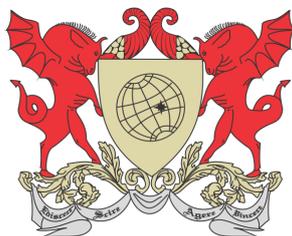


UNIVERSIDADE FEDERAL DE VIÇOSA  
DISSERTAÇÃO DE MESTRADO



RODRIGO CÉSAR MESQUITA GOMES

DESMISTIFICANDO A INTELIGÊNCIA  
ARTIFICIAL PARA ALUNOS DO ENSINO MÉDIO:  
EXPLORANDO CONCEITOS E MATEMÁTICA  
FUNDAMENTAIS

FLORESTAL – MINAS GERAIS  
2024

**RODRIGO CÉSAR MESQUITA GOMES**

**DESMISTIFICANDO A INTELIGÊNCIA ARTIFICIAL PARA  
ALUNOS DO ENSINO MÉDIO: EXPLORANDO CONCEITOS E  
MATEMÁTICA FUNDAMENTAIS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação Mestrado Profissional em Matemática em Rede Nacional, para obter o título de *Magister Scientiae*.

Orientador: Alexandre Alvarenga Rocha

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de  
Viçosa - Campus Florestal

T

G633d  
2024  
Mesquita Gomes, Rodrigo César, 1993-  
Desmistificando a inteligência artificial para alunos do Ensino  
Médio : explorando conceitos e matemática fundamentais / Rodrigo  
César Mesquita Gomes. - Florestal, MG, 2024.  
1 dissertação eletrônica (60 f.): il.

Inclui apêndices.

Orientador: Alexandre Alvarenga Rocha.

Dissertação (mestrado) - Universidade Federal de Viçosa, Instituto  
de Ciências Exatas e Tecnológicas, 2024.

Referências bibliográficas: f. 57-58.

DOI: <https://doi.org/10.47328/ufvcaf.2024.005>

Modo de acesso: World Wide Web.

1. Matemática (Ensino médio). 2. Redes neurais (Computação).  
3. Inteligência Artificial - aplicações educacionais. 4. Ensino. I.  
Rocha, Alexandre Alvarenga, 1984-. II. Universidade Federal de  
Viçosa. Instituto de Ciências Exatas e Tecnológicas. Programa de Pós-  
Graduação em Educação em Ciências e Matemática. III. Título.

CDD 23. ed. 511

**RODRIGO CÉSAR MESQUITA GOMES**

**DESMISTIFICANDO A INTELIGÊNCIA ARTIFICIAL PARA  
ALUNOS DO ENSINO MÉDIO: EXPLORANDO CONCEITOS E  
MATEMÁTICA FUNDAMENTAIS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação Mestrado Profissional em Matemática em Rede Nacional, para obter o título de *Magister Scientiae*.

APROVADA: 14 de março de 2024.

Assentimento:

---

Rodrigo César Mesquita Gomes  
Autor

---

Alexandre Alvarenga Rocha  
Orientador

# Dedicatória

---

Dedico este trabalho aos meus amados filhos, Luísa Vaz Mesquita Gomes e Lucas Vaz Mesquita Gomes.

# Agradecimentos

---

Agradeço primeiramente a Deus por ter me dado força e sabedoria para superar mais essa difícil batalha em minha vida, mas que só engrandece o meu sucesso. Obrigado, Senhor por ter iluminado meus caminhos e minha vida para que nunca desistisse.

Aos meus amados pais, Jeane e Celso que sempre me apoiaram e ajudaram com muito amor, carinho, compreensão e que em nenhum momento mediram esforços para que eu chegasse até aqui. Obrigado por tudo, amo vocês infinitamente.

Ao meu irmão, Renan pelo companheirismo, ajuda e disposição em me ajudar nesses anos que passei e por todos que ainda vamos passar. Conto sempre com seu amor e carinho, para juntos, superarmos todas as etapas das nossas vidas.

Gostaria de agradecer especialmente a minha esposa Carolina Pereira Vaz Mesquita, por toda ajuda, companheirismo, que sempre esteve ao meu lado nos momentos mais difíceis, sempre me apoiando e me dando forças para que eu nunca pensasse em desistir. Muito obrigado por tudo, te amo.

Aos meus queridos filhos Luísa Vaz Mesquita Gomes e Lucas Vaz Mesquita Gomes, que são o propósito de tudo.

Ao orientador Prof. Dr. Alexandre Alvarenga Rocha pela paciência, disponibilidade, pelos ensinamentos de vida e matemáticos que jamais esquecerei. Pela dedicação e tempo gasto nesse trabalho e principalmente por ser o professor que me inspira a ser um professor brilhante como ele. Gostaria de agradecer também ao professor coordenador, Luís e aos demais professores, que ao longo desses anos se tornaram pessoas muito importantes na minha vida. Muito obrigado por todos os ensinamentos transmitidos por vocês, pelo apoio, paciência e confiança depositadas em mim.

Jamais chegaria aqui sem todos vocês. A minha caminhada ainda não terminou, mas foi concluída mais uma grande batalha. Minha eterna gratidão a todos que me apoiaram, colaboraram, incentivaram para que esse grande sonho se tornasse realidade.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

# Resumo

---

GOMES, Rodrigo César Mesquita, M.Sc., Universidade Federal de Viçosa, março de 2024. **Desmistificando a Inteligência Artificial para Alunos do Ensino Médio: Explorando Conceitos e Matemática Fundamentais**. Orientador: Alexandre Alvarenga Rocha.

Este estudo se concentra na simplificação dos princípios matemáticos subjacentes à Inteligência Artificial (IA) para alunos do ensino médio e sua influência no campo educacional. Inicialmente, aborda-se o impacto da tecnologia na sociedade, destacando a relevância da IA no processo de transformação da aprendizagem. Os conceitos fundamentais da IA, como aprendizado de máquina e redes neurais artificiais, são explorados junto com sua aplicabilidade no contexto educacional. Diferentes tipos de IA, desde a IA estreita até a consciente artificialmente, são discutidos, enfatizando suas implicações e desafios éticos. Além disso, são apresentados casos reais de uso da IA em diversas áreas, como comunicação, marketing, medicina e indústria, ressaltando os benefícios e impactos na sociedade contemporânea. O estudo busca fornecer uma visão abrangente da IA e sua importância para o ensino de matemática, preparando os alunos para os desafios e oportunidades futuras na interseção entre matemática, tecnologia e educação.

A pesquisa evidencia conceitos importantes no desenvolvimento, treinamento e operação de sistemas de IA, explorando também a interseção entre Matemática e IA. São apresentados os princípios essenciais para compreender a IA, como Cálculo Diferencial e Integral, Probabilidade, Estatística e Geometria Analítica. Além disso, discute-se a evolução histórica das redes neurais, desde seus conceitos iniciais inspirados no cérebro humano até o desenvolvimento de modelos de redes neurais mais complexos.

Por fim, são apresentadas aplicações para desenvolver o aprendizado de máquina e as redes neurais que são ferramentas de extrema importância para o funcionamento de uma IA. Tais aplicações a serem desenvolvidas com alunos do ensino médio tem

o intuito de desmistificar e facilitar o entendimento do que está "por trás" de uma inteligência artificial.

Palavras-chave: Matemática. Ensino. Ciência. Inteligência Artificial. Rede Neural. Educação.

# Abstract

---

GOMES, Rodrigo César Mesquita, M.Sc., Universidade Federal de Viçosa, March, 2024. **How to Write a PROFMAT Dissertation.** Adviser: Alexandre Alvarenga Rocha.

This study focuses on simplifying the underlying mathematical principles of Artificial Intelligence (AI) for high school students and its influence in the educational field. Initially, it addresses the impact of technology on society, highlighting the relevance of AI in the learning transformation process. Fundamental concepts of AI, such as machine learning and artificial neural networks, are explored along with their applicability in the educational context. Different types of AI, from narrow AI to artificially conscious, are discussed, emphasizing their implications and ethical challenges. Additionally, real-world cases of AI usage in various areas such as communication, marketing, medicine, and industry are presented, highlighting the benefits and impacts on contemporary society. The study seeks to provide a comprehensive view of AI and its importance for teaching mathematics, preparing students for future challenges and opportunities at the intersection of mathematics, technology, and education.

The research highlights important concepts in the development, training, and operation of AI systems, also exploring the intersection between Mathematics and AI. Essential principles to understand AI, such as differential and integral calculus, probability, statistics, and analytical geometry, are presented. Additionally, the historical evolution of neural networks is discussed, from their initial concepts inspired by the human brain to the development of more complex neural network models.

Finally, applications are presented to develop machine learning and neural networks, which are extremely important tools for the operation of AI. Such applications to be developed with high school students aim to demystify and facilitate the understanding of what is "behind" artificial intelligence.

Keywords: Mathematics. Teaching. Science. Artificial Intelligence. Neural Network.  
Education.

# Lista de Figuras

---

4.1	Modelo Perceptron . . . . .	37
4.2	Modelo do neurônio artificial . . . . .	37
4.3	Rede não recorrente feedforward. . . . .	38
4.4	Exemplo Separação de Um Neurônio . . . . .	42
4.5	Exemplo Classificação com Um Neurônio . . . . .	43
4.6	Exemplo Funções ReLU . . . . .	44
4.7	Exemplo Uma Camada e Três Neurônios . . . . .	44
4.8	Exemplo Duas Camadas e Sete Neurônios . . . . .	45
5.1	Exemplo Inicial NIM . . . . .	48
5.2	Exemplo Montagem Caixa de Fósforos . . . . .	49
5.3	Gráfico dos Pontos Gerados Pelos Dados dos Motoristas . . . . .	51
5.4	Gráfico com Um Neurônio . . . . .	52
5.5	Gráfico com mais Neurônios . . . . .	52
A.1	Montagem Caixas de Fósforos . . . . .	59
B.1	Tabela de Dados dos Motoristas . . . . .	60

# Sumário

---

<b>1</b>	<b>Introdução</b>	<b>13</b>
1.1	Inteligência Artificial . . . . .	14
1.2	Conceitos . . . . .	15
1.3	Tipos de IA . . . . .	18
1.4	Aplicações . . . . .	20
<b>2</b>	<b>Fundamentos Matemáticos Para a Compreensão da IA</b>	<b>22</b>
2.1	Cálculo . . . . .	22
2.1.1	Gradiente . . . . .	24
2.2	Probabilidade e Estatística . . . . .	25
2.3	Geometria Analítica . . . . .	26
2.3.1	Equações de Retas . . . . .	27
2.3.2	Equações do Plano . . . . .	30
<b>3</b>	<b>O Uso da Tecnologia Voltada ao Aprendizado e a Educação</b>	<b>31</b>
3.1	Aplicações de IA para Educação . . . . .	31
<b>4</b>	<b>Redes Neurais</b>	<b>35</b>
4.1	Contexto Histórico . . . . .	39
4.2	Fundamentos Matemáticos de uma Rede Neural . . . . .	40
4.2.1	Definindo Uma Rede Neural Artificial . . . . .	41
4.2.2	Playground do Tensorflow . . . . .	44
4.2.3	Modelando uma função a partir de dados . . . . .	45

<b>5</b>	<b>Aplicação no Ensino Médio</b>	<b>47</b>
5.1	O Jogo NIM . . . . .	47
5.2	Desenvolvendo um Computador para Jogar NIM . . . . .	48
5.3	A Aplicação . . . . .	49
5.4	Aplicação de Redes Neurais na Precificação de Seguros Automotivos .	50
<b>6</b>	<b>Considerações Finais</b>	<b>55</b>
<b>7</b>	<b>Referências Bibliográficas</b>	<b>57</b>
<b>A</b>	<b>Apêndice I</b>	<b>59</b>
<b>B</b>	<b>Apêndice II</b>	<b>60</b>

# Introdução

---

A maneira como nos envolvemos com o mundo ao nosso redor passou por uma transformação significativa devido ao rápido avanço da tecnologia, afetando também nosso processo de aprendizagem e educação. Nos últimos anos, um campo emergente tem se destacado no cenário educacional: a inteligência artificial (IA). Originada da integração entre matemática complexa, tecnologia da informação e análise científica de dados gerados, a IA está redefinindo a forma como abordamos problemas e soluções. Encorajando seus usuários a participar ativamente do processo criativo para as máquinas inteligentes, ela oferece ferramentas poderosas para resolver uma variedade de desafios.

No campo educacional, a inteligência artificial está se mostrando como uma ferramenta poderosa e capaz de transformar o futuro da aprendizagem. O objetivo principal deste projeto é analisar e divulgar os aspectos matemáticos essenciais à compreensão da inteligência artificial em um contexto adequado ao nível educacional do ensino médio.

O objetivo específico é simplificar a complexidade matemática da inteligência artificial e das redes neurais artificiais e a transmitir de maneira clara e compreensível. Os princípios básicos dos algoritmos de aprendizado de máquina, as operações matemáticas fundamentais que sustentam o desempenho desses sistemas inteligentes serão explorados neste trabalho.

Ao adquirir conhecimento sobre os princípios matemáticos da inteligência artificial, buscamos que os estudantes se tornem capazes não só de interpretar o mundo digital à sua volta, mas também de se engajarem na evolução tecnológica presente na sociedade. Esta tarefa proporcionará uma visão abrangente do emocionante campo da inteligência artificial e permitirá compreender os elementos complicados que constituem essa importante ferramenta técnica.

Ao trazer à tona na sala, esperamos prepará-los para os desafios e oportunidades futuras através do despertar do interesse pela intersecção entre matemática, tecnologia e educação.

No capítulo 2, oferecemos um breve relato histórico juntamente com conceitos e fundamentos essenciais para o estudo da área. No capítulo 3, exploramos o uso da tecnologia na educação, destacando sua relevância e impacto. Em seguida, no

capítulo 4, aprofundamos o tema das Redes Neurais, discutindo seus princípios e aplicações.

Finalmente, no capítulo 5, apresentamos duas aplicações práticas. Primeiramente, abordamos o Jogo NIM como um meio para explorar o aprendizado de máquina, proporcionando uma experiência interativa e educativa. Em seguida, demonstramos uma aplicação de Redes Neurais em um contexto de seguradora, ilustrando como essa tecnologia pode ser aplicada em cenários do mundo real. Essas aplicações visam enriquecer a experiência dos alunos do ensino médio, promovendo o entendimento e a prática dos conceitos aprendidos ao longo do trabalho.

## 1.1 Inteligência Artificial

Por mais de 250 anos, o crescimento econômico tem sido impulsionado principalmente por inovações tecnológicas. Entre elas, destacam-se as conhecidas pelos economistas como tecnologias de uso geral, que incluem a máquina a vapor, eletricidade e o motor de combustão interna. Cada uma dessas inovações desencadeou ondas de progresso e oportunidades.

Hoje, a tecnologia de uso geral mais crucial é a inteligência artificial (IA). Esta área da ciência da computação concentra-se no desenvolvimento de máquinas e sistemas computacionais capazes de imitar diversas facetas da inteligência humana. A IA capacita sistemas a planejar, aprender, comunicar-se em linguagem natural, conceber, raciocinar sobre ideias complexas, tirar conclusões mesmo em situações de incerteza e interagir com outros seres, além de agir de forma autônoma para executar essas tarefas sem necessidade de programação ou comandos prévios.

Para John McCarthy, conhecido como criador do título Inteligência artificial:

*[...] esta é a ciência e engenharia de fazer máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionada a tarefa semelhante de usar computadores, para estender a inteligência humana, mas a IA não precisa se limitar a métodos biologicamente observáveis.*  
(MCCARTHY, 2012)

Na década de 1950, um artigo foi publicado por Alan Turing sobre o “teste de Turin”, originalmente conhecido como “o jogo da imitação” (the imitation game), no qual ele propôs a avaliação da capacidade das máquinas em exibir comportamento inteligente. Enquanto participa do teste, um jogador humano interage em linguagem natural com outro ser humano e uma máquina projetada para produzir respostas indistinguíveis das de um ser humano. Se o interrogador não puder distinguir entre a máquina e o humano, então considera-se que ela passou no teste de Turing. O objetivo inicial do cientista era responder à questão: É possível para as máquinas pensar? No entanto, devido à dificuldade em definir claramente o pensamento, ele mudou a pergunta para: “Será que um computador digital pode se sair bem no jogo da imitação”. Finalmente, concluiu-se que se um computador conseguisse enganar um terço de seus interlocutores, fazendo-os acreditar que se tratava de um humano, isso demonstraria que o computador estaria pensando por si próprio.

