

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO ECONÔMICO
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS
CURSO CIÊNCIAS ECONÔMICAS

Fábio de Paula Sividanes

Inovação, Inteligência Artificial e Mercado de Trabalho

Florianópolis

2020

Fábio de Paula Sividanes

Inovação, Inteligência Artificial e Mercado de Trabalho

Trabalho Conclusão do Curso de Graduação em Ciências Econômicas do Centro de Economia e Relações Internacionais da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Economia.

Orientador: Prof. Gilson Geraldino da Silva Junior, Dr.(a)

Florianópolis

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Sividanes, Fabio de Paula
Inovação, inteligência artificial e mercado de trabalho
/ Fabio de Paula Sividanes ; orientador, Gílson Geraldino
da Silva Jr, 2020.
42 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro Sócio
Econômico, Graduação em Ciências Econômicas, Florianópolis,
2020.

Inclui referências.

1. Ciências Econômicas. 2. Inovação. 3. Mercado de
trabalho. 4. Inteligência artificial. I. da Silva Jr,
Gílson Geraldino. II. Universidade Federal de Santa
Catarina. Graduação em Ciências Econômicas. III. Título.

Fábio de Paula Sividanes

Inovação, Inteligência Artificial e Mercado de Trabalho

Florianópolis, 10, de dezembro de 2020.

O presente Trabalho de Conclusão de Curso foi avaliado e aprovado pela banca examinadora composta pelos seguintes membros:

Prof.(a) Gilson Geraldino da Silva Junior, Dr.(a)

Universidade Federal de Santa Catarina

Prof.(a) Armando de Melo Lisboa, Dr.(a)

Universidade Federal de Santa Catarina

Prof.(a) Fred Leite Siqueira Campos, Dr.(a)

Universidade Federal de Santa Catarina

Certifico que esta é a **versão original e final** do Trabalho de Conclusão de Curso que foi julgado adequado para obtenção do título de Bacharel em Economia por mim e pelos demais membros da banca examinadora.

Prof.(a) Gilson Geraldino da Silva Junior, Dr.(a)

Orientador(a)

Florianópolis, 2020.

AGRADECIMENTOS

Agradeço primeiramente a Deus, a minha mãe, familiares, orientador e amigos.

RESUMO

O presente trabalho busca encontrar a relação entre a evolução da tecnologia (sobretudo na quarta era tecnológica, respaldada pela inteligência artificial e robótica) e o mercado de trabalho. Baseia-se em pesquisa exploratória (buscando através da discussão entre artigos acadêmicos indícios de fatores que moldam esse novo paradigma, como o caráter das tarefas que mais são afetadas por tal evolução, o aumento de produtividade, quesitos de desigualdade) e um modelo prático, no qual o autor mostra uma tarefa que já foi preenchida pelo uso de automação, mostrando que a máquina consegue substituir o ser humano em tarefas intelectuais com um nível de precisão alto, como no caso do exemplo citado, acima de 99%. Além de elucidar conceitos específicos de aprendizagem profunda (tecnologia de ponta, atualmente, no que se refere inteligência artificial), como seu contexto dentro do escopo de inteligência artificial e detalhes de seu funcionamento. Chegando à conclusão da inevitabilidade da substituição de determinadas tarefas e a criação de outras.

Palavras-chave: Inovação. Inteligência Artificial. Mercado de Trabalho.

ABSTRACT

The present work seeks to find the relationship between the evolution of technology (especially in the fourth technological age, supported by artificial and robotic intelligence) and the job market. It is based on exploratory research (looking through the discussion between academic articles for evidence of factors that shape this new paradigm, such as the character of the tasks that are most affected by such evolution, increased productivity, inequality issues) and a practical model, in which the author shows a task that has already been completed through the use of automation, showing that the machine can replace human beings in intellectual tasks with a high level of precision, as in the case of the example mentioned, above 99%. In addition to elucidating specific concepts of deep learning (cutting edge technology, currently, with regard to artificial intelligence), such the context within the scope of artificial intelligence and details of its operation. Coming to the conclusion of the inevitability of replacing certain tasks and creating others.

Keywords: Innovation. Artificial Intelligence. Labor Market.

LISTA DE FIGURAS

Figura 1 - Diferença entre o fluxo da programação clássica e do aprendizado de máquina....	22
Figura 2 - Conjunto da Inteligência Artificial	23
Figura 3 - Diferença entre características de subconjuntos da Inteligência Artificial	25
Figura 4 - Fluxograma de um modelo de Aprendizagem Profunda	28
Figura 5 - Fluxograma de uma rede neural	29
Figura 6 - Processo que ocorre em cada unidade da rede neural.....	30
Figura 7 - Representação visual das características obtidas em cada camada de uma rede neural	32
Figura 8 - Representação da camada convolucional.....	32
Figura 9 - Representação da camada MaxPooling2D.....	33
Figura 10 - Exemplo de elementos do banco de dados MNIST	36
Figura 11 - Camadas utilizadas no modelo e seus parâmetros.....	36
Figura 12 – Treinamento do modelo.....	37
Figura 13 - Matriz de confusão do modelo	39

LISTA DE GRÁFICOS

Gráfico 1 - Função de ativação ReLU	30
Gráfico 2 - Evolução da <i>loss</i> do modelo	38
Gráfico 3 - Evolução da precisão do modelo	38

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVO GERAL	12
1.2	OBJETIVOS ESPECÍFICOS	12
1.3	JUSTIFICATIVA	12
2	REVISÃO DE LITERATURA.....	14
2.1	IMPACTOS DIRETOS NO MERCADO DE TRABALHO	14
2.2	EFEITOS NA DISTRIBUIÇÃO DE RENDIMENTOS NO MERCADO DE TRABALHO	17
2.3	PROBLEMAS ESTRUTURAIS	18
2.4	DINÂMICA NO MERCADO DE TRABALHO AO LONGO DA DIFUSÃO DE NOVAS TECNOLOGIAS	19
2.5	DIFERENÇA DESSA REVOLUÇÃO TECNOLÓGICA PARA AS ANTERIORES	20
2.6	ESCOPO DA INTELIGÊNCIA ARTIFICIAL.....	21
3	METODOLOGIA	26
3.1	EM RELAÇÃO À METODOLOGIA ACADÊMICA.....	26
3.2	EM RELAÇÃO AO MODELO	26
3.3	CAMADAS (LAYERS)	28
4	MODELO DE DÍGITOS ESCRITOS À MÃO	35
5	CONCLUSÃO.....	40
	REFERÊNCIAS.....	41

1 INTRODUÇÃO

Não é novidade que avanço tecnológico dita as transformações dos paradigmas da sociedade; paradigmas esses os quais moldam tanto os padrões de produtividade, quanto a interação entre trabalho humano e capital. No cenário atual, a humanidade encontra-se diante das mudanças advindas da quarta era tecnológica, respaldada pela inteligência artificial, robótica e realidade virtual. As eras anteriores foram pautadas na produção em massa, eletricidade e nas telecomunicações, como exprimem Autor, Mindell e Reynolds (2019), no relatório sobre o futuro do trabalho, do Instituto de Tecnologia de Massachusetts.

Toda revolução dessa magnitude traz consigo um potencial disruptivo para transformar a vida do homem, aumentando a riqueza geral e a qualidade de vida da sociedade. As incertezas, provocada por essas mudanças, causam insegurança para os trabalhadores, que não sabem se os ofícios que possuem serão impactados e substituídos, ou não, por capital. Como Autor, Mindell e Reynolds (2019) mostram, por mais que haja aumento na produtividade ao longo dos anos, esses ganhos não são distribuídos de forma igualitária. Em seu relatório, mostram que o aumento das rendas acompanhou o ganho em produtividade até 1973; posteriormente, apenas as tarefas que exigiam maior nível de educação aumentaram de rendimento concomitantemente com a produtividade; por outro lado, os trabalhos com pouco nível de instrução estagnaram, ou até mesmo diminuíram, em relação ao poder de compra, comparados com décadas anteriores. Tal fato faz com que a memória dos trabalhadores sobre esse assunto tenha um cunho pessimista, e com razão.

Na comunidade acadêmica, em relação a esse assunto em evidência, o pessimismo sobre o futuro do mercado de trabalho é nítido. Contudo, como veremos a seguir, essa perspectiva pessimista de substituição do trabalho humano pelo trabalho de robôs é apenas um lado da moeda. A dinâmica dessa transformação social passa por assuntos mais profundos do que a apreensão dos trabalhadores. A inteligência artificial, que é o foco desse trabalho, tem o potencial não de apenas substituir a mão-de-obra humana, como também complementá-la e até mesmo, gerar novas ocupações. A efervescência do tema deve-se à três fatores específicos: o avanço nos hardwares (sobretudo o avanço nas placas gráficas), a melhoria nos algoritmos de inteligência artificial e na quantidade exorbitante de dados que é gerado a cada segundo, definido como “big data”; essa enorme quantidade de dados advém do crescimento da internet, além da maior capacidade de armazenamento dos hardwares atuais, Chollet (2017).

Na seção de referencial teórico será abordado o debate acadêmico acerca dessa nova dinâmica imposta pela evolução da inteligência artificial, buscando os principais pontos, que afetam diretamente o mercado de trabalho, discutidos atualmente, além de elucidar conceitos específicos de aprendizagem profunda, um subcampo de inteligência artificial em destaque atualmente, servindo de arcabouço para um exemplo prático em que o aprendizado de máquina substitui mão-de-obra de um trabalhador.

