



UNIVERSIDADE DO ESTADO DO AMAZONAS
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
MESTRADO PROFISSIONAL EM MATEMÁTICA EM
REDE NACIONAL



RENATO FERNANDES DOS SANTOS

**Fundamentos Matemáticos do Ensino Básico na Construção de
Redes Neurais Artificiais**

MANAUS, AGOSTO
2024

RENATO FERNANDES DOS SANTOS

**Fundamentos Matemáticos do Ensino Básico na Construção de
Redes Neurais Artificiais**

Dissertação de Mestrado apresentada à
Universidade do Estado do Amazonas como
parte dos requisitos necessários para
obtenção do título de mestre no programa de
Mestrado Profissional em Matemática em
Rede Nacional - PROFMAT/UEA.

Orientador(a): Prof. Dr. João Batista
Ponciano

Co-orientador(a): Prof^ª Dra. Kelly Alves
Marães de Almeida

MANAUS, AGOSTO

2024

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).
Sistema Integrado de Bibliotecas da Universidade do Estado do Amazonas.

S237f

Santos, Renato Fernandes dos

Fundamentos matemáticos do ensino básico na construção de redes neurais artificiais / Renato Fernandes dos Santos . Manaus : [s.n.], 2024
94 f.: color.; 21,0 cm.

Dissertação - Mestrado Profissional em Matemática em Rede Nacional- Universidade do Estado do Amazonas, Manaus, 2024.

Inclui Apêndice.

Orientador: João Batista Ponciano.

Coorientador: Kelly Alves Marães de Almeida.

1. Educação Matemática. 2. Tecnologia. 3. Redes Neurais Artificiais. 4. Ensino Básico. I. João Batista Ponciano (Orient.) II . Kelly Alves Marães de Almeida (Coorient.) III. Universidade do Estado do Amazonas. IV. Título

CDU(1997)51(043.3)

**ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO PROFISSIONAL EM
MATEMÁTICA EM REDE NACIONAL - PROFMAT DA UNIVERSIDADE DO
ESTADO DO AMAZONAS**

Ata de defesa de Dissertação de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT da Universidade do Estado do Amazonas, no município de Manaus-AM, do aluno **RENATO FERNANDES DOS SANTOS**, matrícula nº **2291940011**.

Em 11 de setembro de 2024, às 14h, na Sala de Videoconferência da Escola Normal Superior no Município de Manaus-AM, na presença da Banca Examinadora composta pelos professores: Prof. Dr. João Batista Ponciano, Prof. Me. Alessandro Monteiro de Menezes e Prof. Dr. Valtemir Martins Cabral, o aluno **RENATO FERNANDES DOS SANTOS** apresentou a defesa de Dissertação de Mestrado Profissional em Matemática em Rede Nacional da UEA, intitulada: "**Fundamentos Matemáticos do Ensino Básico na Construção de Redes Neurais Artificiais**".

A Banca Examinadora deliberou e decidiu pela **APROVAÇÃO** do referido trabalho, divulgando o resultado ao aluno e aos demais presentes.

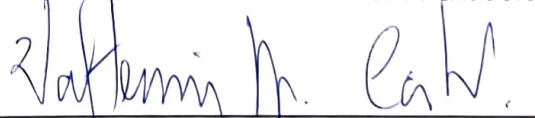
Manaus, 11 de setembro de 2024.



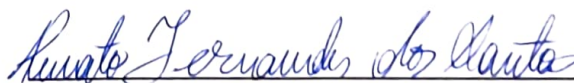
Orientador



Membro Interno da Banca Avaliadora



Membro Externo da Banca Avaliadora



Mestrando

Dedico esta dissertação de mestrado à minha querida tia Maria Tereza, que com amor e coragem assumiu o papel de mãe em minha vida, mesmo diante das maiores adversidades que a vida nos apresentou. Sua presença e seu exemplo de humildade e força foram e sempre serão fontes de inspiração para mim. Mesmo não estando mais entre nós para testemunhar essa conquista, sei que sua luz e seu amor continuam a me guiar em cada passo que dou. Essa vitória também é sua, para sempre em meu coração, tia Tereza.

Agradecimentos

Agradeço primeiramente a Deus;

Aos meus familiares, principalmente minhas tias Maria Tereza e Ana Lúcia que tiveram papel preponderante em minha criação e busca por evolução;

Aos meus colegas de curso que foram um esteio nos momentos difíceis e grandes parceiros no compartilhamento das vitórias, em especial Anderson, Cleo, Laércio, Jacson, Joilson, Lúcio, Luciana, Ruy, Kelvin, Mayara e Yeda.

Agradeço ao corpo docente do PROFMAT da Universidade do Estado do Amazonas (UEA), em particular ao meu orientador Prof. Dr. João Batista Ponciano e minha coorientadora Profa. Dra. Kelly Alves Marães de Almeida pela paciência e inspiração em suas aulas.

Oi, Meu Filho, bom dia!
Maria Tereza Fernandes

Resumo

Este trabalho aborda a intersecção entre Educação Matemática e Tecnologia, com foco na aplicação de conceitos Matemáticos no contexto das Redes Neurais Artificiais (RNAs). O trabalho é fundamentado na necessidade de integrar a Matemática ao Ensino Básico, preparando os alunos para um mundo cada vez mais digital e tecnológico. Destaca-se a importância da Educação Matemática na formação de cidadãos críticos e informados e a necessidade de programas de formação continuada para professores de Matemática do Ensino Básico. A alfabetização digital entre os professores é crucial para a adoção de metodologias pedagógicas modernas e a falta de recursos tecnológicos e a exclusão digital, como os vistos especialmente no contexto da pandemia, acentuam as desigualdades e o acesso à educação moderna e de qualidade. A dissertação é baseada em uma revisão bibliográfica abrangente, que inclui livros, artigos e teses sobre Educação Matemática, formação de professores, tecnologias digitais e fundamentos de Matemática Básica, mapeando o estado da arte sobre o Ensino de Matemática e a aplicação de conceitos matemáticos em contextos práticos, como as RNAs.

Palavras-chave: Educação Matemática, Tecnologia, Redes Neurais Artificiais, Ensino Básico.

Abstract

This paper addresses the intersection between mathematics education and technology, focusing on the application of mathematical concepts in the context of artificial neural networks (ANNs). It is based on the need to integrate mathematics into basic education, preparing students for an increasingly digital and technological world. The importance of mathematics education in the formation of critical and informed citizens and the need for continuing education programs for elementary school mathematics teachers are highlighted. Digital literacy among teachers is crucial for the adoption of modern pedagogical methodologies, and the lack of technological resources and digital exclusion, as seen especially in the context of the pandemic, accentuate inequalities and access to modern and quality education. The dissertation is based on a comprehensive literature review, which includes books, articles and theses on mathematics education, teacher training, digital technologies and foundations of basic mathematics, mapping the state of the art on mathematics teaching and the application of mathematical concepts in practical contexts, such as ANNs.

Keywords: Mathematics Education, Technology, Artificial Neural Networks, Basic Education.

Lista de ilustrações

Figura 1 – Representação do vetor $\vec{v} = (6, 8)$	41
Figura 2 – Representação gráfica da adição $\vec{u} + \vec{v}$	42
Figura 3 – Comutatividade da adição de vetores	43
Figura 4 – Associatividade da adição de vetores	43
Figura 5 – Pontos médios dos lados de um quadrilátero qualquer	44
Figura 6 – Diagrama de Venn exemplo 3.22	46
Figura 7 – Diagrama de Venn exemplo 3.28	49
Figura 8 – Diagrama de Venn exemplo 3.29	50
Figura 9 – Diagrama de Venn exemplo 3.30	50
Figura 10 – Estrutura de um neurônio biológico	59
Figura 11 – Estrutura de um Perceptron de camada única com 3 entradas	60
Figura 12 – Exemplo RNA com 5 neurônios	60
Figura 13 – Primeiro nó da primeira camada oculta	61
Figura 14 – Segundo nó da primeira camada oculta	62
Figura 15 – Terceiro nó da primeira camada oculta	62
Figura 16 – Dígitos manuscritos	69
Figura 17 – Exemplo 1: estrutura do Perceptron de camada única	70
Figura 18 – Exemplo 2: estrutura do perceptron	73
Figura 19 – Exemplo 3: estrutura da RNA	77

Lista de gráficos

Gráfico 1 – Representação cartesiana de uma função real	48
Gráfico 2 – Função $f(x) = 2x + 1$	52
Gráfico 3 – Função $f(x) = 2^x$	53
Gráfico 4 – Função $f(x) = \left(\frac{1}{2}\right)^x$	53
Gráfico 5 – Gráfico da função seno	54
Gráfico 6 – Gráfico da função cosseno	55
Gráfico 7 – Gráfico da função degrau	63
Gráfico 8 – Gráfico da Função ReLU	64
Gráfico 9 – Gráfico da função Sigmoides	65
Gráfico 10 – Gráfico da Função Tangente Hiperbólica	66
Gráfico 11 – Gráfico da função de custo MSE	67
Gráfico 12 – Gradiente descendente	68
Gráfico 13 – Exemplo 1: Época 0 da aprendizagem da rede	74
Gráfico 14 – Exemplo 1: Época 10 da aprendizagem da rede	74
Gráfico 15 – Exemplo 1: Época 100 da aprendizagem da rede	75
Gráfico 16 – Exemplo 1: Época 990 da aprendizagem da rede	76
Gráfico 17 – Exemplo 2: Época 0 da aprendizagem da rede	78
Gráfico 18 – Exemplo 2: Época 1000 da aprendizagem da rede	78
Gráfico 19 – Exemplo 2: Época 10000 da aprendizagem da rede	79
Gráfico 20 – Exemplo 2: Época 20000 da aprendizagem da rede	79
Gráfico 21 – Exemplo 2: Época 50000 da aprendizagem da rede	80

Sumário

1	INTRODUÇÃO	13
2	PROCEDIMENTOS METODOLÓGICOS	15
3	FUNDAMENTAÇÃO TEÓRICA	16
3.1	Ensino de Matemática e Tecnologia	17
3.2	Matemática para a construção de RNAs	28
3.2.1	Matrizes	28
3.2.1.1	Noção de matriz	29
3.2.2	Matrizes especiais	30
3.2.2.1	Operações com matrizes	31
3.2.2.1.1	<i>Igualdade</i>	31
3.2.2.1.2	<i>Adição</i>	32
3.2.2.1.3	<i>Matriz oposta</i>	33
3.2.2.1.4	<i>Produto de um número por uma matriz</i>	33
3.2.2.1.5	<i>Produto de matrizes</i>	33
3.2.2.2	Matriz transposta	36
3.2.2.3	Matriz simétrica e matriz antissimétrica	37
3.2.2.4	Matrizes inversíveis	39
3.2.3	Vetores	40
3.2.3.1	Equipolência de segmentos orientados	40
3.2.3.2	Vetores no plano	41
3.2.3.3	Operações com vetores	42
3.2.3.3.1	<i>Adição</i>	42
3.2.3.3.2	<i>Multiplicação de vetores por escalares</i>	44
3.2.4	Funções reais de uma variável	45
3.2.4.1	Produto cartesiano	45
3.2.4.2	Relação binária	45
3.2.4.3	Conceito de função	46
3.2.4.3.1	<i>Notação das funções</i>	47
3.2.4.3.2	<i>Imagem de um elemento</i>	47
3.2.4.3.3	<i>Domínio e imagem de uma função</i>	47
3.2.4.3.4	<i>Zero de uma função</i>	48
3.2.4.3.5	<i>Composição de funções</i>	48
3.2.4.3.6	<i>Função injetora e sobrejetora</i>	49
3.2.4.3.7	<i>Inversão de funções</i>	50
3.2.4.4	Função afim	51

3.2.4.5	Função exponencial	52
3.2.4.6	Funções trigonométricas	53
3.2.4.6.1	<i>Função seno</i>	54
3.2.4.6.2	<i>Função cosseno</i>	54
4	REDE NEURAL ARTIFICIAL	56
4.1	Histórias das RNAs	56
4.2	Neurônio artificial	57
4.2.1	Funcionamento do neurônio biológico	58
4.2.2	Perceptron	59
4.2.3	Funcionamento de um perceptron	60
4.2.3.1	Pesos e bias	61
4.2.3.2	Função de ativação	63
4.2.3.3	Função de custo	66
5	USO DE RNAs NO ENSINO BÁSICO	69
5.1	RNA para tomada de decisão simples	69
5.2	RNA para representação de uma reta no plano.	72
5.3	RNA para representação de uma curva qualquer no plano	76
6	CONSIDERAÇÕES FINAIS	81
7	REFERÊNCIAS BIBLIOGRÁFICAS	83
	APÊNDICE	86

1 INTRODUÇÃO

A popularização das chamadas Inteligências Artificiais (IA's) traz uma série de questionamentos a respeito de seu funcionamento, estrutura e uso por parte do público consumidor em geral, mais especificamente o público no âmbito escolar, entre alunos e professores do ensino básico.

Sabemos que existe uma convergência entre a Matemática e o desenvolvimento das IA's, e isso é um fator significativo e representa um marco no desenvolvimento científico contemporâneo.

Por mais que as IA's estejam em destaque nos últimos anos com o surgimento, por exemplo, das Inteligências Artificiais Generativas como Chatgpt, Midjourney, entre outros, a idéia de uma inteligência computacional não é nova, e vem sendo estruturada há décadas através de modelos matemáticos e noções de Estatística e Probabilidade.

A base Matemática das Redes Neurais Artificiais (RNA) é bastante complexa e avançada, envolvendo conceitos de Álgebra Linear, Números e Funções, Cálculo Diferencial e Teoria de Probabilidade. É importante ter um bom conhecimento desses conceitos para projetar, implementar e treinar redes neurais eficazes, embora muitos dos conceitos mais avançados não sejam contemplados na base nacional comum curricular (BNCC) do Ensino Médio, abordados apenas no ensino superior em cursos de graduação em Matemática, Ciência da Computação e áreas afins.

Em teoria seria possível fornecer uma base sólida a respeito de muitos desses tópicos no Ensino Médio, já que as bases matemáticas fundamentais na construção de RNA, como estudo de Matrizes, estudo das Funções, Estatística e Probabilidade estão inclusos na BNCC do Ensino Médio, contudo, é importante salientar que apesar da inclusão desses tópicos que constituem as bases matemáticas fundamentais para a construção de RNA, dados de avaliações em larga escala, como o PISA (Programme for International Student Assessment), mostra em seu relatório de 2022 que apenas 27% dos estudantes alcançaram nível 2 de proficiência em Matemática, “onde esses estudantes podem interpretar e reconhecer, sem instruções diretas, como uma situação simples pode ser representada matematicamente (por exemplo, comparar distâncias total de duas rotas ou converter preços em moeda diferente)” (notas sobre o PISA 2022, 2023, p. 10), enquanto a média entre outros países é de 69%, deixando o Brasil em uma posição muito baixa nos ranks internacionais.

O Novo Ensino Médio, aprovado em 2017 e obrigatoriamente implementado nas escolas públicas a partir de 2022, tem o pressuposto de oferecer um ensino mais integrado e flexível, dando a oportunidade para os estudantes escolherem entre itinerários formativos (unidades curriculares eletivas como Projeto de Vida, Educação Financeira, Programação, etc) de acordo com suas preferências e habilidades, contudo, a transição para esse novo modelo de ensino encara desafios, como a falta de estrutura das Escolas e de professores

capacitados para atenderem às novas diretrizes.

O modelo de currículo escolar do Novo Ensino Médio foi inspirado em sistemas educacionais de países como a Inglaterra, que já adota a flexibilização curricular, inserindo disciplinas opcionais a livre escolha do estudante, como Programação, Robótica, Gastronomia, entre outros. Essas disciplinas opcionais adotadas no sistema de ensino inglês são os chamados itinerários formativos do Novo Ensino Médio no Brasil, e propõem oportunizar um aprofundamento dos conhecimentos por parte dos estudantes em áreas específicas, como Linguagens, Matemática, Ciências da Natureza, Ciências Humanas e Formação Técnica e Profissional.

Assim, a escolha de um itinerário formativo de Ciências Exatas e Engenharias pode incluir disciplinas e tópicos de Matemática mais avançada, dando condições dos estudantes terem uma base mais sólida para entenderem como as RNA são construídas, além disso, podem ser incluídas abordagens práticas de construção e implementação de RNA em disciplinas de programação, utilizando linguagens como Python. No entanto, há barreiras e desafios a serem vencidos até que um currículo tão abrangente seja implementado nas Escolas, pois depende da disponibilidade de recursos tecnológicos como laboratórios, materiais didáticos adequados, além de professores capacitados.

É abordado neste trabalho quais os fundamentos matemáticos utilizados na construção de RNA e a profunda interligação entre os princípios matemáticos e a arquitetura das mesmas, delineando a influência e a importância dos fundamentos matemáticos (dentro do contexto do Ensino Básico) na eficácia dessas estruturas computacionais e como esses fundamentos são empregados no design, treinamento e operação de redes neurais. Portanto, busca-se abordar como os elementos matemáticos, extraídos de diferentes áreas da Matemática, são traduzidos e incorporados nas complexas estruturas das redes neurais, em simultâneo, questionando-se como essa integração contribui para o avanço do conhecimento matemático e para melhoria do ensino e aprendizagem da Matemática no Ensino Básico.

2 PROCEDIMENTOS METODOLÓGICOS

Para este trabalho foi realizado uma revisão bibliográfica, estruturada para garantir uma análise abrangente e fundamentada sobre a integração de fundamentos matemáticos no Ensino Básico, especialmente no contexto da construção de RNAs.

A revisão bibliográfica é um procedimento que envolve a coleta, análise e síntese de informações provenientes de diversas fontes acadêmicas e científicas. Essa abordagem permite mapear o estado da arte sobre o Ensino de Matemática, as tecnologias educacionais e a aplicação de conceitos matemáticos em contextos práticos, como as redes neurais. Segundo GIL (2008)

A pesquisa bibliográfica é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos. Embora em quase todos os estudos seja exigido algum tipo de trabalho desta natureza, há pesquisas desenvolvidas exclusivamente a partir de fontes bibliográficas. (GIL, 2008, p. 50)

O trabalho iniciou-se com uma busca sistemática por literatura relevante, utilizando bases de dados acadêmicas, artigos, livros e teses que abordam temas como a Educação Matemática, a formação de professores e a utilização de tecnologias digitais no ensino, tópicos de Matemática Básica e estudos sobre inteligência artificial, permitindo uma análise crítica e fundamentada sobre a intersecção entre Matemática, Tecnologia e Educação.

Essa abordagem não apenas enriquece o trabalho, mas também oferece uma base sólida para futuras pesquisas e práticas pedagógicas no campo Educação Matemática e uso de recursos tecnológicos, como RNAs.

3 FUNDAMENTAÇÃO TEÓRICA

A interseção entre Matemática e Tecnologia tem se tornado cada vez mais relevante no contexto educacional contemporâneo. A construção de Redes Neurais Artificiais (RNAs) não apenas representa um avanço significativo na área da Inteligência Artificial, mas também oferece uma oportunidade única para a aplicação de conceitos matemáticos fundamentais no Ensino Básico. Para Sá e Machado,

O uso das tecnologias na sala de aula vem se tornando uma ferramenta de grande importância, pois consegue auxiliar tanto o professor quanto o aluno na explicação e na compreensão dos conteúdos. Com a tecnologia na aula os alunos sentem-se mais motivados a aprender e a partir disso o docente consegue ensinar de forma mais dinâmica e criativa (Sá; Machado, 2017, p. 1).

A incorporação de tecnologias digitais no Ensino de Matemática tem o potencial de transformar a prática pedagógica. Ferramentas como softwares de simulação, plataformas de aprendizado online e recursos interativos podem facilitar a compreensão de conceitos matemáticos complexos. A utilização de RNAs como ferramenta de aprendizagem matemática no Ensino Básico é um exemplo de como a tecnologia pode ser utilizada para engajar os alunos e promover uma aprendizagem mais dinâmica.

Apesar das oportunidades que a tecnologia oferece, a implementação de novas metodologias de ensino enfrenta desafios significativos. A resistência à mudança por parte de educadores, a falta de formação adequada e a escassez de recursos tecnológicos são barreiras que precisam ser superadas. É fundamental que os professores sejam capacitados para integrar a tecnologia em suas práticas pedagógicas de forma eficaz. Para Coan, Viseu e Moretti,

A falta de recursos tecnológicos, como é o caso de laboratórios de informática que estejam em plenas condições de uso, ainda é um grande obstáculo para professores e alunos, pois quando há o espaço, faltam máquinas ou elas não funcionam. Esta é sem dúvida um dos fatores que desmotiva o professor a apostar na integração das TIC no processo de ensino e planejar suas atividades pedagógicas numa dinâmica que atenda o uso de tecnologias, no caso computador com acesso a Internet. (Coan, Viseu, Moretti, 2013, p. 240)

A pesquisa em RNAs e sua aplicação no Ensino de Matemática representam uma oportunidade valiosa para a Educação. À medida que a tecnologia avança, novas ferramentas e recursos estarão disponíveis para apoiar o ensino e a aprendizagem. A formação contínua de professores e a colaboração entre educadores e especialistas em tecnologia serão essenciais para aproveitar ao máximo essas oportunidades. Interseção entre Matemática e Tecnologia não apenas enriquece o aprendizado, mas também prepara os alunos para os desafios do século XXI. Ao promover uma educação mais dinâmica e aplicada, é possível capacitar os alunos com habilidades essenciais para o futuro.