*“A Inteligência artificial pode ser considerada um guarda-chuva abrangente. Ela se refere a programas de computadores capazes de ‘pensar’, comportar-se e fazer coisas como um ser humano poderia fazê-las” (FOGG, 2017).*

## 1.2 Conceitos

Os conceitos fundamentais da Inteligência Artificial (IA) abrangem várias áreas e técnicas. A sua importância reside na base fundamental que eles proporcionam para o desenvolvimento, compreensão e aprimoramento dos sistemas inteligentes. Alguns aspectos destacam essa relevância:

- **Fundamentação Teórica:** Conceitos sólidos fornecem uma base teórica robusta, permitindo que pesquisadores e desenvolvedores compreendam os princípios subjacentes da IA.

- **Desenvolvimento de Algoritmos:** Conceitos são essenciais para a criação de algoritmos eficientes. A compreensão profunda de conceitos matemáticos e estatísticos é crucial para projetar algoritmos de aprendizado de máquina e outras técnicas empregadas na IA.

- **Interpretabilidade:** A compreensão dos conceitos facilita a interpretabilidade dos modelos de IA. Entender como os modelos tomam decisões é crucial para garantir transparência e confiança, especialmente em aplicações críticas.

- **Resolução de Problemas Complexos:** A IA lida com problemas complexos, e os conceitos ajudam na formulação e resolução eficaz desses desafios, permitindo a construção de sistemas mais inteligentes e adaptáveis.

- **Inovação Contínua:** A IA está em constante evolução, e a compreensão aprofundada dos conceitos é crucial para acompanhar as inovações e contribuir para o avanço da área.

- **Ética e Responsabilidade:** Compreender os conceitos na IA é vital para abordar questões éticas e de responsabilidade associadas ao uso dessas tecnologias. Isso inclui a consideração de preconceitos nos modelos e a avaliação dos impactos sociais.

Os conceitos mais importantes em Inteligência Artificial (IA) são diversos, variando conforme o contexto e a aplicação específica. No entanto, alguns pilares fundamentais são essenciais para compreendermos a natureza e o funcionamento dessa área em constante evolução.

A Inteligência Artificial é um campo interdisciplinar que se baseia em uma variedade de técnicas e abordagens para simular, ampliar ou replicar a capacidade cognitiva humana em sistemas computacionais. Estes conceitos fundamentais desempenham um papel crucial em muitas aplicações da IA, desde sistemas autônomos de condução de veículos até assistentes virtuais em dispositivos móveis.

Vamos explorar alguns desses conceitos-chave:

**Aprendizado de Máquina (Machine Learning):** Machine Learning é um campo da inteligência artificial que capacita os computadores a aprender padrões nos dados e fazer previsões ou tomar decisões com base nessas informações, sem precisar de programação explícita para cada tarefa específica. Os algoritmos de aprendizado

de máquina, ao invés disso, passam por treinamento em conjuntos de dados para identificar padrões e obter informações.

Diversas abordagens podem ser adotadas no aprendizado de máquina, como o aprendizado supervisionado, em que os algoritmos são treinados com conjuntos de dados rotulados, o aprendizado não supervisionado, em que os algoritmos exploram a estrutura dos dados sem rótulos e o aprendizado por reforço onde os algoritmos aprendem a realizar ações em um ambiente para maximizar recompensas.

O processo de aprendizado de máquina geralmente envolve coleta de dados, pré-processamento, seleção e treinamento do modelo, avaliação do modelo e implantação em um ambiente de produção. Este campo tem aplicações em diversas áreas.

Redes Neurais Artificiais (RNAs): As RNAs, compostas por neurônios artificiais dispostos em camadas, têm sua estrutura baseada no cérebro humano e servem como modelos de IA. Em uma variedade de contextos são utilizadas para realizar diversas tarefas: desde a detecção e classificação visual até o tratamento da língua nativa.

Iremos nos aprofundar melhor nesse tópico no capítulo 4, visando uma maior compreensão das Redes Neuras Artificiais que são de extrema importância para a ideia principal deste trabalho, além de serem fundamentais para a compreensão matemática envolvida na IA.

Processamento de Linguagem Natural (PLN): A área de foco do Processamento de Linguagem Natural (PLN) é a interação entre computadores e linguagem humana, sendo esta um campo da inteligência artificial. Capacitar os computadores para compreender, interpretar e até mesmo produzir linguagem humana de maneira eficaz é a sua essência.

O processamento de linguagem natural tem aplicações vastas e afeta vários aspectos da vida diária e do mundo dos negócios. Sistemas de chatbots e assistentes virtuais, como exemplo, utilizam o PLN para compreender e responder perguntas dos usuários. As empresas utilizam processamento de linguagem natural para monitorar a opinião dos clientes em redes sociais e feedbacks sobre os produtos. As ferramentas de tradução automática utilizam o processamento natural da linguagem para tornar mais fácil a comunicação em várias línguas.

O Processamento de Linguagem Natural é utilizado em tarefas como a sumariação de texto, que consiste em resumir grandes volumes de informações de forma concisa e compreensível, e o reconhecimento da fala, possibilitando aos computadores entenderem e transcreverem a linguagem humana.

Muitas maneiras estão sendo transformadas pelo PLN na forma como interagimos com a tecnologia e como os computadores entendem e processam a linguagem humana - essas são apenas algumas delas.

Visão Computacional: A Visão Computacional é uma disciplina da inteligência artificial que permite aos computadores entender e interpretar o conteúdo visual do mundo real. Em vez de depender exclusivamente de dados numéricos, como em muitas outras áreas da computação, a visão computacional lida diretamente com imagens e vídeos, buscando extrair significado e informações úteis dessas mídias visuais.

Além disso, a visão computacional desempenha um papel importante em aplicações

de entretenimento, como jogos de realidade aumentada e filtros de fotos. Também é essencial em automação industrial, onde é utilizada para inspeção de qualidade em linhas de produção e controle de processos. Em resumo, a visão computacional abre um mundo de possibilidades para a interação entre máquinas e o ambiente visual, oferecendo soluções inovadoras para uma variedade de problemas do mundo real.

**Lógica Fuzzy:** A Lógica Fuzzy trata a incerteza e imprecisão de forma mais flexível, sendo uma extensão da lógica clássica (ou lógica booleana). Na lógica fuzzy, os valores de verdade variam em uma escala contínua entre 0 e 1, representando o grau de pertinência de uma afirmação, ao contrário da lógica booleana onde os valores são estritamente definidos como verdadeiros ou falsos (1 ou 0).

O professor Lotfi Zadeh introduziu a lógica fuzzy em 1965 como uma ferramenta para lidar com a ambiguidade e incerteza que são frequentes em muitos problemas do mundo real. A lógica fuzzy permite uma representação mais precisa de conceitos vagos ou imprecisos, em vez de categorizar as coisas de forma binária como na lógica clássica.

A flexibilidade torna a lógica fuzzy particularmente útil em situações onde os dados não são completos, imprecisos ou ambíguos. Ela é usada extensivamente em sistemas de controle, tomada de decisão, reconhecimento de padrões e outras aplicações de Inteligência Artificial e engenharia.

O uso comum da lógica fuzzy é em controladores de sistemas de ar condicionado. Em lugar de só ligar ou desligar o ar condicionado em torno de uma temperatura única, um controlador fuzzy pode gradualmente ajustar a potência do ar condicionado conforme o nível desejado de conforto e as condições ambientais, considerando fatores como temperatura corrente, humidade e preferências dos usuários.

**Algoritmos Genéticos:** Os Algoritmos Genéticos (AG) são uma categoria de algoritmos que buscam otimizar e foram inspirados pelo processo de evolução biológica. Foram criados para solucionar problemas de otimização, busca e aprendizado de máquina. Os algoritmos genéticos têm como ideia fundamental simular a seleção natural e evolução para encontrar soluções ótimas ou aproximadamente ótimas em problemas complexos.

**Sistemas Especialistas:** Sistemas Especialistas são programas de computador que emulam a capacidade de um especialista humano em um domínio específico para resolver problemas complexos. Eles são uma aplicação da inteligência artificial que utiliza um conjunto de regras e conhecimentos específicos para tomar decisões ou fornecer recomendações em uma área particular.

Esses sistemas são construídos com base em um conhecimento especializado fornecido por especialistas humanos no campo de interesse. Esse conhecimento é codificado em um formato que o computador pode entender e manipular. Os sistemas especialistas são projetados para imitar o raciocínio humano e a tomada de decisões, aplicando heurísticas e algoritmos para resolver problemas complexos de maneira eficiente.

Uma característica chave dos sistemas especialistas é a capacidade de explicar o raciocínio por trás de suas recomendações ou decisões. Eles podem fornecer justificativas para suas conclusões, permitindo que os usuários entendam o processo

de pensamento que levou a uma determinada saída.

Esses sistemas encontram aplicações em uma variedade de áreas, incluindo medicina, engenharia, finanças, suporte técnico, entre outros. Por exemplo, um sistema especialista em medicina pode ajudar médicos no diagnóstico de doenças com base em sintomas apresentados pelos pacientes. Da mesma forma, um sistema especialista em finanças pode ajudar na tomada de decisões de investimento, analisando dados do mercado e oferecendo recomendações personalizadas.

Em resumo, os sistemas especialistas são uma ferramenta poderosa que combina conhecimento especializado com técnicas de inteligência artificial para resolver problemas complexos em uma variedade de domínios. Eles ajudam a automatizar tarefas, fornecer insights valiosos e melhorar a eficiência em muitos aspectos da vida e do trabalho humano.

**Robótica Inteligente:** Robótica Inteligente é uma área interdisciplinar que combina conceitos da robótica e da inteligência artificial para desenvolver sistemas robóticos avançados capazes de interagir com o ambiente de maneira autônoma e inteligente. O objetivo da robótica inteligente é criar robôs que possam perceber, entender e responder ao ambiente ao seu redor de forma adaptativa e eficiente, realizando tarefas complexas de maneira autônoma.

Esses sistemas robóticos inteligentes são equipados com sensores avançados, como câmeras, sensores de profundidade, sensores de proximidade, entre outros, que permitem ao robô perceber e interpretar o ambiente. Além disso, eles são dotados de capacidades de processamento de dados poderosas, que utilizam algoritmos de inteligência artificial, como aprendizado de máquina e visão computacional, para entender e tomar decisões com base nas informações sensoriais.

Um aspecto fundamental da robótica inteligente é a capacidade de os robôs aprenderem com a experiência. Isso significa que eles podem melhorar seu desempenho ao longo do tempo, adaptando-se a novas situações e refinando suas habilidades com base nas interações com o ambiente e nos feedbacks recebidos.

Portanto podemos perceber como o conceito de IA é amplo e pode ser explorado em várias partes. Entendemos que para trazer este assunto para alunos do ensino médio requer uma abordagem mais específica, visando propiciar uma maior compreensão e interesse. No presente trabalho, escolhemos nos aprofundarmos nas redes neurais, pois entendemos ser um campo que gera uma maior compreensão e se mostra de grande importância para o funcionamento de uma IA.

### 1.3 Tipos de IA

A Inteligência Artificial (IA) abrange uma variedade de abordagens e técnicas para simular a inteligência humana em máquinas. Os tipos de IA podem ser categorizados de acordo com suas capacidades e funcionalidades. Algumas categorias principais:

#### 1) IA Fraca (ou Estreita):

- Definição: Refere-se a sistemas de IA especializados em realizar tarefas específicas limitadas.

- Características: São altamente direcionados e não possuem a capacidade de realizar atividades fora de sua área específica de competência.

- Exemplos: Chatbots, reconhecimento de voz, recomendações de produtos.

### **2) IA Forte (ou Geral):**

- Definição: Representa sistemas de IA com a capacidade de compreender, aprender e executar tarefas em uma ampla gama de domínios, de forma semelhante à inteligência humana.

- Características: Esses sistemas podem lidar com tarefas complexas, adaptar-se a diferentes contextos e aprender com experiências variadas. Exemplos: Ainda em fase de desenvolvimento, os exemplos mais próximos seriam sistemas hipotéticos capazes de realizar qualquer tarefa cognitiva humana.

Esse tipo de IA ainda é um desafio para os pesquisadores.

### **3) IA Reativa:**

- Definição: Sistemas de IA que respondem a situações específicas com base em regras programadas, sem a capacidade de aprendizado ou adaptação.

- Características: Eficientes em tarefas específicas, mas limitados em sua capacidade de lidar com situações não previstas.

- Exemplos: Xadrez ou jogos de estratégia, onde o programa segue um conjunto fixo de regras.

### **4) IA Baseada em Aprendizado de Máquina (Machine Learning - ML):**

- Definição: Utiliza algoritmos que permitem aos sistemas aprenderem padrões a partir de dados e melhorarem seu desempenho ao longo do tempo.

- Características: Adaptação a novos dados e capacidade de fazer previsões ou tomar decisões com base em experiências anteriores.

- Exemplos: Classificação de emails como spam, reconhecimento facial, sistemas de recomendação.

### **5) IA Cognitiva:**

- Definição: Busca replicar funções cognitivas humanas, como percepção sensorial, reconhecimento de fala, aprendizado e resolução de problemas.

- Características: Envolvem processamento avançado de linguagem natural, visão computacional e raciocínio semelhante ao humano.

- Exemplos: Assistência virtual avançada, diagnóstico médico baseado em IA.

### **6) IA Simbólica:**

- Definição: Utiliza representações simbólicas e lógicas para modelar o raciocínio e o conhecimento humano.

- Características: Baseia-se em regras lógicas e manipulação de símbolos para tomar decisões e resolver problemas.

- Exemplos: Sistemas especialistas que emulam a expertise humana em áreas específicas.

Essas categorias representam uma visão geral dos tipos de IA, sendo importante observar que muitos sistemas atuais podem incorporar características de múltiplas categorias, contribuindo para a evolução contínua dessa área.

Na atualidade, a maioria das implementações concretas de IA corresponde à categoria fraca (AI restrito), com sistemas altamente especializados para tarefas

particulares e sem capacidade geral nem consciente. Tanto a IA forte (General AI) quanto a IA artificialmente consciente (Artificial Consciousness), embora sejam tópicos fascinantes, não podem ser completamente alcançadas pela tecnologia atual. Assim, é comum debater sobre seu potencial no futuro.

Trabalharemos no presente trabalho com a IA fraca e com o Aprendizado de Máquina, para que sejamos mais pontuais na abordagem pedagógica visando uma compreensão mais profunda.

## 1.4 Aplicações

A Inteligência Artificial (IA) tornou-se uma presença ubíqua em quase todas as atividades do nosso cotidiano, simplificando e aprimorando a vida no século XXI. Seu avanço significativo ocorreu entre 2006 e 2010, com o desenvolvimento do deep learning ou aprendizado profundo. Essa abordagem revolucionária permitiu que as máquinas aprendessem com exemplos, transformando os dados gerados exponencialmente (conhecidos como *big data*) em recursos extremamente valiosos para uma variedade de setores, desde diagnósticos médicos até recomendações personalizadas de entretenimento e vestuário.

A sociedade contemporânea pode ser caracterizada como uma sociedade de dados, na qual a IA impacta profundamente a forma como agimos, pensamos, consumimos e nos relacionamos, como argumentado por Kaufman (2018). Embora muitos possam associar IA a máquinas capazes de pensar e executar tarefas autonomamente, sua presença vai além disso, permeando dispositivos cotidianos como smartphones, TVs inteligentes e assistentes de voz como Alexa, além de sistemas de busca na internet.

Na comunicação e no marketing, a IA desempenha um papel vital ao analisar o comportamento do consumidor com base em dados de navegação, permitindo que empresas otimizem suas estratégias de recomendação e vendas. Os chatbots são um exemplo dessa aplicação, utilizando *machine learning* para simular interações humanas e personalizar a experiência do usuário.

Na educação, a IA possibilita uma aprendizagem mais eficiente e personalizada, com conteúdos adaptados às necessidades individuais dos alunos. Recursos como realidade aumentada, computação em nuvem e gamificação são empregados para criar experiências educacionais envolventes e eficazes.

Na medicina, a IA auxilia na análise de grandes volumes de dados para propor diagnósticos e soluções precisas e rápidas. Desde cirurgias assistidas por robôs até sistemas de monitoramento de pacientes, os benefícios da IA têm revolucionado a prática médica.

Nas empresas, a IA impulsiona a eficiência operacional ao processar grandes volumes de dados e oferecer *insights* valiosos sobre o mercado e os consumidores. Ela também está transformando o recrutamento de talentos e otimizando processos industriais, garantindo segurança e precisão.

Na área jurídica, a IA está reduzindo custos e tempo em processos legais, tornando

a justiça mais acessível e eficiente para todos os indivíduos envolvidos.