Na seção seguinte, a metodologia do presente trabalho será descrita, englobando os fatores acadêmicos de pesquisa para este trabalho, além dos conceitos técnicos, os quais sustentarão o exemplo prático apresentados na seção seguinte.

Por fim, os resultados modelo em si e as conclusões.

1.1 OBJETIVO GERAL

Apontar os impactos do avanço da Inteligência Artificial no mercado de trabalho.

1.2 OBJETIVOS ESPECÍFICOS

- Apresentar as perspectivas para um futuro próximo.
- Apontar as consequências para além do mercado de trabalho.
- Mostrar os desdobramentos dessa onda tecnológica.
- Replicar um exemplo concreto de uma tarefa que substituiu mão-de-obra humana.

1.3 JUSTIFICATIVA

A dinâmica da evolução constante da sociedade traz consigo uma discussão contínua acerca de desenvolvimento e tecnologia. O aumento de produtividade que evolui a cada revolução tecnológica muda a relação entre países e transforma variáveis macroeconômicas. Silva Jr e Cavalcante (2020) explicam que a inovação tecnológica fez com que alguns países conseguissem alcançar o patamar, tanto de renda quanto produtividade, de países líderes; sugestionando que uma saída para países em desenvolvimento (para se aproximarem de países de primeiro mundo) seria o investimento em pesquisa e desenvolvimento, assim como treinamento de mão-de-obra. Além disso, Silva Jr e Cavalcante

(2020) mostram a evolução das patentes produzidas pelos países “principais” (contrastando com o Brasil), reforçando o argumento dos autores.

Hall (2004) salienta a importância da difusão da tecnologia na sociedade, mostrando que o impacto da inovação só ocorre se ela é difundida. O autor comenta sobre esse processo, que envolve uma sequência de etapas, em que a tecnologia é aprendida, imitada e novas adaptações são descobertas, melhorando as inovações iniciais.

Acemoglu e Restrepo (2018) mostram um problema estrutural da inovação, que é a falta de trabalhadores capacitados para conduzir as melhorias tecnológicas, fazendo com que haja um tempo de adaptação entre o início da inovação e a consolidação da mesma.

Todos os argumentos acima reforçam a importância de uma discussão aprofundada do assunto, pois não envolve apenas questões técnicas, como também o futuro do mercado de trabalho, que impacta inúmeras vidas, além das vantagens competitivas em âmbito internacional, com produtos e serviços com maior agregação de valor.

A mudança da dinâmica impacta não só a produtividade das empresas como um todo, como também impacta vidas, e isso faz com que seja imprescindível o estudo dessa relação, a fim de estar munido de possíveis soluções pro futuro que essa nova realidade traz consigo.

2 REVISÃO DE LITERATURA

2.1 IMPACTOS DIRETOS NO MERCADO DE TRABALHO

Os avanços na tecnologia, sobretudo os avanços na Inteligência Artificial e automação, causam impacto direto no mercado de trabalho; a imprevisibilidade do futuro faz com que os possíveis cenários causados pelo aprimoramento tecnológico tragam um sentimento de incerteza para a classe trabalhadora. Acemoglu e Restrepo (2018) mapearam os possíveis cenários para horizontes não tão distantes. A análise feita pelos autores possui um caráter baseado nas tarefas, contrastando com outras correntes econômicas, como a vertente *mainstream*, a qual considera que o aumento de tecnologia serve como um aumento na produtividade dos fatores de produção. Considerando a análise baseada em tarefas, o efeito de substituição de trabalho humano por trabalho automatizado foi mapeado, assim como os efeitos compensadores (aqueles os quais são benéficos para o mercado de trabalho, do ponto de vista social), o qual aumenta a demanda por trabalho em certas tarefas, ou até mesmo aumentando a produtividade do trabalhador. Quatro efeitos foram tirados desse estudo, sendo que apenas o efeito de deslocamento traz impactos negativos no âmbito trabalhista, substituindo, de fato, o trabalho humano pelo trabalho automatizado de robôs; diminuindo a participação da renda do trabalho no agregado nacional, mesmo com o aumento de produtividade do capital humano. Como frisam Acemoglu e Restrepo (2018), existem tarefas executadas por trabalhadores, e outras que podem ser executadas tanto por trabalhadores quanto por capital. E a produtividade desses dependem das vantagens comparativas nas diferentes tarefas que permeia uma organização. É vantajoso para o capitalista substituir trabalho humano por capital se essa troca reduzir seu custo, fazendo com que a automatização cause o efeito de deslocamento.

Apesar do efeito de deslocamento possuir um caráter direto, e negativo, em relação aos impactos dos avanços da tecnologia no mercado de trabalho, Acemoglu e Restrepo (2018) apresentam os efeitos, chamando-os de “efeitos compensadores”, que impactam positivamente o emprego e a taxa salarial; sendo eles: o efeito produtividade, a acumulação de capital e o aprofundamento em automação.

Ao reduzir o custo de produção de bens, através de certas tarefas que ficam mais baratas com a automatização, os bens finais também ficam mais baratos, deixando a sociedade mais “rica”, no quesito poder de compra, aumentando a demanda por todos os produtos e

serviços da sociedade, como ressalta Autor (2015). Esse aumento real na renda aumenta a demanda por trabalho em outras indústrias, contrapondo o efeito de deslocamento. Um exemplo desse ponto aconteceu na mecanização da agricultura tanto os Estados Unidos quanto países europeus; em que os alimentos ficaram mais baratos, fazendo com que os consumidores ficassem mais ricos, em relação ao poder de compra, e demandassem mais produtos feitos por indústrias, Brynjolfsson e McAfee (2014).

No quesito intensificação de capital, a automação implica no maior acúmulo de capital, e quanto maior a demanda por capital, maior a demanda por trabalho para operar tal capital. Nessa abordagem mais neoclássica de Acemoglu e Restrepo (2018), os autores comentam que tal acumulação pode ser suficientemente impactante no aumento dos salários no longo prazo. Efeitos observados, por exemplo, na economia britânica durante a revolução industrial, em que o acúmulo de maquinário aumentou a necessidade de operadores para tais máquinas, aumentando a demanda por trabalho, Allen (2009).

O efeito de deslocamento substitui o trabalho de um homem pelo trabalho de uma máquina/robô. Mas, o aprofundamento de automação não gera tal efeito, pois substitui o trabalho de um robô por um robô de tecnologia mais recente. Essa evolução tecnologia aumenta a produtividade da tarefa, trazendo o benefício do efeito mencionado anteriormente, o efeito produtividade. Olmstead e Rhode (2001) exemplificam tal substituição de tecnologia na agricultura americana, ao substituir o arado puxado a cavalo por tratores movidos à diesel.

Além dos efeitos citados anteriormente, a criação de novas tarefas tem um grande impacto no crescimento do emprego. Acemoglu e Restrepo (2018) consideram o efeito das novas tarefas como efeito de reintegração da mão-de-obra no mercado de trabalho, aumentando a demanda por emprego e a participação da renda do trabalho na renda nacional.

Nas revoluções anteriores, tarefas como engenheiros e gerentes foram criadas para suprir as novas necessidades criadas pelas novas tecnologias, Landes (1969). Especificamente em relação à inteligência artificial, novas tarefas são criadas como: treinadores de modelos (pessoas que treinam modelos de inteligência artificial), pessoas que explicam modelos (intermediando a linguagem técnica por trás do modelo para os clientes), e as pessoas que sustentam os modelos (monitoram a performance e fazem reparos nos modelos), Acemoglu e Restrepo (2018).

Por mais que as tarefas executadas pela inteligência artificial esteja cada evoluindo à medida que o tempo passa, essas tarefas são de caráter lógico, em que o poder computacional

supera drasticamente a capacidade humana, como previsão, tomadas de decisão e logística, Nilsson (2009). Entretanto, há diversos aspectos da natureza humana que a inteligência artificial não consegue replicar, como: empatia, resolução abstrata de problemas, aprendizado baseado em analogias, como ressaltam Acemoglu e Restrepo (2018).

Outro aspecto de grande relevância para o tema é o tipo de tarefa que a inteligência artificial é capaz de substituir; Agrawal, Gans e Goldfarb (2019) definem que a inteligência artificial impactará as tarefas de caráter preditivo, as quais os humanos conseguem executar. Afirmam que as tarefas de cunho preditivo serão substituídas pela inteligência artificial, enquanto as demais tarefas tornar-se-ão orientadas através de predições feitas, aumentando a qualidade das tarefas. Os autores deixam claro a distinção entre tarefas preditivas e tarefas de ação, em que a tarefa preditiva tem caráter auxiliar nas tomadas de decisões. Outro ponto é a capacidade limitada do corpo humano. A inteligência artificial pode substituir, de forma eficaz, tarefas que exigem reflexo e precisão cirúrgica, por exemplo. A inteligência artificial, tem a capacidade de substituir ou auxiliar tarefas, de forma que aumente a produtividade do trabalho, aumentando, em segunda instância, a remuneração do trabalho.