3.1 Ensino de Matemática e Tecnologia

O isolamento social determinado pela Organização Mundial de Saúde – OMS, em decorrência da Pandemia de COVID-19 (infecção viral causada pelo SARS-CoV-2), promoveu sérias transformações econômicas imediatas, com o fechamento obrigatório de setores variados, alterando experiências de contato e interação humanas pela ausência de compartilhamento direto destas. Na educação, especificamente, experimentou-se um momento que favoreceu a desconstrução do modelo de ensino-aprendizagem até então estabelecido, bem como a forma como ele é visto socialmente.

No que se refere ao ensino de Matemática, especificamente, Santos e Santos (2020) descrevem que, “embora haja uso da TDIC pelos professores, ela ocorreu de forma muito limitada, abordando atividades mais simples”, como atualização da agenda de aula, preparação de slides para apresentação de conteúdo, redação de relatórios diversos, entre outras ações consideradas mínimas. Assim, a seu ver, as TDIC não têm o seu uso muito difundido para o ensino em si, servindo mais como apoio.

Durante a pandemia, os professores tiveram que adaptar suas técnicas de ensino típicas para aulas que agora, às vezes, acontecem online. Os professores estão eliminando os padrões de aprendizagem, abandonando os testes de obtenção de respostas e voltando-se para jogos e aplicativos de Matemática para complementar o ensino. Eles também estão pedindo apoio aos pais e encontrando maneiras de envolver os alunos nas telas, incluindo aulas sobre como Matemática e justiça social se cruzam (CAFARDO, 2020).

O Ministério da Educação, por meio da Secretaria de Educação a Distância (SEED), agia como um agente de inovação tecnológica nos processos de ensino e aprendizagem, fomentando a incorporação das tecnologias de informação e comunicação, e das técnicas de Educação a Distância aos métodos didático-pedagógicos. Além disso, promovia a pesquisa e o desenvolvimento, voltados para a introdução de novos conceitos e práticas nas Escolas Públicas brasileiras (PORTAL MINISTÉRIO DA EDUCAÇÃO, 2020). O Decreto nº 5.622 de 2005, em seu Art. 1º, caracteriza a educação a distância como:

Modalidade educacional em que a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorre com a utilização de meios e tecnologias de comunicação e informação, com estudantes e professores desenvolvendo atividades educativas em lugares ou tempos diversos.

Por meio desse decreto, comprova-se que a Educação a Distância busca conceder uma interação direta entre as tecnologias e os procedimentos de ensino aplicados dentro dos ambientes educativos, observando as formas e os métodos mais eficazes no que se refere ao alinhamento de Tecnologia e Educação. Algo que vem sendo apresentado como uma relação muito eficiente e produtiva, tanto para as instituições de ensino como para os alunos.

Nesse sentido, a utilização das novas ferramentas tecnológicas na Educação permite não apenas aproximar as práticas sociais dos alunos, suas experiências e conhecimentos

anteriores, mas também a oportunidade de desenvolver um modo de aprendizagem baseado na cooperação entre pares. O aprendizado colaborativo pode ocorrer, por exemplo, ao implementar-se e manter-se uma comunidade virtual.

É fato que se vive em uma nova realidade onde a difusão do conhecimento acontece em alta velocidade. “Esta é uma exigência social e organizacional que o educador deve dirigir aos seus alunos para querer conhecer e tenha a oportunidade de utilizar os recursos informáticos disponíveis e utilizá-los como suporte para o desenvolvimento educacional”. Evidencia-se ainda que a LDB- Lei nº 9.394/96 regulamenta a eventualidade de substituição de aulas presenciais pelo ensino à distância nos casos em que se fizer necessário. É o que se destaca mediante o art. 32, § 4º,

Fica autorizada, em caráter excepcional, a substituição das disciplinas presenciais, em andamento, por aulas que utilizem meios e tecnologias de informação e comunicação, por instituição de educação superior integrante do sistema federal de ensino, de que trata o art. 2º do Decreto nº 9.235, de 15 de dezembro de 2017. (BRASIL, 2020).

A portaria em referência foi alterada e decreta a respeito da substituição antes retratada no que diz respeito somente às instituições de ensino superior, incumbindo aos Estados e Municípios a negativa ou aprovação de práticas de ensino à distância em seus territórios, tendo como base as necessidades e situações locais.

As ações ora em destaque, aplicáveis a todos os níveis de ensino, possuíam caráter excepcional e poderiam ser utilizadas enquanto durasse a emergência de saúde pública por conta da corrente pandemia. No entanto, devido às limitações de infraestrutura de algumas escolas e até mesmo a falta de capacitação por parte de certos docentes, foi necessário que se fizesse um planejamento com vistas ao alcance dos resultados para que houvessem diretrizes a serem seguidas depois.

As estratégias de ensino à distância deveriam, portanto, desempenhar uma função crucial na diminuição das consequências negativas da interrupção presencial das aulas. No entanto, a evidência sugere que surgiram lacunas de vários tipos. Como resultado, as regulamentações sobre sua equivalência para fins de conclusão do ano letivo precisaram ser objeto da atenção dos reguladores educacionais no Brasil.

Segundo Ribeiro (2017), a exclusão digital é o primeiro obstáculo para a introdução do ensino à distância para alunos e educadores. Em um momento em que o estudo no domicílio deve ser sustentado, os estudantes brasileiros convivem com problemas de saneamento e de acesso à água e alimentação, além da ausência de um ambiente de qualidade para o estudo e falta de apoio dos pais e responsáveis.

Tal afirmação é vivenciada por muitos discentes, já que o isolamento social igualmente incrementa maior visibilidade da desigualdade presente no país, não apenas no tocante ao acesso à internet ou aos recursos que lhe permitem esta conexão, bem como a uma desigualdade social, cultural e educacional, sobretudo nos interiores municipais, onde a carência de recursos financeiros e pessoais é ainda mais significativa. Desse modo,

Quinzani (2020) relata que,

A pandemia deixa ainda mais visível uma ideia já há muito tempo conhecida, mas agora não mais como preconceito, mas sim como um conceito bastante evidente. O crescimento da desigualdade é a união de problemas econômicos, da insegurança no trabalho, problema aumentado em razão da pandemia, e ainda as disparidades de renda e a falta de oportunidades no país. (QUINZANI, 2020, p. 44).

Além disso, o contexto atual proporcionou um crescimento da ansiedade e receio com relação às questões futuras, por exemplo, sobre como e quando será a reabertura das escolas, com a possibilidade de retorno dos alunos, isso em razão, da sensação de incerteza e instabilidades a longo prazo criarem, lentamente, um trauma, um sentimento de medo, aumentando os níveis de estresse, ocasionando riscos à saúde mental. Conforme posicionamento de Ornell et al (2020):

Além de um medo concreto da morte, a pandemia da COVID-19 tem implicações para outras esferas: organização familiar, fechamento de escolas, empresas e locais públicos, mudanças nas rotinas de trabalho, isolamento, levando a sentimentos de desamparo e abandono. (ORNELL *et al.*, 2020, p. 3)

No contexto de pandemia, o ensino à distância foi a única alternativa de auxílio à situação imposta. Nesse sentido, foram necessárias intervenções alternativas para que os educadores possam usar as ferramentas digitais para resolver os dilemas potenciais mencionados acima. Constata-se, portanto, que a prática recorrente ainda se constitui como um contexto empírico que se evidenciou de forma incapacitante para dar uma explicação satisfatória para o conhecimento matemático. Incapaz, porque o empirismo parte do pressuposto de que a verdade Matemática – como, aliás qualquer verdade – está inscrita nos objetos sendo retirada destes por uma abstração simples ou empírica (CAFARDO, 2020).

Ribeiro (2017) explica que o ensino híbrido combina elementos do ensino presencial e à distância. Os alunos participam de aulas presenciais em parte do tempo e realizam atividades de aprendizagem online em outras ocasiões. Esse modelo permite uma maior flexibilidade e personalização do aprendizado, em simultâneo, em que mantém a interação face a face entre alunos e professores durante as aulas presenciais. Durante a pandemia da COVID-19, tanto o ensino a distância quanto o híbrido tornaram-se alternativas essenciais para garantir a continuidade do ensino, uma vez que as restrições de distanciamento social impediram o funcionamento normal das Escolas e Universidades. Muitas instituições de ensino tiveram que se adaptar rapidamente, desenvolvendo e implementando estratégias de ensino remoto para atender às necessidades dos alunos.

Embora as microgerações difiram, há fortes características comuns. Elas realizam várias tarefas e desejam informações rapidamente usando recursos visuais, sons e cores de várias fontes multimídia de uma maneira nova ou útil. Nossos alunos atuais cresceram em um mundo digital de computadores, internet e mídias sociais, como Instagram, Facebook, Twitter, Google e outras. Eles aprendem interagindo com outras pessoas online e gostam de ser ativos e colaborar usando as mais recentes tecnologias e oportunidades de visualização (MOREIRA; HENRIQUE; BARROS, 2020).

Os alunos de hoje preferem buscar suas próprias informações em vez de serem apresentados a elas. Preferem o acesso sob demanda ao conhecimento, disseminado pela internet, e absorvem o conhecimento rapidamente em diferentes canais. Usando redes para compartilhar e criar conhecimentos, eles estão em contato frequente com seus amigos (ARRUDA, 2020).

A mudança nos nossos alunos implica que a forma de ensino deve ser totalmente adaptada para enfrentar estes desafios e responder às novas solicitações. Nas últimas décadas, uma Web mais social e conectada se desenvolveu, apoiando a idéia de aprendizagem em rede. Ambientes de aprendizagem em rede aberta são ambientes digitais que capacitam os alunos a conduzirem redes sociais, organizar conteúdos sociais e gerenciar atos sociais conectando pessoas, recursos e ferramentas, integrando ferramentas de internet para criar ambientes transparentes (AVELINO; MENDES, 2020).

Moreira, Henrique e Barros (2020) descrevem que os primeiros dados de testes parecem confirmar que a pandemia atingiu os alunos mais em Matemática do que em Leitura. A Matemática pode ser mais sensível a interrupções na Escola relacionadas à pandemia por alguns motivos, dizem os especialistas:

- Ao contrário da leitura, a Matemática quase sempre é aprendida formalmente na escola. Os pais costumam estar menos preparados para ajudar os filhos com a Matemática, numa época em que o apoio dos pais pode ser ainda mais crucial para o progresso do aluno;
- Estresse e trauma mais amplos relacionados à pandemia podem piorar a ansiedade matemática existente em alguns alunos, e a ansiedade matemática pode exacerbar outro estresse dos alunos durante as aulas;
- Pode ser mais desafiador para os professores se envolverem em práticas eficazes de Ensino de Matemática por meio de plataformas remotas. Embora as escolas híbridas e virtuais tenham ganhado força na última década, a mudança repentina e completa para o aprendizado virtual na maioria dos distritos na primavera passada e as mudanças contínuas no formato foram altamente prejudiciais ao ensino (BARRETO; ROCHA, 2020).

Há um movimento para identificar padrões de poder ou os elementos mais críticos de aprendizagem em cada série. Mas alguns especialistas em Matemática alertam sobre essas abordagens porque os currículos de Matemática mais recentes tendem a ser espirais, com os conceitos introduzidos nas primeiras séries reaparecendo. Isso significa que nas séries posteriores o currículo presumirá que os alunos aprenderam coisas que talvez não tenham aprendido (AVELINO; MENDES, 2020).

Um dos temas mais comuns na pesquisa de perda de aprendizagem é a importância do apoio dos pais na aprendizagem do aluno. Alunos com pais altamente envolvidos que relatam participar de atividades educacionais durante o verão tendem a perder menos terreno. E os primeiros estudos de respostas à pandemia descobriram que as escolas estão cada vez mais dependentes das famílias para facilitar a instrução durante a crise atual (CAFARDO, 2020).

Estudos sugerem que as famílias podem ter menos probabilidade de se envolver em atividades Matemáticas do que em Leitura com seus filhos devido à ansiedade matemática e aos novos métodos de ensino para ensinar os Padrões Estaduais do Núcleo Comum que

diferem de como os próprios pais aprenderam Matemática (SANTOS; SANTOS, 2020).

Os educadores podem sugerir jogos e aplicativos relacionados à Matemática que incentivem as famílias a integrar conversas matemáticas em suas vidas domésticas. Alguns estudos descobriram que filhos de pais ansiosos por Matemática, que regularmente usavam um aplicativo de Matemática voltado para a família, mostraram mais progresso em Matemática do que alunos de pais semelhantes que não usavam o aplicativo (ARRUDA, 2020).

Analisar a estrutura dos currículos, ser flexível, motivar os alunos a ver a Matemática como útil e estimulante e planejar para continuar enfrentando os desafios da pandemia são algumas ações que ajudarão a ter um Ensino de Matemática mais equitativo (AVELINO; MENDES, 2020).

O uso de recursos tecnológicos no Ensino de Matemática tem sido cada vez mais explorado como uma estratégia para enriquecer e aprimorar a experiência de aprendizagem dos alunos. Esses recursos incluem uma variedade de ferramentas, como softwares educacionais, aplicativos móveis, simulações interativas e ambientes virtuais de aprendizagem. Um dos principais benefícios dessas tecnologias é a capacidade de tornar conceitos matemáticos abstratos mais concretos e acessíveis aos alunos. Por meio de representações visuais e interativas, os estudantes podem visualizar gráficos, modelos tridimensionais e processos matemáticos de uma maneira que antes não seria possível apenas com métodos tradicionais de ensino. Além disso, os recursos tecnológicos oferecem oportunidades para a personalização do aprendizado, permitindo que os alunos progridam em seu próprio ritmo e recebam feedback imediato sobre seu desempenho. Isso pode aumentar a motivação e o engajamento dos alunos, pois eles se sentem mais envolvidos no processo de aprendizagem (ZANELLA, MACHADO, MARCON, 2019).

No entanto, o uso eficaz dessas tecnologias requer não apenas acesso aos recursos adequados, mas também uma compreensão sólida de como integrá-los de forma significativa ao currículo e às práticas pedagógicas. Os professores precisam receber formação adequada para utilizar essas ferramentas de maneira eficaz, além de terem tempo e suporte para experimentar e desenvolver novas abordagens de ensino. Outro desafio é garantir que o acesso às tecnologias seja equitativo, especialmente em contextos em que há disparidades socioeconômicas. Os educadores precisam estar cientes dessas disparidades e trabalhar para encontrar maneiras de garantir que todos os alunos tenham acesso igualitário aos recursos necessários para aprender (BARRETO; ROCHA, 2020).

O papel da mídia social está se tornando cada vez mais importante neste contexto, movendo o processo educacional de uma abordagem de 'empurrar' para uma abordagem de 'puxar' centrada no aluno, onde os alunos se tornam parte integrante de muitas facetas do processo. Além disso, a introdução de recursos e ferramentas digitais questiona a eficácia das práticas atuais dos professores e das salas de aula tradicionais (SANTOS; SANTOS; SANTOS, 2020).

Os aspectos sociais da internet tornam-se cada vez mais relevantes e noções como 'humanos com a mídia' enfatizam que, se a mídia for alterada, todo o processo de aquisição de conhecimento pode mudar. Além disso, à medida que os humanos desenvolvem e constroem novas mídias, essas mídias parecem transformar e 'construir' um novo ser humano (CAFARDO, 2020).

Nesse processo de desenvolvimento, a sala de aula, como a conhecemos, pode mudar inteiramente de uma área física com limites definidos para um ambiente virtual incluindo vários componentes que provavelmente serão determinados pelo aluno e não apenas pelo professor. A tecnologia móvel, os ambientes de aprendizagem pessoal, os objetos de aprendizagem digitais e outros artefatos estão 'esticando' a sala de aula, transformando a sala de aula, a ponto de dificilmente ser reconhecida como tal (ARRUDA, 2020). Atualmente parece claro que a tecnologia digital está 'desconstruindo' a noção de sala de aula. Salas de aula invertidas mudam a noção do que está dentro e fora da sala de aula e mudam os papéis dos alunos e professores. Há uma profusão de recursos online (por exemplo, widgets, vídeos), projetados com relação a conteúdo matemático específico, que transforma a apresentação do conteúdo e permite que os alunos tenham acesso à resolução de tarefas matemáticas e compartilhem suas explorações matemáticas. Os alunos também devem avaliar a qualidade do conhecimento disseminado pela internet, eles precisam ser capazes selecionar recursos válidos (AVELINO; MENDES, 2020).

Por outro lado, o desenvolvimento de tecnologias digitais móveis, como fóruns, wikis, e redes sociais está permitindo a nova geração diferentes tipos de oportunidades de aprendizagem, apoiadas em novos espaços de interação social. Essas tecnologias interativas permitem que os alunos colaborem com seus colegas na utilização de multimídia e da internet, possibilitando novas formas sociais de construção do conhecimento. Finalmente, diferentes combinações estão sendo forjadas entre a educação presencial e a educação online à distância (BARRETO; ROCHA, 2020).

Barreto; Rocha (2020) destacou que, no campo educacional, as ferramentas facilitam o desempenho de uma tarefa, estendendo o poder das capacidades humanas e ampliando o poder da mente. Embora essas ferramentas imponham restrições às atividades de aprendizagem, elas também geram novas oportunidades de aprendizagem. Essas oportunidades de aprendizagem estão ligadas a novas maneiras de buscar informações, que moldam as experiências matemáticas dos alunos e determinam como os professores em exercício podem tomar decisões curriculares (ARRUDA, 2020).

Os estudos de Barreto; Rocha (2020), mostram três abordagens diferentes para ferramentas digitais: descrever práticas inovadoras; tornar explícitos os princípios do design e desenvolver uma nova etapa de ensino. As práticas inovadoras incluem o uso das Artes e da Tecnologia Digital para a criação de instrumentos virtuais como softwares musicais ou a produção de vídeos como recurso para produzir sentido e mudar a dinâmica em sala de aula. Outra prática inovadora é como os alunos dos cursos de Matemática podem

se beneficiar ao buscar assistência na internet para obter ajuda na resolução de seus problemas matemáticos (AVELINO; MENDES, 2020).

Em particular, essas perspectivas influenciam o design de tarefas de avaliação para avaliação baseada em computador para avaliar a aprendizagem matemática dos alunos, ou influenciam a produção de ferramentas como o vídeo como recursos para apoiar a aprendizagem profissional dos professores, em que os recursos existentes apoiam aprendizagem profissional de professores (CAFARDO, 2020).

Rech e Pescador (2022) descrevem a necessidade de superar abordagens tradicionais e autoritárias de ensino, que podem ser exacerbadas pelo ambiente virtual, e buscar práticas que promovam a participação ativa dos alunos, o engajamento e a autonomia. Isso implica em criar espaços de diálogo e reflexão, onde os estudantes se sintam respeitados e valorizados como agentes ativos de seu próprio processo de aprendizagem. Os autores ainda ressaltam a importância de oferecer suporte e recursos adequados aos professores para poderem adaptar suas práticas pedagógicas ao ensino remoto de forma eficaz. Isso inclui investir em formação continuada, tecnologia acessível e estratégias de ensino que considerem as necessidades específicas de cada contexto e grupo de estudantes e sugerem que a perspectiva freireana pode servir como um guia valioso para repensar e reestruturar o ensino remoto de maneira mais humanizada, inclusiva e centrada no aluno. Ao priorizar o diálogo, a cooperação e a valorização da experiência individual, podemos criar ambientes virtuais de aprendizagem que promovam não apenas o desenvolvimento acadêmico, mas também o crescimento pessoal e social dos estudantes.

A transição para o ambiente virtual trouxe desafios como a falta de acesso à internet e tecnologia por parte dos alunos, a dificuldade de manter a atenção e o engajamento dos estudantes em aulas virtuais e a necessidade de repensar a avaliação e acompanhamento do progresso dos alunos. Os professores precisaram encontrar formas de estabelecer conexões significativas com seus alunos, promover a interação e o diálogo em ambientes virtuais e garantir um ambiente de aprendizagem acolhedor e inclusivo. De forma que é importante garantir a formação continuada dos educadores para lidar com os desafios do ensino remoto e aproveitar as oportunidades oferecidas pela tecnologia. A capacitação dos professores em ferramentas digitais e metodologias de ensino online tornou-se essencial para garantir a eficácia do processo educacional nesse novo contexto (GOMES et al.2022).

Hoje em dia, os rápidos desenvolvimentos no campo das Tecnologias de Informação e Comunicação (TIC) precisam de uma mudança na educação como em todos os campos. Essa situação exige a atualização dos programas de ensino e a reorganização dos ambientes educacionais.