Em resumo, a presença e o impacto da Inteligência Artificial são inegáveis em nossa sociedade atual, moldando significativamente nossas vidas e impulsionando avanços em diversos campos. A capacidade de aplicar a IA em uma variedade de contextos é evidente, e um exemplo concreto é o seu uso em empresas. No capítulo 5, exploraremos como as Redes Neurais podem ser aplicadas em um contexto empresarial específico, mais precisamente em uma seguradora de veículos. Isso ilustra como a IA pode ser aproveitada para melhorar processos, otimizar decisões e oferecer soluções inovadoras em ambientes corporativos.

# Fundamentos Matemáticos Para a Compreensão da IA

---

A Matemática para a vida prática em sociedade é de suma relevância. Seus princípios e técnicas são fundamentais em muitas situações do dia a dia, abrangendo várias atividades profissionais e áreas de conhecimento. Contribui para o desenvolvimento de competências fundamentais e desempenha um papel crucial na formação do indivíduo, além disso, organiza o pensamento e raciocínio dedutivo.

O funcionamento específico da Inteligência Artificial envolve uma variedade de algoritmos e técnicas.

Durante o treinamento de modelos de IA, são empregados métodos de otimização, como o gradiente descendente, que ajusta os pesos das conexões entre as unidades da rede neural. Esse processo envolve o cálculo das derivadas parciais e a atualização iterativa dos parâmetros, com o objetivo de minimizar a discrepância entre as previsões do modelo e os resultados desejados.

Além disso, a probabilidade e a estatística desempenham um papel crucial, especialmente em algoritmos de aprendizado de máquina, onde a inferência estatística é usada para fazer previsões e tomar decisões com base em dados.

Em resumo, o cálculo matemático na IA é diversificado, incorporando conceitos de álgebra linear, cálculo diferencial, probabilidade e estatística para desenvolver e treinar modelos que podem realizar tarefas específicas, aprender com dados e tomar decisões.

Vamos desenvolver melhor os conceitos de cálculo, estatística e probabilidade e geometria analítica utilizado pela IA.

## 2.1 Cálculo

O cálculo matemático, denominado cálculo diferencial e integral, é uma ferramenta poderosa que impulsiona muitos dos algoritmos e técnicas utilizados na Inteligência Artificial (IA). Proporcionando ferramentas poderosas para entender, modelar e otimizar sistemas complexos. Da otimização de parâmetros ao tratamento da incerteza, redes neurais, os princípios do cálculo permeiam todos os aspectos da IA.

As redes neurais, como os blocos de construção fundamentais da maioria das IAs, são compostas por neurônios interconectados em camadas. Cada neurônio aplica uma função de ativação aos seus *inputs* ponderados, o que determina sua saída. Existem várias funções de ativação comumente usadas, como a função sigmoide, ReLU (Rectified Linear Unit) e a função tangente hiperbólica, que introduzem não linearidades nos cálculos, permitindo que as redes aprendam e representem relações complexas nos dados. Trataremos especificamente deste assunto mais tarde.

Durante o treinamento de uma rede neural, o cálculo é usado para propagar a informação da entrada (propagação de forward) até a saída, onde é gerada uma previsão. Em seguida, é calculada a diferença entre essa previsão e o valor real (um processo chamado de cálculo do erro). Esse erro é então retropropagado através da rede (propagação de backward), permitindo ajustes nos pesos das conexões entre os neurônios por meio do algoritmo de otimização, como o gradiente descendente. Iremos desenvolver posteriormente esse assunto no decorrer do trabalho.

Além do treinamento de redes neurais, o cálculo é usado em várias outras tarefas de IA, como o processamento de dados. Isso inclui operações matriciais para manipular e transformar conjuntos de dados, técnicas estatísticas para análise exploratória e pré-processamento de dados, e algoritmos de aprendizado de máquina para extrair padrões e insights dos dados.

O cálculo diferencial desempenha um papel crucial na otimização, um componente fundamental do treinamento de modelos de IA. As derivadas calculam a taxa de variação de uma função em relação a uma variável, enquanto os gradientes estendem esse conceito para funções multivariadas. Em algoritmos de aprendizado de máquina, os gradientes são frequentemente utilizados para ajustar os parâmetros do modelo, permitindo a convergência para mínimos locais ou globais.

O cálculo integral é empregado para calcular áreas sob curvas e é essencial na modelagem de probabilidades. Em IA, as integrais são usadas em métodos como as redes Bayesianas para calcular distribuições de probabilidade. Além disso, o cálculo integral tem aplicações na suavização de funções, fundamental em técnicas como kernel density estimation (KDE) e suavização de curvas de aprendizado.

Séries de Taylor são ferramentas poderosas para aproximar funções complexas por meio de somas infinitas de termos. Na IA, essas séries são frequentemente utilizadas para linearizar funções não lineares, facilitando análises e otimizações. Algoritmos de otimização, como o método de Newton, baseiam-se na expansão de Taylor para encontrar mínimos ou máximos de funções.

As equações diferenciais são fundamentais em modelos dinâmicos na IA. Em sistemas que evoluem ao longo do tempo, como em processos estocásticos, equações diferenciais estocásticas (EDS) são aplicadas.

O cálculo também desempenha um papel na teoria da informação, em que conceitos como entropia são fundamentais. A entropia é utilizada em algoritmos de compressão e na avaliação da incerteza em modelos de aprendizado de máquina, especialmente na teoria da informação de Shannon.

### 2.1.1 Gradiente

O treinamento é primordial parte dos algoritmos de aprendizado de máquina, pois é durante esse processo que os algoritmos internalizam os dados. No contexto do aprendizado supervisionado, o treinamento de redes neurais, frequentemente pode ser representado como um processo de otimização, onde buscamos minimizar uma função de perda,  $L(\alpha)$ . Aqui,  $\alpha$  denota o vetor de parâmetros da Rede Neural.

Os pontos críticos ou estacionários de uma função multivariada  $L(\alpha)$  são os pontos  $\alpha$  que satisfazem a equação  $\nabla L(\alpha) = 0$ . Tais pontos críticos incluem mínimos locais, máximos locais e pontos de sela, os quais podemos distinguir através do teste da segunda derivada (verificando a positividade da matriz Hessiana).

Para funções simples em baixas dimensões, é ocasionalmente possível encontrar analiticamente mínimos locais ou mesmo mínimos globais comparando os valores de  $L(\alpha)$  nesses pontos. Contudo, para funções de perda de redes neurais, tal abordagem analítica é geralmente impraticável: As expressões analíticas para  $L(\alpha)$ ,  $\nabla L(\alpha)$  e  $\nabla^2 L(\alpha)$  não são simples. Além disso, a dimensão  $\mu$  do espaço de parâmetros  $\mathbb{R}^\mu$  é frequentemente grande demais para permitir a obtenção de uma solução exata para o sistema  $\nabla L(\alpha) = 0$ . Portanto, na prática, minimizar esse tipo de função com um grande número de parâmetros é um desafio para o qual várias abordagens numéricas têm sido propostas.

Vamos abordar alguns métodos comuns de otimização numérica usados para treinar redes neurais, todos baseados na ideia da descida do gradiente. A ideia central desses métodos é simples: uma função multivariada tende a diminuir mais rapidamente na direção oposta ao gradiente negativo. Simplificadamente, isso pode ser comparado aos passos que alguém daria para descer uma encosta arborizada quando não consegue ver o fundo da colina por entre as árvores. Nesse caso, a pessoa busca minimizar sua altitude, dando um passo de cada vez na direção que parece levar mais rapidamente para baixo localmente, sem considerar os passos futuros em direção ao mínimo global. Como essa analogia sugere, garantir a convergência dos métodos gradiente para um mínimo global pode ser desafiador, especialmente quando há mínimos locais presentes.

Vamos considerar métodos gradientes determinísticos e estocásticos (SGD), além de mencionar brevemente uma generalização do gradiente chamada gradiente com momentum ou método da bola pesada. A ideia é adicionar inércia para evitar "ficar preso" ou "desacelerar" em mínimos locais devido a oscilações na função de perda  $L(\alpha)$ . Tanto o SGD quanto o gradiente com momentum ajudam a reduzir o tempo de computação e lidar com mínimos locais e pontos de sela.

Começamos com o método de descida do gradiente determinístico (GD), que é um processo iterativo no qual atualizamos os parâmetros a cada passo na direção do gradiente da função de perda. O método GD parece ter sido introduzido inicialmente por Cauchy no meio do século XIX e redescoberto por Hadamard no início do século XX. Desde então, tornou-se um dos algoritmos numéricos mais amplamente estudados e demonstrou sua utilidade para encontrar mínimos locais de funções. No caso específico de funções convexas, várias extensões foram propostas para melhorar

a taxa de convergência, permanecendo de primeira ordem (por exemplo, envolvendo apenas o gradiente da função a ser minimizada em vez de derivadas de ordem superior). Considere um ponto  $\alpha$  junto com uma direção  $\gamma$ . Dizemos que uma função  $L$  está diminuindo na direção de  $\gamma$  se a derivada de  $L$  na direção de  $\gamma$  for negativa, isto é,

$$\nabla L(\alpha) \cdot \gamma < 0.$$

Recordamos a identidade elementar para o produto escalar em  $\mathbb{R}^\mu$

$$\nabla L(\alpha) \cdot \gamma = \|\nabla L(\alpha)\| \|\gamma\| \cos\theta$$

onde  $\theta$  é o ângulo entre os vetores  $\nabla L(\alpha)$  e  $\gamma$ . Portanto, a derivada direcional é maximizada quando  $\theta = 0$  ( $\gamma$  está na direção do gradiente) e minimizada quando  $\theta = \pi$  ( $\gamma$  está na direção do gradiente negativo). Isso pode ser usado para encontrar outro ponto  $\tilde{\alpha}$  a partir de  $\alpha$  onde  $L$  é menor. Pela fórmula de Taylor, desde que  $L$  seja suave,

$$L(\alpha + \tau\gamma) = L(\alpha) + \tau\nabla L(\alpha) \cdot \gamma + O(\tau^2).$$

Damos a definição precisa da notação big  $O(\cdot)$ . Se  $\tau$  for suficientemente pequeno, nossa melhor “suposição” para um novo ponto  $\tilde{\alpha}$  naturalmente consiste em escolher

$$\tilde{\alpha} = \alpha + \tau\gamma, \quad \text{com } \gamma = -\nabla L(\alpha).$$

Essa é a ideia por trás dos algoritmos de descida do gradiente (GD): a partir de um ponto inicial aleatório no espaço de parâmetros, mova-se para o próximo ponto escolhendo a direção apropriada do gradiente. Continue iterativamente até que a precisão desejada na minimização seja alcançada.

Em termos mais matemáticos, começamos com um parâmetro inicial escolhido aleatoriamente vetor  $\alpha^{(0)} \in \mathbb{R}^\mu$ . Em seguida, calculamos iterativamente uma sequência de parâmetros  $(\alpha^{(n)})_{n \in \mathbb{N}}$  de acordo com a seguinte regra:

$$\alpha^{(n)} = \alpha^{(n-1)} - \tau\nabla L(\alpha^{(n-1)}).$$

## 2.2 Probabilidade e Estatística

A probabilidade e estatística são ferramentas cruciais para a Inteligência Artificial, ou seja, desempenham um papel fundamental no embasamento teórico e na prática da (IA). Desde a modelagem da incerteza até a inferência sobre padrões e relações nos dados, esses campos desempenham um papel central no desenvolvimento e na operação de sistemas.

A probabilidade é a base sobre a qual a IA lida com a incerteza. Na modelagem probabilística, eventos e resultados são atribuídos a probabilidades, permitindo que os sistemas de IA avaliem a confiabilidade das previsões. A teoria das probabilidades é utilizada em algoritmos de aprendizado de máquina para calcular a probabilidade de um evento ocorrer, informando decisões e inferências.

Os modelos probabilísticos são uma categoria crucial em IA. Modelos como redes

Bayesianas e processos estocásticos são fundamentais para representar a incerteza em sistemas complexos. Eles incorporam distribuições de probabilidade para capturar a variabilidade nos dados e ajudam na tomada de decisões sob incerteza.

A estatística descritiva é utilizada para resumir e descrever características essenciais dos dados. Médias, medianas e desvios padrão são métricas comuns na análise exploratória de dados. Por outro lado, a estatística inferencial envolve fazer inferências sobre uma população com base em uma amostra, sendo fundamental na validação e generalização de modelos de IA.

A regressão é uma técnica estatística utilizada para modelar a relação entre variáveis. Em IA, é comumente empregada para prever valores contínuos. A análise de correlação, por sua vez, ajuda a entender a força e a direção das relações entre variáveis, fornecendo insights valiosos sobre padrões nos dados.

A inferência Bayesiana é uma abordagem que atualiza as probabilidades de hipóteses à medida que mais dados são disponibilizados. Na IA, isso é especialmente útil quando lidamos com a incerteza e precisamos ajustar nossas crenças com base em novas evidências. A inferência Bayesiana é usada em sistemas de recomendação, processamento de linguagem natural e diagnósticos médicos.

No cenário industrial e comercial do futuro, a integração contínua de estatística, probabilidade e IA em sistemas de IDP promete uma transformação radical. Os benefícios já percebidos em termos de eficiência e automação são apenas o começo, com sistemas cada vez mais sofisticados e adaptáveis para lidar com um espectro amplo de documentos e situações. A capacidade de tomada de decisão informada, respaldada pela análise estatística e probabilística, será fundamental para empresas que buscam permanecer competitivas em um ambiente empresarial em constante evolução.

A estatística e a probabilidade são pilares fundamentais na construção e no treinamento de modelos de IA, incluindo machine learning, deep learning, redes neurais e Intelligent Document Processing (IDP). Estes conceitos não somente permitem compreender a incerteza nos dados, mas também facilitam a tomada de decisões informadas. À medida que a IA evolui, a importância da estatística e da probabilidade só aumenta, impulsionando a capacidade de desenvolver modelos cada vez mais poderosos e precisos.

## 2.3 Geometria Analítica

A geometria analítica é uma disciplina que combina conceitos da álgebra e da geometria, amplamente utilizada nos fundamentos teóricos da Inteligência Artificial (IA) e faz conexões entre as equações matemáticas e os algoritmos que impulsionam as máquinas inteligentes.

Na IA, os dados muitas vezes residem em espaços multidimensionais, onde cada dimensão representa uma característica específica. A geometria analítica entra em cena ao representar essas características como coordenadas em um sistema de eixos. Se tivermos um conjunto de dados com, por exemplo, três características, podemos visualizá-lo em um espaço tridimensional.

A representação de dados em forma de vetores é uma aplicação direta da geometria analítica na IA. Vetores são utilizados para expressar informações como pontos no espaço de características, e operações lineares sobre esses vetores possuem grande importância em algoritmos de aprendizado de máquina. Transformações lineares, rotações e escalonamentos são operações fundamentais que moldam os dados durante o processo de treinamento.

Ao modelar fenômenos complexos, as equações paramétricas são frequentemente empregadas. Na IA, os modelos podem ser representados por equações paramétricas, onde os parâmetros são ajustados durante o treinamento para melhor se adequar aos dados. A otimização desses parâmetros é, essencialmente, um problema de geometria analítica, onde estamos procurando encontrar o ponto no espaço de características que minimiza ou maximiza uma determinada função de custo.

A teoria dos hiperplanos é uma extensão natural da geometria analítica na IA. Em problemas de classificação, onde o objetivo é separar dados em diferentes categorias, os hiperplanos são usados para criar fronteiras de decisão. Métodos como máquinas de vetor de suporte (SVM) dependem fortemente da geometria analítica para encontrar o hiperplano que melhor separa os dados.

A geometria diferencial, uma área que trata das propriedades geométricas de curvas e superfícies, é utilizada na otimização de funções na IA. Algoritmos de otimização utilizam conceitos como gradientes e derivadas parciais para encontrar os mínimos ou máximos de funções de custo, desempenhando um papel crucial no treinamento de modelos.

A geometria analítica proporciona a base matemática que sustenta muitos dos algoritmos e modelos utilizados na Inteligência Artificial. Desde a representação de dados até a otimização de parâmetros, a geometria analítica oferece ferramentas poderosas para entender e aprimorar os sistemas de IA.

### 2.3.1 Equações de Retas

As equações de retas são uma parte fundamental da geometria analítica e desempenham um papel crucial em diversas áreas da matemática e suas aplicações práticas. Vamos explorar as diferentes formas de representar as equações de retas e discutir suas propriedades e usos.

Existem algumas diferentes formas de representação de equações de retas, uma delas é a forma geral da equação de uma reta no plano cartesiano, e é dada por:

$$Ax + By + C = 0$$

onde  $A$ ,  $B$  e  $C$  são constantes, e  $A$  e  $B$  não são ambos iguais a zero. Esta forma é útil porque é muito geral e pode representar qualquer reta.