O grau de predição de determinada tarefa é o que vai dizer se ela vai ser substituída ou não, segundo Agrawal, Gans e Goldfarb (2019), fazendo com que o valor da ocupação diminua ou seja substituída. Efeitos indiretos, como a demanda por trabalhos na cadeia dessas modificações, podem acontecer, aumentando ou diminuindo o valor do trabalho. De acordo com o tipo de tarefa que é executada, quatro possíveis impactos com o implemento de inteligência artificial nas empresas: automação das tarefas de predição, automação nas tarefas de decisão, aumento do desempenho do trabalho nas tarefas de decisão, aumento do desempenho do trabalho em outras tarefas. A automação em determinadas tarefas aumenta a produtividade do trabalho, pois a máquina executa de forma melhor tarefas que são executadas por trabalhadores, mas possuem a limitação do ser humano, como o reflexo para a tomada de decisão, sejam essas tarefas predições na área do direito, medicina, tecnologia, ou em qualquer área que a predição seja um elemento fundamental. No que tange a automação das tarefas decisórias, a inteligência artificial também possui arcabouço para o aumento de produtividade; esse tipo de automação quando ambas as tarefas, tanto preditiva quanto decisiva, são automatizadas, substituem o fator de produção humano. Em ambientes mais simples, as decisões baseadas em máquinas são de caráter binário, e tentam replicar o comportamento

humano. Com o avanço da inteligência artificial, ambientes mais complexos tornam-se capazes de serem previstos.

Agrawal, Gans e Goldfarb (2019) exprimem esse impacto como o aumento de produtividade nas tarefas preditivas que auferem um aumento de produtividade nas tarefas decisivas, diminuindo erros cometidos pelos trabalhadores humanos, elevando as remunerações do trabalho. Ao aumentar a produtividade das tarefas preditivas com a inteligência artificial, as tarefas decisivas tornam-se melhores também. Esse aumento de produtividade em ambas essas tarefas faz com que tarefas anteriores ou posteriores também tenham aumentos de produtividade, elevando a produtividade da cadeia como um todo.

2.2 EFEITOS NA DISTRIBUIÇÃO DE RENDIMENTOS NO MERCADO DE TRABALHO

Por mais que a produtividade da cadeia como um todo aumente, novos padrões ao longo da distribuição de rendimento são percebidos, mudando totalmente a dinâmica extraempresa e intraempresa.

Stokey (2016) comenta sobre a participação da evolução tecnológica no aumento da desigualdade salarial, sobretudo nas últimas três décadas. Tal desigualdade baseia-se em dois fatores:

- A polarização no mercado de trabalho, um padrão que expande tanto o nível de emprego quanto os salários relativos nos extremos da distribuição do mercado de trabalho, em contraste com o centro dessa distribuição. Assim como Autor, Mindell e Reynolds (2019) ressaltam, essa polarização de tal distribuição faz com que as tarefas de capacidade técnica média sejam substituídas por inteligência artificial, enquanto há um aumento na demanda pelos trabalhos com pouca qualificação ou com muita qualificação. Entretanto, a polarização acontece de forma não uniforme, uma vez que os empregos com pouca qualificação não têm sua remuneração acompanhada pelo aumento de produtividade que o avanço tecnológico traz, enquanto os cargos que exigem mais nível educacional, acompanham o aumento de produtividade, aumentando a desigualdade entre os trabalhadores nos dois extremos da distribuição.

- A desigualdade salarial entre firmas, no que tange cargos semelhantes, aumenta mais do que a desigualdade salarial entre cargos da mesma empresa; mostrando que a desigualdade

não se dá apenas de uma forma vertical, hierárquica, como também há uma desigualdade dentro da própria indústria.

Stokey (2016) ainda salienta que a mudança em uma tecnologia, faz com que toda cadeia envolvida com determinado avanço seja realocada de alguma forma, alterando assim, os preços dos bens e os salários. Tal transformação do paradigma das tarefas presentes no mercado impactam toda a cadeia produtiva, ao passo que tarefas antes feita por humanos sejam executadas por robôs, e melhorando o desempenho de tarefas anteriores ou posteriores do processo produtivo, aumentando a demanda por trabalhadores e mudando as proporções salariais em toda cadeia.

2.3 PROBLEMAS ESTRUTURAIS

Ao abordar em que a sociedade deve focar-se, no que tange à investimento em educação, Autor, Mindell e Reynolds (2019) deixam claro que não se deve fazer investimentos intensivos na qualificação de ocupações na faixa média da distribuição do mercado de trabalho, uma vez que existe o fator substituição de mão-de-obra pelo aposento dos trabalhadores nessa função, prática que as firmas fazem constantemente. Analisando o mercado americano, a instituição The Bureau of Labor Statistics estimou que, entre 2016 e 2026, cerca de quatrocentos mil operários serão substituídos por robôs, ao passo que perto de dois milhões de operários estarão perto de se aposentar; tal fato cria um déficit de, aproximadamente, um milhão e meio de trabalhadores. O envelhecimento da força de trabalho parece ser um padrão global, tal fato deve-se a diminuição expressiva da taxa de crescimento do mercado de trabalho. Através da abordagem demográfica, Autor, Mindell e Reynolds (2019) vão em sentido oposto à narrativa convencional, antecipando que nos países industrializados, nas próximas décadas, existirão mais vagas de trabalho do que trabalhadores para preenche-las. Com isso, a automação tem um papel crucial para um futuro à médio prazo.

Além do problema demográfico discutidos anteriormente, Acemoglu e Restrepo (2018) apontam outro problema estrutural que é o descasamento entre esse avanço e as habilidades dos trabalhadores disponíveis no mercado. Essas novas tarefas, criadas por novas tecnologias, necessitam de trabalhadores com habilidades específicas, e, a falta de profissionais habilitados para tais tarefas faz com que esse ajuste seja comprometido, segurando assim os potenciais de ganho e atrasando os avanços na produtividade. Essa falta de trabalhadores

qualificados eleva os salários das pessoas com altos níveis de estudo apontados por Stokey (2016).

2.4 DINÂMICA NO MERCADO DE TRABALHO AO LONGO DA DIFUSÃO DE NOVAS TECNOLOGIAS

Por mais que esse aumento de produtividade advinda da melhora tecnológica reduza custos, aumentando a produção como um todo, duas características são inerentes à difusão da tecnologia na sociedade: a lentidão da difusão de forma geral e o nível de aceitação das diferentes invenções, Rosenberg (1972); uma vez que inovação deriva de uma invenção que foi transformada num produto ou serviço, tendo fins comerciais, Kupfer e Hasenclever (2013). Essa lentidão citada por Rosenberg (1972) é causada pelo desconhecimento do mercado em relação as novas invenções, uma vez que uma decisão errada ao implementar uma nova tecnologia, e caso está não seja bem sucedida no mercado, causaria custos irrecuperáveis para a empresa. Além desse tópico, à medida que a tecnologia vai se estabelecendo no mercado, novos ambientes adotam o novo paradigma, reformulando tal avanço tecnológico, adaptando e melhorando a produtividade em outros setores, Hall (2004).

Esse aumento faz com que diminua a relação entre trabalho e unidade de produto, ou seja, menos trabalho é necessário na obtenção de uma unidade de bem. Como Bessen (2018) ressalva, o preço dos bens reduzidos por causa da automatização aumenta a demanda. A automatização não só reduz os preços, como aumenta a qualidade dos bens. Tais fatores aumentam a demanda pelos produtos, fazendo com que a necessidade de trabalho aumente, mesmo com a automatização. Com isso a demanda por trabalhador aumenta nos estágios iniciais do avanço tecnológico, compondo a primeira parte da curva, “U”, de demanda por trabalho.

Ao satisfazerem suas necessidades, os consumidores, por estarem mais ricos, no quesito poder de compra, migram suas preferências para bens superiores, fazendo com que a demanda pelos bens beneficiados pelo aumento de produtividade, diminua, tornando a curva de demanda por trabalhadores decrescente, na segunda parte do processo. Com isso, Bessen (2018) descreve o padrão de “U” invertido que a curva do mercado de trabalho tem como característica nesse movimento de avanço tecnológico.

2.5 DIFERENÇA DESSA REVOLUÇÃO TECNOLÓGICA PARA AS ANTERIORES

Um ponto na discussão acadêmica sobre o tema na atualidade é o quão diferente a revolução tecnológica que passamos, em relação às transformações passadas. Autor, Mindell e Reynolds (2019) deixam claro que, nas eras anteriores, a automação e mecanização acabaram por remover trabalhos prejudiciais, seja por esforço repetitivo ou por características insalubres, ao passo que aumentaram a produtividade e os padrões de vida da sociedade como um todo. Entretanto, dois pontos divergem dos impactos passados, são eles: a polarização do mercado de trabalho e as tecnologias “mais ou menos”, como denominaram.

Ao comparar a era de produção em massa com a era tecnológica atual, os autores mostram a melhora nos rendimentos para os operários, que segundo a literatura, estão categorizados como empregos de habilidades medianas, dentro da distribuição de empregos no mercado de trabalho; contrastando com a quarta revolução tecnológica, em que a inteligência artificial tem um impacto significativo na substituição nessa parte da distribuição, criando uma polarização da curva de distribuição de empregos. Essa polarização remete ao aumento dos trabalhos de habilidades mais básicas, as quais possuem características que os robôs não conseguem reproduzir (como empatia e flexibilidade nas tomadas de decisão sob condições diversas), e o aumento dos trabalhos que exigem um nível maior de proficiência. Esse aumento nas pontas da distribuição tem como contraparte a diminuição nas tarefas de habilidades intermediárias. Contudo, em relação à maior igualdade na distribuição desse aumento de produtividade não acompanha essa polarização de forma uniforme; enquanto a remuneração dos cargos mais habilidosos acompanha os avanços na produtividade, os trabalhos de baixa complexidade não tiveram esse mesmo direcionamento, fazendo com que a renda desse tipo de ocupação permanecesse estagnada. Essa dinâmica traz à tona uma das principais indagações contemporâneas, que é: como contornar a desigualdade nesse ambiente que o atual cenário mundial se encontra?