A consciência da complexidade da integração das TIC no Ensino da Matemática foi, na década de 1990, com a emergência de uma nova perspectiva teórica situando os recursos tecnológicos como parceiros ativos dos processos de aprendizagem. A compreensão do papel crucial da mediação para o desenvolvimento humano, é claro, não constitui em si uma

novidade. A nova geração de estudantes, chamada de “nativos digitais”, está envolta aos mais variados tipos de tecnologia, logo, promover o uso de recursos tecnológicos durante as aulas é atividade essencial para gerar uma boa interação e promover conhecimentos de forma contextualizada. Desta forma, é importante investigar a utilização de aplicativos para dispositivos móveis, que são recursos tão acessíveis na atualidade, nas práticas pedagógicas do profissional da Educação Matemática e averiguar os saberes docentes que estão presentes neste processo (GIORGETTI; LIBERATTI, 2013; COLLING; RICHIT, 2019).

O baixo nível de alfabetização digital entre os professores de Matemática pode dificultar a adoção de abordagens pedagógicas e tecnológicas modernas no ensino e aprendizagem da Matemática. Muitos professores estão começando a adotar a perspectiva de que a aprendizagem só pode ser bem-sucedida se os alunos construírem seu conhecimento e não apenas estenderem seu conhecimento por meio da memorização. Os professores devem, portanto, ajudar os alunos a usar e reforçar os conhecimentos que possuem e a produzir novos conhecimentos. Quando o fazem, a abordagem de ensino muda de uma abordagem centrada no professor para uma abordagem mais centrada no aluno. As abordagens de ensino centradas no aluno desempenham um papel essencial neste processo. Várias dessas abordagens de ensino (por exemplo, Aprendizagem Baseada em Problemas) foram desenvolvidas e começaram a ganhar força na Educação Matemática.

Para Figueiredo (2019) as tecnologias possuem como vantagens a capacidade de serem interativas, o que traz ao professor e aos seus alunos possibilidades como a exploração dinâmica viabilizando, assim, discussões que propiciem a construção de conhecimentos. Nesse cenário, onde estão inseridas as tecnologias, com toda sua capacidade de exploração o professor tem o papel de mediar o processo de construção da aprendizagem.

Integração de tecnologia significa incorporar tecnologia e práticas baseadas em tecnologia em todos os aspectos do ensino e aprendizagem, incorporando a tecnologia apropriada para seus objetivos, lições e avaliação dos resultados da aprendizagem. A tecnologia no contexto do ensino e aprendizagem da Matemática inclui computadores com software matemático apropriado, fóruns de discussão online, entre outros. Outros exemplos incluem softwares de geometria dinâmica e análise de dados, como o GeoGebra, que permite a interatividade, calculadora gráfica visual. Há também o palm pilots, que permitem a coleta, análise e representações de dados em tempo real, e alguns softwares aplicativos, como planilhas, que permitem para Álgebra e Análise de Dados (PONTES; BARBOZA, 2020; MOTTA, 2017).

Sant’ana e Amaral (2012) explicam ainda que um problema comum da formação inicial de professores identificado na literatura é a falta de relevância da área de conteúdo e a transferência inadequada de habilidades tecnológicas para estratégias instrucionais.

Para Motta (2017) a inserção das tecnologias na Educação demonstra a necessidade de preparar os professores, na formação inicial e continuada, para lidarem com situações em que as tecnologias digitais sirvam de ferramentas para tornar a aprendizagem dos

alunos significativa.

Educar e apoiar os professores na adoção de tecnologias educacionais precisa, portanto, ser fortemente adaptado às necessidades individuais e aos recursos e infraestrutura de apoio institucional disponíveis.

A questão da apropriação dos recursos vai necessariamente ao encontro da formação de professores, aparecendo de fato como uma questão chave para a integração das tecnologias informáticas. Os professores precisam perceber o potencial da tecnologia digital em suas práticas diárias e usá-las bem. No entanto, os professores precisam de treinamento e orientações para desenvolver sua expertise ao usar esse recurso para ensinar e aprender. A falha em fazer isso pode resultar em alunos sem as habilidades de enfrentamento necessárias para sua vida futura na era da informação.

No que se refere ao tema de como utilizar as TDICs no ambiente escolar, que o processo de ensino/aprendizagem terá mais significado, e contando com a possibilidade real de que os fortes avanços das tecnologias digitais poderão trazer muitos benefícios e ser aplicados no processo de educação, promoção da inclusão social e a disseminação das tecnologias às populações mais carentes, podem caminhar em direção à construção de uma sociedade do conhecimento, ultrapassando e qualificando a sociedade da informação em que se vive hoje (PONTES; BARBOZA, 2020; MOTTA, 2017).

Valente (2013) pondera que o equilíbrio entre posturas otimistas e céticas diante das influências das tecnologias digitais na Educação é algo necessário, visto que elas não podem ser consideradas a panaceia para os problemas, ao mesmo tempo que têm muito a contribuir para mudanças necessária.

Hoje em dia, o professor enfrenta um grande desafio para implementar as tecnologias digitais em suas atividades, com o intuito de favorecer os processos de ensino e aprendizagem. Percebe-se que o fato de querer modificar a prática docente ou parte dela, não é a tarefa mais fácil que o professor enfrenta, mas se faz necessário, pois é primordial que a tecnologia esteja contemplada na prática pedagógica do professor, de modo a incentivá-lo a agir e interagir no mundo com critério e visão transformadora. Os professores necessitam ser os protagonistas dessa mudança e não apenas implantadores de sistemas e meros expectadores. Se isso não acontecer, colocar a tecnologia à frente dos alunos não vai fazer diferença alguma.

O conhecimento é universal e acessível em poucos cliques e hoje, em uma mesma sala de aula, é possível vislumbrar alunos com diferentes níveis de conhecimento. O professor deixou de ser o único detentor do conhecimento. Por isso, cabe a ele estimular a curiosidade dos alunos para buscarem informações úteis e as transformem em conhecimento, além de servir de mediador desses conteúdos, a fim de garantir que a instrução ocorra adequadamente. O papel do professor diante das novas tecnologias educacionais é ter a mente aberta e abraçar os avanços tecnológicos, utilizando-os em proveito próprio e em benefício de seus alunos.

A partir do momento em que as novas ferramentas foram introduzidas no ensino, há uma crença equivocada de que o professor seria uma profissão extinta em breve, pois se antes o professor era o mestre que dominava o conhecimento e só através dele se chegava a um determinado conteúdo, hoje disponível na internet à distância de um clique. E essa democratização de acesso a dados e informações vem transformando a atuação e as expectativas relacionadas ao trabalho dos professores.

Ser professor que utiliza ativamente tais recursos implica em coordenar, orientar, incentivar a aprendizagem colaborativa e cada vez mais personalizada. Não se trata mais de uma mesma tarefa para todos dentro de um espaço temporal-local. O professor agora é aquele que coordena as atividades em torno de algum problema ou de alguns problemas. Assim, muitos grupos, em diferentes espaços e tempos, podem trabalhar em conjunto.

Se faz necessário, também, que os professores desenvolvam habilidades de comunicação no geral, mas em especial, em mídias sociais. Compartilhar informações, saber dar e receber feedbacks em público e entender a natureza de cada mídia social vai ajudá-los a interagir de forma mais adequada em cada um desses ambientes virtuais. As tradicionais habilidades de ler, falar e escrever de forma coerente permanecem, porém, é preciso pensar a comunicação de um jeito mais amplo, em diferentes contextos, formatos e mídias.

Com relação a pandemia, Souza e Fernandes (2023) constataram que existe uma série de preocupações e consequências para a saúde dos professores. Em primeiro lugar, destacam-se os riscos relacionados à exposição ao vírus da COVID-19, especialmente para os professores que continuaram trabalhando presencialmente durante a pandemia. Esses profissionais enfrentaram um aumento significativo no estresse e na ansiedade devido ao medo de contaminação e à preocupação com sua própria saúde e a de seus familiares. Além dos riscos físicos, os autores também identificam uma série de impactos na saúde mental dos professores. O isolamento social, a sobrecarga de trabalho, as dificuldades na adaptação ao ensino remoto e a pressão por resultados acadêmicos contribuíram para o aumento dos níveis de estresse, ansiedade e exaustão emocional entre os profissionais da Educação. Outro aspecto relevante abordado pelo artigo é a falta de suporte institucional e recursos adequados para lidar com os desafios enfrentados pelos professores durante a pandemia. Muitos profissionais relataram a falta de apoio emocional e psicológico por parte das instituições de ensino, bem como a escassez de recursos tecnológicos e treinamento para o ensino remoto.

É possível ainda perceber como a Tecnologia Digital pode impactar a Educação de maneira favorável tanto para os alunos quanto para os professores e a sociedade. Porém, esse avanço não atinge todas as pessoas. No Brasil ainda há uma parcela da população que não tem acesso a essa tecnologia. Conforme os dados, a Pesquisa Nacional por Amostra de Domicílios (PNAD) (IBGE, 2020) mostrou que uma em cada quatro pessoas não tem acesso à internet no Brasil, o que representa cerca de 46 milhões de brasileiros. 41,6% das pessoas que não acessam a rede não o fazem por não saber usar e não utilizam por esse

motivo, mas também por condição financeira, para 11,8% o serviço de acesso à internet é caro e para 5,7% o equipamento necessário para acessar a internet (celular, computador, tablet) também é caro. Para 4,5% das pessoas em todo o país que não acessam a internet, o serviço não está disponível nos locais que frequentam. Os dados confirmam a eliminação de uma grande parcela da população com direitos ao conhecimento proporcionado pela internet, e não apenas a supressão intelectual, surge com isso a redução de oportunidades de trabalho, de acordo com esse pensamento de (Bonilla, 2011.p.30) “impede que se reduza a exclusão social, uma vez que as principais atividades econômicas, governamentais e boa parte da produção cultural da sociedade vão migrando para a rede”

O distanciamento da sala de aula se tornou gritante com o surgimento da pandemia devido à necessidade de distanciamento social. Dessa forma, o ensino remoto foi implantado, ocorrendo da seguinte forma: os estudantes que possuem notebooks, tablets, internet de qualidade conseguiram prosseguir com seus estudos, enquanto boa parte dos alunos de instituições públicas não mantiveram seus estudos por falta de equipamentos. Muitos se viram desestimulados e evadiram da comunidade escolar.

A garantia do direito à educação implica construir caminhos concretos para o desenvolvimento pleno dos estudantes, caminhos para acesso e permanência de todos os estudantes nos sistemas de ensino com a aquisição das aprendizagens, competências e habilidades adequadas ao percurso completo da educação básica. É consenso entre pesquisadores brasileiros que, desde a segunda metade do século XX, o Brasil experimentou um processo intenso de massificação da oferta escolar pública (HASENBALG, 2003).

A tecnologia que invadiu todo o mundo, chegou à Educação, no entanto, são inegáveis as facilidades que permitem e viabilizam um ensino de qualidade, do qual estudantes e professores se beneficiam. Com inúmeras vantagens, este tipo de ensino, se expande cada dia mais, e com isso, não resta uma alternativa para os profissionais de Educação, a não ser a adaptação a essa nova realidade.

Apesar de alguns educadores ditos mais tradicionais relutarem para incorporar a tecnologia em suas aulas. Sobre isso Castro (2016, p.11) diz: “Um professor do ensino presencial pode utilizar diversos recursos tecnológicos e assim obter feedbacks através do que os alunos estão falando mesmo que indiretamente”.

A incorporação das Tecnologias de Informação e Comunicação (TIC) na escola contribui para expandir o acesso à informação atualizada e, principalmente, para promover a criação de comunidades colaborativas de aprendizagem que privilegiam a construção do conhecimento, a comunicação, a formação continuada e a gestão articulada entre as áreas administrativa, pedagógica e informacional da escola (RIOS, 2011).

3.2 Matemática para a construção de RNAs

Com os avanços tecnológicos e a crescente implementação da Inteligência Artificial em diversas áreas, o estudo e compreensão da Matemática empregada na construção de RNAs desempenham um papel fundamental. A Matemática por trás das Redes Neurais abrange diversos conceitos e técnicas provenientes da Álgebra Linear, estudo de Funções, Cálculo Diferencial e Integral, Teoria da Probabilidade e Estatística. A compreensão desses fundamentos matemáticos é essencial para projetar, treinar e otimizar eficientemente redes neurais capazes de aprender e generalizar a partir de dados.

Em um cenário onde as redes neurais são amplamente utilizadas em aplicações práticas, como reconhecimento facial, recomendação de conteúdo, diagnóstico médico assistido por computador e veículos autônomos, o domínio da Matemática associada a esses sistemas se torna imprescindível para garantir resultados confiáveis e seguros.

Portanto, esta seção do trabalho é direcionada ao estudo e entendimento da Matemática empregada na construção de RNAs que, depois, será usada na elaboração e construção de uma RNA, explicitando os processos matemáticos envolvidos, principalmente os fundamentos matemáticos empregados no Ensino Básico, assim mostrando na prática como os conteúdos, especificamente de Matemática, ensinados nas Escolas de Ensino Básico do Brasil estão diretamente relacionados com algo tão em voga na atualidade, como a chamada Inteligência Artificial.

3.2.1 Matrizes

O estudo das matrizes têm grande relevância no campo da Matemática Aplicada e possuem uma alta aplicabilidade no contexto das tecnologias modernas como na construção de RNAs. Uma matriz é uma estrutura matemática bidimensional composta por elementos dispostos em linhas e colunas, representada de forma geral por uma tabela de m linhas e n colunas (mais detalhes irão ser vistos no decorrer do capítulo, como definições, propriedades e operações).

A representação e manipulação eficiente de dados por meio de matrizes permitem o desenvolvimento de algoritmos complexos que são a base de muitas aplicações tecnológicas atuais, principalmente no contexto de construção e treinamento de inteligências artificiais, e os fundamentos matemáticos das matrizes são amplamente utilizados na construção de RNAs, dado que os *inputs* (entradas), os pesos das conexões entre os neurônios e os *outputs* (saídas) são representados por matrizes.

Os Parâmetros Curriculares Nacionais (PCN), o qual é uma coleção de documentos que auxilia o professor na tarefa de reflexão e discussão de aspectos do cotidiano da prática pedagógica, diz que:

Aprender Matemática de uma forma contextualizada, integrada e relacionada a outros conhecimentos traz em si o desenvolvimento de competências e habilidades

que são essencialmente formadoras, à medida que instrumentalizam e estruturam o pensamento do aluno, capacitando-o para compreender e interpretar situações para se apropriar de linguagens específicas, argumentar, analisar e avaliar, tirar conclusões próprias, tomar decisões, generalizar e para muitas outras ações necessárias à sua formação (BRASIL, 2002, p. 111).

Portanto, trazer essa contextualização do estudo dos fundamentos das matrizes, principalmente em um âmbito tão em alta como no contexto de Inteligência Artificial, beneficia em muito o aluno, pois o faz compreender estruturas e aplicações desses conceitos, além disso, desenvolve habilidades e competências essenciais para seu desenvolvimento cognitivo.

A Secretaria de Educação do Amazonas, através do Referencial Curricular Amazonense do Ensino Médio (RCN), estipula uma série de habilidades que o aluno deve desenvolver durante o Ensino Médio. Um exemplo é a habilidade (**EM13MAT301AM**) que propõe a resolução e elaboração de problemas do cotidiano, que envolvam matrizes e determinantes, com ou sem uso de tecnologias digitais.

Além dessa, há também a habilidade (**EM13MAT405**) que propõe a utilização de conceitos iniciais de uma linguagem de programação na implementação de algoritmos escritos em linguagem matemática.

Portanto, o estudo e compreensão dos fundamentos matemáticos das matrizes tem grande importância e relevância no desenvolvimento dessas e de outras habilidades por parte dos alunos, habilidades essas estipulados nas Matrizes Curriculares do Ensino Médio, na esfera estadual e federal.

3.2.1.1 Noção de matriz

Para fins didáticos, toda matriz introduzida nesta sessão pertence ao conjunto das matrizes $M_{m \times n}(\mathbb{R})$.

Definição 3.1 Dados dois números, m e n , naturais e não nulos, chama-se matriz m por n (indica-se $m \times n$), toda tabela M formada por números reais a_{ij} distribuídos em m linhas e n colunas, onde $1 \leq i \leq m$ e $1 \leq j \leq n$.

Assim o elemento a_{ij} está localizado no cruzamento da i -ésima linha com a j -ésima coluna. Dessa forma a matriz M de ordem $m \times n$ é representada por:

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \text{ ou } M = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Uma matriz M de ordem $m \times n$ pode ser indicada por: $A = (a_{ij}); i \in \{1, 2, 3, \dots, m\}$ e $j \in \{1, 2, 3, \dots, n\}$ ou simplesmente $M = (a_{ij})_{m \times n}$.

3.2.2 Matrizes especiais

Algumas matrizes, por possuírem determinadas características e recebem nomes especiais, das quais destacamos:

i. Matriz linha é toda matriz de ordem $1 \times n$, isto é, uma matriz com apenas uma linha.

Exemplo 3.1. $\begin{bmatrix} 0 & 1 & 3 & 6 \end{bmatrix}$ é uma matriz linha.

ii. Matriz coluna é toda matriz de ordem $m \times 1$, isto é, uma matriz com apenas uma coluna.

Exemplo 3.2. $\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ é uma matriz coluna.

iii. Matriz nula é toda matriz que possui todos os seus elementos iguais a zero.

Exemplo 3.3. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ é uma matriz nula de ordem 2×2 .

iv. Matriz quadrada de ordem n é toda matriz de ordem $n \times n$, ou seja, uma matriz que possui o número de linhas igual ao número de colunas.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

A **diagonal principal** de uma matriz quadrada de ordem n é indicada pelos elementos que possuem os dois índices iguais, isto é, a_{ij} onde $i = j$.

A **diagonal secundária** de uma matriz quadrada de ordem n é indicada pelos elementos que possuem a soma de seus índices igual a $n + 1$, isto é, a_{ij} onde $i + j = n + 1$.

Exemplo 3.4 Tomemos a matriz $\begin{bmatrix} 5 & -1 & 2 & 6 \\ 1 & 4 & -6 & 0 \\ 0 & -9 & 12 & -3 \\ -7 & -1 & 2 & 0 \end{bmatrix}$ de ordem 4.

Sua diagonal principal é 5, 4, 12 e 0 e sua diagonal secundária 6, -6, -9 e -7.

v. Matriz diagonal é uma matriz quadrada onde os elementos que não pertencem à diagonal principal são iguais a zero.

Exemplo 3.5. A matriz quadrada $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ de ordem 4 e um exemplo de matriz diagonal.

vi. Matriz identidade de ordem n (indica-se $I_n = (\delta_{ij})_{n \times n}$) é uma matriz diagonal em que os elementos da diagonal principal são iguais a 1.

δ_{ij} é o chamado **delta de Kronecker** e é definida por $\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}$.

Exemplo 3.6. A matriz quadrada $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ é uma matriz identidade de ordem 3.

3.2.2.1 Operações com matrizes

3.2.2.1.1 Igualdade

Definição 3.2. As matrizes $A = (a_{ij})$ de ordem $m \times n$ e $B = (b_{ij})$ de ordem $m \times n$ são iguais quando $a_{ij} = b_{ij}$ para todo $i \in \{1, 2, 3, \dots, m\}$ e todo $j \in \{1, 2, 3, \dots, n\}$. Dessa forma, para que a matriz A e a matriz B sejam iguais, A e B devem ser de mesma ordem e todos os elementos correspondentes iguais.

Exemplo 3.7. Sabendo-se que $\begin{bmatrix} 2x & 3y \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} x+1 & 2y \\ 3 & y+4 \end{bmatrix}$, os valores de x e y devem satisfazer o sistema

$$\begin{cases} 2x = x + 1 \\ 3y = 2y \\ 4 = y + 4 \end{cases}, \text{ então, } x = 1 \text{ e } y = 0.$$

3.2.2.1.2 Adição

Definição 3.3. Sejam as matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$, chama-se soma $A + B$ a matriz $C = (c_{ij})_{m \times n}$, tal que $c_{ij} = a_{ij} + b_{ij}$, para todo $i \in \{1, 2, 3, \dots, m\}$ e todo $j \in \{1, 2, 3, \dots, n\}$. A soma das matrizes A e B de ordem $m \times n$ é um matriz C também de ordem $m \times n$ em que cada elemento da matriz C é a soma dos elementos correspondente nas matrizes A e B .

$$A + B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Exemplo 3.8.
$$\begin{bmatrix} 3 & 7 \\ 4 & -5 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 11 & 4 \\ -4 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3+11 & 7+4 \\ 4-4 & -5+0 \\ 0+3 & 3+1 \end{bmatrix} = \begin{bmatrix} 14 & 11 \\ 0 & -5 \\ 3 & 4 \end{bmatrix}$$

Teorema 3.1. A adição de matrizes de mesma ordem ($m \times n$) apresenta as seguintes propriedades:

- i. Comutatividade: $A + B = B + A$
- ii. Associatividade: $(A + B) + C = A + (B + C)$, para quaisquer que sejam as matrizes A , B e C de ordem $m \times n$.
- iii. Elemento neutro: $\exists N \mid A + N = A$, para qualquer que seja a matriz A do tipo $m \times n$.
- iv. Elemento simétrico: para toda matriz A do tipo $m \times n$, $\exists A' \mid A + A' = N$.

3.2.2.1.3 Matriz oposta

Definição 3.4. Dada a matriz $A = (a_{ij})_{m \times n}$, chama-se matriz oposta (indica-se $-A$) a matriz A' tal que $A + A' = 0$, onde 0 é matriz nula de ordem $m \times n$.