Outra forma, é a forma reduzida que é particularmente útil e intuitiva, sendo expressa como:

$$y = mx + b$$

onde  $m$  é o coeficiente angular (a inclinação da reta) e  $b$  é o coeficiente linear (a interseção da reta com o eixo  $y$ ). Esta forma é fácil de usar quando se conhece a

inclinação da reta e onde ela corta o eixo  $y$ .

Temos ainda a equação segmentária, quando a reta intercepta os eixos  $x$  e  $y$  em pontos distintos, ela é dada por:

$$\frac{x}{a} + \frac{y}{b} = 1$$

onde  $a$  e  $b$  são os interceptos da reta nos eixos  $x$  e  $y$ , respectivamente. Esta forma é útil quando os interceptos são conhecidos ou fáceis de determinar.

Por último, a forma paramétrica expressa a reta em termos de um parâmetro  $t$ :

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases}$$

onde  $(x_0, y_0)$  é um ponto conhecido da reta e  $(a, b)$  é um vetor diretor da reta. Esta forma é útil em aplicações que envolvem movimento ao longo de uma reta ou em geometria computacional.

As equações de retas são amplamente utilizadas em várias áreas, incluindo:

- **Engenharia e Física:**

- Análise de trajetórias de objetos.
- Projetos de estruturas e sistemas de transporte.

- **Economia:**

- Modelagem de funções de custo e receita.

- **Geometria Computacional:**

- Detecção de colisões e construção de gráficos.

- **Ciências da Computação:**

- Algoritmos de renderização gráfica e jogos.

As equações de retas têm um papel significativo no aprendizado de máquinas, especialmente em áreas como aprendizado de máquina e visão computacional. Vamos explorar algumas das aplicações mais importantes.

Uma das aplicações mais diretas da equação de reta no aprendizado de máquinas é a regressão linear. A regressão linear é uma técnica estatística usada para modelar a relação entre uma variável dependente  $y$  e uma ou mais variáveis independentes  $x$ . O modelo de regressão linear simples é representado pela equação da reta:

$$y = mx + b$$

onde  $m$  é o coeficiente angular e  $b$  é o intercepto. No aprendizado de máquinas, a regressão linear é utilizada para previsões e inferências. Por exemplo, pode-se

prever o preço de uma casa com base em suas características, como área e número de quartos.

Algoritmo de Gradiente Descendente é utilizado para encontrar os valores ótimos de  $m$  e  $b$  que minimizam o erro entre a linha prevista e os dados reais, usamos algoritmos de otimização como o gradiente descendente. O objetivo é minimizar uma função de custo, geralmente o erro quadrático médio (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - (mx_i + b))^2$$

Um outro tipo de aprendizado de máquinas é denominado Support Vector Machine (SVM). Na classificação, especialmente com o SVM, as retas são utilizadas para encontrar o hiperplano que melhor separa as classes de dados. Em duas dimensões, esse hiperplano é uma reta. O objetivo do SVM é encontrar a reta que maximiza a margem entre as duas classes:

$$w \cdot x + b = 0$$

onde  $w$  é um vetor de pesos e  $b$  é o bias. A margem máxima é alcançada resolvendo o problema de otimização com restrições.

Embora as redes neurais sejam modelos mais complexos, em certos casos, uma camada de rede neural pode atuar como uma regressão linear. A saída de um neurônio é uma combinação linear dos inputs:

$$y = w \cdot x + b$$

onde  $w$  é um vetor de pesos e  $b$  é o bias. Esta função linear pode ser usada como uma função de ativação em redes neurais profundas, contribuindo para a previsão de valores contínuos.

Na visão computacional, as retas são usadas para várias tarefas, como detecção de bordas e segmentação de imagem. Algoritmos como a Transformada de Hough detectam linhas em imagens, representando-as na forma:

$$\rho = x \cos \theta + y \sin \theta$$

onde  $\rho$  é a distância perpendicular da linha à origem e  $\theta$  é o ângulo de inclinação da linha.

A equação da reta é uma ferramenta fundamental na inteligência artificial. Seja na modelagem de relações lineares em dados através da regressão linear, na classificação de dados com SVMs, ou na análise de imagens na visão computacional, as retas fornecem uma base sólida para muitas técnicas avançadas. A compreensão profunda dessas aplicações permite que os cientistas de dados e engenheiros de IA desenvolvam modelos mais precisos e eficientes.

Compreender as equações de retas e suas diferentes formas de representação é essencial para a aplicação de conceitos matemáticos em situações práticas. As propriedades das retas, como coeficiente angular e intercepto, permitem análises

detalhadas e precisas em várias disciplinas, tornando este tópico uma peça central na educação matemática e nas suas aplicações no mundo real.

### 2.3.2 Equações do Plano

As equações do plano são fórmulas matemáticas que descrevem um plano em um espaço tridimensional. A forma geral de uma equação de plano é

$$Ax + By + Cz + D = 0,$$

onde  $A$ ,  $B$ , e  $C$  não são todos zero e representam os coeficientes que definem a orientação do plano no espaço, enquanto  $D$  é a constante que determina a posição do plano em relação à origem.

Também podemos descrever o plano em sua forma vetorial:

$$\mathbf{r} \cdot \mathbf{n} = d,$$

onde  $\mathbf{r}$  é o vetor posição,  $\mathbf{n}$  é o vetor normal ao plano, e  $d$  é a distância do plano à origem.

Outra possível forma é a paramétrica:

$$\mathbf{r}(u,v) = \mathbf{r}_0 + u\mathbf{v}_1 + v\mathbf{v}_2,$$

onde  $\mathbf{r}_0$  é um ponto no plano, e  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  são vetores diretores no plano.

Em IA, essas representações matemáticas são aplicadas para modelar problemas complexos. As redes neurais convolucionais, por exemplo, elas usam uma abordagem de filtragem que pode ser vista como a definição de um hiperplano em um espaço de alta dimensão para identificar características em imagens.

As máquinas de vetores de suporte, encontram o hiperplano que maximiza a margem entre as classes de dados. A equação deste hiperplano é central para o algoritmo e é derivada durante o processo de treinamento. Esses métodos dependem fortemente da compreensão e manipulação de equações de plano para realizar tarefas como classificação, regressão e agrupamento em IA. O que veremos mais a frente no capítulo 4.

# O Uso da Tecnologia Voltada ao Aprendizado e a Educação

---

A inteligência artificial (IA) está revolucionando o campo da educação, oferecendo oportunidades sem precedentes para personalizar a aprendizagem, otimizar processos educacionais e abrir novos caminhos para a descoberta do conhecimento.

## 3.1 Aplicações de IA para Educação

A cada dia novas aplicações fazem uso de recursos de *machine learning* e inteligência artificial para resolver problemas, fazer análise de grandes quantidades de dados, indicar o melhor caminho, recomendar produtos e serviços baseados em suas preferências armazenadas em aplicativos. O uso dessas tecnologias também tem suas aplicações no campo da arte e educação. A seguir apresentarei alguns casos de uso dessas tecnologias nesses dois campos.

### Museu do Amanhã: IRIS+

O Museu do Amanhã, localizado na cidade do Rio de Janeiro, tem como objetivo gerar reflexões sobre o nosso modo de viver e como podemos transformar o nosso futuro a partir dos valores éticos de sustentabilidade e convivência. Trata-se de um museu de ciências que por meio de jogos, ambientes audiovisuais e instalações interativas conduz o visitante olhar o passado e pensar em futuros possíveis.

O projeto denominado IRIS+ foi desenvolvido em 2017, e atualizado em 2020, para o Museu do Amanhã. Consiste em uma assistente virtual que tem como objetivo responder dúvidas e fazer perguntas sobre as temáticas de sustentabilidade e convivência, dois dos principais eixos temáticos do museu. Como vimos na seção 1.3 são utilizados aqui a IA Fraca pois temos uma assistente virtual para respostas específicas, além do uso da IA Cognitiva para tentar replicar funções cognitivas humanas.

Durante a visita ao museu, são capturadas informações dessa pessoa que está visitando e, ao final, é perguntado ao visitante sobre como ele pode modificar sua preocupação com o ambiente e, a partir de sua resposta - que pode ser por voz ou digitando em um tablet - são apresentados projetos de diferentes organizações

e instituições, registrados pela equipe do museu do amanhã, que podem levá-lo a comprometer-se com as ações sociais.

Para treinar a tecnologia, aproximadamente 800 pessoas apresentaram temas relevantes à realidade da população e diversas possibilidades de respostas que os visitantes do museu poderiam utilizar foram cadastradas. A tecnologia empregada foi o IBM Watson, uma plataforma de inteligência artificial para negócios, utilizando a API (interfaces de programação de aplicativos) do Watson Conversation Services (WCS), hospedada na IBM Cloud. No caso da aplicação para o Museu do Amanhã, de acordo com Fabrício Barth, líder técnico do IBM Watson em entrevista ao canal Museu do Amanhã no *youtube*, embora a tecnologia seja capaz tecnicamente de buscar informações em fontes online, esse recurso não é adequado para os museus, pois essas instituições têm a responsabilidade de controlar o que é transmitido aos visitantes, tornando a curadoria do conteúdo indispensável.

Fabrício também menciona que “na maioria das vezes, treinamos o sistema para entender a intenção de uma pergunta, enquanto no Museu do Amanhã isso foi feito para entender a intenção por trás de uma resposta. A dificuldade está em que a resposta é geralmente um conteúdo mais longo e menos objetivo que uma pergunta”.

## **BIA**

O projeto da assistente virtual BIA (Bradesco Inteligência Artificial) foi implantado em 2016 e inicialmente tinha o foco em sanar as dúvidas dos mais de 60 mil funcionários das agências do banco Bradesco. O problema que o Bradesco queria resolver inicialmente era dar agilidade e conhecimento de seus mais de 65 produtos e serviços aos seus funcionários para que o atendimento oferecido por eles pudesse melhorar a experiência dos clientes que usavam as suas agências. Essa versão foi denominada BIA Agências.

Para os funcionários fazerem uso da BIA Agências era necessário realizar um login na ferramenta e enviar a dúvida via voz e em seguida a BIA devolvia a resposta por texto. O funcionário classificava a resposta como “sim” (se a resposta atendeu a pergunta) ou “não” (caso a resposta não foi suficiente). Como a BIA funciona por níveis de interação, caso a classificação seja não, ela apresentava alternativas com a mesma opção de classificação em “sim” e “não”. Se nesse próximo nível a classificação da resposta também foi “não”, a BIA buscava as respostas em mais um nível de dados e se nesse último nível a classificação da resposta ainda fosse “não”, a conversa era encaminhada para um atendente humano.

Em 2017, a BIA foi implantada para uso também pelos clientes do banco com a capacidade de responder a dúvidas em linguagem natural, como vimos ser uma função da IA Cognitiva e com um vocabulário simples e adequado ao seu público alvo, que mostra ser uma característica da IA Fraca, reduzindo o tempo de resposta, em alguns casos, de 10 minutos para 3 segundos.

A assistente virtual pode ser acessada pelos aplicativos do banco nas plataformas Android e IOS e ajuda os clientes com diversas necessidades (pagamentos, transferências bancárias, consulta de saldo, acesso a serviços etc.). Para que a construção desse projeto, foi utilizado o IBM Watson Assistant, Watson Discovery e outras

ferramentas relacionadas.

### **Pinacoteca de São Paulo: A Voz da Arte**

A Pinacoteca do Estado de São Paulo é o mais antigo museu de artes visuais de São Paulo e um dos mais importantes do Brasil. É voltado para a apresentação de obras de arte brasileira, abrigando um dos mais representativos acervos nessa linha.

Em 2017, realizada em parceria com a IBM, a Pinacoteca apresentou a experiência denominada “A Voz da Arte” que ficou disponível para o público de abril a dezembro de 2017. A experiência consistia em conversar com as 7 obras do acervo e para isso foi desenvolvido um aplicativo para *smartphone* por meio do qual poderiam ser colocadas as questões.

Ao entrar na Pinacoteca, o visitante que optava por conhecer a experiência, recebia um *smartphone* e fone de ouvido, mediante assinatura de um termo de responsabilidade e a disponibilização de um documento pessoal. Ao passar por uma das 7 obras mapeadas, o visitante era informado sobre a possibilidade de interagir com o dispositivo, permitindo que ele comece a fazer perguntas. Os visitantes podiam perguntar por digitação ou por voz, e o aplicativo, que utilizava o IBM Watson como ferramenta de inteligência artificial, apresentava as respostas.

O sistema por traz do aplicativo funciona em etapas. A primeira envolvia converter a entrada fornecida pelo visitante em texto. Em seguida, iniciava-se o processo de interpretação do que foi dito, identificando a intenção por trás da pergunta. Métodos de aprendizado de máquina foram utilizados, onde uma pergunta ou entrada é dada ao sistema, que é então classificado em diferentes categorias por um algoritmo de aprendizado. A partir delas, o sistema extrai as respostas que melhor correspondem às categorias pré-existentes. Se nenhuma resposta cadastrada foi inserida, uma mensagem de erro era enviada e a pergunta é armazenada no banco de dados para futuras respostas. O sistema de classificação foi projetado para permitir que o visitante escolhesse se a resposta foi satisfatória, selecionando “gostei” ou “não gostei” no aplicativo. Se a última opção fosse a escolhida, ela era enviada para uma equipe dedicada de historiadores e curadores que registram novas respostas.

Para preparar a base de informações que seria utilizada pelo aplicativo, uma curadoria da Pinacoteca selecionou 7 obras em exposição no museu. Em seguida foi criada uma plataforma para recolhimento de perguntas no qual apenas usuários autorizados poderiam trazer suas dúvidas, em forma de perguntas, sobre as obras em questão. O próximo passo foi organizar as questões por similaridade para que uma única resposta pudesse ser dada para diferentes questões que tratassem daquela temática. Segundo Fabricio Barth, foram 6 meses para realizar o treinamento da IA para que ela pudesse responder as perguntas do público.

### **Museu Catavento: Ensinando Robôs**

O Museu Catavento é uma instituição de ciência e tecnologia situada na cidade de São Paulo. Inaugurado em março de 2009, seu objetivo principal é aproximar o público do mundo científico, incentivando a curiosidade e transmitindo conhecimentos básicos e valores sociais. Isso é realizado por meio de exposições interativas e envolventes, que utilizam uma linguagem simples e acessível para alcançar pessoas de todas as

idades e níveis de conhecimento.

Em parceria com a IBM, o museu criou a exposição Ensinando Robôs que tem por objetivo demonstrar, de forma prática, como se dá o funcionamento do aprendizado de máquinas que vimos na seção 1.3 e qual a importância dos humanos nesse processo. Para isso, são demonstrados, de forma lúdica e interativa, os conceitos de IA.

Segundo o Museu, a jornada proposta está estruturada para que para o público interaja com os 3 robôs que lá estão, com o objetivo de aprimorar as habilidades desses robôs por intermédio de um jogo de perguntas e repostas, nas quais os visitantes, em grupos e orientados por monitores, colaboram para esse aprendizado. É uma característica importante da inteligência artificial a habilidade de treinar uma rede neural até que ela tenha capacidade de entregar a resposta certa.

Por meio da reestruturação das suas conexões internas, as redes neurais aprendem com as informações que são disponibilizadas. O objetivo é dar aos visitantes uma compreensão tangível de como o aprendizado de máquina funciona em redes neurais. Os robôs foram previamente treinados com temas relacionados a ciências naturais e o público pode interagir com eles por meio de uma tela. Nessa tela são apresentadas 3 perguntas, previamente estabelecidas, que o público pode selecionar. Uma vez selecionada da pergunta, o público a lê em voz alta e os robôs iniciam uma conversa. Nessa conversa, fica demonstrado que um robô tem mais conhecimento sobre aquela pergunta que os outros dois e então é feito o paralelo com a necessidade de treinamento da IA e o papel dos humanos nesse processo. A interação dura aproximadamente 3 minutos.

Para que o processo de interação entre os robôs e os visitantes aconteça, assistentes virtuais fazem parte da composição desses robôs possibilitando a interação por voz e o entendimento da linguagem humana. A IBM utilizou o Watson Assistant com recursos de processamento de linguagem natural e “*text to speech*” para disponibilização dos robôs.

Conhecido como tecnologia educacional ou EdTech, vem transformando significativamente os métodos de ensino e proporcionado oportunidades inovadoras para alunos, educadores e instituições educacionais. Essa revolução tecnológica na educação tem vários aspectos positivos e desafios associados, pois ao mesmo tempo em que essa mudança amplia nossas possibilidades de aprender e criar, ela impõe algumas responsabilidades para o universo da educação.

Agora que abordamos os conteúdos existentes em uma IA, vamos nos aprofundar no funcionamento de um neurônio e de uma rede neural no próximo capítulo, para compreendermos como são utilizados os conceitos definidos até aqui.