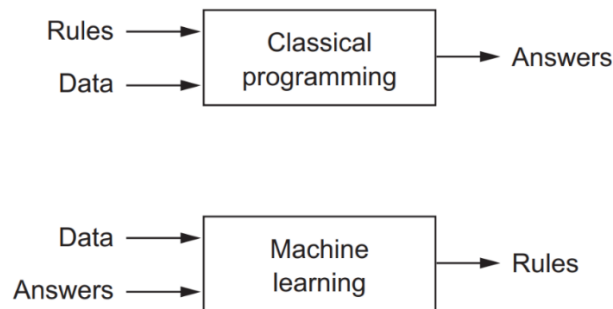
Em relação às tecnologias “mais ou menos”, Autor, Mindell e Reynolds (2019) deixam claro que no mercado americano, essa prática parece ter mais envolvimento com o sistema fiscal do país do que com o mercado de trabalho em si. O custo, em relação a taxas e impostos, de um trabalhador é substancialmente maior do que o custo do capital, fazendo com que empresas tomem algumas manobras para contornar tais medidas governamentais, substituindo trabalho humano por capital. As tecnologias são chamadas de “mais ou menos”

pois não acrescentam incremento algum na sociedade, pelo contrário, são implementadas ao custo de empregos de vidas humanas, causando transtornos sociais. Para a empresa e pro consumidor em si, tais mudanças são bem-vindas, uma vez que reduz o custo, para a empresa, e o preço, para o consumidor final; porém num contexto social geral, como os próprios autores sugerem, uma transformação na lei fiscal poderia contornar esse problema, trazendo mais vagas empregatícias para a população como um todo.

2.6 ESCOPO DA INTELIGÊNCIA ARTIFICIAL

Chollet (2017) define inteligência artificial como “o esforço para automatizar tarefas normalmente performadas por humanos”. Através da criação de algoritmos (que, segundo o dicionário da Google consiste numa “sequência finita de regras, raciocínios ou operações que, aplicada a um número finito de dados, permite solucionar classes semelhantes de problemas”), uma unidade computacional tem a capacidade de exercer funções para qual foi programada. COMO REFERENCIAR? Entretanto, a abordagem algorítmica decompõe e inteligência artificial (IA) em conjuntos distintos, como Godfellow, Bengio e Courville (2016) expõem, são dois grandes grupos: a IA tradicional e o aprendizado de máquina. A grande diferença entre esses subgrupos consiste em sua lógica, enquanto na IA tradicional, os dados de entrada e as regras são fornecidos para uma unidade computacional, a qual retorna uma saída, na IA baseada em aprendizado de máquina, tanto os dados de entrada quanto os de saída são fornecidos, assim, a unidade computacional, através de algoritmos, gera o conjunto de regras que estabelece relações entre as entradas e saída; depois de obter as regras, o algoritmo é capaz de prever as saídas relacionadas a novos dados inseridos. A **figura 1** mostra a diferença de paradigma entre os subconjuntos da inteligência artificial, recebendo dados rotulados, ou seja, o criador do modelo especifica qual resposta está atrelada a cada dado de entrada, e o algoritmo de aprendizado de máquina encontra as regras que mapeiam a conexão entre esses dados.

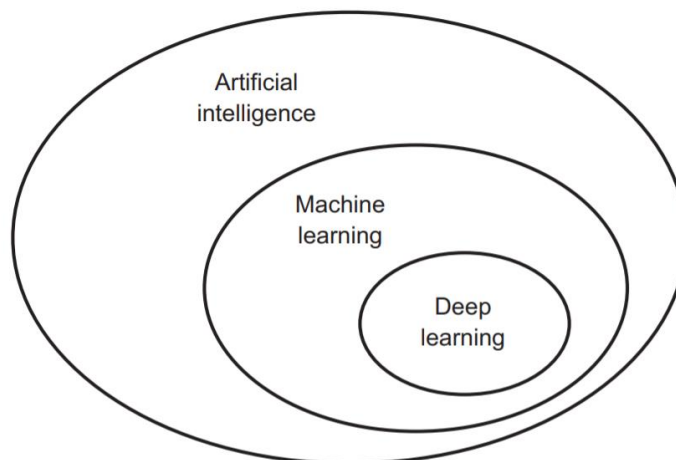
Figura 1 - Diferença entre o fluxo da programação clássica e do aprendizado de máquina



Fonte: Deep Learning with Python (2017).

Dentro do conjunto de aprendizado de máquina encontra-se o subcampo “aprendizagem profunda”. O aprofundamento dentro do grupo de inteligência artificial, chegando no aprendizado profundo, traz conjuntos para resolver problemas mais complexos. Chollet (2017) relata que o algoritmo (de aprendizado de máquina), *gradient boosting machines* “é um dos melhores, se não o melhor” para tarefas não perceptuais, isto é, aprendizado no qual os dados inseridos no modelo estão num formato estruturado, enquanto a aprendizagem profunda tem a especialidade de trabalhar com dados não-perceptuais, exercendo tarefas mais complexas para as máquinas, porém costumeiras para o seres humanos, como funções que usam o artifício da visão, por exemplo; esse ponto faz com que as máquinas consigam performar melhor que humanos num escopo maior de tarefas. A **figura 2** exibe um diagrama mostrando os subgrupos de inteligência artificial até chegar em aprendizado profundo; o diagrama apresentado não exibe o subgrupo *representation learning*, que está entre aprendizado de máquina (*machine learning*) e aprendizagem profundo (*deep learning*), pois o exemplo prático envolverá aprendizado profundo, que na comunidade tecnológica, é contrastado primordialmente com aprendizado de máquina.

Figura 2 - Conjunto da Inteligência Artificial



Fonte: Deep Learning with Python (2017).

Dois fatores fazem com que a aprendizagem profunda seja mais robusta que o conjunto o qual se encontra, fazendo com que esse tipo de algoritmo consiga concluir tarefas mais complexas, do ponto de vista computacional, são eles:

- Representação: o algoritmo é capaz de criar suas próprias representações dos dados, encontrando padrões entre eles sem a ajuda de um especialista. No aprendizado de máquina, a maior parte do tempo de um engenheiro de modelo consiste em estruturar os dados para o treino do modelo; a qualidade dos dados é tão importante quanto o modelo em si. Já em *deep learning*, o próprio algoritmo encontra representações do modelo, que muitas vezes está fora da compreensão do especialista por trás do modelo, fazendo com que o desenvolvedor passe mais tempo otimizando o modelo, e não os dados os quais serão inseridos, Krohn, Beyleveld e Bassens (2019)

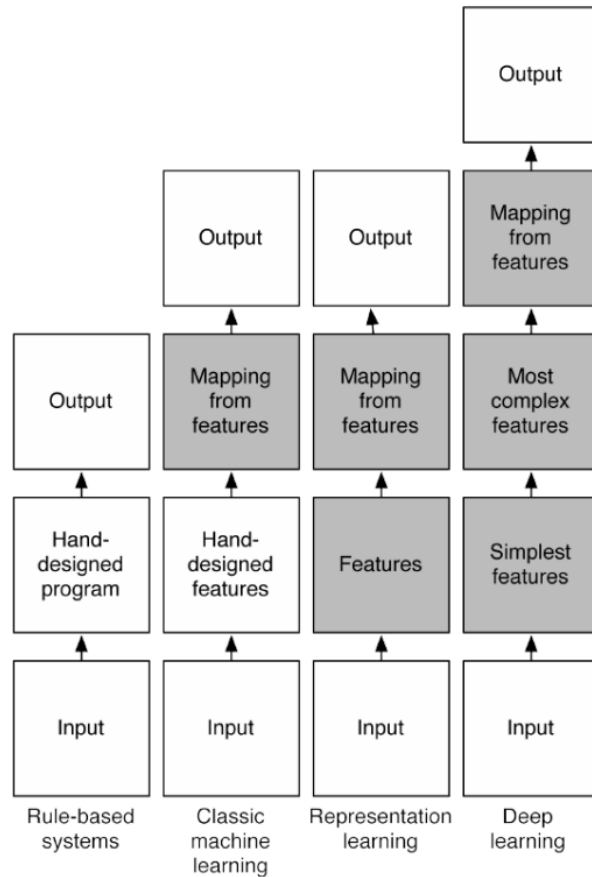
- Profundidade: o outro fator para a robustez do modelo é o número de camadas que o modelo pode ter, fazendo com que o modelo tenha uma quantidade enorme de parâmetros, que se ajustam de forma mais precisa aos dados, Chollet (2017). A sequência de camadas também possui uma lógica: as camadas do início do modelo extraem dos dados atributos mais simples, e a medida que avança na corrente de camadas, os atributos extraídos pelo modelo possuem mais complexidade, Godfellow, Bengio e Courville (2016).