Exemplo 3.9. Dada a matriz $A = \begin{bmatrix} 3 & 1 \\ 7 & -11 \end{bmatrix}$, então $-A = \begin{bmatrix} -3 & -1 \\ -7 & 11 \end{bmatrix}$, pois $A + (-A) = \begin{bmatrix} 3 & 1 \\ 7 & -11 \end{bmatrix} + \begin{bmatrix} -3 & -1 \\ -7 & 11 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

3.2.2.1.4 Produto de um número por uma matriz

Dada uma constante real k ($k \in \mathbb{R}$) e uma matriz $A = (a_{ij})_{m \times n}$, chama-se produto kA a matriz $B = (b_{ij})_{m \times n}$ tal que $b_{ij} = k \cdot a_{ij}$ para todo $i \in \{1, 2, 3, \dots, m\}$ e todo $j \in \{1, 2, 3, \dots, n\}$. Isto é, multiplicar a matriz A pela constante k significa construir uma matriz B onde todos os seus elementos são formados multiplicando os elementos da matriz A pela constante k .

Exemplo 3.10. $3 \cdot \begin{bmatrix} 0 & 4 & 1 \\ 6 & -3 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 12 & 3 \\ 18 & -9 & 30 \end{bmatrix}$

Teorema 3.2. Dado os números reais a e b e as matrizes A e B de ordem $m \times n$, então valem as seguintes propriedades:

- i. $a \cdot (b \cdot A) = (a \cdot b) \cdot A$
- ii. $a \cdot (A + B) = a \cdot A + a \cdot B$
- iii. $(a + b) \cdot A = a \cdot A + b \cdot A$
- iv. $1 \cdot A = A$

3.2.2.1.5 Produto de matrizes

Definição 3.5. Sejam as matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{jk})_{n \times p}$, chama-se produto AB a matriz $C = (c_{ik})_{m \times p}$ tal que $c_{ik} = a_{i1} \cdot b_{1k} + a_{i2} \cdot b_{2k} + a_{i3} \cdot b_{3k} + \dots + a_{in} \cdot b_{nk} = \sum_{j=1}^n a_{ij} b_{jk}$ para todo $i \in \{1, 2, \dots, m\}$ e todo $k \in \{1, 2, \dots, p\}$.

Esta definição garante a existência do produto AB se, e somente se, o número de colunas da matriz A for igual ao número de linhas da matriz B , isto é, a matriz A é de ordem $m \times n$ e a matriz B é de ordem $n \times p$. Além disso, o produto AB é de ordem $m \times p$, tendo o mesmo número de linhas da matriz A e o mesmo número de colunas da matriz B .

Exemplo 3.11. Dadas as matrizes $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ e a matriz $B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$, para

calcular o produto AB , é preciso seguir o seguinte procedimento:

i. verifica-se a condição de existência do produto AB .

No exemplo em questão, como a matriz A é do tipo 2×3 e a matriz B é do tipo 3×1 , decorre que existe a matriz AB e é do tipo 2×1 , ou seja, $AB = C = \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix}$.

ii. calcula-se o elemento c_{11} .

Toma-se a linha 1 da matriz A :

$$a_{11} \quad a_{12} \quad a_{13}$$

Toma-se a coluna 1 da matriz B :

$$b_{11}$$

$$b_{21}$$

$$b_{31}$$

Multiplica-se o primeiro elemento da linha 1 da matriz A com o primeiro elemento da coluna 1 da matriz B :

$$a_{11} \cdot b_{11}$$

Multiplica-se o segundo elemento da linha 1 da matriz A com o segundo elemento da coluna 1 da matriz B :

$$a_{12} \cdot b_{21}$$

Fazendo o mesmo procedimento com o terceiro elemento da linha 1 da matriz A com o terceiro elemento da coluna 1 da matriz B :

$$a_{13} \cdot b_{31}$$

De posse desses resultados, o valor de c_{11} é a soma de cada um deles:

$$c_{11} = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31}$$

iii. Repete-se o mesmo procedimento para a linha 2 da matriz A e a coluna 1 da matriz B para calcular o valor de c_{21} .

Assim tem-se:

$$AB = C = \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix} = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} \end{bmatrix}$$

Teorema 3.3. Se $A = (a_{ij})_{m \times n}$, então $AI_n = A$ e $I_m A = A$.

Demonstração: Sendo $I_n = (\delta_{ij})_{n \times n}$ e $B = AI_n = (b_{ij})_{m \times n}$, temos:

$$b_{ij} = a_{i1} \cdot \delta_{1j} + a_{i2} \cdot \delta_{2j} + a_{i3} \cdot \delta_{3j} \cdots + a_{ii} \cdot \delta_{ii} + \cdots + a_{in} \cdot \delta_{ni} \implies$$

$$b_{ij} = a_{i1} \cdot 0 + a_{i2} \cdot 0 + a_{i3} \cdot 0 \cdots + a_{ii} \cdot 1 + \cdots + a_{in} \cdot 0 = a_{ii} \text{ para todos } i \text{ e } j, \text{ então}$$

$$AI_n = A$$

A demonstração de $I_m A = A$ é feita de forma análoga.

Teorema 3.4. O produto de matrizes, estabelecido conforme a *definição 3.5*, apresenta as seguintes propriedades:

i. Associatividade: $(AB)C = A(BC)$ quaisquer que sejam $A = (a_{ij})_{m \times n}$, $B = (b_{jk})_{n \times p}$ e $C = (c_{kl})_{p \times r}$.

Demonstração: Fazendo $D = AB = (d_{ik})_{m \times p}$, $E = (AB)C = (e_{il})_{m \times r}$ e $F = BC = (f_{jl})_{n \times r}$, temos:

$$\begin{aligned} e_{il} &= \sum_{k=1}^p d_{ik} \cdot c_{kl} \\ &= \sum_{k=1}^p \left(\sum_{j=1}^n a_{ij} \cdot b_{jk} \right) \cdot c_{kl} \\ &= \sum_{k=1}^p \left(\sum_{j=1}^n a_{ij} \cdot b_{jk} \cdot c_{kl} \right) \\ &= \sum_{j=1}^n a_{ij} \cdot \left(\sum_{k=1}^p b_{jk} \cdot c_{kl} \right) \\ &= \sum_{j=1}^n a_{ij} \cdot f_{jl} \end{aligned}$$

$$\text{então } (AB)C = A(BC).$$

ii. Distributividade à direita com relação à adição: $(A + B)C = AC + BC$ quaisquer que sejam as matrizes $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$ e $C = (c_{jk})_{n \times p}$.

Demonstração: Fazendo $D = (A + B)C = (d_{ik})_{m \times p}$, temos:

$$\begin{aligned}
d_{ik} &= \sum_{j=1}^n (a_{ij} + b_{ij}) \cdot c_{jk} \\
&= \sum_{j=1}^n (a_{ij} \cdot c_{jk} + b_{ij} \cdot c_{jk}) \\
&= \sum_{j=1}^n a_{ij} \cdot c_{jk} + \sum_{j=1}^n b_{ij} \cdot c_{jk}
\end{aligned}$$

$$\text{então } (A + B)C = AC + BC.$$

iii. Distributividade à esquerda com relação à adição: $C(A + B) = CA + CB$ para quaisquer que sejam as matrizes $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$ e $C = (c_{ki})_{n \times p}$.

Demonstração: análoga a (ii).

iv. $(kA)B = A(kB) = k(AB)$ quaisquer que sejam o número k e as matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{jk})_{n \times p}$.

Demonstração: Fazendo $C = kA = (c_{ij})_{m \times n}$, $D = kB = (d_{jk})_{n \times p}$ e $E = AB = (e_{ik})_{m \times p}$, temos:

$$\begin{aligned}
\sum_{j=1}^n c_{ij} \cdot b_{jk} &= \sum_{j=1}^n (k \cdot a_{ij}) \cdot b_{jk} \\
&= k \sum_{j=1}^n a_{ij} \cdot b_{jk} = \sum_{j=1}^n a_{ij} \cdot d_{jk} \\
&= \sum_{j=1}^n a_{ij} \cdot (k \cdot b_{jk}) = k \sum_{j=1}^n a_{ij} \cdot b_{jk}
\end{aligned}$$

$$\text{então } (kA)B = A(kB) = k(AB).$$

3.2.2.2 Matriz transposta

Definição 3.6. Seja uma matriz $A = (a_{ij})_{m \times n}$, chama-se transposta de A a matriz $A^t = (a'_{ji})_{n \times m}$ tal que $a'_{ji} = a_{ij}$, para todo i e todo j .

Dessa forma, podemos dizer que primeira linha da matriz A equivale à primeira coluna da matriz A^t , e que a primeira coluna da matriz A equivale à primeira linha da matriz A^t .

Exemplo 3.12. A matriz $A = \begin{bmatrix} -4 & 1 \\ 3 & 7 \\ 0 & -5 \end{bmatrix}$ possui a transposta $A^t = \begin{bmatrix} -4 & 3 & 0 \\ 1 & 7 & -5 \end{bmatrix}$.

Teorema 3.5. A matriz transposta apresenta as seguintes propriedades:

i. $(A^t)^t = A$ para toda matriz $A = (a_{ij})_{m \times n}$.

Demonstração: Fazendo $(A^t)^t = (a''_{ij})_{m \times n}$, então $a''_{ij} = a'_{ji} = a_{ij}$ para todo i, j .

ii. Se $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$, então $(A + B)^t = A^t + B^t$.

Demonstração: Fazendo $A + B = C = (c_{ij})_{m \times n}$ e $(A + B)^t = C^t = (c'_{ji})_{n \times m}$, temos:
 $c'_{ji} = c_{ij} = a_{ij} + b_{ij} = a'_{ji} + b'_{ji}$, para todo i, j .

iii. Se $A = (a_{ij})_{m \times n}$ e $k \in \mathbb{R}$, então $(kA)^t = kA^t$.

Demonstração: Fazendo $(kA)^t = (a''_{ji})_{n \times m}$, temos:
 $a''_{ji} = ka_{ji} = ka'_{ij}$, para todo i, j .

iv. Se $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times p}$, então $(AB)^t = B^t A^t$.

Demonstração: Fazendo $A = C = (c_{ik})_{m \times p}$ e $(AB)^t = C^t = (c'_{ki})_{p \times m}$, temos:

$$\begin{aligned} c'_{ki} &= c_{ik} = \sum_{j=1}^n a_{ij} \cdot b_{jk} \\ &= \sum_{j=1}^n b_{jk} \cdot a_{ij} = \sum_{j=1}^n b'_{kj} \cdot a'_{ji} \end{aligned}$$

então $(AB)^t = B^t A^t$.

Exemplo 3.13. Dadas as matrizes $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$, verifica-se a validade da propriedade (iv).

$$\begin{aligned} \text{Solução: } AB &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \Rightarrow \\ (AB)^t &= \begin{bmatrix} ae + bg & ce + dg \\ af + bh & cf + dh \end{bmatrix} = \begin{bmatrix} e & g \\ f & h \end{bmatrix} \cdot \begin{bmatrix} a & c \\ b & d \end{bmatrix} = B^t A^t \end{aligned}$$

3.2.2.3 Matriz simétrica e matriz antissimétrica

Definição 3.7. Toda matriz quadrada A , de ordem n , tal que $A^t = A$ é chamada de matriz simétrica.

Desta forma, se $A = (a_{ij})$ é uma matriz simétrica, temos:

$$a_{ij} = a_{ji} \text{ para todo } i, j \in \{1, 2, 3, \dots, n\}.$$

Exemplo 3.14. Dada a matriz $A = \begin{bmatrix} 1 & x & 5 \\ 2 & 7 & -4 \\ y & z & -3 \end{bmatrix}$, vamos determinar x , y e z de modo que a matriz A seja simétrica:

Solução: Temos $A^t = \begin{bmatrix} 1 & 2 & y \\ x & 7 & z \\ 5 & -4 & -3 \end{bmatrix}$ e para que a matriz A seja simétrica, segundo a **definição 3.7**, devemos ter $A^t = A$.

$$\text{Assim, } A^t = A \Rightarrow \begin{bmatrix} 1 & 2 & y \\ x & 7 & z \\ 5 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 1 & x & 5 \\ 2 & 7 & -4 \\ y & z & -3 \end{bmatrix} \Rightarrow x = 2, y = 5 \text{ e } z = -4.$$

Definição 3.8. Toda matriz quadrada A , de ordem n , tal que $A^t = -A$ é chamada de matriz antissimétrica.

Desta forma, se $A = (a_{ij})$ é uma matriz antissimétrica, temos:

$$a_{ij} = -a_{ji} \text{ para todo } i, j \in \{1, 2, 3, \dots, n\}.$$

Exemplo 3.15. Dada a matriz $A = \begin{bmatrix} 0 & -4 & 2 \\ x & 0 & 1-z \\ y & 2z & 0 \end{bmatrix}$, vamos determinar x , y e z de modo que a matriz A seja antissimétrica:

Solução: Temos $A^t = \begin{bmatrix} 0 & x & y \\ -4 & 0 & 2z \\ 2 & 1-z & 0 \end{bmatrix}$, para que a matriz A seja antissimétrica, segundo a **definição 3.8**, devemos ter $A^t = -A$.

$$\begin{aligned} \text{Assim, } A^t = -A \Rightarrow \begin{bmatrix} 0 & x & y \\ -4 & 0 & 2z \\ 2 & 1-z & 0 \end{bmatrix} &= - \begin{bmatrix} 0 & -4 & 2 \\ x & 0 & 1-z \\ y & 2z & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 & -2 \\ -x & 0 & z-1 \\ -y & -2z & 0 \end{bmatrix} \\ \Rightarrow x &= 4, y = -2 \text{ e } z = -1. \end{aligned}$$

3.2.2.4 Matrizes inversíveis

Definição 3.9. Uma matriz A , de ordem n , é inversível se existir uma matriz B tal que $AB = BA = I_n$. Caso contrário, A é uma matriz singular.

Teorema 3.6. Se a matriz A é inversível, então a matriz B tal que $AB = BA = I_n$ é única.

Demonstração: como A é inversível, então existe a matriz B tal que $AB = BA = I_n$. Admita que exista uma matriz C tal que $AC = CA = I_n$, temos:

$$C = I_n C = (BA)C = B(AC) = BI_n = B$$

então a matriz B é única.

Seja matriz A é inversível, então existe uma matriz chamada de matriz inversa de A (que é única) indicada por A^{-1} tal que $AA^{-1} = A^{-1}A = I_n$.

Note que, como A^{-1} comuta com A , A^{-1} deve também ser quadrada de ordem n .

Exemplo 3.16. Sabendo que a matriz $A = \begin{bmatrix} 3 & 7 \\ 5 & 11 \end{bmatrix}$ é inversível e, vamos determinar A^{-1} .

Solução: como A é inversível, existe A^{-1} tal que $AA^{-1} = A^{-1}A = I_n$

Fazendo $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, temos:

$$A^{-1}A = I_n \Rightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 5 & 11 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3a+5b & 7a+11b \\ 3c+5d & 7c+11d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

pela definição de igualdade de matrizes, decorre:

$$\begin{cases} 3a+5b=1 \\ 7a+11b=0 \end{cases} \Rightarrow a = -\frac{11}{2} \text{ e } b = \frac{7}{2}$$

e

$$\begin{cases} 3c+5d=0 \\ 7c+11d=1 \end{cases} \Rightarrow c = \frac{5}{2} \text{ e } d = -\frac{3}{2}$$

$$\text{portanto } A^{-1} = \begin{bmatrix} -\frac{11}{2} & \frac{7}{2} \\ \frac{5}{2} & -\frac{3}{2} \end{bmatrix}.$$

3.2.3 Vetores

Os vetores conseguem representar grandezas físicas e abstratas de forma compacta e eficiente. Suas aplicações estão em diversas áreas da Ciência e Engenharia, utilizados para representar forças, deslocamentos, velocidades, campos elétricos, entre outras grandezas, incluindo a construção e o funcionamento de RNA. Além disso, no campo da Matemática Aplicada, os vetores são essenciais para o desenvolvimento de modelos matemáticos que descrevem fenômenos complexos.

No contexto das RNA, os vetores desempenham um papel fundamental na representação e manipulação dos dados de entrada e dos parâmetros do modelo. As entradas são frequentemente representadas por vetores numéricos processados por meio de operações matriciais e operações não-lineares nas camadas do neurônio artificial.

Em resumo, o uso inteligente dos conceitos vetoriais é fundamental para avanços no campo da inteligência artificial e no desenvolvimento de sistemas computacionais cada vez mais sofisticados e eficazes.

3.2.3.1 Equipolência de segmentos orientados

Definição 3.10. Diz-se que os segmentos orientados AB e CD são equipolentes (escreve-se $AB \equiv CD$), quando possuem a mesma direção, o mesmo sentido e o mesmo comprimento.

Decorre da definição o seguinte critério para verificação de equipolência entre dois segmentos:

$$AB \equiv CD \iff \text{ponto médio de } AD = \text{ponto médio de } BC.$$

Em termos de coordenadas a caracterização da equipolência, considerando um sistema de eixos ortogonais OXY no plano e $A = (a_1, a_2)$; $B = (b_1, b_2)$; $C = (c_1, c_2)$ e $D = (d_1, d_2)$ pontos do plano expressos em coordenadas com relação ao sistema dado, tem-se:

$$AB \equiv CD \iff b_1 - a_1 = d_1 - c_1 \text{ e } b_2 - a_2 = d_2 - c_2$$

Demonstração: Considere os segmentos equipolentes AB e CD , assim temos:

$AB \equiv CD \iff$ ponto médio de AD = ponto médio de BC

$$\iff \left(\frac{a_1+d_1}{2}, \frac{a_2+d_2}{2} \right) = \iff \left(\frac{b_1+c_1}{2}, \frac{b_2+c_2}{2} \right)$$

$$\iff (a_1 + d_1, a_2 + d_2) = (b_1 + c_1, b_2 + c_2)$$

$$\iff a_1 + d_1 = b_1 + c_1 \text{ e } a_2 + d_2 = b_2 + c_2$$

$$\iff b_1 - a_1 = d_1 - c_1 \text{ e } b_2 - a_2 = d_2 - c_2$$

como queríamos demonstrar.

Exemplo 3.17. Vamos determinar o ponto D tal que $CD \equiv AB$, sabendo que $A = (-1, -1)$, $B = (2, 3)$ e $C = (2, 1)$:

Solução: Fazendo $D = (x, y)$, como $CD \equiv AB \iff x - 2 = 2 + 1$ e $y - 1 = 3 + 1$
 $\iff x = 5$ e $y = 5$.

Fazendo a verificação de equipolência, temos:

$$\text{Ponto médio de } CB = \left(\frac{2+2}{2}, \frac{3+1}{2} \right) = (2, 2)$$

$$\text{Ponto médio de } DA = \left(\frac{5-1}{2}, \frac{5-1}{2} \right) = (2, 2)$$

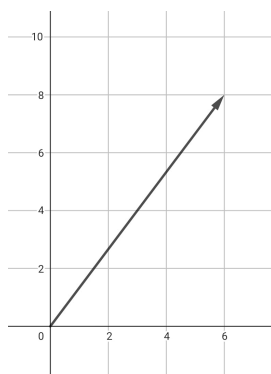
como ponto médio de CB = ponto médio de DA , concluímos que $CD \equiv AB$.

3.2.3.2 Vetores no plano

Definição 3.11. Dados os pontos A e B no plano. O vetor $\vec{v} = \overrightarrow{AB}$ é o conjunto de todos os segmentos orientados e equipolentes a AB .

Exemplo 3.18. Dados $A = (a_1, a_2)$ e $B = (b_1, b_2)$, os números $b_1 - a_1$ e $b_2 - a_2$ são as coordenadas do vetor $\vec{v} = \overrightarrow{AB}$ (escreve-se $\vec{v} = (b_1 - a_1, b_2 - a_2)$).

Figura 1 – Representação do vetor $\vec{v} = (6, 8)$



Fonte: elaborado pelo autor

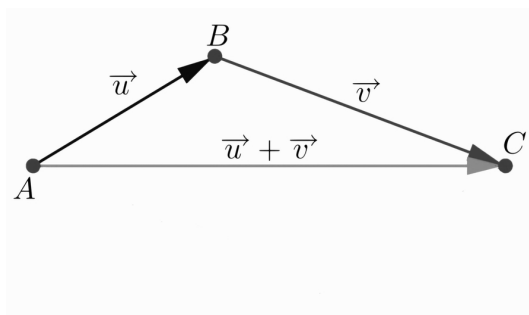
3.2.3.3 Operações com vetores

3.2.3.3.1 Adição

Definição 3.12. A cada vetor $\vec{u}$ e $\vec{v}$ associa-se um novo vetor $\vec{u} + \vec{v}$ chamado soma dos vetores $\vec{u}$ e $\vec{v}$, definido:

Se $\vec{u} = \overrightarrow{AB}$, seja C o único ponto tal que $\vec{v} = \overrightarrow{BC}$. O vetor soma de $\vec{u}$ com $\vec{v}$ é o vetor $\overrightarrow{AC}$, isto é, $\vec{u} + \vec{v} = \overrightarrow{AC}$.