# Redes Neurais

---

Agora que estabelecemos uma compreensão mais sólida do que é Inteligência Artificial, dos conceitos matemáticos subjacentes e das aplicações tanto profissionais quanto educacionais, vamos nos aprofundar no funcionamento dos neurônios e das redes neurais. Nosso objetivo é proporcionar aos alunos uma explicação mais detalhada para que possam conectar de forma mais significativa com as informações teóricas apresentadas neste capítulo. Isso permitirá que eles compreendam melhor as aplicações práticas discutidas no capítulo 5 e reconheçam como esses conceitos fundamentais da IA se traduzem em soluções do mundo real.

Os conceitos principais aqui apresentados são baseados em 8 e 9.

Nesse capítulo apresentamos o conceito Matemático por trás de um rede neural profunda. A seção 4.2, acrescida do exemplo desenvolvido na seção 5.4, irá compor um produto na forma de artigo a ser enviado para publicação em alguma revista da área de Ensino de Matemática.

O cérebro humano é frequentemente comparado a um computador altamente complexo, caracterizado por um sistema de processamento de informações intrincado, não-linear e paralelo (HAYKIN, 2001). Sua notável capacidade de reconhecer, associar e generalizar padrões tem sido uma inspiração fundamental para o desenvolvimento das Redes Neurais Artificiais (RNA).

As RNA, por sua vez, são sistemas paralelos compostos por unidades de processamento simples, chamadas neurônios artificiais, que executam cálculos de funções matemáticas específicas. Essas redes são capazes de abordar uma ampla variedade de problemas, devido à sua capacidade de aprendizado e adaptação, espelhando em certo sentido a flexibilidade e eficiência do cérebro humano.

As redes neurais, inspiradas pelo intrincado sistema neural do cérebro humano, emergiram com um papel de destaque da revolução na Inteligência Artificial (IA). Estas estruturas computacionais complexas têm transformado a capacidade da IA em lidar com tarefas complexas, desde reconhecimento de padrões até tomada de decisões sofisticadas.

No cerne das redes neurais está a representação abstrata do processamento de informações em camadas. Composta por neurônios artificiais, ou nós, organizados em camadas de entrada, ocultas e de saída, essa arquitetura espelha a capacidade

do cérebro em realizar cálculos em diferentes níveis de abstração. Cada conexão ponderada entre os neurônios é ajustada durante o treinamento, permitindo à rede aprender padrões complexos.

Um dos processos cruciais para um rede neural é o seu treinamento. Utilizando algoritmos de otimização, como o gradiente descendente, a rede ajusta seus pesos para reduzir a diferença entre as previsões e os rótulos reais, adaptando-se assim a novos dados. A combinação de aprendizado supervisionado e não supervisionado contribui para a flexibilidade e capacidade de generalização.

O advento das redes neurais profundas, também conhecido como *deep learning*, marcou um ponto crucial na inteligência artificial. Com múltiplas camadas ocultas, essas redes têm a capacidade de aprender representações hierárquicas e complexas dos dados. Essa capacidade tem impulsionado avanços significativos em áreas como reconhecimento de imagem, processamento de linguagem natural e jogos.

Apesar dos triunfos, as redes neurais e o *deep learning* enfrentam desafios. A tendência ao sobreajuste e a demanda por grandes conjuntos de dados são questões críticas. Estratégias como regularização e arquiteturas avançadas, incluindo redes recorrentes e convolucionais, surgiram para mitigar esses desafios, consolidando a eficácia das redes neurais.

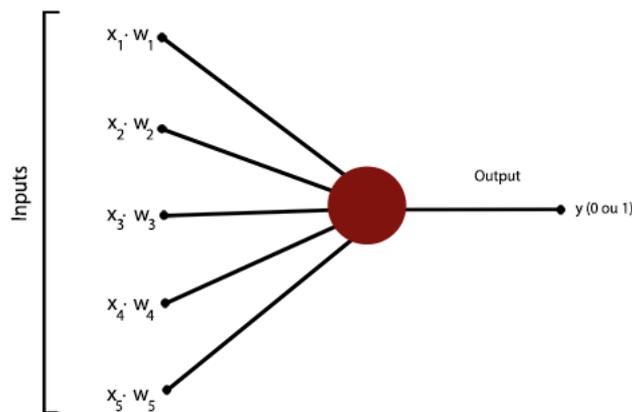
À medida que avançamos, a interseção entre redes neurais e IA continua a moldar o panorama tecnológico. Aplicações diversificadas, desde diagnósticos médicos precisos até assistentes virtuais inteligentes, destacam o papel central das redes neurais. O constante aprimoramento dessas estruturas e a busca por modelos mais eficientes representam um caminho promissor no desenvolvimento contínuo da IA. A jornada das redes neurais na IA é, sem dúvida, uma narrativa evolutiva que continua a redefinir o potencial e os limites da inteligência artificial.

Um perceptron é um dos blocos fundamentais de construção em redes neurais artificiais. Ele foi concebido como um modelo simplificado de um neurônio biológico, com o objetivo de aprender e realizar tarefas de classificação binária como pode ser observado em [13](#).

O perceptron consiste em várias entradas, cada uma ponderada por um peso correspondente, somadas e passadas através de uma função de ativação para produzir uma saída. A função de ativação é geralmente uma função degrau simples, que retorna 1 se a soma ponderada das entradas exceder um determinado limiar e 0 caso contrário.

Durante o treinamento, os pesos das entradas são ajustados de acordo com o erro entre a saída prevista e a saída desejada. Isso é feito usando um algoritmo de aprendizado.

Embora perceptrons simples sejam limitados a tarefas de classificação linearmente separáveis, ou seja, problemas nos quais as classes podem ser divididas por uma linha reta ou hiperplano, eles serviram como base para o desenvolvimento de redes neurais mais complexas e poderosas. Redes neurais multicamadas, como as redes neurais profundas, que consistem em múltiplos perceptrons organizados em camadas interconectadas, podem aprender e resolver problemas muito mais complexos e não lineares.

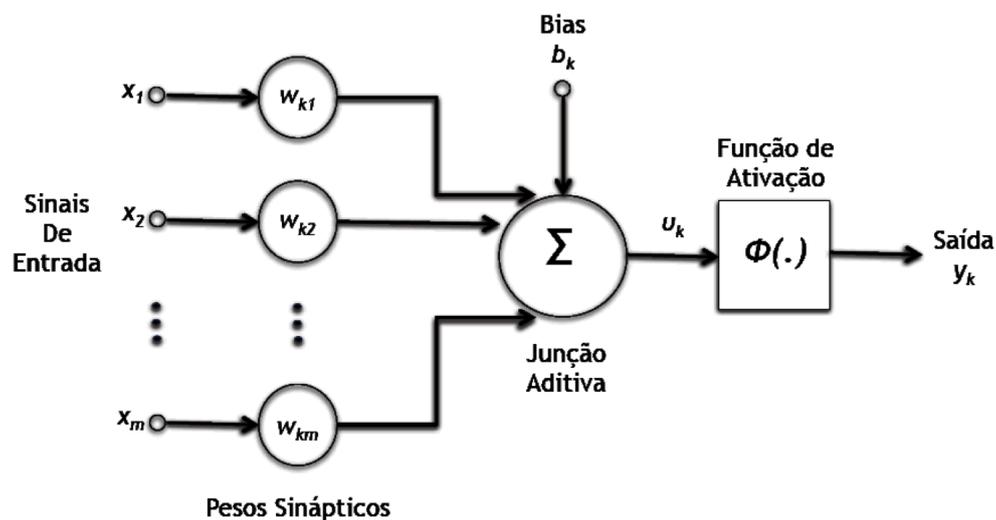


**Figura 4.1:** Modelo Perceptron

Fonte: O Autor

A unidade básica de uma Rede Neural Artificial (RNA) é o neurônio artificial, cujo modelo é ilustrado na Figura 3.1. O modelo de um neurônio  $k$  qualquer apresenta três elementos fundamentais (HAYKIN, 2001):

- Um conjunto de **sinapses** (conexões), onde, a cada uma delas está associado um peso sináptico  $w_{kj}, j = 1, 2, 3, \dots, m$ ;
- Um **somador** que na verdade é um combinador linear dos sinais de entrada;
- Uma função de ativação para restringir ou limitar a amplitude de saída.



**Figura 4.2:** Modelo do neurônio artificial

Fonte: adaptado de (HAYKIN, 2001)

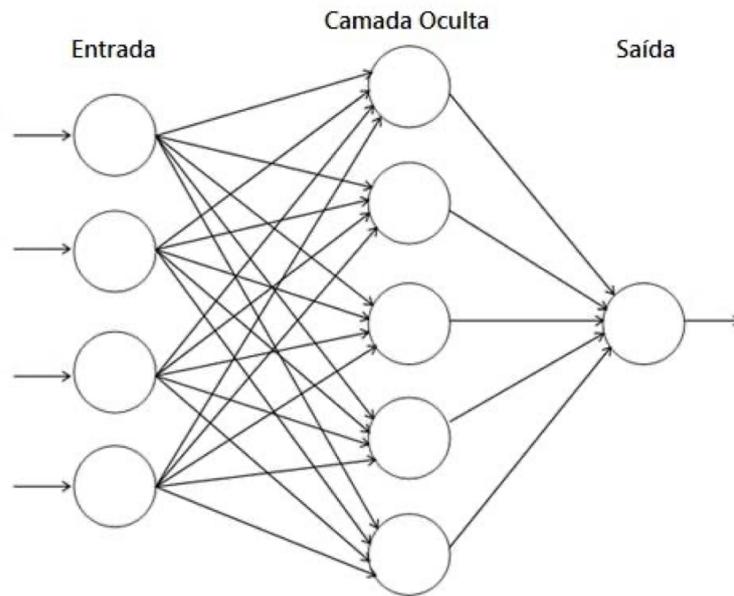
Este modelo também inclui um viés  $b_k$  (ou polarização), que desempenha o papel de ajustar a entrada líquida da função de ativação, podendo ser positivo ou negativo. Em termos matemáticos, podemos descrever o neurônio  $k$  por meio das seguintes equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j$$

$$y_k = \varphi(u_k + b_k)$$

Onde:  $x_1, x_2, \dots, x_m$  representam os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  representam os pesos sinápticos;  $u_k$  representa a saída do combinador linear devido aos sinais de entrada;  $b_k$  representa o viés;  $\varphi(\cdot)$  representa a função de ativação; e  $y_k$  representa o sinal de saída do neurônio.

O termo arquitetura em redes neurais refere-se à organização dos neurônios e suas conexões, sendo fundamental para sua capacidade de processamento (ZANINI, 2000). As redes neurais podem ser classificadas em dois tipos: recorrentes e não recorrentes. As redes não recorrentes não possuem laços de realimentação de suas saídas para suas entradas, sendo consideradas “sem memória”. A estrutura dessas redes pode consistir em uma única camada ou em múltiplas camadas, como ilustrado na Figura 3.3. Uma rede de uma única camada não possui camadas com neurônios ocultos.



**Figura 4.3:** Rede não recorrente feedforward.

Fonte: MEDEIROS, Romero

Um exemplo de redes não recorrentes são as redes *feedforward*, nas quais o sinal é sempre propagado para frente.

Uma rede neural recorrente se diferencia de uma rede não recorrente pelo fato de possuir pelo menos um laço de realimentação. Devido à presença de realimentação, essas redes respondem dinamicamente a estímulos. Após a aplicação de uma nova entrada, a saída é calculada e então realimentada para modificar a entrada, tornando essas redes “com memória” (VELLASCO, 2007).

A presença de laços de realimentação tem um impacto significativo na capacidade

de aprendizagem e no desempenho da rede. Além disso, os laços de realimentação geralmente envolvem o uso de elementos de atraso unitário,  $Z^{-1}$ , resultando em um comportamento dinâmico não linear, especialmente quando a rede neural contém unidades não lineares (HAYKIN, 2001).

## 4.1 Contexto Histórico

O surgimento da neurocomputação teve início na década de 40. Em 1943, Warren McCulloch e Walter Pitts publicaram um artigo descrevendo um cálculo lógico das redes neurais que combinava estudos de neurofisiologia e lógica matemática (MCCULLOCH; PITTS, 1943). Esse trabalho foi tão significativo que é considerado o marco inicial dos estudos em inteligência artificial e redes neurais (HAYKIN, 2001).

McCulloch e Pitts basearam-se em três fontes principais: o conhecimento da fisiologia básica e da função dos neurônios no cérebro, uma análise formal da lógica proposicional desenvolvida por Russel e Whitehead, e a teoria da computação de Turing. Eles demonstraram que qualquer função computável poderia ser calculada por uma rede de neurônios conectados e que todos os conectivos lógicos poderiam ser implementados por estruturas de redes simples. Além disso, sugeriram que redes definidas adequadamente seriam capazes de aprender (NORVIG; RUSSELL, 2004).

Em 1949, Donald Hebb demonstrou uma regra simples de atualização (aprendizagem de Hebb) para modificar as intensidades das conexões entre os neurônios (HEBB, 1949). Embora já houvesse diversos trabalhos relacionados à área de inteligência artificial (IA) até então, foi Alan Turing, em 1950, quem primeiro condensou a ideia de IA em seu artigo, apresentando o teste de Turing, a aprendizagem de máquina, algoritmos genéticos e aprendizagem por reforço.

O primeiro neurocomputador, SNARK, foi criado em 1951 por Marvin Minsky. O perceptron, uma das primeiras redes neurais artificiais (RNA) com apenas um neurônio e aprendizagem supervisionada, foi apresentado por Rosenblatt em 1958. Dois anos depois, Widrow e Hoff desenvolveram o algoritmo dos mínimos quadrados. No entanto, em 1969, o livro de Minsky e Papert demonstrou limitações fundamentais dos perceptrons de camada única (MINSKY, 1986), o que resultou em uma desaceleração da pesquisa nessa área.

Na década de 80, as redes neurais ressurgiram com base na teoria de perceptrons de múltiplas camadas (MLP) treinados com o algoritmo de retropropagação *backpropagation*, descoberto por Bryson e Ho em 1969 e posteriormente desenvolvido por Rumelhart, Hinton e Williams em 1986. O aumento da capacidade de processamento e armazenamento dos computadores também contribuiu para esse ressurgimento, possibilitando simulações que eram computacionalmente impossíveis duas décadas antes. Em 1987, ocorreu a primeira conferência de redes neurais, a IEEE International Conference on Neural Networks, e dois anos depois foi fundada a INNS (International Neural Networks Society) (HAYKIN, 1994).

## 4.2 Fundamentos Matemáticos de uma Rede Neural

Classificar em duas categorias é uma das tarefas essenciais no campo de aprendizado de máquina, cujo objetivo é designar uma entre duas possíveis categorias a um conjunto específico de dados. Vários setores, como diagnóstico médico, detecção de spam em e-mails, reconhecimento facial e outros têm esse tipo de problema comum.

No aprendizado de máquina, um modelo de neurônio artificial é uma abstração baseada no funcionamento do cérebro humano. Esses modelos podem aprender padrões complexos nos dados e realizar tarefas como classificação, regressão, reconhecimento de padrões, entre outras.

Um **perceptron**, que também é conhecido como neurônio artificial, é uma unidade básica desse modelo. Ele executa várias entradas, realiza uma combinação linear dessas entradas ponderadas por pesos específicos e aplica uma função de ativação para gerar um resultado. Durante o treinamento, os pesos do neurônio são ajustados repetidamente para reduzir algum critério de erro, como a diferença entre a saída prevista e a saída real.

No âmbito da classificação binária, o neurônio artificial passa por treinamento para dividir os dados em duas categorias diferentes. Encontra-se a fronteira de decisão no espaço de entrada para separar efetivamente os exemplos positivos dos nulos. A complexidade do problema determina a forma da fronteira de decisão, que pode ser uma linha reta em espaços bidimensionais ou superfícies mais complexas em espaços de alta dimensão.

Depois de ser treinado, o modelo de neurônio artificial pode ser usado para prever ou classificar novos exemplos. Basta fornecer as características do exemplo como entrada para o neurônio, e ele produzirá uma saída indicando a categoria prevista.

De forma sucinta, a classificação binária em aprendizado de máquina usando um modelo de neurônio artificial é uma poderosa e versátil abordagem para resolver uma ampla gama de problemas de classificação em diversas áreas.

O aprendizado de máquina pode ser contextualizado para descrever o processo de classificação binária usando um modelo de neurônio artificial. Em resumo:

**Dados Conhecidos:** Pontos no plano representam exemplos previamente rotulados. Os pontos azuis representam as situações em que a informação é “si”, ao passo que os pontos alaranjados sinalizam as situações em que a informação é “não”.