Essa mudança de paradigma dentro das inteligência artificial faz com que os profissionais envolvidos com o desenvolvimento de modelo passem mais tempo criando e melhorando o modelo em si, ao invés de passar a maior parte do tempo trabalhando com os

dados (seja limpando os dados, estabelecendo relações entre eles, solucionando problema com dados omissos), através de suas próprias expertises, a fim de inserir dados de alta qualidade no modelo; além do mais, o modelo pode encontrar representações nos dados que não fazem parte do escopo da ciência atualmente, tendo a possibilidade de superar o desempenho humano nesse quesito, Krohn, Beyleveld e Bassens (2019).

A **figura 3** mostra a diferença dos conjuntos dentro do escopo da inteligência artificial, em relação às informações extraídas pelo modelo. O modelo básico de regras feitas pelos humanos recebe os dados, passa pelo modelo, e retorna uma saída. No segundo modelo, aprendizagem de máquina, o engenheiro organiza os dados e insere as características que deseja que o modelo relacione os dados com seus rótulos, descobre as regras e consegue encontrar saídas para novos dados. O terceiro modelo, aprendizado por representações, o próprio modelo gera as características que relacionam os dados com seus rótulos e é capaz de prever saídas para novos dados. Enquanto no último modelo, aprendizagem profunda, o modelo gera características simples nas primeiras camadas, passando para características mais complexas ao longo do modelo, relaciona os dados de entrada com os dados de saída (estabelecendo as regras entre os dados e seus rótulos) e consegue prever novas saídas para novos dados inseridos no modelo já treinado.

Figura 3 - Diferença entre características de subconjuntos da Inteligência Artificial



Fonte: Deep Learning (2016).

O avanço dentro da inteligência artificial só foram possíveis pela evolução das placas gráficas, pautada fundamentalmente pelos avanços da indústria jogos eletrônicos, placas as quais possuem característica de executar cálculos de forma paralela, ao contrário das unidades de processadores, que calculam de forma sequencial, exigindo que uma conta seja executada somente após o término do cálculo anterior; tal fato torna possível a criação de modelos robustos com diversas camadas e de complexidade superior aos modelos de aprendizado de máquina. A evolução nessa indústria está em tanta evidência que a Google desenvolveu uma unidade de processamento especificamente para o desenvolvimento de modelos muito grandes, e batizou de unidade de processamento de tensores (pelo fato do *deep learning* ser baseado em tensores – matrizes com qualquer número de dimensões), recomendando para modelos para os quais até mesmo as placas-de-vídeo demorem para rodar.

3 METODOLOGIA

3.1 EM RELAÇÃO À METODOLOGIA ACADÊMICA

Este trabalho apresenta, como metodologias de análise, a pesquisa exploratória. Segundo Prodanov e Freitas (2013, p.52), “possui planejamento flexível, o que permite o estudo do tema sob diversos ângulos e aspectos.” A metodologia foi escolhida, também, pelo fato de o tema em questão possuir uma natureza dinâmica, fazendo com que as informações acerca dessa indústria estejam sempre em constantes mudanças.

A análise possui um caráter qualitativo, trazendo números apenas no que tange o modelo na penúltima seção desse texto.

Quanto as fontes usadas, a pesquisa bibliográfica é utilizada no referencial teórico, enquanto a pesquisa documental é usada na parte aplicada do projeto. Essa escolha se dá pela diferença entre os dois tipos, que segundo Gil (2008), está no fato da contribuição ter recebido ou não um tratamento sobre o assunto analítico de outros autores.

As limitações do trabalho são basicamente duas: o escopo do objeto estudado e o dinamismo do assunto. Quanto ao escopo, os impactos da inteligência artificial como um todo não são analisadas, porém, um modelo específico de aprendizagem profunda, que está na vanguarda do tema, foi utilizado. Como a inovação e transformação dentro desse assunto são muito aceleradas, esse trabalho pode tornar-se obsoleto, em relação à tecnologia, num futuro não tão distante, ao contrário dos impactos sociais apresentados no referencial teórico, que possuem um espectro duradouro.

3.2 EM RELAÇÃO AO MODELO

O modelo utilizado no exemplo é baseado em aprendizagem profunda, como visto anteriormente, um subconjunto de aprendizado de máquina. A criação de modelos de *deep learning* consistem em três etapas:

- Transformar os dados no formato que serão inseridos dentro do modelo. Ao contrário do *machine learning*, essa transformação a expertise de encontrar relação entre os dados, apenas manipulá-los para que o formato dos tensores seja compatível com o formato de entrada aceito pelo modelo.

- Definição das camadas: os avanços em aprendizagem profunda proporcionaram a criação de diversas camadas distintas para problemas distintos, sejam eles problemas de visão computacional, processamento de linguagem natural, problemas sequenciais, reconhecimento de fala, tradução, sistema de recomendação, entre demais funções as quais podem ser contempladas por modelos de redes neurais (esse nome deve-se à comparação dos modelos com o funcionamento do cérebro humano, em que a célula cerebral- o neurônio- serviu de inspiração para a criação da ‘célula’ de inteligência das máquinas, em que cada unidade computacional representa um neurônio), Krohn, Beyleveld e Bassens (2019).

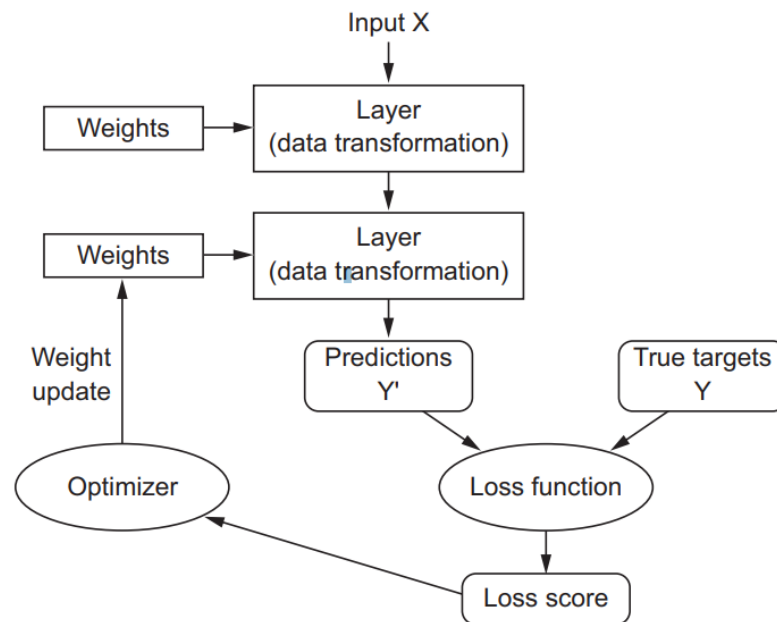
- Definições das:

- o Função de perda do modelo;
- o Otimizador;
- o E medidores de desempenho.

Com o modelo definido, os dados de treino são utilizados para treinar o modelo, seguido da validação com dados distintos, a fim de generalizar o modelo para dados ainda não vistos antes, para não cair no viés de *overfitting*, isto é, ‘decorar’ os padrões dos dados de treino e não performar bem nos demais dados, Géron (2019).

Já o funcionamento do modelo consiste numa série de loops de melhoria contínua (composto por duas etapas em cada passagem do loop), como detalha Chollet (2017), em que os dados de entrada (inputs), passam pelas camadas (*layers*) onde são computados (cada camada tem seus próprios parâmetros), gerando uma saída, a qual é comparada com um valor rotulado pelo criador do modelo através da função de perda (*loss function*) – nas redes neurais existem diversas funções de perda, que são utilizadas de acordo com o problema específico. A distância entre o valor dado pelo desenvolvedor e o valor encontrado pelo modelo é usado para ver o quão preciso o modelo está. Essa primeira etapa é chamada *forward propagation*. A segunda etapa, denominada *back propagation*, faz o processo inverso; capta o valor obtido pela *loss function* e, calcula a derivada (através da regra da cadeia, a essência das redes neurais) dos parâmetros do modelo (camada a camada), e através do otimizador, ajusta os parâmetros do modelo para uma nova tentativa dentro do loop, criando uma melhoria contínua no modelo. A **figura 4** mostra um diagrama com as etapas de uma passagem. É possível visualizar os pesos das diferentes camadas do modelo, assim como o *loss score*, que é a medida de precisão do modelo usada pelo otimizador para ajustar os parâmetros e prosseguir numa nova passagem dentro do loop, o qual o número de passagens é definido pelo criador do modelo.

Figura 4 - Fluxograma de um modelo de Aprendizagem Profunda



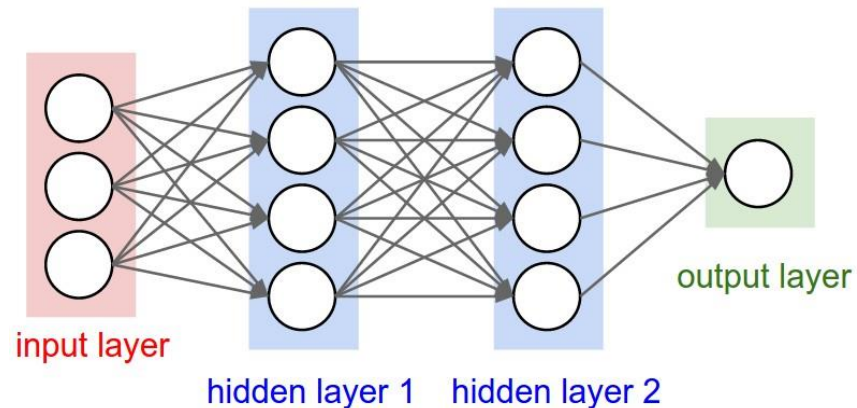
Fonte: Deep Learning with Keras (2017).