Figura 2 – Representação gráfica da adição $\vec{u} + \vec{v}$



Fonte: elaborado pelo autor

Na prática, a representação é feita por meio de coordenadas em sistema de eixos ortogonais da seguinte forma:

Dados $\vec{u} = (u_1, u_2)$ e $\vec{v} = (v_1, v_2)$ vetores no plano expressos em termos de coordenadas em um sistema de eixos ortogonais OXY , então $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2)$.

Exemplo 3.19. Dados os vetores $\vec{u} = (2, 4)$ e $\vec{v} = (5, -3)$, vamos calcular, em termo de coordenadas, o vetor $\vec{u} + \vec{v}$.

Solução: Sabe-s, que $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2)$, assim:

$$\vec{u} + \vec{v} = (2 + 5, 4 + (-3)) = (7, 1)$$

Teorema 3.7. Sejam os vetores $\vec{u}$, $\vec{v}$ e $\vec{w}$ no plano, a adição de vetores, estabelecida conforme **definição 3.12**, apresenta as seguintes propriedades:

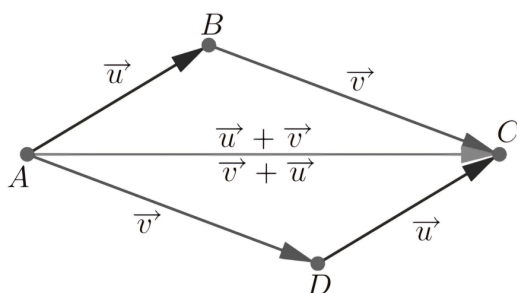
i. Comutatividade: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

Demonstração: Tomemos os vetores $\vec{u} = \overrightarrow{AB}$ e $\vec{v} = \overrightarrow{BC}$. Pela **definição 5.20**, tem-se $\vec{u} + \vec{v} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

Agora tomemos os vetores $\vec{v} = \overrightarrow{AD}$ e $\vec{u} = \overrightarrow{DC}$, então $\vec{v} + \vec{u} = \overrightarrow{AD} + \overrightarrow{DC} = \overrightarrow{AC}$.

Geometricamente podemos representar esse fato em paralelogramo.

Figura 3 – Comutatividade da adição de vetores



Fonte: elaborado pelo autor

Portanto $\vec{u} + \vec{v} = \overrightarrow{AC} = \vec{v} + \vec{u}$.

ii. Associatividade: $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

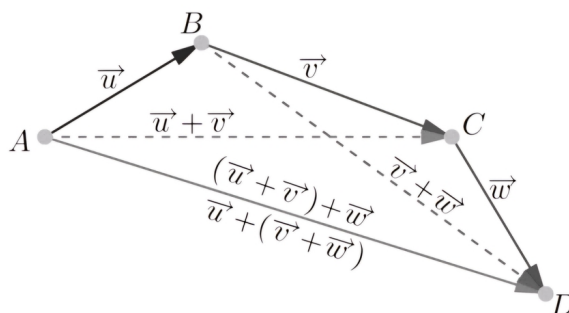
Demonstração: Considere os vetores $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{BC}$ e $\vec{w} = \overrightarrow{CD}$.

Assim $\vec{u} + (\vec{v} + \vec{w}) = \overrightarrow{AB} + (\overrightarrow{BC} + \overrightarrow{CD}) = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}$.

por outro lado $(\vec{u} + \vec{v}) + \vec{w} = (\overrightarrow{AB} + \overrightarrow{BC}) + \overrightarrow{CD} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD}$.

A representação geométrica fica da seguinte forma:

Figura 4 – Associatividade da adição de vetores



Fonte: elaborado pelo autor

iii. Elemento neutro aditivo: o vetor $\vec{0}$ é tal que $\vec{v} + \vec{0} = \vec{v}$.

Demonstração: Fazendo $\vec{v} = \overrightarrow{AB}$ e $\vec{0} = \overrightarrow{AA} = \overrightarrow{BB}$, temos:

$\vec{v} + \vec{0} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = \vec{v}$ e $\vec{0} + \vec{v} = \overrightarrow{AA} + \overrightarrow{AB} = \overrightarrow{AB} = \vec{v}$

iv. Inverso aditivo: para cada vetor $\vec{v}$ existe um único vetor, designado $-\vec{v}$, tal que $\vec{v} + (-\vec{v}) = \vec{0}$.

3.2.3.3.2 Multiplicação de vetores por escalares

Definição 3.13. Dados $A = (a_1, a_2)$, $B = (b_1, b_2)$, $\lambda \in \mathbb{R}$ então, $\lambda \overrightarrow{AB} = \overrightarrow{AC}$, onde $C = (a_1 + \lambda(b_1 - a_1), a_2 + \lambda(b_2 - a_2))$.

Consequentemente $\lambda \overrightarrow{AB} = (\lambda(b_1 - a_1), \lambda(b_2 - a_2))$.

Cabe a observação de que as semirretas $\overrightarrow{AB}$ e $\overrightarrow{AC}$ coincidem quando $\lambda > 0$ e são opostas quando $\lambda < 0$. Portanto AB e AC têm o mesmo sentido se $\lambda > 0$ e sentido oposto se $\lambda < 0$.

Teorema 3.8. Sejam os vetores $\vec{u}$ e $\vec{v}$ no plano e $\lambda, \mu \in \mathbb{R}$, a multiplicação de vetores por escalares, estabelecida conforme **definição 3.13**, apresenta as seguintes propriedades:

- i. Elemento neutro multiplicativo: $1 \in \mathbb{R}$ satisfaz $1\vec{u} = \vec{u}$.
- ii. Distributividades: $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$ e $(\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}$.

Demonstração: Tomando os vetores $\vec{u} = (a, b)$ e $\vec{v} = (a', b')$, e dados $\lambda, \mu \in \mathbb{R}$, temos:

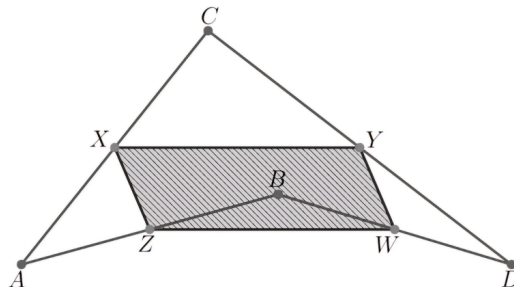
$$\begin{aligned}\lambda(\vec{u} + \vec{v}) &= \lambda[(a, b) + (a', b')] = \lambda(a + a', b + b') \\ &= [\lambda(a + a'), \lambda(b + b')] = (\lambda a + \lambda a', \lambda b + \lambda b') \\ &= (\lambda a, \lambda b) + (\lambda a', \lambda b') = \lambda(a, b) + \lambda(a', b') = \lambda\vec{u} + \lambda\vec{v}.\end{aligned}$$

A segunda parte da propriedade distributiva é demonstrada de forma análoga.

Exemplo 3.20. Mostre que os pontos médios dos lados de um quadrilátero são os vértices de um paralelogramo.

Solução: Seja $ABCD$ um quadrilátero qualquer e sejam X, Y, W e Z pontos médios dos lados AC, CD, DB e BA , respectivamente, como na figura.

Figura 5 – Pontos médios dos lados de um quadrilátero qualquer



Fonte: elaborado pelo autor

nesse caso, temos:

$$\begin{aligned}\overrightarrow{XY} &= \overrightarrow{XC} + \overrightarrow{CY} = \frac{\overrightarrow{AC}}{2} + \frac{\overrightarrow{CD}}{2} = \frac{1}{2}(\overrightarrow{AC} + \overrightarrow{CD}) = \frac{1}{2}\overrightarrow{AD} \\ \text{e } \overrightarrow{ZW} &= \overrightarrow{ZB} + \overrightarrow{BW} = \frac{\overrightarrow{AB}}{2} + \frac{\overrightarrow{BD}}{2} = \frac{1}{2}(\overrightarrow{AB} + \overrightarrow{BD}) = \frac{1}{2}\overrightarrow{AD}\end{aligned}$$

ou seja $\overrightarrow{XY} = \frac{1}{2}\overrightarrow{AD} = \overrightarrow{ZW}$, então $XY \equiv ZW$ e, portanto, $XYZW$ é um paralelogramo.

3.2.4 Funções reais de uma variável

A importância do estudo das funções reais reside na sua capacidade de descrever e analisar uma ampla variedade de fenômenos naturais e artificiais de forma precisa e concisa, além do desenvolvendo modelos do comportamento de neurônios artificiais.

O Referencial Curricular Amazonense do Ensino Médio (RCN), estipula algumas habilidades que o aluno deve desenvolver relacionadas ao estudo das funções reais, tais como (**EM13MAT101**) que propõe análise de gráficos das funções com ou sem apoio de tecnologias digitais.

Há também as habilidades do RCN, de códigos EM13MAT302, EM13MAT304, EM13MAT305, EM13MAT306, EM13MAT401, EM13MAT402, EM13MAT403, EM13MAT404, EM13MAT502 e EM13MAT503, as quais propõem a utilização dos conceitos do estudo das funções reais durante o Ensino Médio nas escolas públicas do Estado do Amazonas.

3.2.4.1 Produto cartesiano

Definição 3.14. Dados dois conjuntos A e B não vazios, denominamos produto cartesiano de A por B o conjunto $A \times B$ (lê-se “ A cartesiano B ”) cujo os elementos são todos os pares ordenados (x, y) , em que o elemento x pertence ao conjunto A e o elemento y pertence ao conjunto B .

$$A \times B = \{(x, y) \mid x \in A \text{ e } y \in B\}$$

Exemplo 3.21. Se $A = \{-1, 3, 5\}$ e $B = \{1, 2\}$, temos:

$$A \times B = \{(-1, 1), (-1, 2), (3, 1), (3, 2), (5, 1), (5, 2)\}$$

e

$$B \times A = \{(1, -1), (2, -1), (1, 3), (2, 3), (1, 5), (2, 5)\}$$

3.2.4.2 Relação binária

Definição 3.15. Dados os conjuntos A e B , chama-se relação binária de A em B todo subconjunto R de $A \times B$.

R é a relação binária de A em $B \iff R \subset A \times B$.

O conjunto A , conjunto de partida da relação R , é chamado Domínio, assim $x \in D \iff \exists y, y \in B \mid (x, y) \in R$.

O conjunto B , conjunto de chegada da relação R , é chamado Contradomínio.

O conjunto de todos os elementos de pares ordenados pertencentes ao contradomínio é chamado de Conjunto Imagem, assim $y \in Im \iff \exists x, x \in A \mid (x, y) \in R$.

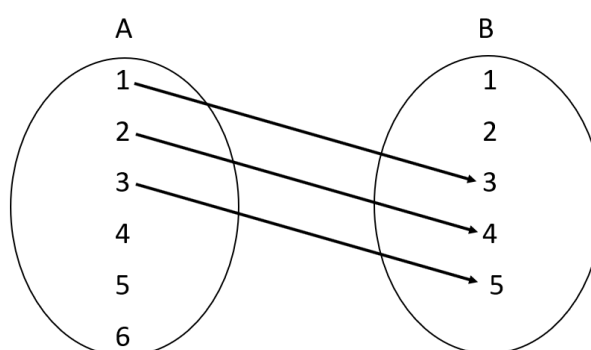
Exemplo 3.22. Se $A = \{1, 2, 3, 4, 5, 6\}$ e $B = \{1, 2, 3, 4, 5\}$, vamos determinar os elementos da relação $R = \{(x, y) \mid y = x + 2\}$.

Solução: Fazem parte da relação todos os pares ordenados (x, y) tais que $x \in A$, $y \in B$ e $y = x + 2$, dessa forma:

$$R = \{(1, 3), (2, 3), (3, 5)\}$$

Utilizando o diagrama de Venn:

Figura 6 – Diagrama de Venn exemplo 3.22



Fonte: elaborado pelo autor

Temos que o domínio é o conjunto A , o contradomínio é o conjunto B e o conjunto imagem é $Im = \{3, 4, 5\}$.

3.2.4.3 Conceito de função

Definição 3.16. Dados os Conjuntos Reais A e B , não vazios, uma relação f de A e B , chamada de função definida em A com imagens em B se, e somente se, para todo $x \in A$ existe um único $y \in B$ tal que $(x, y) \in f$.

$$f \text{ é aplicação de } A \text{ em } B \iff \{\forall x \in A, \exists \mid y \in B \mid (x, y) \in f\}$$

Em outras palavras, para que a relação f seja uma função, é necessário que f satisfaça as seguintes condições:

- 1º) todo elemento $x \in A$ precisa fazer parte de pelo menos um par $(x, y) \in f$.
- 2º) cada elemento $x \in A$ precisa participar de um único par $(x, y) \in f$.

3.2.4.3.1 Notação das funções

Como toda função é um conjunto de pares ordenados, existe uma sentença aberta $y = f(x)$ que expressa a lei a qual, dado $x \in A$, determina-se $y \in B$ tal que $(x, y) \in f$, então:

$$f = \{(x, y) \mid x \in A, y \in B \text{ e } y = f(x)\}.$$

Usaremos a notação $f : A \rightarrow B$ tal que $y = f(x)$.

Exemplo 3.23. $f = \mathbb{R}_+ \rightarrow \mathbb{R}$ tal que $y = \sqrt{x}$ é uma função que faz corresponder a cada $x \in \mathbb{R}_+$ um $y \in \mathbb{R}$ tal que $y = \sqrt{x}$.

3.2.4.3.2 Imagem de um elemento

Se $(a, b) \in f$, o elemento b é chamado imagem de a pela f e indicamos $f(a) = b$.

Exemplo 3.24. Seja a função $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $y = 3x - 1$, então:

- $f(0) = 3 \cdot 0 - 1 = -1$, isto é, a imagem de 0 pela aplicação f é -1 .
- $f(1) = 3 \cdot 1 - 1 = 2$, isto é, a imagem de 1 pela aplicação f é 2.

Exemplo 3.25. A função $f : \mathbb{R} \rightarrow \mathbb{R}$ é tal que, para todo $x \in \mathbb{R}$, $f(3x) = 3 \cdot f(x)$. Se $f(9) = 45$, calcule $f(1)$.

Solução: Fazendo $3x = 9 \implies x = 3$

$$f(9) = f(3 \cdot 3) = 3 \cdot f(3) = 45 = 3 \cdot 15 \implies f(3) = 15$$

Fazendo $3x = 3 \implies x = 1$

$$f(3) = f(3 \cdot 1) = 3f(1) = 15 = 3 \cdot 5 \implies f(1) = 5$$

Portanto, $f(1) = 5$.

3.2.4.3.3 Domínio e imagem de uma função

Domínio

Dados os conjuntos A e B não vazios, Domínio é o conjunto D formado pelos elementos $x \in A$ para os quais existe $y \in B$ tal que $(x, y) \in f$. Pela **definição 3.16**, todo elemento de A possui essa propriedade, temos $D = A$.

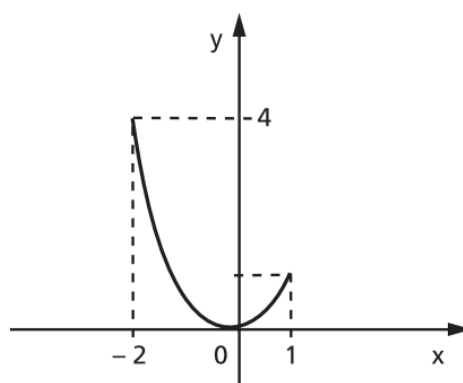
Imagem

Dados os conjuntos A e B não vazios, imagem é o conjunto Im formado pelos elementos $y \in B$ para os quais existe $x \in A$ tal que $(x, y) \in f$.

Na representação cartesiana, Im é o conjunto formado por todas ordenadas dos pontos do gráfico de f .

Exemplo 3.26.

Gráfico 1 – Representação cartesiana de uma função real



Fonte: Fundamentos de matemática elementar, 1; conjuntos e funções, pág 89

No exemplo acima, temos:

$$D = \{x \in \mathbb{R} \mid -2 \leq x \leq 1\} \text{ e } Im = \{y \in \mathbb{R} \mid 0 \leq y \leq 4\}.$$

3.2.4.3.4 Zero de uma função

Zero de uma função real é todo x cuja imagem é nula, isto é:

$$x \text{ é zero de } y = f(x) \iff f(x) = 0.$$

Exemplo 3.27. Dada a função $f(x) = 2x + 1$, temos que $x = -\frac{1}{2}$ é o zero de $f(x)$, pois $f(-\frac{1}{2}) = 2 \cdot (-\frac{1}{2}) + 1 \implies f(-\frac{1}{2}) = 0$.

3.2.4.3.5 Composição de funções

Sejam os conjuntos A , B e C não vazios e as funções $f : A \rightarrow B$ (função f de A em B) e $g : B \rightarrow C$ (função g de B em C), a função f composta com g ($g \circ f$) é função de A em C , em que a imagem de cada x é obtida da seguinte forma:

1º) aplica-se a x a função f de forma a se obter $f(x)$;

2º) aplica-se a $f(x)$ a função g de forma a se obter $g(f(x))$.

Tem-se $g \circ f = g(f(x))$ para todo $x \in R$.

Exemplo 3.28. Sejam os conjuntos $A = \{-1, 0, 1, 2\}$, $B = \{0, 1, 2, 3, 4\}$, $C = \{1, 3, 5, 7, 9\}$ e as funções $f : A \rightarrow B$ definida por $f(x) = x^2$ e $g : B \rightarrow C$ definida por $g(x) = 2x + 1$, vamos obter $(g \circ f)(x)$ e calcular $(g \circ f)(2)$

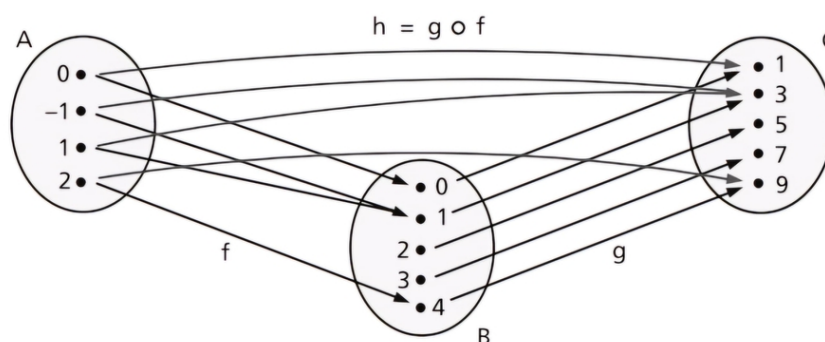
Solução: Aplicando a $f(x)$ a função $g(x)$, temos:

$$g \circ f = g(f(x)) = 2 \cdot x^2 + 1 = 2x^2 + 1$$

$$\text{Assim } (g \circ f)(2) = 2 \cdot 2^2 + 1 = 9$$

Visualizando no diagrama de Venn:

Figura 7 – Diagrama de Venn exemplo 3.28



Fonte: Fundamentos de matemática elementar, 1; conjuntos e funções, pág 213

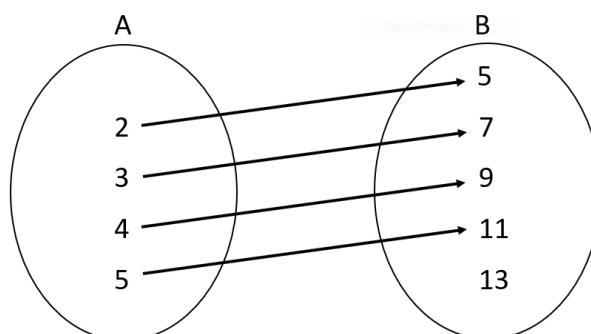
3.2.4.3.6 Função injetora e sobrejetora

Função Injetora é uma função f de A em B se, e somente se, quaisquer que sejam x_1 e x_2 de A , se $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$.

De forma equivalente, uma função f de A em B é injetora se, e somente se, quaisquer que sejam x_1 e x_2 de A , se $f(x_1) = f(x_2) \implies x_1 = x_2$.

Exemplo 3.29. A função f de $A = \{2, 3, 4, 5\}$ em $B = \{5, 7, 8, 9, 11, 13\}$ definida pela lei $f(x) = 2x + 1$ é injetora, pois se tomarmos dois elementos distintos de A têm como imagem dois elementos distintos de B .

Figura 8 – Diagrama de Venn exemplo 3.29



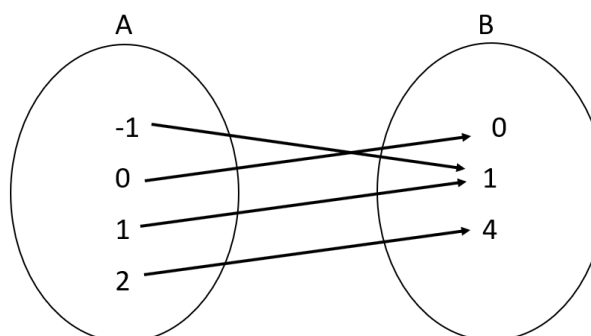
Fonte: elaborado pelo autor

Função sobrejetora é uma função f de A em B se, e somente se, para todo $y \in B$, existe $x \in A$ tal que $f(x) = y$.