**Modelo de Neurônio:** Este modelo foi treinado para distinguir e classificar os dados em duas categorias diferentes com base nos exemplos fornecidos. Durante o treinamento, o modelo ajusta seus parâmetros internos para maximizar a capacidade de distinguir entre as classes.

**Separação de Regiões:** O objetivo do modelo é encontrar uma fronteira de decisão que divida o espaço de entrada em áreas correspondentes às diferentes classes de forma eficiente. Esta fronteira pode ser linear ou complexa, dependendo da natureza dos dados e da arquitetura do modelo.

**Diagnóstico:** Após o treinamento, é possível que o modelo faça previsões ou diagnósticos para novos pontos no plano. Para um ponto qualquer, o modelo determina sua categoria prevista com base na qual região do plano ele pertence:

Se estiver na área azul, será considerado “sim”; se estiver na área alaranjada, será considerado “não” como mostra a figura 4.4.

O modelo de neurônio exemplifica a essência do processo de classificação binária em aprendizado de máquina, demonstrando uma abordagem para resolver esse tipo de problema. É importante destacar que existem muitas outras técnicas e algoritmos para classificação, cada um com suas próprias vantagens e aplicações específicas.

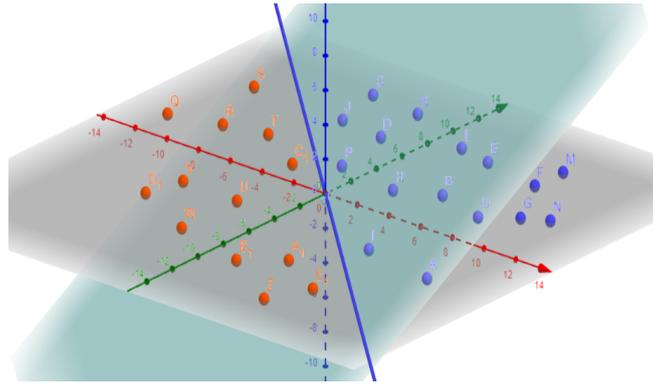
Apresentamos neste trabalho a essência da Matemática por trás do funcionamento de um modelo muito usado nos dias de hoje, conhecido como rede neural artificial. Esse modelo é fundamental para o entendimento de uma Inteligência Artificial, a qual utilizando-se um computador é possível por exemplo, reconhecer imagens, avaliar risco de empréstimos e seguros, ajudar no diagnóstico médico, etc. O objetivo aqui é dar uma noção intuitiva e algumas definições para motivar o entendimento de uma rede neural artificial. O livro 16 é uma excelente referência que aborda a Matemática de uma rede neural artificial de modo bem detalhado.

### 4.2.1 Definindo Uma Rede Neural Artificial

Embora esse assunto tenha se popularizado nos dias de hoje, os estudos desse modelo acontecem desde os anos 60. Com a melhoria do poder computacional e o armazenamento de grandes quantidades de dados, foi possível aplicar todo o conhecimento adquirido no passado. De fato esse é um exemplo interessante de como funciona a pesquisa na Matemática. O modelo é criado em uma época que não tinha como aplicar (não existiam computadores com capacidade suficiente), o estudo perde um pouco da importância e depois de algum tempo é possível aplicá-lo e, neste caso, se torna muito importante para diversas áreas da ciência.

**Definição 1:** Um neurônio é uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  da forma  $f(x) = \lambda(\alpha \cdot \mathbf{x} + \beta)$ , onde  $\lambda : \mathbb{R} \rightarrow \mathbb{R}$  é uma função contínua chamada função de ativação,  $\alpha \in \mathbb{R}^n$  é o vetor de pesos e o escalar  $\beta \in \mathbb{R}$  é o viés. O produto  $\alpha \cdot \mathbf{x}$  representa o produto interno em  $\mathbb{R}^n$ .

Um exemplo simples de um neurônio é quando  $\lambda$  é a função identidade, ou seja  $\lambda(x) = x$  e o neurônio uma função afim, cujo gráfico é um hiperplano  $G = \{(\mathbf{x}, \alpha \cdot \mathbf{x} + \beta) : \mathbf{x} \in \mathbb{R}^n\}$  do espaço  $\mathbb{R}^{n+1}$ . Neste caso a interseção de  $G$  com o domínio  $\mathbb{R}^n$  o divide em duas regiões. Região 1: pontos  $\mathbf{x} \in \mathbb{R}^n$  tais que  $\alpha \cdot \mathbf{x} + \beta \geq 0$  e Região 2: pontos  $\mathbf{x} \in \mathbb{R}^n$  tais que  $\alpha \cdot \mathbf{x} + \beta < 0$ . Na prática um neurônio faz uma tomada de decisão entre o sim ou não. Pontos que são de uma determinada região são associados ao SIM. E os da região complementar são associados ao NÃO.



**Figura 4.4:** Exemplo Separação de Um Neurônio

A figura acima exemplifica um cenário hipotético onde os pontos coloridos são os dados já conhecidos. Digamos que os pontos azuis são os pontos onde a informação é SIM e os alaranjados são os pontos onde a informação é NÃO. Na prática o que esse neurônio faz é separar todos os dados em duas regiões. Dessa forma, quando se quer fazer um diagnóstico com relação a um ponto qualquer do plano basta verificar se esse ponto está na região dos azuis (SIM) ou dos alaranjados (NÃO). Nesse caso o neurônio é dado por  $f(x,y) = x + y$  e seu gráfico está esboçado na figura acima. A reta obtida pela interseção de seu gráfico com o plano  $z = 0$  faz o serviço de separação dos dados e portanto, neste caso de distribuição dos dados, onde os pontos são ditos *separáveis*, um neurônio é suficiente para a tomada de decisão.

Um algoritmo de aprendizado de máquinas usado para encontrar tais neurônios é chamado *algoritmo do perceptron* e pode ser encontrado em 17.

Na maioria dos problemas aplicados os dados não são separáveis. Então uma conclusão natural é que precisamos encontrar mais neurônios para fazer tal classificação. Para isso definimos uma camada de neurônios:

**Definição 2:** Uma função camada é uma aplicação  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  da forma

$$g(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})),$$

onde cada  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função neurônio da forma  $f_i(x) = \lambda(\alpha_i \cdot \mathbf{x} + \beta_i)$  com vetor de parâmetros  $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$  e vieses  $\beta_i, i = 1, \dots, m$ .

**Observação 1:** Observe que se  $m = 1$  a camada é um neurônio.

**Observação 2:** A função de ativação  $\lambda : \mathbb{R} \rightarrow \mathbb{R}$  é a mesma para todos os neurônios da camada. Um exemplo muito usado como função de ativação é a função *ReLU*, que tem a função de ativar ou desativar o estado de um neurônio, e é definida por

$$ReLU(x) = \begin{cases} x, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases}$$

Um elemento do domínio  $\mathbb{R}^n$  de uma função camada representa a entrada de um objeto (com  $n$  entradas) que queremos classificar, mas observe que a dimensão  $m$  do

contradomínio da camada representa o número de neurônios que essa camada possui, que não necessariamente coincide com o número de classes possíveis de classificação. Por exemplo, se queremos classificar um ponto como SIM ou NÃO, o contradomínio deve ser  $\mathbb{R}$ . Os valores positivos representam o SIM e os nulos representam o NÃO. Neste caso, precisamos de uma composição de camadas, obtendo uma rede neural artificial com duas funções camadas.

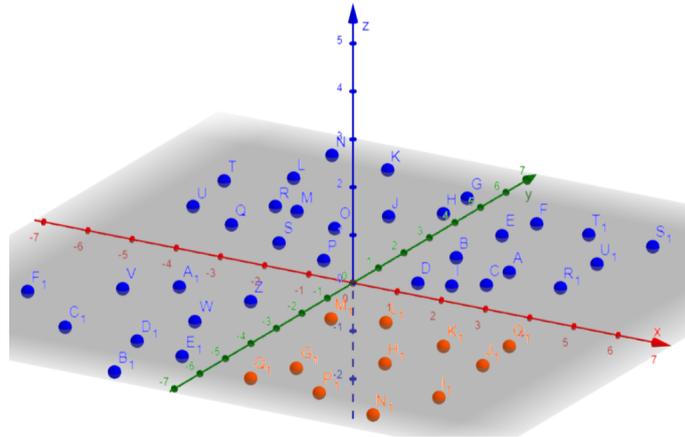
**Definição 3:** Uma rede neural artificial (RNA) é uma aplicação  $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$  escrita como composição de funções camadas, isto é,

$$h = h_M \circ \dots \circ h_2 \circ h_1.$$

Neste caso,  $k$  é o número de classes e  $M$  o número de camadas da RNA. Uma RNA que possui mais de uma camada é chamada de rede neural profunda (RNP) e as camadas que ficam entre a entrada e a saída da RNA são chamadas camadas ocultas.

**Observação 3:** Uma camada intermediária se chama camada escondida pelo fato de que sua saída é uma entrada de uma outra camada, sem que o usuário a conheça. Por exemplo, em uma RNP com três camadas  $h = h_3 \circ h_2 \circ h_1$ , a saída (imagem) de  $h_1$  será a entrada (domínio) de  $h_2$  e a saída de  $h_2$  será entrada de  $h_3$ . Portanto  $h_1$  e  $h_2$  são as camadas escondidas da RNP.

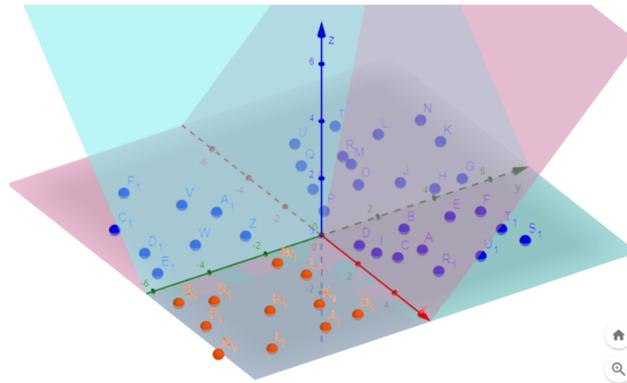
Vejamos um exemplo de distribuição de pontos (dados) no plano o qual não é possível separar (classificar) em duas classes com apenas um neurônio:



**Figura 4.5:** Exemplo Classificação com Um Neurônio  
Fonte: O Autor

Na imagem abaixo, o gráfico na cor azul é o gráfico da função neurônio  $f(x,y) = ReLU(y)$  e o da cor rosa é o gráfico da função  $g(x,y) = ReLU(x)$

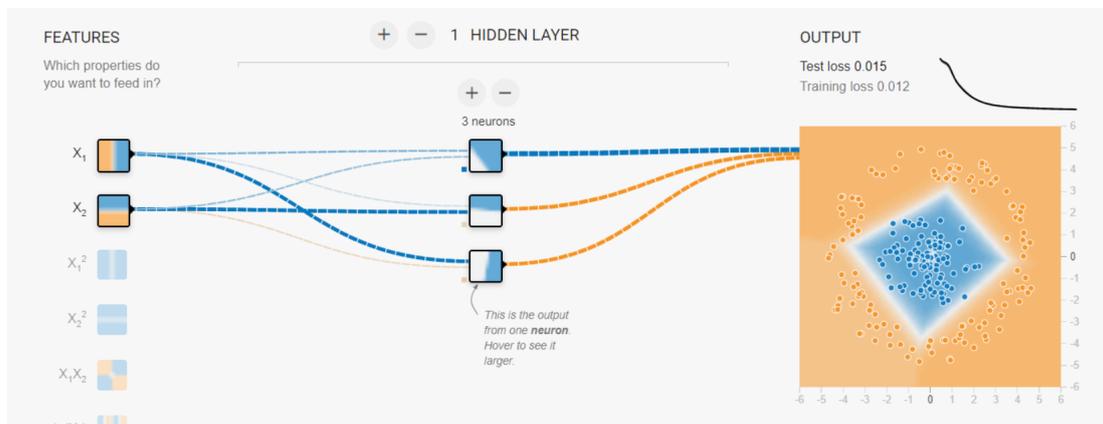
Então, se considerarmos as camadas  $h_1(x,y) = (f(x,y), g(x,y))$  e  $h_2(f,g) = ReLU(f + g)$ , teremos que a RNP  $h(x,y) = h_2 \circ h_1$  vai classificar nossos dados em duas classes.



**Figura 4.6:** Exemplo Funções ReLU  
Fonte: O Autor

### 4.2.2 Playground do Tensorflow

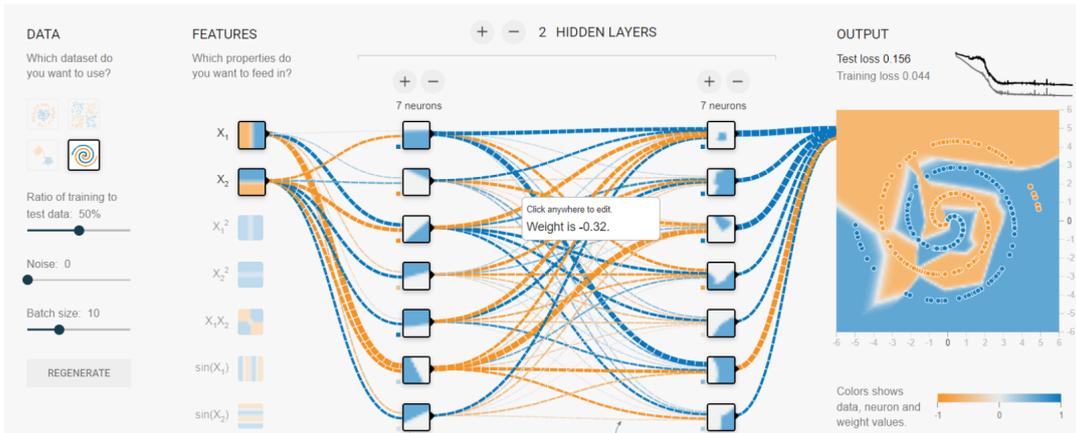
Outros tipos de distribuição mais complexas dos pontos vão exigir um número maior de neurônios na camada escondida. É aí que entra o poder computacional atual. Uma plataforma bem didática que nos dá uma ideia dessas possibilidades é o playground do Tensorflow, disponível em: [playground.tensorflow.org](https://playground.tensorflow.org). Nela é possível ver a partir de alguns exemplos de distribuição no plano que, diante da impossibilidade com determinado número de neurônios, ir aumentando até se chegar numa separação adequada. A configuração abaixo exige 3 neurônios em uma camada escondida:



**Figura 4.7:** Exemplo Uma Camada e Três Neurônios  
Fonte: O Autor

Já o exemplo abaixo exige um cálculo computacional muito maior. Foi possível separar com 7 neurônios em cada uma das duas camadas escondidas. Mesmo assim, foi necessário um tempo bem maior do que o exemplo anterior e houve um pequeno erro.

Observe na rede neural anterior que ela não foi capaz de separar todos os pontos. Houve um pequeno erro. Isso é devido a minimização de uma função chamada *função perda*, a qual não entraremos em detalhes mas comentaremos um pouco mais a respeito na próxima seção.



**Figura 4.8:** Exemplo Duas Camadas e Sete Neurônios  
Fonte: O Autor

Nessa plataforma há outras configurações de pontos e o usuário pode fazer mudanças nas quantidades de camadas e neurônios. Ela é uma ferramenta muito interessante para ser usada no ensino desse tema.

### 4.2.3 Modelando uma função a partir de dados

O principal motivo de um modelo de RNP ser tão eficiente é o fato de que toda função contínua, não-negativa e definida num conjunto  $K$  limitado e fechado pode ser aproximada por uma rede neural com apenas duas funções camadas. Esse é o Teorema de aproximação universal e sua demonstração pode ser encontrada em 19

**Teorema 1:** (Teorema da Aproximação Universal): Sejam  $K \subset \mathbb{R}^n$  um conjunto limitado e fechado e  $f : K \rightarrow \mathbb{R}^m$  função contínua cujas componentes são não negativas, isto é  $f_i(x) \geq 0$  para todo  $x \in K$  e  $i = 1, \dots, m$ . Então dado  $\epsilon > 0$  existe uma rede neural com duas funções camadas  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  tal que:

$$\max_{x \in K} \|h(x) - f(x)\| < \epsilon$$

Esse Teorema é muito importante porque, ao modelar um problema de dados, o que na prática você quer é uma função de modo que estes dados estejam no gráfico desta função. Isso seria o que chamamos de interpolação (por ex. interpolação polinomial). Mas como o volume de dados é muito grande, interpolar é inviável e daí aproximar é uma boa opção. E graças a este Teorema, aproximar com uma RNP é uma excelente opção.