3.3 CAMADAS (LAYERS)

Como o exemplo final do trabalho é um modelo de visão computacional, três tipos distintos de camadas serão usados: o *convolution layer*, o *maxpooling layer*, o *dense layer* e o *flatten layer*. Uma camada pode ser interpretada como um filtro para os dados, como argumenta Chollet (2017).

A espinha dorsal de uma rede neural é o *dense layer*. Segundo Krohn, Beyleveld e Bassens (2019), essa camada também é chamada de “totalmente conectada” pois liga todas as unidades um *dense layer* com todas unidades da próxima camada de *dense layer*. As camadas intermediárias são como caixas-pretas para os desenvolvedores, então são chamados de unidades escondidas, sendo as unidades de representação do modelo, Chollet (2017). As unidades escondidas correspondem a uma dimensão no espaço de representação daquela camada. A **figura 5** exemplifica uma rede neural composta por três *dense layers* (os dados de entrada não contam como uma camada) com quatro, quatro e uma unidades (ou neurônios) respectivamente. Percebe-se que são *dense layers* pois todas as unidades de uma camada são conectadas com todas as unidades da camada seguinte.

Figura 5 - Fluxograma de uma rede neural



Fonte: Stackoverflow.com

As unidades dentro de um *dense layer* são denominadas *threshold logic units* (TLU), e segundo Géron (2019), são computadas em duas partes, sendo a primeira o somatório dos dados de entrada (cada um multiplicado por um peso – parâmetros do modelo) e um fator de viés, denominada de fase linear da TLU.

$$z = \left(\sum_{i=1}^{nx} \omega_i x_i \right) + b \quad (1)$$

ω_i : peso atribuído ao dado de entrada na i -ésima posição

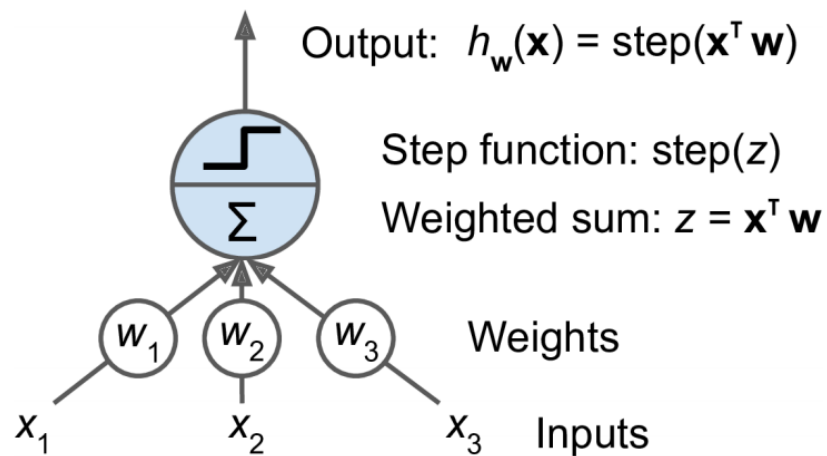
x_i : dado de entrada da i -ésima posição

b : fator de erro

nx : total de variáveis de entrada

A segunda parte consiste na ativação do TLU, isto é, aplica-se uma transformação no resultado da fase linear. Segundo Chollet (2017), ao aplicar uma função de ativação (tornando a saída não-linear), o modelo pode beneficiar-se com a adição de mais camadas, pois aumenta o espaço de hipótese do modelo, o que não seria possível com apenas a fase linear, restringindo assim, os resultados possíveis de ser obtido. A **figura 6** ilustra uma unidade dentro de um *dense layer* e suas fases.

Figura 6 - Processo que ocorre em cada unidade da rede neural



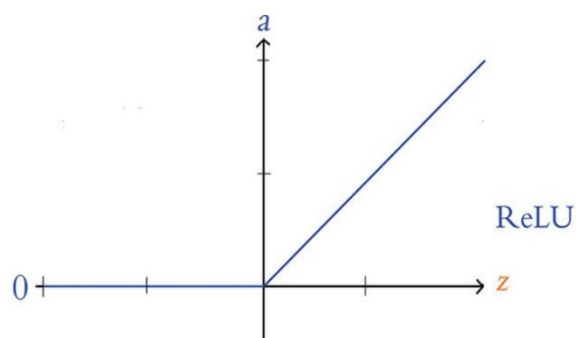
Fonte: Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow (2019).

O modelo exemplo utiliza duas funções de ativações distintas. A primeira delas, chama-se Rectified Linear Unit (ReLU), e, segundo Géron (2019), essa função tem uma derivada que computacionalmente tem um baixo custo, ajudando a reduzir problemas na fase de *back propagation*. Goodfellow, Bengio e Courville (2016) a definem como:

$$a = \max(0, z) \quad (2)$$

O **gráfico 1** mostra a representação gráfica da função ReLU.

Gráfico 1 - Função de ativação ReLU



Fonte: Deep Learning Illustrated (2019).

A segunda função de ativação é chamada de *softmax*. Goodfellow, Bengio e Courville (2016) ressaltam que essa função de ativação retorna um vetor como saída, ao contrário da

função ReLU, definindo-a como “a distribuição das probabilidades de um conjunto finito”, isto é, retorna as probabilidades saída ser determinado rótulo. Sua expressão é:

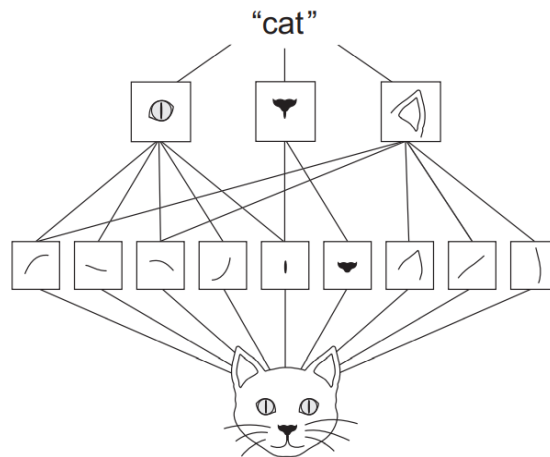
$$\text{softmax}(z) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (3)$$

Sendo o somatório do vetor *softmax* igual a um, pois trata-se de todas as probabilidades (que somadas equivalem a cem por cento), e que cada elemento do vetor é maior que zero.

As outras três camadas do modelo compõe o início da rede neural, e são utilizadas para problemas de visão computacional (um campo em que a máquina consegue, agora, realizar tarefas que antes só era possível para humanos). Krohn, Beyleveld e Bassens (2019) descrevem essa camada como um conjunto de filtros (pequenas janelas) que “scaneiam” a imagem e aplicam uma multiplicação com os determinados pesos (os quais são os parâmetros desse tipo de camada). Se houver, ao menos uma aplicação de filtros nas imagens (que numericamente é composta por uma matriz de pixels), a rede neural é uma rede neural convolucional, Godfellow, Bengio e Courville (2016). Chollet (2017) frisa que as características chaves desse tipo de camada são:

- depois de aprender um padrão num ponto da imagem, consegue encontrar padrões semelhantes em qualquer lugar.
- A sequência de camadas convolucionais seguem uma hierarquia, em que as camadas iniciais aprendem padrões mais simples, e à medida que aumenta a profundidade da rede neural, vai aprendendo padrões mais complexos. A **figura 7** mostra essa característica da camada.

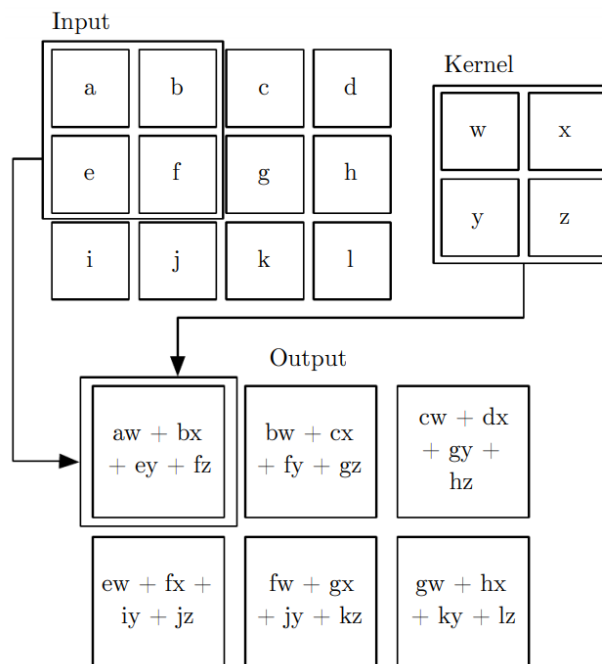
Figura 7 - Representação visual das características obtidas em cada camada de uma rede neural



Fonte: Deep Learning with Keras (2017).

A **figura 8** mostra as operações matemáticas de uma camada convolucional; cada elemento da matriz “scaneada” é multiplicado pelo valor na posição correspondente dentro do núcleo, então soma-se as multiplicações, gerando uma nova matriz (convolução). Nota-se que a dimensão do tensor produzido é menor que a dimensão do tensor de input. Uma camada pode ter diversos filtros, sendo cada filtro uma representação dos dados numa forma única, Krohn, Beyleveld e Bassens (2019).

