04/technology/self-driving-cars-aurora.html>. Acesso em: 13 maio 2020.

do índice de acidentes e danos, no aumento de índices de qualidade e expectativa de vida, na melhoria do trânsito e redução de tempo gasto em deslocamento (SOUZA; OLIVEIRA, 2019).

Na área da segurança também podem ser apontados alguns importantes resultados alcançados pelas pesquisas em IA. Exemplos de aplicações são: *Drones* de rastreamento¹⁰, armas autônomas para fins militares (PEREZ et al., 2017) e uso de reconhecimento facial para detecção de criminosos¹¹ (LAUTERBACH, 2019). Quem defende seu uso no setor de segurança, aponta como benefícios “redução de custos, preservação da vida humana, melhor precisão.” (SOUZA; OLIVEIRA, 2019, p. 69). Cabe ressaltar, entretanto, que essa pesquisa concorda com as críticas realizadas por autores como Ruha (2021), que são contrários à utilização da IA para reconhecimento facial na segurança pública tendo em vista o potencial discriminatório e de reforço ao encarceramento de determinados grupos sociais, à seletividade penal e à hipervigilância.

Nos setores consumerista, de seguros e financeiro, a IA tem sido usada tanto para ofertar e recomendar produtos e serviços adequados às necessidades do usuário (GORDON-MURNANE, 2018), quanto para elaboração de pontuação de crédito (*credit-score*), para análise de concessão de benefícios e seguros e na detecção de fraudes, por bancos e instituições financeiras (SADATRASOUL et al., 2013). Tais análises permitem a redução de riscos e aumento da rentabilidade para empresas, além de aprimorar a experiência das necessidades de clientes (TZIRULNIK; BOAVENTURA, 2019). Na agricultura, há avanços de AM para monitoramento de campo, controle de pragas, análise de imagem por satélite, robotização da produção e colheita, que aumenta a eficiência no setor alimentício (LAUTERBACH, 2019).

Na área da educação, algumas aplicações são: sistemas tutoriais inteligentes, formas de pesquisas avançadas e detalhadas de referenciais teóricos e aprendizagem personalizada (KAY, 2012). Tais utilizações

10. Uma reportagem do jornal The New York Times anuncia com preocupação a criação de Drone capaz de seguir usuários.

11. Sistemas de reconhecimento facial vêm sendo usados ao redor do mundo, principalmente com fins de colaborar com a segurança. No Brasil, inclusive, a Bahia foi o primeiro Estado a usar essa tecnologia. Disponível em: <<https://www.correio24horas.com.br/noticia/nid/entenda-como-funciona-o-reconhecimento-facial-que-ajudou-a-prender-mais-de-100-na-ba/>>. Acesso em: 14 maio 2020.