O problema computacional é que, ao aproximar de uma função, pode ser necessário ter na camada escondida da RNP um número muito grande de neurônios, o que inviabiliza o cálculo. Por isso, a depender do que se quer modelar, pode ser necessário mais camadas. De outro modo, podemos pensar que diminuir o erro de aproximação é o mesmo que minimizar a função perda.

Minimizar a função perda é o grande desafio computacional por trás de uma RNA quando se tem um volume muito grande de dados. Isso porque a função perda tem como variáveis de entrada todos os parâmetros (pesos) e encontrar os parâmetros é equivalente a encontrar todos os neurônios. Além disso ela depende de todos os pontos (dados) e, por ser uma função convexa, possui um único mínimo global. Portanto, encontrar esse mínimo é equivalente a resolver um sistema de equações: derivada (gradiente) igual a zero. Para derivar usa-se um algoritmo chamado *Backpropagation*. Mas resolver a equação é uma tarefa quase impossível, visto que seu número de equações depende do número de parâmetros. Então é lançado mão de um método numérico, chamada **gradiente descendente**, com objetivo de aproximar desse ponto de mínimo. O modelo matemático usado neste método foi abordado na seção 2.1.1 e os demais detalhes podem ser encontrados em 16. 16

## Aplicação no Ensino Médio

---

Após o desenvolvimento teórico abordado nos capítulos anteriores, partimos para a prática, afim de concluirmos esta sequência didática que busca trazer conhecimento e incentivar os alunos a se interessarem pelas novidades tecnológicas e pela IA.

Por meio de um simples projeto envolvendo uma máquina programável que pode realizar atividades impressionantes e por um desenvolvimento de um neurônio artificial, adaptado para diferentes níveis educacionais, este projeto busca promover o interesse pela Matemática e Computação, ao mesmo tempo em que destaca a relevância da Lógica Matemática no progresso tecnológico e nas possibilidades que nos aguardam.

Os Parâmetros Curriculares Nacionais salientam a relevância de temas transversais que estabeleçam conexões entre diferentes áreas do conhecimento no Ensino Médio e Fundamental. Embora possua sua própria linguagem, é indispensável que a Matemática estabeleça diálogo com outras disciplinas, incluindo a Inteligência Artificial (IA), visando à abordagem interdisciplinar.

Pela BNCC, EM03IA03: Saber treinar uma IA por meio da técnica de *Machine Learning*.

Podemos perceber a importância do desenvolvimento do trabalho com IA para alunos da terceira série do Ensino Médio.

Finalmente, procuramos com isso em levantar questões pertinentes sobre como os computadores estão integrados ao ensino, defendendo a implementação das bases científicas associadas à computação e sua aplicabilidade para ampliar o raciocínio computacional no Ensino Médio.

### 5.1 O Jogo NIM

Na Antiga China surgiu o Nim, um antigo jogo de estratégia para dois jogadores. Um fascinante desafio intelectual que tem cativado mentes ao longo dos séculos. Com base em um conjunto simples de regras, o jogo Nim requer dos jogadores não apenas habilidades lógicas, mas também uma compreensão aguçada das dinâmicas estratégicas envolvidas.



**Figura 5.1:** Exemplo Inicial NIM

Em sua forma mais básica, o jogo consiste em uma formação de palitos, sendo dispostos inicialmente em colunas como mostra a figura 5.1. Os jogadores alternam turnos, escolhendo uma coluna e removendo um número qualquer de palitos dessa mesma coluna. O objetivo é ser o jogador a remover o último objeto, pois o jogador que retirar o último palito alcança a vitória.

A simplicidade do jogo esconde uma complexidade estratégica profunda. Os jogadores devem não apenas considerar suas próprias ações, mas também antecipar e responder às jogadas do oponente. Uma estratégia central é garantir que se deixe sempre um número ímpar de colunas contendo um número ímpar de palitos. Isso se deve ao fato de que, ao fazer isso, você restringe as opções do oponente e o coloca em uma posição onde ele pode ser forçado a fazer uma jogada que leva à derrota.

Quando há um número ímpar de colunas contendo um número ímpar de palitos, isso limita as escolhas do oponente. Ele não pode simplesmente equilibrar as colunas retirando o mesmo número de palitos de cada uma delas, o que restringe suas opções estratégicas. Isso permite que você mantenha o controle do jogo e planeje suas jogadas de forma a manter a vantagem tática.

Em resumo, essa estratégia é eficaz porque cria uma situação na qual o oponente tem menos flexibilidade e pode ser forçado a fazer jogadas que favorecem sua própria vitória.

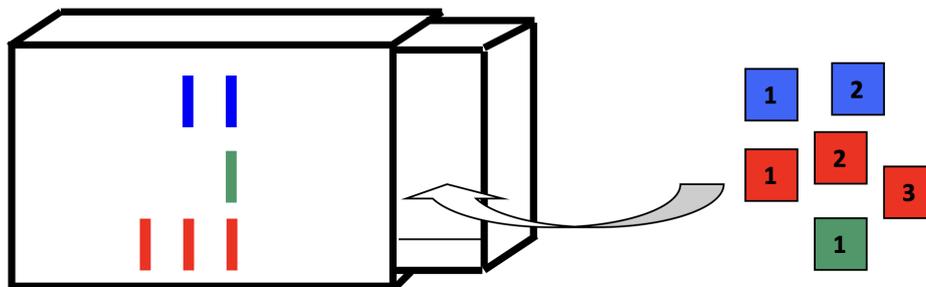
Além de sua aplicação puramente lúdica, o jogo Nim será utilizado aqui como uma ferramenta educacional para demonstrar como o aprendizado de máquina pode ser desenvolvido sem o uso de eletrônicos. Sua simplicidade inicial dá lugar a uma riqueza de desafios estratégicos e cognitivos, tornando-o um valioso recurso pedagógico em ambientes educacionais.

## 5.2 Desenvolvendo um Computador para Jogar NIM

A ideia é desenvolver um computador capaz de jogar o jogo NIM. A fim de concluir esse projeto, é necessário obter 18 caixas vazias de fósforo junto com pequenos pedaços das folhas grossas do papel cartão nas tonalidades azul, vermelha e verde.

As caixas de fósforos com cores distintas (por exemplo: azul, verde e vermelho) serão usadas para identificar todas as potenciais jogadas no processo da construção do jogo. Numerando-os do número um ao número três, dentro das caixas serão

colocados pequenos quadrados feitos em papel cartão nas mesmas cores usadas fora. Os palitos que devem ser retirados durante a partida serão indicados pelos números correspondentes. Por exemplo, dentro da caixa específica estarão os cartões correspondentes à jogada ao lado dela:



**Figura 5.2:** Exemplo Montagem Caixa de Fósforos

No jogo, agitamos a caixa e, sem espiar, retiramos uma das seis pequenas fichas. O número e a cor selecionados orientam nossa jogada subsequente.

Por exemplo, se retirarmos a ficha azul com o número 2, devemos remover dois palitos da primeira fileira, representada pela cor azul. Se a ficha retirada for azul e mostrar o número 1, retiramos um palito da primeira fileira (azul). Se a ficha mostrar uma cor diferente, devemos retirar da linha correspondente à cor o número indicado. Agora, o aspecto mais interessante em relação à aprendizagem: caso o computador perca uma jogada, identificamos a última ficha que levou à derrota e retiramos ela da caixa. Se ele vencer, reforçamos toda a sequência vencedora desde a primeira jogada, permitindo que o computador 'aprenda' e repita esse padrão para vencer novamente.

Portanto, nas primeiras tentativas é provável que os alunos percam o jogo com frequência. Porém à medida em que continua jogando, ele aprende a vencer até se tornar imbatível.

Os alunos terão a oportunidade de formar grupos para o desenvolvimento do computador e, em seguida, iniciarão o treinamento jogando com ele. Ao longo de algumas rodadas, será perceptível o processo de aprendizado se iniciando. Após varias rodadas, o computador em caixas de fósforos estará bem treinado e pronto para ser desafiado por outros alunos. Para concluir, eles poderão apresentar o computador "invencível" para outras séries do colégio afim de que tentem vence-lô.

Podemos encontrar as caixas necessárias para montagem do jogo no Apêndice 1.

### 5.3 A Aplicação

A elaboração de um computador capaz de jogar e ganhar o jogo de NIM com alunos do ensino médio oferece uma abordagem prática e envolvente para o aprendizado. Ela combina teoria e prática, permitindo que os alunos compreendam os conceitos por trás do computador e de uma IA, enquanto jogam e o experimentam de maneira interativa.

Podemos ressaltar alguns pontos positivos na aplicação, como:

**Experiência Interativa:** A aplicação oferece uma experiência prática, permitindo que os alunos joguem e experimentem diferentes estratégias.

**Relevância Educativa:** Ajuda os alunos a desenvolver habilidades lógicas, estratégicas e matemáticas enquanto se divertem.

**Customização de Dificuldade:** A possibilidade de ajustar o nível de dificuldade do computador permite que os alunos enfrentem desafios adequados ao seu nível de habilidade.

**Aprendizado Iterativo:** A capacidade de aprender com as jogadas e reforçar sequências vencedoras cria um ciclo de aprendizado contínuo.

Também existem algumas possíveis melhorias a serem realizadas na aplicação, como:

**Ampliação de Recursos Didáticos:** Adição de recursos educacionais complementares, como tutoriais interativos ou quizzes, para reforçar ainda mais o aprendizado.

**Feedback Aprofundado:** Oferecer *feedback* mais detalhado sobre as estratégias utilizadas pelos alunos ou *insights* sobre as decisões do computador.

**Incorporação de Elementos Visuais:** Elementos visuais mais detalhados ou animações para ilustrar as jogadas e tornar a experiência mais envolvente.

Em resumo, essa aplicação se mostra como uma importante ferramenta no apoio aos alunos ao entenderem as bases lógicas que fundamentam tanto o funcionamento dos computadores quanto das aplicações da IA, pois demonstra uma das ferramentas mais utilizadas por ela, o aprendizado de máquina. Propiciando então ao estudante aprender mediante experiências concretas.

Partiremos agora para a segunda aplicação a ser desenvolvida com os alunos.

## 5.4 Aplicação de Redes Neurais na Precificação de Seguros Automotivos

Além da primeira aplicação, desenvolveremos agora outra dinâmica com os alunos, afim de que possam compreendam melhor o funcionamento prático das redes neurais, mais uma ferramenta essencial da Inteligência Artificial (IA). Vamos explorar um exemplo didático e simples, que possa facilitar a compreensão dos conceitos fundamentais.

Este exemplo foi elaborado com o objetivo de simplificar o entendimento das redes neurais para os alunos. Os dados foram cuidadosamente selecionados e desenvolvidos de forma a tornar o exemplo acessível, utilizando apenas o plano cartesiano e representações em duas dimensões.

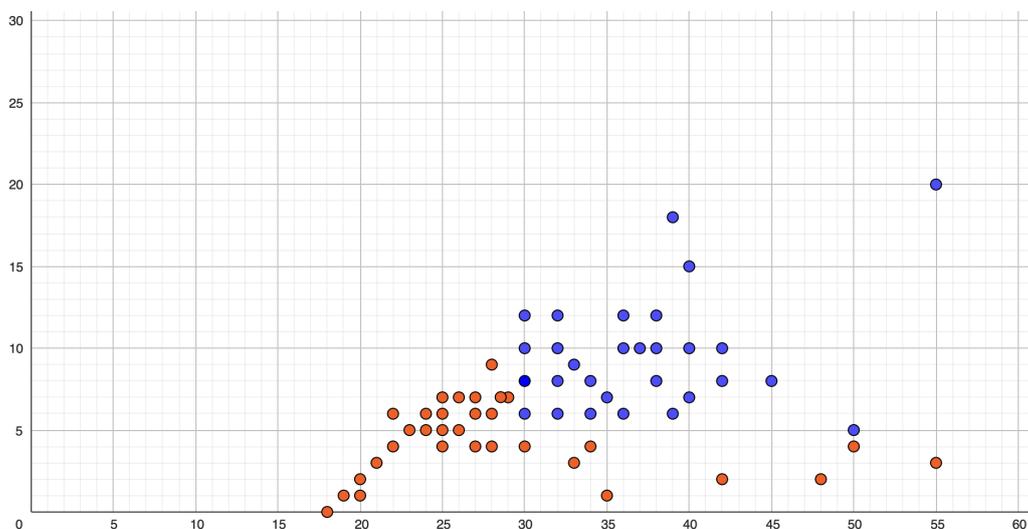
A indústria de seguros automotivos tem evoluído consideravelmente nas últimas décadas, impulsionada pela necessidade de métodos mais eficientes e personalizados na determinação dos valores dos seguros. Uma abordagem inovadora nesse cenário é a aplicação de redes neurais, que permite uma análise mais aprofundada das características dos segurados.

Determinar preços justos e competitivos para os seguros automotivos é um desafio complexo. Uma seguradora enfrenta a necessidade de equilibrar a atratividade de

seus produtos com a necessidade de cobrir adequadamente os riscos associados aos diferentes perfis de motoristas. A idade do motorista e o tempo de posse da carteira são fatores críticos na avaliação desses riscos.

Para análise inicial foram colhidos dados, que estão representados na tabela presente no apêndice 2, de 60 possíveis novos assegurados.

A escolha de uma arquitetura de rede neural adequada é crucial para o sucesso do modelo. Portanto, a seguradora implementou uma rede neural, utilizando funções de ativação específicas para lidar com a natureza dos dados. A entrada da rede consiste na idade do motorista e no tempo de posse da carteira, enquanto a saída representa o preço do seguro.



**Figura 5.3:** Gráfico dos Pontos Gerados Pelos Dados dos Motoristas

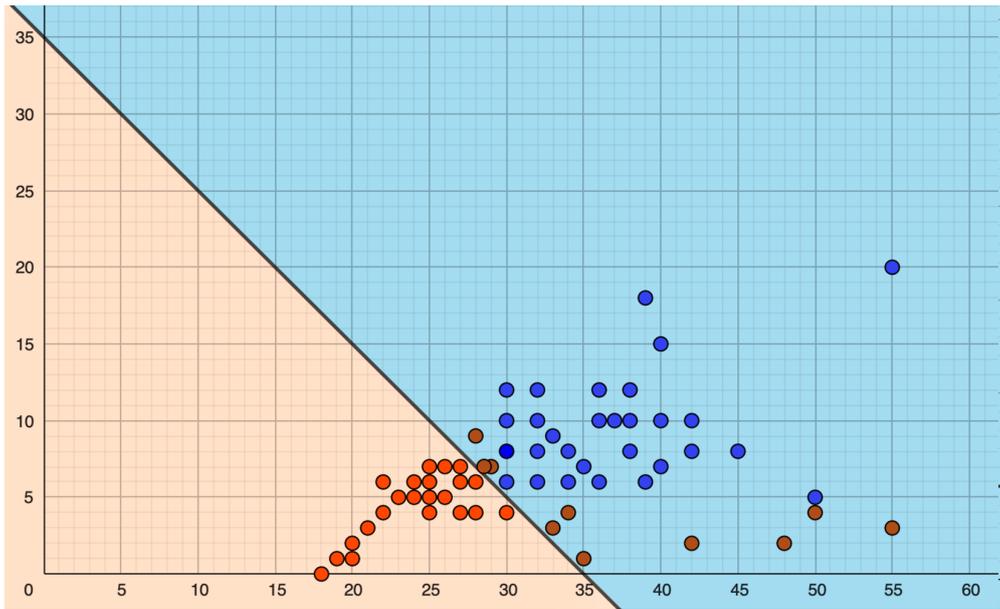
Na Figura 5.3, podemos observar que o eixo Y representa os anos de habilitação que o associado possui, enquanto o eixo X representa a idade do condutor.

O treinamento da rede neural envolveu a otimização dos pesos dos neurônios para minimizar a diferença entre os preços previstos e os preços reais observados no conjunto de dados de treinamento.

Os alunos serão desafiados a descrever esta função de um neurônio para tentar classificar os dados em duas categorias: SIM (representados por pontos azuis), indicando que o segurado apresenta um menor risco e, conseqüentemente, terá um desconto no valor da apólice; ou NÃO (pontos laranjas), indicando que o segurado apresenta um risco maior e, portanto, terá um valor mais alto para a apólice. que minimize ao máximo o erro e consiga separa todos esses pontos de forma mais eficiente.