Importante notar que $f : A \rightarrow B$ é sobrejetora se, e somente se, $Im(f) = B$.

Exemplo 3.30. A função f de $A = \{-1, 0, 1, 2\}$ em $B = \{0, 1, 4\}$ definida pela lei $f(x) = x^2$ é sobrejetora, pois para todo elemento $y \in B$, existe $x \in A$ tal que $f(x) = y$.

Figura 9 – Diagrama de Venn exemplo 3.30



Fonte: elaborado pelo autor

3.2.4.3.7 Inversão de funções

Definição 3.17 Se f é uma função bijetora de A em B , a relação inversa é uma função de B em A , denominada função inversa de f e indicada por f^{-1} .

Dizemos que f e f^{-1} são inversas entre si, ou seja $(f^{-1})^{-1} = f$.

Exemplo 3.31. Vamos determinar a sentença aberta que define f^{-1} da função f definida por $f(x) = 2x + 1$.

Solução: Primeiro fazemos uma mudança de variável, isto é, trocamos x por y e y por x , assim:

$$x = f(y) \implies x = 2y + 1 \implies y = \frac{x-1}{2}$$

Fazendo $y = f^{-1}$, temos que $f^{-1} = \frac{x-1}{2}$.

Há vários tipos e classificações de funções reais de uma variável abordadas no Ensino Médio, tais como função afim, função quadrática, função exponencial, função logarítmica, funções trigonométricas, etc. É abordado aqui três dessas funções em específico, função afim, função exponencial e funções trigonométricas que possuem grandes aplicações na construção e treinamento de RNAs.

3.2.4.4 Função afim

Definição 3.18 Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = ax + b$, com $a, b \in \mathbb{R}$ e $a \neq 0$, para todo $x \in \mathbb{R}$, é chamada função afim.

a e b são os coeficientes da função afim, e são chamados de coeficiente angular e coeficiente linear, respectivamente.

O gráfico da função afim é uma reta, e o coeficiente angular define o crescimento ou decréscimo da reta, isto é, caso $a > 0$ a reta é crescente e caso $a < 0$ a reta é decrescente.

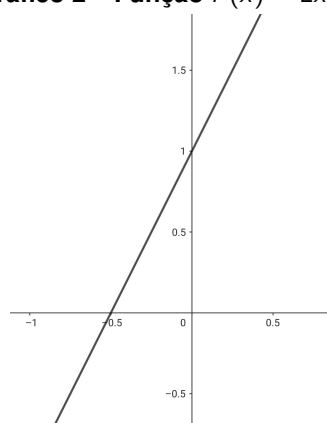
Exemplo 3.32. Dada a função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = 2x + 1$, vamos construir o gráfico da função f .

Solução: Há várias maneiras de construir o gráfico de uma função real, nesse caso específico, como o gráfico é uma reta, basta definirmos dois pontos dessa reta para podermos construí-la.

Fazendo $y = 0$, assim $2x + 1 = 0 \implies x = -\frac{1}{2}$, portanto o primeiro ponto da reta é $(-\frac{1}{2}, 0)$.

Agora fazendo $x = 0$, assim $y = 2 \cdot 0 + 1 \implies y = 1$, o segundo ponto da reta é $(0, 1)$.

Definidos os pontos da reta e esboçando no plano cartesiano, temos:

Gráfico 2 – Função $f(x) = 2x + 1$ 

Fonte: elaborado pelo autor

O zero da função afim é dado por $x = -\frac{b}{a}$, assim, no exemplo 3.32, o zero da função f é $x = -\frac{1}{2}$.

3.2.4.5 Função exponencial

Definição 3.19. Dado o número real a , tal que $0 < a \neq 1$, chama-se função exponencial de base a a função f de $\mathbb{R}$ em $\mathbb{R}$ que associa a cada x real o número a^x .

A função exponencial $f(x) = a^x$ possui as seguintes propriedades:

i. $x = 0 \implies f(0) = a^0 = 1$, isto é, para todo $a \in \mathbb{R}_+^* - \{1\}$, o gráfico cartesiano da função exponencial corta o eixo y no ponto ordenada 1.

ii. A função exponencial $f(x) = a^x$ é crescente se, e somente se, $a > 1$ e decrescente se, e somente se, $0 < a < 1$, isto é:

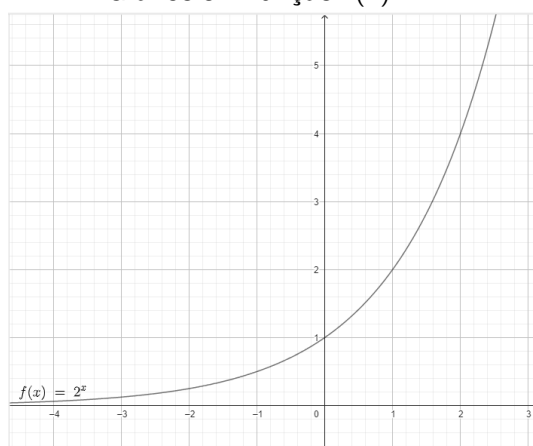
$$\text{quando } a > 1; \text{ tem-se } x_1 < x_2 \implies f(x_1) < f(x_2),$$

e

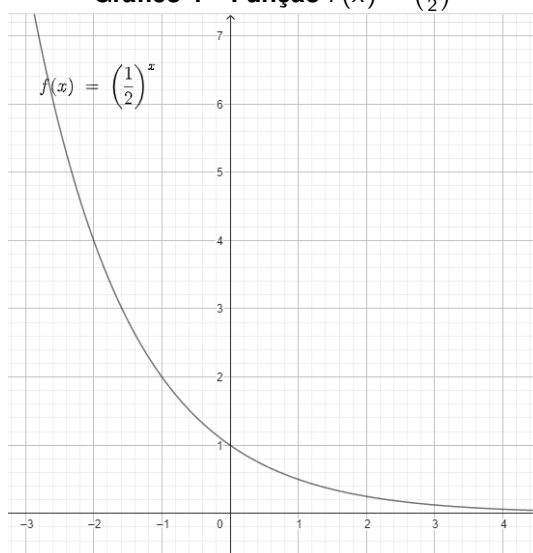
$$\text{quando } 0 < a < 1; \text{ tem-se } x_1 < x_2 \implies f(x_1) > f(x_2)$$

iii. A função exponencial $f(x) = a^x$, com $0 < a \neq 1$, é injetora, pois, dados x_1 e x_2 tais que $x_1 \neq x_2$ tem-se $f(x_1) \neq f(x_2)$.

Exemplo 3.33. Veja abaixo os gráficos da função crescente $f(x) = 2^x$ (**Gráfico 3**) e da função decrescente $f(x) = \left(\frac{1}{2}\right)^x$ (**Gráfico 4**).

Gráfico 3 – Função $f(x) = 2^x$ 

Fonte: elaborado pelo autor

Gráfico 4 – Função $f(x) = \left(\frac{1}{2}\right)^x$ 

Fonte: elaborado pelo autor

3.2.4.6 Funções trigonométricas

As funções trigonométricas fazem parte de uma classe de funções chamadas de funções periódicas. Uma função $f : A \rightarrow B$ é chamada periódica se existir um número $p > 0$ tal que $f(x + p) = f(x)$, $\forall x \in A$.

O gráfico da função periódica se caracteriza por apresentar elementos de curvas que se repetem em um determinado período, e tal período é medido no eixo dos x .

O Círculo Trigonométrico, também chamado de Ciclo ou Circunferência Trigonométrica, é uma representação gráfica que auxilia no cálculo das razões trigonométricas.

De acordo com a simetria do círculo trigonométrico temos que o eixo vertical corresponde ao seno e o eixo horizontal ao cosseno. Cada ponto dele está associado aos valores dos ângulos. Radianos do Círculo Trigonométrico

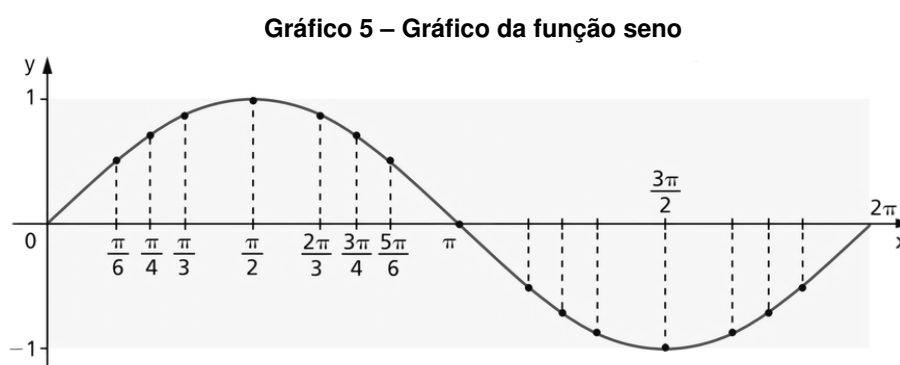
Focaremos em dois tipos de funções trigonométricas, a função seno e a função cosseno, já que essas curvas são usadas como exemplo de aprendizagem de máquina mais à frente neste trabalho.

3.2.4.6.1 Função seno

Definição 3.20. A função seno é uma função periódica e seu período é 2π .

A função seno $f : \mathbb{R} \rightarrow \mathbb{R}$, é expressa por $f(x) = \sin x$.

O gráfico da função seno é denominado senoide, e indica como varia a função $f(x) = \sin x$.



Fonte: Fundamentos de matemática elementar, 3: trigonometria / Gerson lezzi; 9. ed., página 94

Importante notar que a imagem da função seno é o intervalo $[-1, 1]$, isto, $-1 < \sin x < 1$ e seu período é 2π .

A senoide continua para a direita de 2π e para a esquerda de 0, já que o domínio da função seno é $\mathbb{R}$.

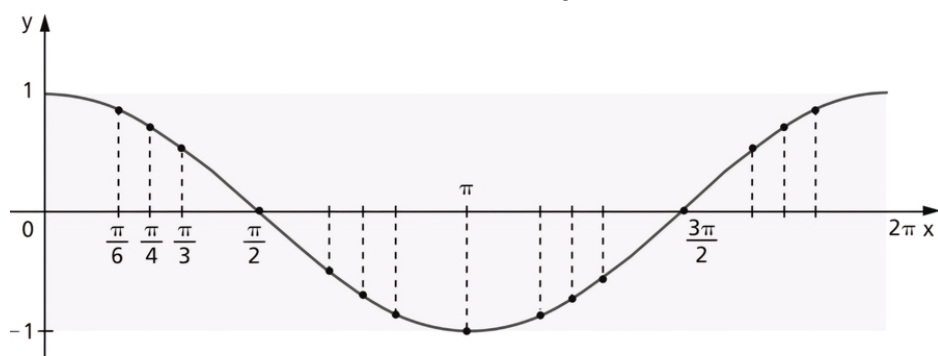
3.2.4.6.2 Função cosseno

Definição 3.21. A função cosseno é uma função periódica e seu período é 2π .

A função cosseno $f : \mathbb{R} \rightarrow \mathbb{R}$, é expressa por $f(x) = \cos x$.

O gráfico da função cosseno é denominado cossenoide, e indica como varia a função $f(x) = \cos x$.

Gráfico 6 – Gráfico da função cosseno



Fonte: Fundamentos de matemática elementar, 3: trigonometria / Gerson Iezzi; 9. ed., página 104

Assim como na função seno a imagem da função cosseno é o intervalo $[-1, 1]$, isto, $-1 < \cos x < 1$.

A cossenoide continua para a direita de 2π e para a esquerda de 0 , já que o domínio da função cosseno é $\mathbb{R}$.

4 REDE NEURAL ARTIFICIAL

4.1 Histórias das RNAs

As redes neurais artificiais, representações matemáticas que buscam replicar a estrutura fisiológica do cérebro humano, têm sua origem nos esforços pioneiros do neurofisiologista Warren McCulloch e do matemático Walter Pitts na década de 1940. Eles conceberam um modelo computacional simplificado de neurônios artificiais, delineando a primeira arquitetura de rede neural sem capacidade de aprendizado. Inicialmente, as limitações computacionais da época impediam avanços significativos (Santos; Farias, 2019).

Na década de 1980, novas arquiteturas e técnicas de treinamento foram introduzidas, reacendendo o interesse nas redes neurais. No entanto, o progresso era lento, e nos anos 90, o surgimento de técnicas de aprendizado de máquina, como o *Support Vector Machine* (SVM), temporariamente relegou as redes neurais a um segundo plano. A virada ocorreu nos anos 2000, com o aumento do poder computacional e a disponibilidade massiva de dados, propiciando o ressurgimento e avanço do *Deep Learning* (Marques; Von Wangenheim; Hauck, 2020).

RNAs são compostas por unidades de processamento chamadas neurônios, conectadas em camadas e altamente interligadas. Essas redes se destacam em diversas aplicações, desde reconhecimento facial e classificação de imagens até jogos complexos como xadrez. Neste contexto, o neurônio artificial corresponde a uma versão computacional do neurônio biológico, composto por dendritos, corpo celular e axônio (Topin, 2020). Os dendritos recebem estímulos nervosos, transmitindo-os ao corpo celular, que processa as informações e gera um novo impulso. Esse impulso é transportado pelo axônio até outros neurônios, ocorrendo nas sinapses, onde o sinal flui de neurônio para neurônio (Filipakis et al., 2021).

Na construção de RNAs, os neurônios artificiais são representados em diagramas, cada um com entradas ponderadas por pesos sinápticos e um termo de bias. A saída do neurônio é obtida aplicando uma função de ativação ao campo local induzido, resultado da soma ponderada das entradas (Batista; Martins, 2022). Segundo Colle, Rodrigues e Silva (2022), as redes neurais, com sua origem histórica e estrutura complexa inspirada no cérebro humano, se tornaram uma ferramenta poderosa no campo do aprendizado de máquina, impulsionando avanços notáveis nas capacidades computacionais e na compreensão de padrões complexos. Também esta é a compreensão de Belter (2023), segundo quem as redes neurais artificiais, conhecidas por sua capacidade distintiva de reconhecimento de padrões através de processos de aprendizagem, têm sido amplamente empregadas em diversas áreas para resolver uma variedade de problemas.

Para Santos, Sant'ana e Sant'ana (2023), seja na engenharia química, onde modelam e otimizam processos e equipamentos, ou na classificação de alimentos para garantir

padrões de qualidade, as redes neurais têm demonstrado sua versatilidade e eficácia. Contudo, conforme Leal Junior, Andrade e Barros (2022), sua aplicação no contexto educacional vai além das tradicionais áreas de atuação. Em um projeto inovador, os autores exploraram o potencial das redes neurais como ferramenta de Aprendizagem Matemática no Ensino Básico. O foco foi direcionado para a observação do comportamento de funções estudadas no ensino médio, com o intuito de proporcionar uma abordagem dinâmica e interativa para os alunos.

A observação minuciosa das simulações das funções matemáticas evidenciou a eficiência das redes neurais em diferentes faixas de entrada. Esse enfoque permitiu identificar situações em que a rede apresentou desempenho otimizado na representação correta da função simulada. Além disso, tornou-se possível delinear cenários nos quais a rede poderia não representar de maneira precisa a função em questão (Leal Junior; Andrade; Barros, 2022). Diante disso, destacaram que o projeto não apenas trouxe uma abordagem inovadora ao Ensino de Matemática, mas também ofereceu uma oportunidade única de compreender as nuances do funcionamento das redes neurais. Com isso, constataram que, ao incorporar as redes neurais no Ensino Básico, não apenas expandimos as possibilidades de aprendizado matemático, mas também exploramos novas fronteiras educacionais, capacitando os alunos com habilidades que transcendem os limites do conhecimento teórico, promovendo uma educação mais dinâmica e aplicada (Leal Junior; Andrade; Barros, 2022).

4.2 Neurônio artificial

Um neurônio artificial, usado em redes neurais, é uma unidade de processamento inspirada em um neurônio biológico, imitando seu funcionamento. O neurônio artificial recebe entradas, realiza cálculos ponderados e gera uma saída com base em uma função de ativação. O perceptron, um tipo específico de neurônio artificial, foi proposto por Frank Rosenblatt em 1957 em seus trabalhos *The Perceptron: A Perceiving And Recognizing Automaton* e *The Perceptron: A Probabilistic Model For Information Storage And Organization in The Brain*. Ele é um modelo simplificado de um neurônio biológico e foi uma das primeiras formas de rede neural desenvolvidas. O perceptron pode tomar decisões simples e consegue aprender ao ser alimentado com dados.

A concepção do perceptron foi influenciada pela compreensão dos processos biológicos do cérebro e pela ideia de criar sistemas que pudessem imitar a capacidade humana de aprendizado. A criação do perceptron marcou o início do estudo e desenvolvimento de redes neurais artificiais e teve um impacto significativo no campo da inteligência artificial.

Esses avanços históricos pavimentaram o caminho para o desenvolvimento de redes neurais mais complexas, levando a aplicações em diversas áreas, como reconhecimento de padrões visuais, processamento de linguagem natural, visão computacional, entre outros.

4.2.1 Funcionamento do neurônio biológico

Um neurônio biológico é uma célula especializada do sistema nervoso responsável por transmitir informações por meio de sinais elétricos e químicos. Os neurônios processam e transmitem informações entre si, permitindo a comunicação entre diferentes partes do corpo e o cérebro. O neurônio possui três partes principais: dendritos, corpo celular e axônio.

Os dendritos são ramificações curtas que se estendem a partir do corpo celular e recebem sinais de outros neurônios ou de células sensoriais. Eles transmitem esses sinais para o corpo celular, ou soma, que é a parte principal do neurônio e onde se encontra o núcleo celular. O núcleo celular abriga o material genético e controla as funções celulares.

O funcionamento do neurônio envolve a transmissão de sinais elétricos ao longo do seu comprimento. Quando um estímulo é recebido pelos dendritos, ele pode desencadear uma mudança no potencial elétrico da célula. Se essa mudança atinge um limiar crítico, um impulso elétrico, chamado potencial de ação neuronal, é gerado e viaja ao longo do axônio até alcançar as sinapses (conexões) com outros neurônios.

O axônio é uma única fibra longa que transmite sinais elétricos a partir do corpo celular em direção aos terminais axonais, extremidades dos axônios, e se comunicam com dendritos de outros neurônios por meio de sinapses (conexões especializadas entre os neurônios no sistema nervoso). As sinapses permitem a transmissão de sinais elétricos e químicos entre os neurônios.

A descarga elétrica no neurônio é mediada por íons, como sódio e potássio, que atravessam a membrana celular em resposta a mudanças no potencial elétrico. Quando um estímulo ativa os dendritos, canais iônicos na membrana celular abrem-se, permitindo a entrada de íons carregados positivamente. Isso gera uma inversão temporária no potencial elétrico da célula, conhecida como despolarização. Essa mudança desencadeia a abertura de mais canais iônicos ao longo do axônio, propagando o potencial de ação.

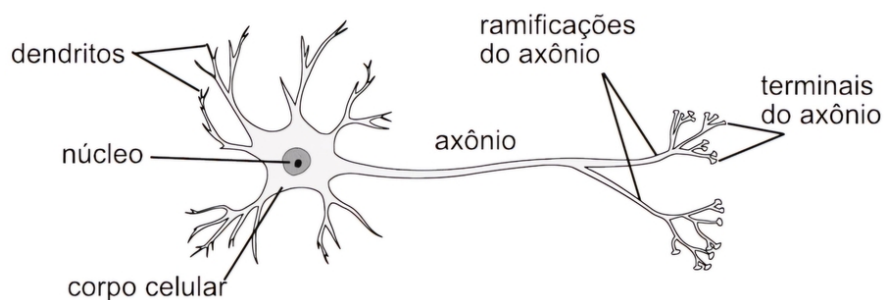
Ao alcançar as sinapses, o potencial de ação desencadeia a liberação de neurotransmissores, substâncias químicas que transmitem o sinal para o próximo neurônio ou célula alvo. Essa combinação de sinais elétricos e químicos permite a comunicação eficaz entre os neurônios e é essencial para funções cerebrais como pensamento, movimento e percepção.

Além dessas partes principais de um neurônio biológico, há também os nódulos de Ranvier, lacunas ao longo do axônio onde a bainha de mielina está ausente, que facilitam a transmissão rápida do impulso nervoso ao longo do axônio, além da bainha de mielina, que é uma camada isolante que cobre muitos axônios, acelerando a transmissão dos impulsos nervosos.

Essas partes estruturais, e muitas outras não especificadas aqui, trabalham juntas para permitir a comunicação eficiente entre os neurônios e desempenham um papel vital no

funcionamento do sistema nervoso geral.

Figura 10 – Estrutura de um neurônio biológico



Fonte: <https://www.goconqr.com/flashcard/14604740/tecido-nervoso>

4.2.2 Perceptron

O perceptron é um modelo de neurônio artificial proposto por Frank Rosenblatt em 1957. Ele é a unidade fundamental de processamento em RNA de uma única camada. O perceptron recebe entradas, aplica pesos a essas entradas e, em seguida, passa o resultado por uma função de ativação para produzir uma saída. “O perceptron é projetado para ilustrar algumas das propriedades fundamentais dos sistemas inteligentes em geral, sem se envolver muito profundamente nas condições especiais e frequentemente desconhecidas que se aplicam a organismos biológicos específicos” (ROSENBLATT, 1958, p. 387).