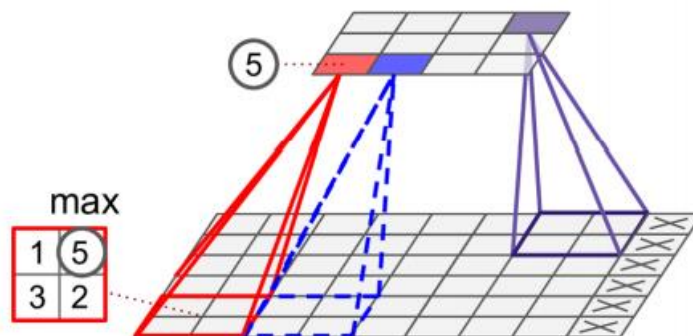
Figura 8 - Representação da camada convolucional



Fonte: Deep Learning (2016).

Uma camada convolucional é seguida, geralmente, por uma camada que reduz a quantidade de parâmetro da camada anterior, Krohn, Beyleveld e Bassens (2019). Existem diferentes tipos de *pooling*, entretanto, será usado o *max-pooling* no modelo do trabalho. Segundo Chollet (2017), *max-pooling* usa como dados de entrada os filtros gerados pela camada convolucional, cria “janelas de extração” dentro desses filtros e retorna os maiores valores dentro de cada “janela”. Seu objetivo é diminuir o consumo de memória computacional (uma vez que reduz o número de parâmetro), além de diminuir o risco de *overfitting*, Géron (2019). A **figura 9** mostra o funcionamento, de forma visual, do *max-pooling layer*, e no exemplo da imagem, a “janela” tem dimensões 2x2, retornando apenas o maior valor, fazendo com que o tamanho da matriz de saída seja 25% do tamanho da matriz de entrada.

Figura 9 - Representação da camada MaxPooling2D



Fonte: Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow (2019).

Por fim, o *flatten layer* transforma um vetor n-dimensional num vetor de apenas uma dimensão, Krohn, Beyleveld e Bassens (2019); essa camada é necessária para passar da parte convolucional do modelo para a parte que contém os *dense layers*.

Outros dois pontos essenciais do modelo são a função de custo e o otimizador.

A função de custo utilizada no modelo é a *cross entropy*, que segundo Géron (2019), serve para estimar a classe determinado exemplo encontra-se. Chollet (2017) escreve que essa função mede o sucesso da tarefa e é nela que se aplica a etapa reversa, sendo a função utilizada para o cálculo das derivadas. Géron (2019) descreve a *cross entropy loss function* como:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log(\hat{p}_k^{(i)}) \quad (4)$$

Em que m é o número total de exemplos utilizados no modelo, K é o número total de classes, i é o i -ésimo exemplo do modelo, k é a classe (ou rótulo) específico.

Segundo Chollet (2017), o otimizador define como os parâmetros da modelo serão atualizados e, de forma gradual, faz com que os parâmetros do modelo atinjam valores que diminuam a *loss function*, aproximando cada vez mais os valores calculados dos valores “alvo” fornecidos. No caso do modelo feito na próxima seção, o otimizador RMSProp será utilizado, que segundo Krohn, Beyleveld e Bassens (2019), serve para ajustar a taxa de atualização dos parâmetros à medida que a precisão do modelo for aumentando, a fim de não fazer um ajuste que possa piorar os parâmetros ao invés de melhorar.

4 MODELO DE DÍGITOS ESCRITOS À MÃO

O modelo que será ilustrado nessa seção do trabalho é uma réplica da rede neural LeNet-5, criada por Yann LeCun e seus companheiros, Krohn, Beyleveld e Bassens (2019). Esse modelo de *deep learning* foi implementado no serviço postal americano, substituindo assim, mão-de-obra humana; trazendo a teoria da inteligência artificial pro contexto social, os quais foram abordados nas seções anteriores, mostrando de uma forma mais geral a capacidade que a inteligência tem impactar no mercado de trabalho. LeCun et al. (1998) mostram com esse modelo que extrair as características das imagens com uma rede neural pode ser mais vantajoso do que extrair essas mesmas características à mão, utilizando redes convolucionais para fazer tal extração de atributos.

A base de dados utilizada para o modelo é famosa (justamente pelo desempenho da rede LeNet-5) na comunidade de inteligência artificial, o banco de dados MNIST (NIST modificado). LeCun et al. (1998) explicam que esse banco é uma modificação do banco NIST (Instituto Nacional de Padrões e Tecnologia), cujos dados de teste treino foram obtidos de duas formas distintas. Os dados para treino foram obtidos através de coletas feitos por alunos de colegial (fazendo com que os dados sejam mais desorganizados), batizando esse banco de dados SD-1. Enquanto o banco de dados para a validação do modelo foi obtido através de empregados da agência censitária americana, batizado de SD-3. O MNIST consiste numa mistura entre os bancos SD-1 e SD-3, possuindo 60000 exemplos como conjunto de treino, e 10000 exemplos para o conjunto de teste, sendo metade de cada valor citado (60000 e 10000) obtido de cada banco de dados. As imagens estão no formato de escala de cinza (preto e branco), e possuem dimensão 28x28 pixels, sendo classificadas em dez categorias distintas (números de 0 a 9); esse exemplo é considerado o problema introdutório na comunidade de inteligência artificial. A **figura 10** mostra exemplos dos dígitos utilizados como entradas no modelo.

Figura 10 - Exemplo de elementos do banco de dados MNIST

```

3 6 8 1 7 9 6 6 9 1
6 7 5 7 8 6 3 4 8 5
2 1 7 9 7 1 2 8 4 5
4 8 1 9 0 1 8 8 9 4
7 6 1 8 6 4 1 5 6 0
7 5 9 2 6 5 8 1 9 7
2 2 2 2 2 3 4 4 8 0
0 2 3 8 0 7 3 8 5 7
0 1 4 6 4 6 0 2 4 3
7 1 2 8 1 6 9 8 6 1

```

Fonte: Gradient-Based Learning Applied to Document Recognition (1998).

A linguagem de programação usada para o desenvolvimento do modelo é a linguagem Python, cuja biblioteca Keras é utilizada para o auxílio de tarefas executadas em *deep learning*. A biblioteca Keras possui o banco de dados MNIST, fornecendo o banco de dados para a execução do modelo.

O modelo conta com cinco camadas próprias para imagem computacional, sendo três *layers* convolucionais e dois de *maxpooling*, intercalados; uma camada *flatten* como intermediária para a progressão pros *dense layers*. A **figura 11** descreve as camadas, os tensores de saída de cada camada além da quantidade de parâmetros totais (e por camadas).

Figura 11 - Camadas utilizadas no modelo e seus parâmetros

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_3 (Conv2D)	(None, 26, 26, 32)	320
max_pooling2d_3 (MaxPooling2D)	(None, 13, 13, 32)	0
conv2d_4 (Conv2D)	(None, 11, 11, 64)	18496
max_pooling2d_4 (MaxPooling2D)	(None, 5, 5, 64)	0
conv2d_5 (Conv2D)	(None, 3, 3, 64)	36928
flatten_1 (Flatten)	(None, 576)	0
dense_2 (Dense)	(None, 64)	36928
dense_3 (Dense)	(None, 10)	650
Total params: 93,322		
Trainable params: 93,322		
Non-trainable params: 0		

Fonte: elaboração própria.

Após treinar por cinco ciclos completos, o modelo atinge uma precisão de 99,41% e um erro de 1,92%, para o conjunto de treino, isto é, os dados que o modelo utilizou para o aprendizado. Entretanto, o modelo, no momento de validação (com os dados de teste), atingiu a acurácia de 99,15%, ligeiramente inferior à marca da performance no conjunto de treino, indicando um pequeno *overfitting*. A **figura 12** ilustra a aprendizagem do modelo, dentro de cada ciclo (*epoch*). Cada ciclo conta com pacote (*batch_size*) de 64 imagens, fazendo com que em cada ciclo, o modelo passe por 938 otimizações de parâmetros por ciclo, otimizando ainda mais o processo.