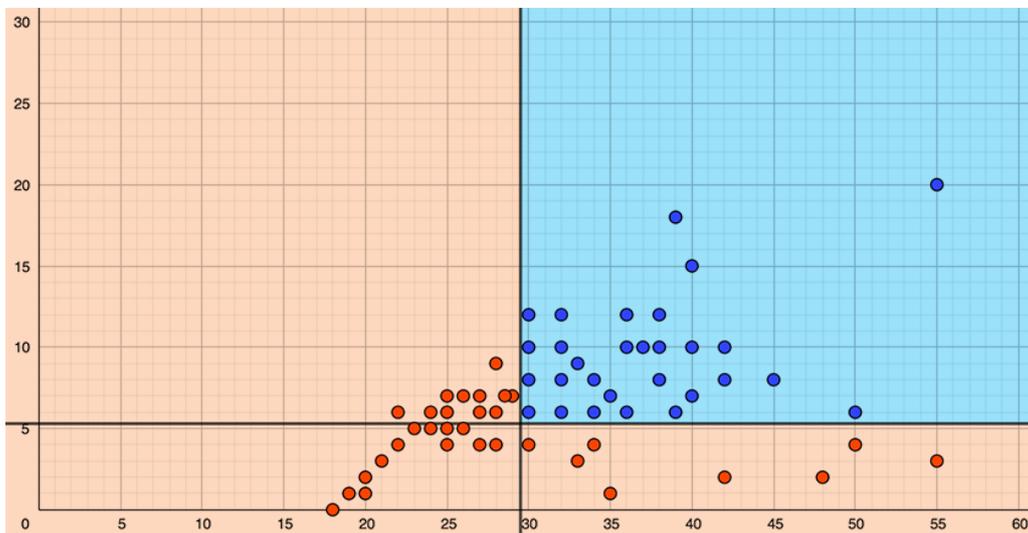
Após a tentativa dos alunos, mostraremos a eles qual seria a opção que seria mais eficiente com apenas uma neurônio, podemos ver a função esboçada na figura 5.4.

Ao analisarmos o gráfico acima, em que foi construída a função que melhor divide os dados  $y = -x + 35$ , os alunos poderão observar um erro considerável decorrente do uso de apenas um neurônio para classificação. No entanto, ao expandirmos a rede neural e empregar múltiplos neurônios para abordar o mesmo problema, notamos



**Figura 5.4:** Gráfico com Um Neurônio

uma significativa melhora na capacidade de separação dos valores. Essa otimização nos permite reduzir substancialmente o erro de classificação. Portanto, ao confiar na rede neural para lidar com dados futuros, podemos ter maior confiança na precisão das previsões e na capacidade do modelo de aprendizado de máquina em lidar com tarefas complexas de classificação.



**Figura 5.5:** Gráfico com mais Neurônios

Os estudantes deverão notar que a implementação da rede neural mais complexa resultou em uma maior precisão na precificação dos seguros, permitindo à seguradora oferecer preços mais competitivos e personalizados. A partir do treinamento da IA com dados conhecidos, novos possíveis assegurados serão calculados de forma automática pela rede neural desenvolvida por nós na aplicação. No entanto, desafios contínuos incluem a necessidade de atualização constante do modelo para se adaptar

a mudanças nas tendências e comportamentos dos segurados.

Assim, com base nos dados disponíveis para o treinamento da rede neural, estamos prontos para implementá-la na seguradora. Isso significa que, ao se cadastrar, o usuário será automaticamente classificado de acordo com o seu perfil de risco, determinado pela rede neural. Com essa classificação precisa, poderemos oferecer descontos por menor risco de forma justa e eficiente.

Para treinar essa rede, utilizamos um conjunto de dados conhecidos como conjunto de treinamento, que consiste em exemplos onde já sabemos a entrada e a saída desejada. Durante o treinamento, ajustamos os pesos das conexões entre os neurônios para minimizar a diferença entre as saídas previstas pela rede e as saídas reais do conjunto de treinamento. Esse processo é repetido várias vezes até que a rede seja capaz de produzir saídas precisas para novos conjuntos de entrada.

É importante salientar que a rede neural não alcançará um erro igual a zero na prática, como exemplificado anteriormente. Isso ocorre porque, na realidade, os dados e os sistemas são suscetíveis a variações e imperfeições. O exemplo dado foi simplificado para facilitar o entendimento didático. Na prática, durante o treinamento da rede neural, buscamos minimizar o erro o máximo possível, mas geralmente não atingimos uma precisão absoluta. Isso é importante para entendimento que a tecnologia, embora avançada, tem suas limitações e que sempre há margem para melhoria e refinamento.

Ao mostrar esse processo de forma visual e prática, os alunos podem compreender melhor como a inteligência artificial, especificamente as redes neurais, são desenvolvidas e treinadas para desempenhar tarefas específicas, como determinar o valor do seguro automotivo com base em diferentes variáveis. Isso os ajuda a entender o funcionamento por trás de tecnologias que cada vez mais impactam o nosso cotidiano.

Na primeira dinâmica, os alunos desenvolvem uma compreensão fundamental do aprendizado de máquina, explorando conceitos-chave e aplicação em uma situação do mundo real. Isso estabelece uma base sólida para entender os princípios subjacentes da IA.

Na segunda dinâmica, os alunos se aprofundam na compreensão das redes neurais, explorando sua estrutura e funcionamento. Eles podem experimentar a construção e treinamento de redes neurais simples em ambientes controlados, permitindo uma compreensão prática de como esses sistemas aprendem e fazem previsões.

Ao combinar essas duas dinâmicas, os alunos poderão ser capazes de conectar os conceitos de aprendizado de máquina com a prática específica das redes neurais. Ver como os princípios aprendidos na primeira dinâmica se manifestam na implementação e treinamento de redes neurais na segunda dinâmica.

Com esta sequência didática, propomos uma abordagem interativa e engajadora para os alunos do ensino médio, explorando a matemática por trás das inteligências artificiais. Ao desvendar como algoritmos aprendem e evoluem, e discutir as perspectivas de crescimento deste campo fascinante nos próximos anos podemos preparar melhor o aluno para o que pode ser a grande nova revolução tecnológica. Através de atividades práticas e discussões teóricas, os estudantes são incentivados a refletir sobre o impacto da IA em nossa sociedade e as habilidades matemáticas necessárias

para contribuir com este avanço tecnológico.

## Considerações Finais

---

Ao longo desta pesquisa acadêmica, mergulhamos em uma jornada para decifrar o mistério por trás da Inteligência Artificial (IA), explorando sua conexão intrínseca com os conceitos fundamentais da matemática e aprofundando no aprendizado de máquina e nas redes neurais, enfatizando seu potencial aplicativo no contexto escolar do ensino médio. Explorar os segredos da IA revelou-se uma demanda árdua, ao mesmo tempo em que nos deparamos com a empolgante jornada na descoberta das sutilezas desta revolucionária política.

O intuito da abordagem de desmistificação era garantir aos estudantes do ensino médio um entendimento consistente dos princípios subjacentes à IA, destacando-se sua natureza como uma ferramenta impulsionada por conceitos matemáticos acessíveis para exploração e compreensão. A combinação da matemática com áreas como álgebra geométrica ou análise do cálculo resulta em uma maior compreensão das aplicações reais desses conceitos.

Através do conhecimento dos princípios matemáticos básicos, tais como álgebra linear para representar dados, cálculo para otimizar modelos e probabilidade para lidar com incertezas, pode-se tornar possível obter uma visão mais clara acerca das bases sobre as quais se assenta a IA.

A implementação de atividades práticas, experimentos e estudos de caso possibilitou aos alunos uma experiência *hands-on* que estimulou o raciocínio crítico e a capacidade resolver problemas. Isso poderá os incentivar tanto academicamente quanto profissionalmente ao explorarem as possibilidades da AI em seu dia-a-dia agora bem como nas suas futuras carreiras.

À medida que damos término a esse trabalho, percebemos claramente que o processo de desvendar os segredos da IA vai além disso; ele também tem o objetivo principal de equipar os estudantes com maior compreensão e soluções inovadoras num mundo movido pela tecnologia. No contexto da educação no ensino médio, é fundamental promover o entendimento da IA nas mentes dos estudantes para que possam moldar seu desenvolvimento de maneira consciente e duradoura.

Consideramos o ensino de IA para alunos do ensino médio mais do que apenas transmitir conhecimento, envolve incentivar sua curiosidade enquanto exploramos juntos a beleza dessa tecnologia. Durante esse processo, os alunos podem perceber

como a inteligência artificial é fundamentada em uma lógica matemática subjacente, que tem o poder de influenciar o presente e moldar o futuro.

## Referências Bibliográficas

---

- 1) DASGUPTA, S. PAPADIMITRIOU, C. H.; VAZIRANI, U.V. *Algorithms* Disponível em: <http://algorithmics.lsi.upc.edu/docs/Dasgupta-Papadimitriou-Vazirani.pdf> 2006. Acesso em: 12 dez. 2023.
- 2) FOGG, Andrew. *A history of deep learning*. 2018. Em: <https://www.import.io/post/history-of-deep-learning/> acesso em 10 jan. 2024.
- 3) GABRIEL, Martha. *Você, eu e os robôs: como se transformar no profissional do futuro*. São Paulo: Atlas, 2a ed. 2021.
- 4) MALAGUTTI, Pedro. *Como Construir Máquinas que Aprendem: I. A. no Ensino Básico*. Disponível em: <https://impa.br/wp-content/uploads/2022/07/AulaPedro-Malagutti-PAPMEM-jul2022.pdf> acesso : 4jan.2024.
- 5) MCCARTHY, John. *What is AI?: Basic Questions*. Stanford, 2012. Disponível em: [encurtador.com.br/yBCJV](http://encurtador.com.br/yBCJV) Acesso em 05 nov. 2023
- 6) WIKIPEDIA. *Teste de Turing*. Disponível em: [https://pt.wikipedia.org/wiki/Teste\\_de\\_Turing](https://pt.wikipedia.org/wiki/Teste_de_Turing). acesso : 10dez.2023.
- 7) BOSE, A.; LIU, H.; WANG, F. *Intelligent Document Processing (IDP): Technology Landscape and Current Market Trends*. In: IEEE International Conference on Robotics and Automation (ICRA), 2019, p. 7679-7684.
- 8) HAYKIN, S. *Neural networks: a comprehensive foundation*. [s.l.] MacMillan Publishing Company, 1994
- 9) HAYKIN, S. S. *Redes neurais - princípios e prática*. 2ª edition ed. Porto Alegre:Bookman, 2001.
- 10) MCCULLOCH, W. S.; PITTS, W. *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Math. Biophys, v. 5, p. 115-133, 1943.

- 11) NORVIG, P.; RUSSELL, S. *Inteligência artificial*. Edição: 2<sup>a</sup> ed. [s.l.] Elsevier Acadêmico, 2004.
- 12) HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
- 13) ROSENBLATT, F. *The perceptron: a probabilistic model for information storage and organization in the brain*. Psychological review, v. 65, n. 6, p. 386, 1958.
- 14) MINSKY, M. *The society of mind*. New York: Simon and Schuster, 1986.
- 15) MORETTO, André Luis. *Desenvolvimento de Modelo Conceitual de Sistema de IA para Formação de Empreendedores*. São Paulo, 2023
- 16) BERLYAND, L., Jabin, P-E , *Mathematics of Deep Learning, An Introduction*, De Gruyter Textbook (2023).
- 17) FARIA, L., Oliveira, F. S., Pinto, P. E. D., Szwarcfiter, J. L., *Ciência de dados: algoritmos e aplicações*, 33<sup>o</sup> Colóquio Brasileiro de Matemática, IMPA (2021).
- 18) FACELI, K., Lorena, A. C. , Gama, J. e Carvalho, A. C. P. L. F. *Inteligência Artificial: uma Abordagem de Aprendizado de Máquina*. Rio de Janeiro, RJ: Livros Técnicos e Científicos (2011).
- 19) HORNIK, K., *Approximation capabilities of multilayer feedforward networks*. Neural Networks (1991).
- 20) MALAGUTTI, P. L., *Inteligência Artificial no Ensino Médio*, UFSCar (2002).

# Apêndice I

---

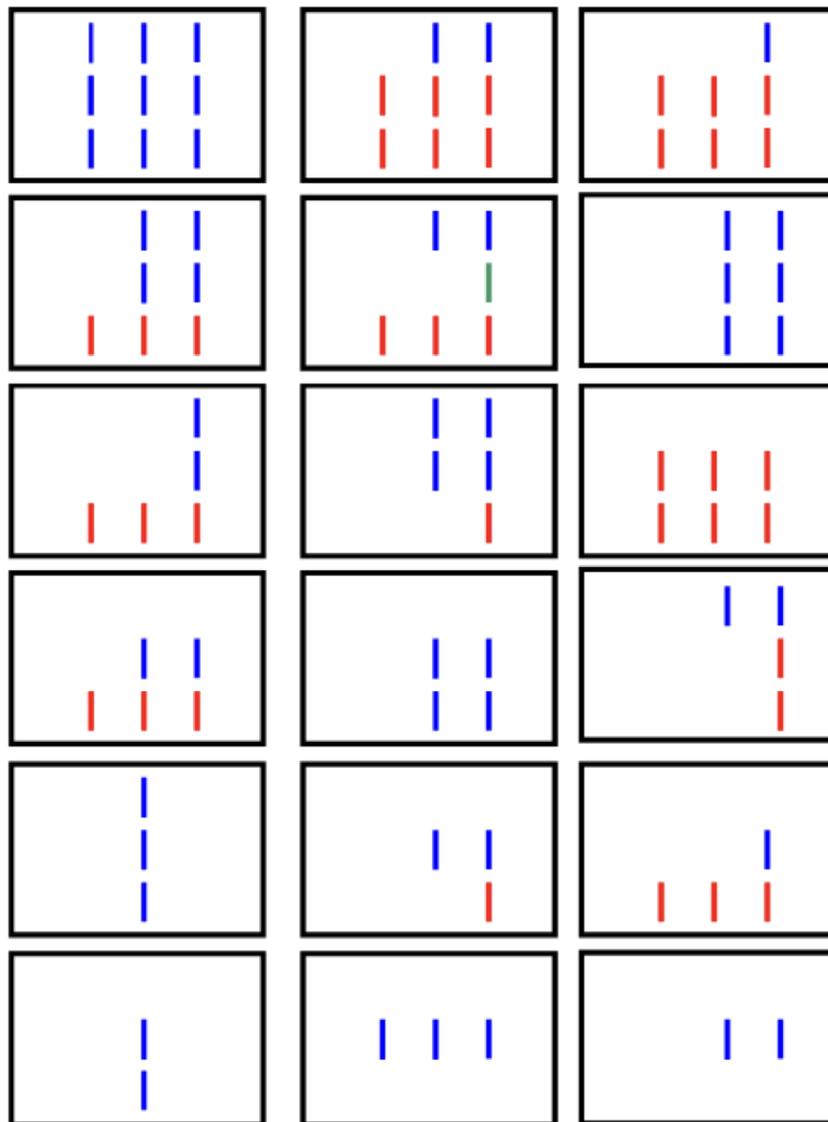


Figura A.1: Montagem Caixas de Fósforos

## Apêndice II

---

	Idade	Tempo de Habilitação
Pessoa 1	22	4
Pessoa 2	30	10
Pessoa 3	42	2
Pessoa 4	55	20
Pessoa 5	18	0
Pessoa 6	23	5
Pessoa 7	30	12
Pessoa 8	33	3
Pessoa 9	39	18
Pessoa 10	25	5
Pessoa 11	25	7
Pessoa 12	25	4
Pessoa 13	26	5
Pessoa 14	19	1
Pessoa 15	20	1
Pessoa 16	20	2
Pessoa 17	21	3
Pessoa 18	28	4
Pessoa 19	55	3
Pessoa 20	48	2
Pessoa 21	50	4
Pessoa 22	50	5
Pessoa 23	40	7
Pessoa 24	45	8
Pessoa 25	33	9
Pessoa 26	37	10
Pessoa 27	35	1
Pessoa 28	39	6
Pessoa 29	34	4
Pessoa 30	24	5
Pessoa 31	24	6
Pessoa 32	25	6
Pessoa 33	26	7
Pessoa 34	22	6
Pessoa 35	35	7
Pessoa 36	29	7
Pessoa 37	27	4
Pessoa 38	27	7
Pessoa 39	28	7
Pessoa 40	28	9
Pessoa 41	30	4
Pessoa 42	30	6
Pessoa 43	32	6
Pessoa 44	28	6
Pessoa 45	27	6
Pessoa 46	36	10
Pessoa 47	38	8
Pessoa 48	36	6
Pessoa 49	32	8
Pessoa 50	32	10
Pessoa 51	30	8
Pessoa 52	34	6
Pessoa 53	34	8
Pessoa 54	40	10
Pessoa 55	42	8
Pessoa 56	42	10
Pessoa 57	38	12
Pessoa 58	36	12
Pessoa 59	32	12
Pessoa 60	40	15

Figura B.1: Tabela de Dados dos Motoristas