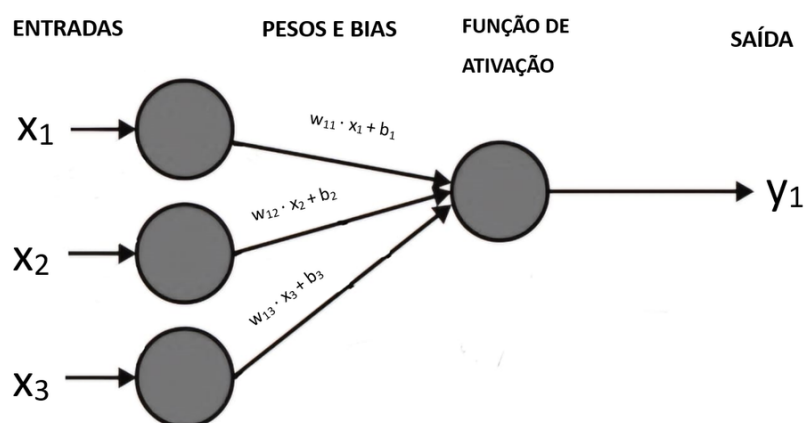
O perceptron é composto por três partes principais: entradas, pesos e função de ativação. As entradas (x_m) são os valores que entram na RNA. Os pesos (w_{jm}) são os parâmetros ajustáveis que a rede aprende durante o treinamento e determinam a importância relativa de cada entrada. A função de ativação (φ) é responsável por introduzir não-linearidade na saída do perceptron. Ela decide se o perceptron deve ser ativado ou não com base na soma ponderada das entradas e pesos (será explicado mais à frente neste trabalho).

O processo de treinamento do perceptron envolve ajustar os pesos das entradas de acordo com os erros de saída obtidos durante o treinamento. Isso é feito usando algoritmos de aprendizado supervisionado, como o *backpropagation*, um algoritmo utilizado em RNA para treiná-las e funciona calculando o gradiente da função de perda em relação aos pesos da rede, e então usando esse gradiente para atualizar os pesos e minimizar a função de perda. Em resumo, o *backpropagation* é responsável por ajustar os pesos da rede neural

durante o treinamento, permitindo que ela aprenda a mapear corretamente as entradas para as saídas desejadas.

O perceptron é capaz de realizar classificação linear simples, ou seja, separar dados em duas classes usando uma linha reta. No entanto, ele tem limitações, como a incapacidade de lidar com problemas não linearmente separáveis.

Figura 11 – Estrutura de um Perceptron de camada única com 3 entradas

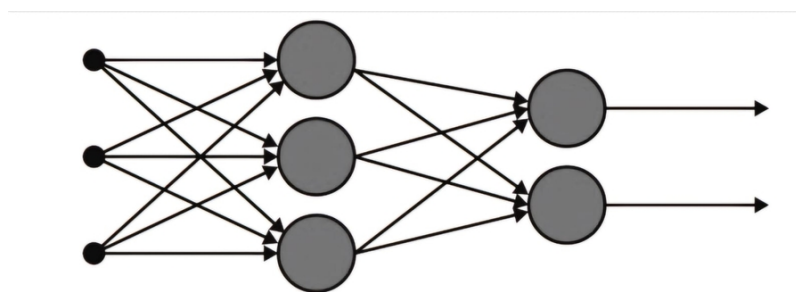


Fonte: elaborado pelo autor

4.2.3 Funcionamento de um perceptron

A rede é composta por camadas, sendo a primeira camada a camada de entrada, a última camada a camada de saída e as camadas intermediárias conhecidas como camadas ocultas. Cada camada contém vários neurônios artificiais que processam os dados em paralelo.

Figura 12 – Exemplo RNA com 5 neurônios



Fonte: elaborado pelo autor

Na **Figura 12** temos um exemplo de um perceptron com três camadas e cinco neurônios, sendo a primeira camada com três entradas, a segunda camada (a única

camada oculta da rede) com três neurônios, e a última camada com dois neurônios é a camada de saída.

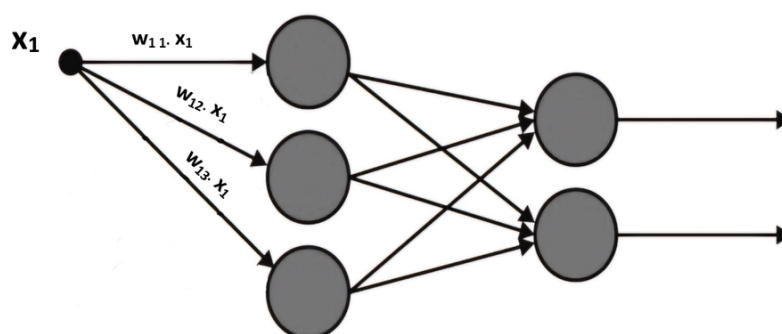
4.2.3.1 Pesos e bias

O funcionamento de um neurônio artificial começa com o processo de *feedforward*, onde os dados de entrada são multiplicados pelos pesos sinápticos e somados. Essa soma é então modificada pela adição do *bias*, que permite ao neurônio ajustar o ponto de ativação. Em seguida, a função de ativação é aplicada ao resultado da soma ponderada, gerando o sinal de saída do neurônio.

O *bias* é um parâmetro aditivo que permite ajustar o ponto de ativação de um neurônio. Ele é adicionado à combinação linear dos sinais de entrada ponderados pelos pesos sinápticos, antes que o resultado seja processado pela função de ativação. Em termos mais simples, o bias é como um ajuste fino que a rede neural pode fazer e considera certas influências ou preconceitos nos dados. Ele ajuda a modelar o comportamento do neurônio e é fundamental para a capacidade da rede neural de aprender e generalizar a partir dos dados e é um parâmetro importante que contribui para a flexibilidade e adaptabilidade dos neurônios artificiais em redes neurais, permitindo que eles capturem padrões complexos nos dados.

Tomando a rede da **Figura 12**, no primeiro nó, a entrada x_1 é multiplicado pelos pesos w_{11} , w_{12} e w_{13} , respectivamente, e na camada oculta os resultados são somados entre si e adicionado o bias b_1 . Chamaremos esse resultado de z_1 , que nos dá $z_1 = w_{11}x_1 + w_{12}x_1 + w_{13}x_1 + b_1$.

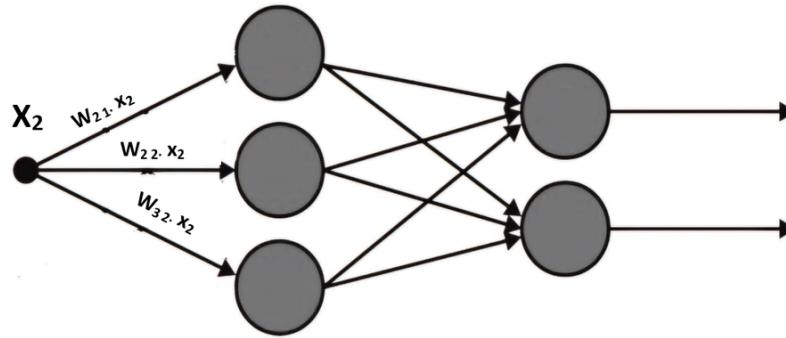
Figura 13 – Primeiro nó da primeira camada oculta



Fonte: elaborado pelo autor

De forma análoga, no segundo nó a entrada x_2 é multiplicada pelos pesos w_{21} , w_{22} e w_{23} . Na camada oculta são somados entre si e adicionado o bias b_2 , nos dando o valor de z_2 , assim $z_2 = w_{21}x_2 + w_{22}x_2 + w_{23}x_2 + b_2$.

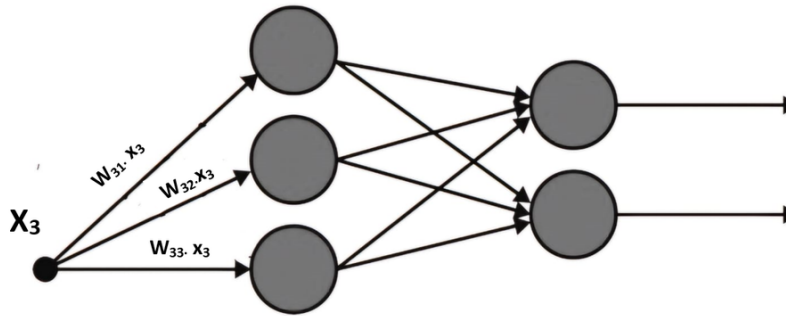
Figura 14 – Segundo nó da primeira camada oculta



Fonte: elaborado pelo autor

No terceiro nó a entrada x_3 é multiplicado pelos pesos w_{31} , w_{32} e w_{33} , na camada oculta são somados entre si e adicionado o bias b_3 , nos dando o valor de z_3 , assim $z_3 = w_{31}x_3 + w_{32}x_3 + w_{33}x_3 + b_3$.

Figura 15 – Terceiro nó da primeira camada oculta



Fonte: elaborado pelo autor

Expressando essas três expressões na forma matricial, temos:

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} w_{11}x_1 + w_{12}x_1 + w_{13}x_1 + b_1 \\ w_{21}x_2 + w_{22}x_2 + w_{23}x_2 + b_2 \\ w_{31}x_3 + w_{32}x_3 + w_{33}x_3 + b_3 \end{bmatrix} = \begin{bmatrix} w_{11}x_1 + w_{12}x_1 + w_{13}x_1 \\ w_{21}x_2 + w_{22}x_2 + w_{23}x_2 \\ w_{31}x_3 + w_{32}x_3 + w_{33}x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} =$$

$$= \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Observe que esse processo se resume em multiplicação e soma de matrizes, isto é:

$$Z = WX + B$$

$$\text{onde } Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}, W = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ e } B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

4.2.3.2 Função de ativação

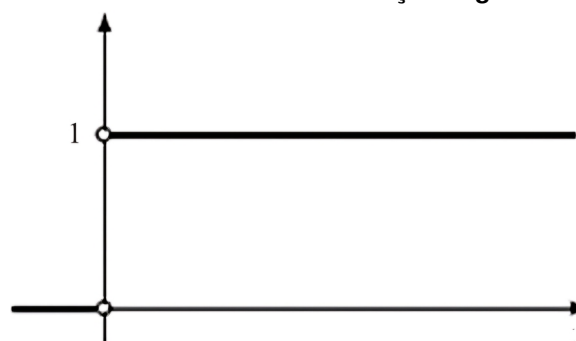
A função de ativação é responsável por introduzir não-linearidade na saída do neurônio, permitindo que a rede aprenda relações complexas nos dados. Exemplos comuns de funções de ativação incluem a Função Degrau, a Função ReLU, a Função Sigmoid e a Função Tangente Hiperbólica.

A **Função Degrau** ou Função de Heaviside é uma função que assume o valor unitário para argumento positivo e assume um valor nulo para argumento negativo, isto é:

$$\varphi(x) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x < 0 \end{cases}$$

A função degrau tem a vantagem de ser simples e é utilizada principalmente em processos com uma resposta binária, inserindo a não-linearidade nos modelos e permitindo que a rede aprenda padrões mais complexos.

Gráfico 7 – Gráfico da função degrau

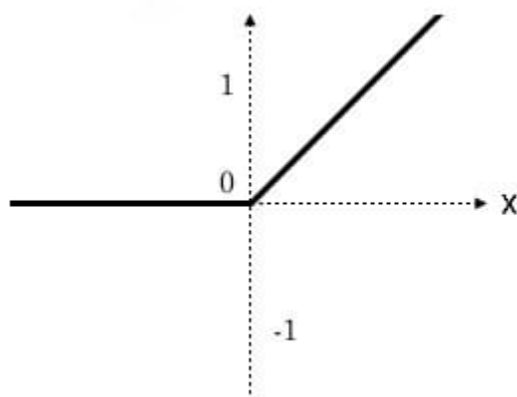


Fonte: elaborado pelo autor

A **Função ReLU**, ou Unidade Linear Retificada, é definida como $\varphi(x) = \max(0, x)$, o que significa que ela retorna o valor de x se $x > 0$, e retorna zero se $x < 0$. Em outras palavras, a função ReLU ativa o neurônio se o sinal de entrada for positivo, e o desativa se for negativo.

$$\text{Ou seja } \varphi(x) = \begin{cases} x & \text{se } x > 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Gráfico 8 – Gráfico da Função ReLU

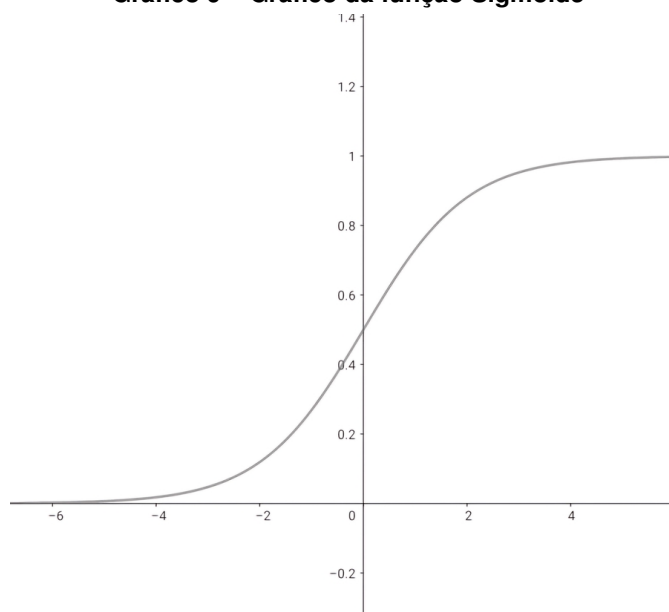


Fonte: elaborado pelo autor

Uma das vantagens da função ReLU é a sua simplicidade computacional, o que a torna eficiente em termos de processamento. Além disso, ela ajuda a mitigar o problema do desaparecimento do gradiente durante o treinamento da RNA, facilitando a convergência do algoritmo de *backpropagation*.

A **Função Sigmoid** é uma função matemática que mapeia qualquer valor real para um valor no intervalo de 0 a 1. Sua estrutura é definida pela fórmula matemática $\varphi(x) = \frac{1}{1+e^{-x}}$, onde “e” representa a base do logaritmo natural e “x” é o valor de entrada. A curva resultante da função sigmoide tem uma forma de “S” invertido, com um ponto de inflexão em $x = 0$. Isso significa que valores muito grandes ou muito pequenos de x resultarão em valores próximos de 1 ou 0, respectivamente, enquanto valores próximos de zero resultarão em um valor próximo de 0,5.

Gráfico 9 – Gráfico da função Sigmoid



Fonte: elaborado pelo autor

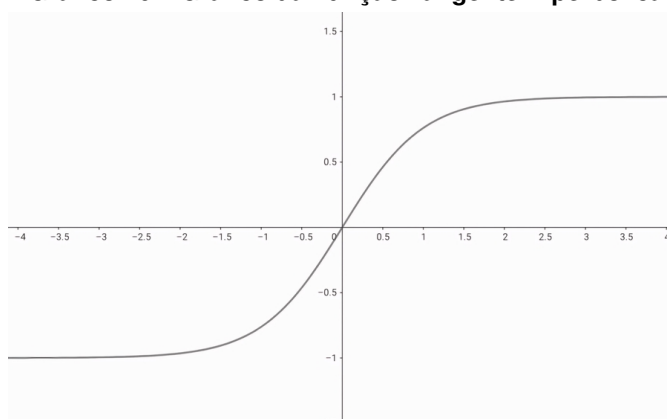
A função sigmoide é diferenciável em todo o seu domínio, o que a torna adequada para o uso em algoritmos de otimização baseados em gradiente, como o algoritmo de *backpropagation*.

A função sigmoide desempenhou um papel importante no desenvolvimento inicial de redes neurais artificiais e ainda é útil em certos contextos, mas foi amplamente substituída por funções de ativação mais eficazes em muitas aplicações modernas de redes neurais, como a função ReLU.

A **Função Tangente Hiperbólica**, denotada como $\tanh(x)$, é uma função matemática que mapeia qualquer número real para um valor no intervalo $[-1, 1]$. Ela é definida pela fórmula $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ onde “e” é a base do logaritmo natural.

Em redes neurais, a função $\tanh(x)$, é comumente usada como uma função de ativação em camadas ocultas. A vantagem da função $\tanh(x)$ em comparação com outras funções de ativação, como a função sigmoide, é que ela produz valores no intervalo $[-1, 1]$, ajudando a reduzir problemas de saturação dos gradientes durante o treinamento da RNA.

Gráfico 10 – Gráfico da Função Tangente Hiperbólica



Fonte: elaborado pelo autor

A função tangente hiperbólica é uma ferramenta fundamental em redes neurais artificiais, proporcionando não-linearidade suave e estabilidade durante o treinamento. Sua capacidade de normalizar e regularizar as saídas dos neurônios a torna uma escolha popular em arquiteturas de redes neurais profundas.

Cabe ao usuário decidir qual função de ativação será usada para inserir a não-linearidade na rede, e diferentes nós podem ter diferentes funções de ativação, mas que, na prática, não é usual.

Ainda utilizando a RNA de 5 neurônios, a função de ativação pode ser aplicada no somatório dos valores na segunda camada (camada escondida com 3 neurônios), e seguir para a camada de saída, multiplicado por novos pesos e somados a novos bias, onde pode ser aplicada outra ou a mesma função de ativação, dependendo do tipo do problema abordado.

4.2.3.3 Função de custo

Para que a rede aprenda, é preciso que ela passe por um processo de regressão (*backpropagation*) com o intuito de identificar o quão longe está da solução ideal. Assim a rede consegue mudar os valores dos pesos w_{jm} e do bias, através da função de custo (*cost function*), o processo envolve a propagação do erro da saída da rede para as camadas anteriores, permitindo que cada neurônio contribua para o erro total de acordo com seu peso na rede. Os pesos e bias são então ajustados conforme o valor do erro.

Existem vários tipos de função de custo, mas vamos usar como exemplo função MSE (*Mean Squared Error*), definida da seguinte forma:

$$MSE = \sum_{i=1}^n \left(\frac{(y_i - \hat{y}_i)^2}{n} \right)$$

onde n é número total de observações, y_i os valores esperados e $\hat{y}_i$ os valores previstos pela rede.

Apesar de parecer ser uma função complicada de se entender, alunos do Ensino Médio estão familiarizados com seu gráfico, pois se trata de uma função quadrática.



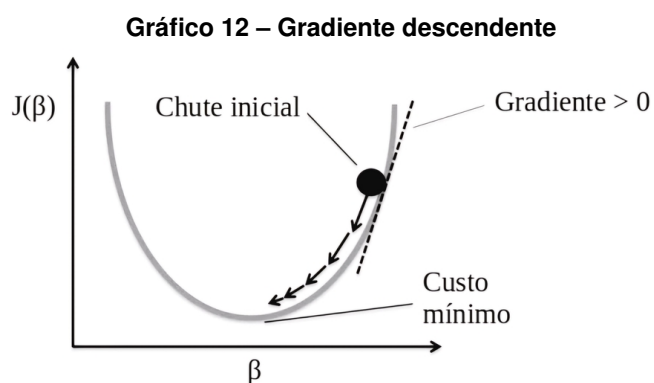
Fonte: <https://medium.com/@DilaneKombou/presentation-of-mse-mean-squared-error-a4ee9b6cff49>

O erro calculado pela rede será um dos pontos desse gráfico, onde o ponto de mínimo (o ponto em destaque no gráfico) define o menor erro possível do problema em questão, e a rede sempre corrigirá os valores dos pesos e dos *bias* conforme o valor do erro.

Quanto maior o erro, maior será a correção dos pesos e *bias*, e esse processo se dá usando uma técnica de otimização, como o gradiente descendente, para minimizar a função de custo.

O uso do gradiente descendente envolve o cálculo de derivadas parciais, tema que foge do escopo deste trabalho, mas em termos simples, o cálculo do gradiente permite descrever o declive de uma função, auxiliando a rede a identificar a região global de menor perda.

A função de custo mostra a direção que a rede deve seguir para que o erro seja minimizado, se aproximando assim da solução ideal para o problema.



Fonte: http://cursos.leg.ufpr.br/ML4all/slides/Gradiente_descendente.pdf

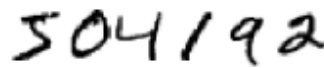
Em resumo, o funcionamento e treinamento de um neurônio artificial envolvem o processamento dos dados por meio do *feedforward*, a aplicação da função de ativação, a organização em camadas na RNA e o ajuste dos parâmetros por meio do *backpropagation*. Esses conceitos fundamentais são essenciais para compreender como as RNA aprendem e realizam tarefas complexas.

5 USO DE RNAs NO ENSINO BÁSICO

O reconhecimento de dígitos manuscritos é um processo complexo que envolve várias áreas do cérebro humano. Embora o cérebro seja altamente adaptado para interpretar o mundo visual, essa tarefa é realizada inconscientemente, dificultando apreciar sua complexidade. Ao tentar programar um computador para reconhecer dígitos manuscritos, a simplicidade percebida se transforma em um desafio. Regras intuitivas sobre formas revelam-se complicadas e cheias de exceções, demonstrando a dificuldade intrínseca do reconhecimento de padrões visuais.