Figura 12 – Treinamento do modelo

```

model.fit(train_images, train_labels, epochs = 5, batch_size = 64)

Epoch 1/5
938/938 [=====] - 4s 5ms/step - loss: 0.1612 - accuracy: 0.9492
Epoch 2/5
938/938 [=====] - 4s 5ms/step - loss: 0.0465 - accuracy: 0.9859
Epoch 3/5
938/938 [=====] - 4s 4ms/step - loss: 0.0329 - accuracy: 0.9901
Epoch 4/5
938/938 [=====] - 4s 5ms/step - loss: 0.0244 - accuracy: 0.9926
Epoch 5/5
938/938 [=====] - 4s 5ms/step - loss: 0.0192 - accuracy: 0.9941
<tensorflow.python.keras.callbacks.History at 0x2012eb05e80>

test_loss, test_acc = model.evaluate(test_images, test_labels)

313/313 [=====] - 1s 2ms/step - loss: 0.0319 - accuracy: 0.9915

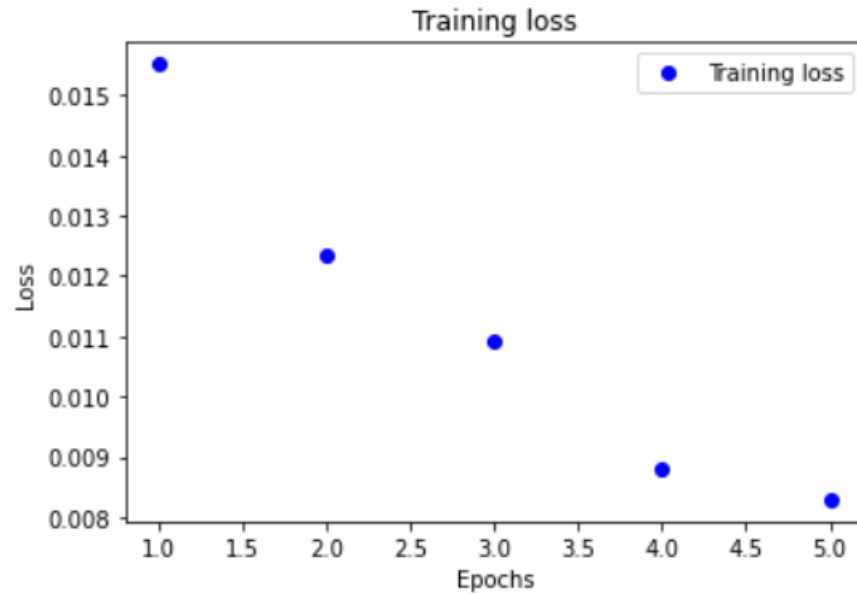
test_acc

0.9915000200271606

```

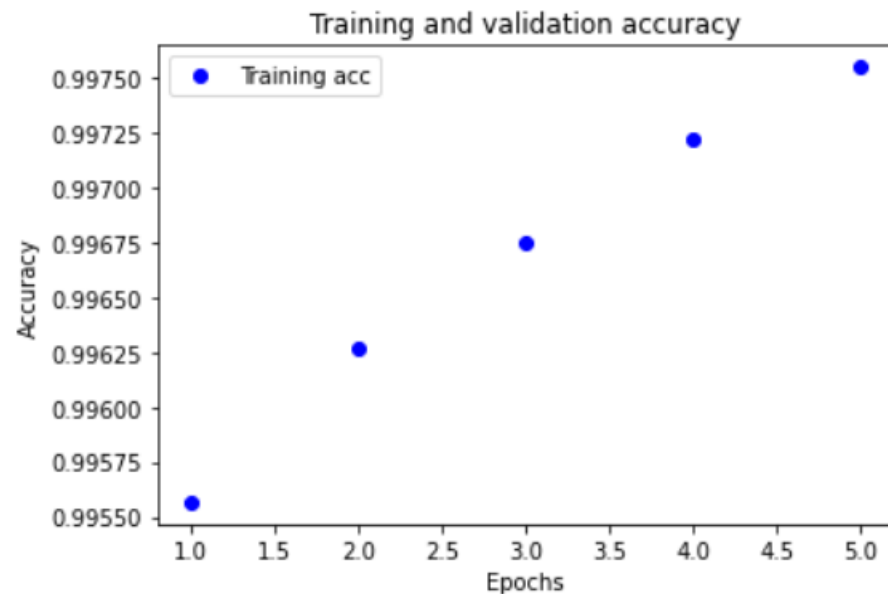
Fonte: elaboração própria.

A evolução das métricas do modelo (em relação a cada *epoch*), pode ser vista nos **gráficos 2 e 3**, em que o **gráfico 2** mostra a evolução do erro a cada ciclo, enquanto o **gráfico 3** exhibe a progressão da precisão do modelo a cada ciclo.

Gráfico 2 - Evolução da *loss* do modelo

Fonte: elaboração própria.

Gráfico 3 - Evolução da precisão do modelo



Fonte: elaboração própria.

A matriz de confusão do modelo associa os valores previstos com os valores reais fornecidos como rótulos ao modelo. Nas colunas da matriz estão os valores reais do modelo, enquanto as linhas são ocupadas pelos valores previstos. A **figura 13** exibe a matriz de confusão do modelo a fim de ilustrar as imprecisões do modelo. Essas imprecisões podem ser usadas (de modo geral) para a melhoria de modelos, uma vez que identifica os problemas que um modelo apresenta.

Figura 13 - Matriz de confusão do modelo

```

array([[ 976,    1,    1,    0,    0,    0,    0,    1,    1,    0],
       [   0, 1133,    0,    1,    0,    1,    0,    0,    0,    0],
       [   2,    0, 1024,    0,    1,    0,    1,    4,    0,    0],
       [   0,    0,    1, 1007,    0,    2,    0,    0,    0,    0],
       [   0,    0,    0,    0,  977,    0,    2,    0,    0,    3],
       [   1,    0,    0,    6,    0,  882,    2,    0,    1,    0],
       [   3,    2,    0,    0,    1,    1,  950,    0,    1,    0],
       [   0,    2,    2,    1,    0,    0,    0, 1020,    1,    2],
       [   1,    0,    2,    1,    1,    0,    0,    0,  965,    4],
       [   1,    0,    0,    0,    3,    3,    0,    0,    1, 1001]],
      dtype=int64)

```

Fonte: elaboração própria.

5 CONCLUSÃO

O modelo, apresentado no trabalho, deixa nítido a produtividade que pode ser alcançada através de robôs, que além de serem mais eficientes que um trabalhador humano, conseguem atingir altos níveis de precisão, tornando o homem obsoleto em diversas áreas.

Tanto na justificativa, através de Silva Jr e Cavalcante (2020), quanto com os argumentos em relação à difusão, Hall (2004), produzir trabalhadores com capacidade técnica para conduzir as novas tecnologias parece ser o caminho mais sensato para qualquer país, e uma forma de se aproximar dos países líderes, no que tange países em desenvolvimento.

Por conta da redução de custo que o aumento de produtividade promove, é inevitável a perda de postos de trabalho para aparatos tecnológicos, mesmo que outras funções sejam criadas, diminuindo a aceleração do aumento da taxa de desemprego nos países.

O investimento na capacitação de mão-de-obra, além da promoção da pesquisa e desenvolvimento para ser o caminho a ser tomado por qualquer país que não queira ser passado para trás.

E, talvez o ponto mais importante do trabalho, a máquina evolui cada vez mais, a ponto de conseguir executar tarefas que antes eram executadas essencialmente por humanos, competindo com o homem em mais e mais tarefas. Mostrando que àqueles que não se adequarem, tornar-se-ão obsoletos, encontrando grandes dificuldades para se encaixar na dinâmica atual da sociedade capitalista.

REFERÊNCIAS

- ACEMOGLU, D.; RESTREPO, P. **Artificial Intelligence, Automation and Work**. NBER Working Papers, 2018.
- AGRAWAL, A.; GANS, J. S.; GOLDFARB, A. **Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction**. Journal of Economic Perspectives, 2019.
- ALLEN, R. C. **Engels' Pause: Technical Change, Capital Accumulation, and Inequality in the British Industrial Revolution**. Explorations in Economic History, 2009.
- AUTOR, D. H. **Why Are There Still So Many Jobs? The History and Future of Workplace Automation**. Journal of Economic Perspectives, 2015.
- AUTOR, D. H.; MINDELL, D. A.; REYNOLDS, E. B. **The Work of the Future: Shaping Technology and Institutions**. MIT Fall Report, 2019.
- BESSEN, J. **AI and Jobs: The Role of Demand**. NBER Working Papers, 2018.
- BRYNJOLFSSON, E.; MCAFEE, A. **The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies**. W. W. Norton & Company, 2014.
- CHOLLET, F. **Deep Learning with Python**. Manning, 2017.
- GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**. 2ed. O'Reilly, 2019.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6.ed. São Paulo: Atlas, 2008.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. The MIT Press, 2016, 2016.
- HALL, B. H. **Innovation and Diffusion**. NBER Working Papers, 2004.
- KROHN, J.; BEYLEVELD, G.; BASSENS, A. **Deep Learning Illustrated: A visual, Interactive Guide to Artificial Intelligence**. Addison-Wesley, 2019.
- KUPFER, D.; HASENCLEVER, L. **Economia Industrial: Fundamentos Teóricos e Práticas no Brasil**. 2ed. Rio de Janeiro: Elsevier, 2013.
- LANDES, D. **The Unbound Prometheus**, Cambridge University Press: New York, 1969.
- LECUNN, Y; BOTTON, L.; BENGIO, J.; HAFFNER, P. **Gradient-Based Learning Applied to Document Recognition**. Proc. of the IEEE, 1998.
- NILSSON, N. J. **The Quest for Artificial Intelligence: A History of Ideas and Achievements**. Cambridge University Press, 2009.

PRODANOV, C; FREITAS, E. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2.ed. Nova Hamburgo: Feevale, 2013.

ROSENBERG, N. **Factors Affecting the Diffusion of Technology**. Explorations in Economic History, 1972.

SILVA Jr, G. G.; CAVALCANTE, B. S. **Pandemia e corrida tecnológica**: o Brasil tem como recuperar o tempo perdido? Evidência a partir dos dados de patentes, 1980-2018. In: Ronaldo Pereira Santos; Marcio Pochmann. (Org.). **Brasil pós-pandemia: reflexões e propostas**. São Paulo: ALEXA Cultural, 2020, p. 103-118.

STOKEY, N. L. **Technology, Skill and the Wage Structure**. NBER Working Papers, 2016.