Considere a sequência de dígitos manuscritos abaixo:

Figura 16 – Dígitos manuscritos



Fonte: <http://neuralnetworksanddeeplearning.com/chap1.html>

A maioria das pessoas consegue identificar que se trata da sequência de dígitos 504192, mas para a RNA se trata de um problema complexo que demanda um esforço computacional alto em treinamento e correções, até que a rede tenha um desempenho aceitável em identificá-los.

Esse exemplo serve para mostrar que uma tarefa que aparentemente é simples para o cérebro humano pode tomar um rumo diferente quando é tratado em um cérebro “digital”, seria necessário uma RNA de multicamadas, com muitos neurônios em cada uma delas, a chamada Deep Learning (Aprendizagem Profunda). Por esse motivo são abordados exemplos mais simples neste trabalho, e que podem ser aplicados em turmas do Ensino Básico nas aulas de Matemática.

Nesta seção são usados conceitos de construção e treinamento de uma rede neural para interagir com problemas Matemáticos do Ensino Básico, e assim visualizar como uma RNA aprende.

Os códigos de programação dos exemplos dados não são abordados nesta seção, contudo a programação é feita em Linguagem Python no Google Colab e os códigos estão disponíveis em anexo para que o leitor implemente-os, caso necessite.

5.1 RNA para tomada de decisão simples

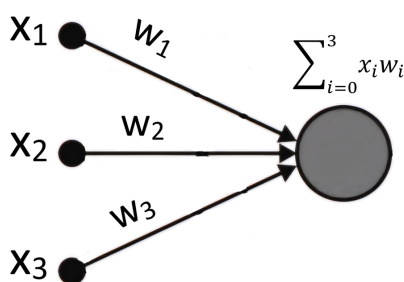
No primeiro exemplo é usado um perceptron de camada única para tomar uma decisão binária (sim ou não). Vale ressaltar que neste exemplo a rede terá apenas o uso de valores de entradas x_i , pesos w_i e um valor de saída binário $\varphi(x_t)$. Não é usado *bias* e a função de ativação é a função degrau.

Segue o problema: Matheus é um aluno da 1ª série do Ensino Médio de uma Escola de Manaus. Seu professor de Matemática organizou um evento fora da Escola e irá atribuir uma nota extra para o aluno que participar. Matheus deverá tomar a decisão se participará do evento ou não, baseado em três fatores:

- i. fator “nota” (ele precisa de nota);
- ii. fator “amigos” (seus amigos irão ao evento);
- iii. fator “distância” (o evento é perto da casa de Matheus).

A rede perceptron de camada única tem a seguinte estrutura:

Figura 17 – Exemplo 1: estrutura do Perceptron de camada única



Fonte: elaborado pelo autor

A entrada x_1 é o fator “nota”, a entrada x_2 é o fator “amigos” e a entrada x_3 é o fator “distância”, e cada um desses fatores assumem valores 0 ou 1. Quando a entrada recebe 0, significa que o fator está “desligado”, ou seja, não será considerado e não contribui com a probabilidade de Matheus participar do evento. Quando recebe valor 1, significa que o fator está “ligado” sendo considerado e contribuindo significativamente com a probabilidade de Matheus participar do evento.

w_1 é o peso atribuído ao fator “nota”, w_2 o peso atribuído ao fator “amigos” e w_3 o peso atribuído ao fator “distância”. Os pesos w_1 , w_2 e w_3 assumem valores de 0 a 8. Quanto maior o valor do peso, maior a probabilidade de Matheus participar do evento do professor de Matemática.

Após atribuídos os valores das entradas (x_1 , x_2 e x_3) e dos pesos (w_1 , w_2 e w_3), é feita a soma ponderada $x_t = x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3$ e o resultado segue para a função de ativação.

A função de ativação utilizada nesse exemplo é a função degrau, dada por:

$$\varphi(x_t) = \begin{cases} 1 & \text{se } x_t \geq \text{limiar} \\ 0 & \text{se } x_t < \text{limiar} \end{cases}.$$

O “limiar” é um valor determinado pelo perceptron e varia de 0 a 11, ou seja, caso a soma $x_t = x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3$ assuma valor igual ou maior que o limiar, a saída nos dá o valor 1, o que significa que Matheus irá ao evento do professor de Matemática, e caso a soma $x_t = x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3$ assuma valor menor que o limiar, a saída nos dá o valor 0,

o que significa que Matheus não participará do evento do professor de Matemática. Quanto maior for o valor do limiar, mais rígida será a rede na tomada de decisão e a probabilidade de Matheus participar do evento é menor.

O código para implementação da rede deste exemplo está em anexo, mas em palavras simples, são dados valores aleatório para as entradas x_1 , x_2 e x_3 de 0 ou 1, sendo 0 significando “fator desligado” e 1 significando “fator ligado”.

Também são dados valores aleatórios aos pesos w_1 , w_2 e w_3 variando de 0 a 8. Temos então:

i. Para o fator “nota”

Se $0 \leq w_1 < 3$, significa que Matheus não se motiva a participar do evento pela nota.

Se $3 \leq w_1 < 6$, significa que Matheus se motiva um pouco a participar do evento pela nota.

Se $6 \leq w_1 < 9$, significa que Matheus se motiva muito a participar do evento pela nota.

ii. Para o fator “amigos”

Se $0 \leq w_2 < 3$, significa que Matheus não se motiva a participar do evento mesmo se seus amigos participarem.

Se $3 \leq w_2 < 6$, significa que Matheus se motiva um pouco a participar do evento caso seus amigos participem.

Se $6 \leq w_2 < 9$, significa que Matheus se motiva muito a participar do evento caso seus amigos participem.

ii. Para o fator “distância”

Se $0 \leq w_3 < 3$, significa que Matheus não se motiva a participar do evento mesmo se o evento for perto de sua casa.

Se $3 \leq w_3 < 6$, significa que Matheus se motiva um pouco a participar do evento caso o evento seja perto de sua casa.

Se $6 \leq w_3 < 9$, significa que Matheus se motiva muito a participar do evento caso o evento seja perto de sua casa.

É dado um valor aleatório para o limiar, variando de 0 a 11, sendo:

Se $0 \leq \text{limiar} < 4$, o limiar escolhido está frouxo e a probabilidade de Matheus ir ao evento é alta.

Se $4 \leq \text{limiar} < 8$, o limiar escolhido está flexível e a probabilidade de Matheus ir ao evento é média.

Se $8 \leq \text{limiar} < 12$, o limiar escolhido está rigoroso e a probabilidade de Matheus ir ao evento é baixa.

Determinando todos os parâmetros e implementado o código, o programa pode rodar e o perceptron toma a decisão se Matheus irá participar ou não do evento.

Importante salientar que esta rede perceptron não “aprende” pois não realiza a

retropropagação para ajuste de seus parâmetros, todos os valores são gerados de forma aleatória e o objetivo é apenas compreender a estrutura e funcionamento de uma RNA de camada única. Exemplos de redes mais complexas com retropropagação e que realmente “aprendem” serão dados mais a frente.

Tomando como exemplo as entradas

$x_1 = 1$ (o fator “nota” está ligado),;

$x_2 = 0$ (o fator “amigos” está desligado) e

$x_3 = 1$ (o fator “distância” está ligado).

Tomando também os pesos

$w_1 = 6$ (Matheus se motiva muito a participar do evento pela nota);

$w_2 = 5$ (Matheus se motiva um pouco a participar do evento caso seus amigos participem) e

$w_3 = 2$ (Matheus não se motiva a participar do evento mesmo se o evento for perto de sua casa).

Considerando $limiar = 9$ (limiar está rigoroso e a probabilidade de Matheus ir ao evento do é baixa), tem-se $x_t = 1 \cdot 6 + 0 \cdot 5 + 1 \cdot 2 \implies x_t = 8$, e como $x_t < limiar = 9$, então $\varphi(x_t) = 0$. Portanto Matheus não participará do evento.

Como são gerados valores aleatórios pela rede, a cada vez que o programa é rodado valores diferentes das entradas, pesos e limiar são gerados e o resultado pode ser diferente a cada rodada.

5.2 RNA para representação de uma reta no plano.

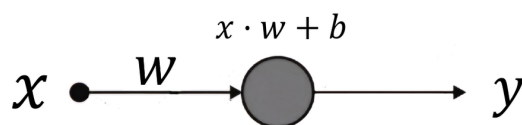
No segundo exemplo é construída e treinada uma RNA em linguagem de programação Python para identificar e prever os pontos da reta definida por $f(x) = 2x + 1$. Vale salientar que, apesar de se saber que $f(x)$ é representada por uma reta, a rede não sabe, e irá adequar seus parâmetros até que encontre uma solução ideal para o problema.

Os códigos para a construção da RNA para este problema, feita na plataforma Google Colab, está disponível em anexo para ser implementada pelo leitor, caso necessite.

Primeiro é criada a estrutura da rede, que neste caso se chama Line Network. Como irá trabalhar com representação de uma reta, apenas um neurônio é o suficiente, pois a entrada x será multiplicada pelo peso w e adicionado ao *bias* b , obtendo a expressão $x \cdot w + b$, ou seja, a rede precisa apenas ajustar os valores de w e b para encontrar uma solução ideal.

A rede para este problema tem a seguinte estrutura:

Figura 18 – Exemplo 2: estrutura do perceptron



Fonte: elaborado pelo autor

Não será preciso usar nenhuma função de ativação neste exemplo, pois se trata de uma função linear.

Com a estrutura da rede neural criada, segue-se para a preparação dos dados de treinamento, para isso usa-se a classe *Dataset* dentro da biblioteca *torch* no Google Colab.

Para os dados de treinamento, escolhemos o conjunto de pares ordenados da reta $f(x) = 2x + 1$ no intervalo $[-10, 10]$, criando 1000 amostras (pontos da reta) como dados de treinamento da rede. Cada dado é um par ordenado.

Feita a preparação dos dados de treinamento, segue-se para a parte onde a rede aprende.

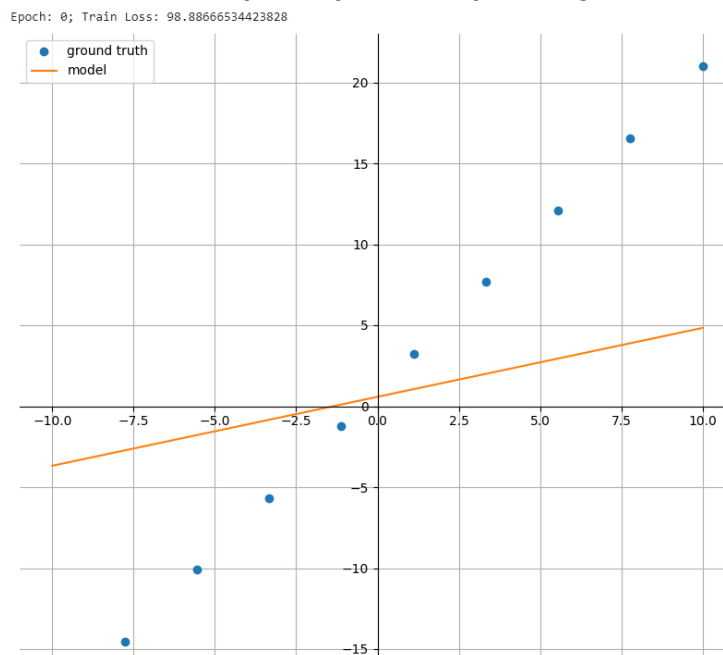
Lembrando que é necessária uma função de custo, usada neste caso a MSE, então com o uso do gradiente descendente a rede identifica o quão longe ela está da solução ideal para o problema e ajusta seus parâmetros w e b . Em palavras simples, a rede irá ajustar os valores de w e b até que identifique que os valores ideais são $w = 2$ e $b = 1$.

Partindo para a parte do treinamento de fato, usa-se a biblioteca *matplotlib* no Google Colab para visualizar o avanço do aprendizado da rede.

Define-se o número de épocas (número de vezes) que a rede irá usar para treinar e ajustar seus parâmetros, que nesse exemplo é de 1000 épocas.

Com tudo pronto, o programa pode ser rodado e visualiza-se através dos gráficos a evolução do aprendizado da rede, depois de 1000 épocas.

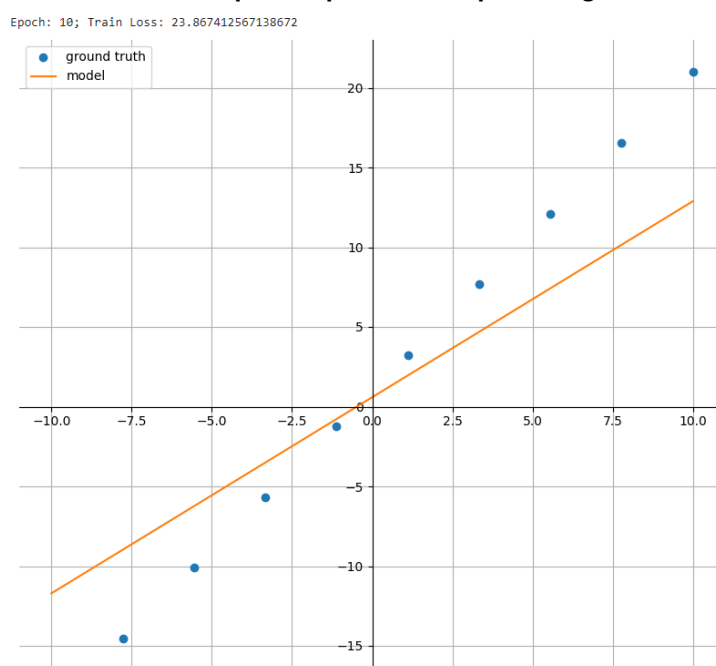
No **Gráfico 13** observa-se a época 0, a primeira tentativa da rede de encontrar a solução ideal para o problema, nota-se que a inclinação do modelo (em laranja) ainda está muito afastado dos valores esperados nos dados de treinamento (*ground truth* - pontos em azul).

Gráfico 13 – Exemplo 1: Época 0 da aprendizagem da rede

Fonte: elaborado pelo autor

Após essa primeira tentativa, a rede identifica o quão longe está da solução ideal e faz a retropropagação ajustando os valores do peso w e do bias b , fazendo então outra tentativa com os novos parâmetros.

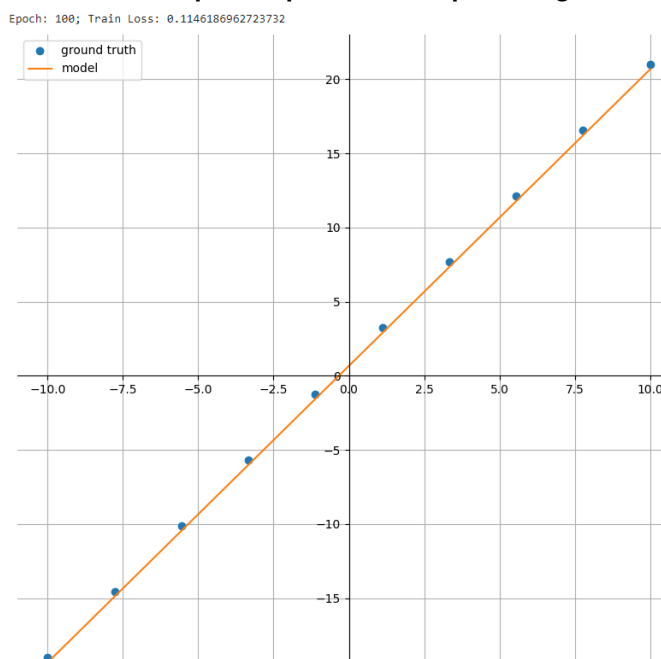
Já na época 10 (**Gráfico 14**), observa-se uma maior aproximação da estimativa da rede, e a cada época ela se aproxima da solução ideal após ajustes de seus parâmetros.

Gráfico 14 – Exemplo 1: Época 10 da aprendizagem da rede

Fonte: elaborado pelo autor

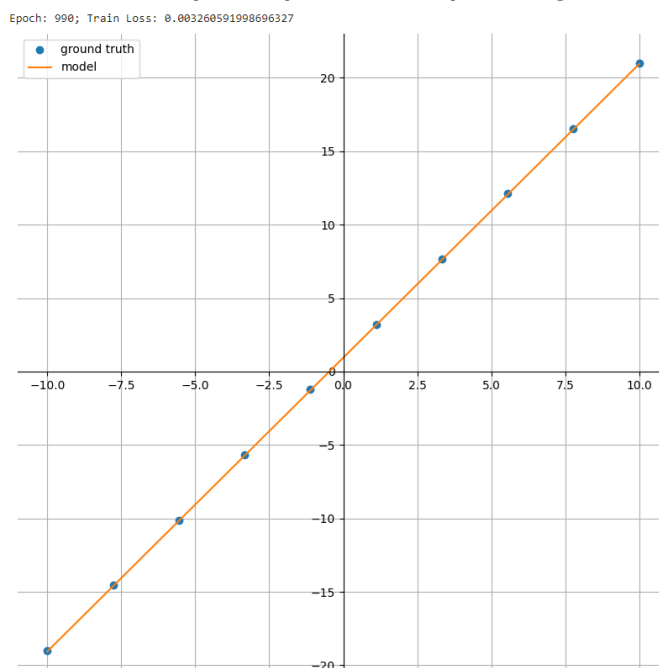
Na época 100 (**Gráfico 15**) a predição está muito próxima da realidade dos dados de treinamento, e o ajuste dos parâmetros ficando cada vez mais finos e próximos do ideal.

Gráfico 15 – Exemplo 1: Época 100 da aprendizagem da rede



Fonte: elaborado pelo autor

Na época 990 (**Gráfico 16**), a previsão da rede e os dados de treinamento são praticamente os mesmos. Pode-se então dizer que a rede aprendeu e pode prever qualquer par ordenado pertencente a reta $f(x) = 2x + 1$.

Gráfico 16 – Exemplo 1: Época 990 da aprendizagem da rede

Fonte: elaborado pelo autor

Esse exemplo simples permite vislumbrar de forma prática como a rede neural artificial é construída, treinada e como ela aprende, além de introduzir temas e conceitos matemáticos abordados no Ensino Médio, como o ensino de funções.

5.3 RNA para representação de uma curva qualquer no plano

Foi visto no exemplo anterior como uma rede neural simples, de apenas uma camada e sem uma função de ativação consegue representar uma reta no plano. Pode-se expandir esta idéia para outras curvas, mas para isso é necessária uma rede neural multicamadas, com mais neurônios e mais épocas de treinamento, ou seja, uma rede mais complexa do que a rede do primeiro exemplo.

Apesar de ser uma rede mais complexa, é usada a mesma arquitetura de construção da rede do exemplo anterior, com algumas modificações.

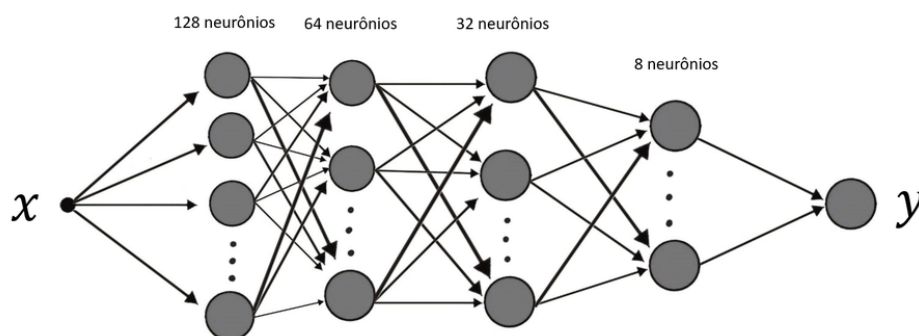
Todos os códigos de programação estão em anexo.

A rede será chamada de *Multi Layer Network* (Rede de Múltiplas camadas), e terá 6 camadas, sendo uma Camada Entrada, quatro Camadas Escondidas e uma Camada de Saída.

Entre a camada de entrada e a primeira camada escondida é usada a função ReLU como função de ativação, e também entre todas outras camadas até que chegue na camada de saída.

A estrutura da rede fica da seguinte forma:

Figura 19 – Exemplo 3: estrutura da RNA



Fonte: elaborado pelo autor

Observe que há uma entrada x na camada de entrada, já na primeira camada escondida há 128 neurônios, cada um com seus pesos e bias. Os 128 valores da primeira camada escondida passam pela função de ativação e seguem para a segunda camada escondida que possui 68 neurônios, cada um com seus pesos e bias e que também passam pela função de ativação.

Os 64 valores seguem para a terceira camada escondida que possui 32 neurônios, passando pelo mesmo processo de multiplicar pelos pesos e adicionar o *bias* e depois passa pela função de ativação.

A quarta camada escondida possui 8 neurônios com seus pesos e *bias*, passando pela função de ativação até que segue para a camada de saída com apenas um valor y de saída.

Trata-se uma rede mais complexa, que exige um pouco mais de esforço computacional para rodar.

Os dados de treinamento usados neste exemplo são pontos da curva $f(x) = \cos\left(\frac{x}{2}\right)$.

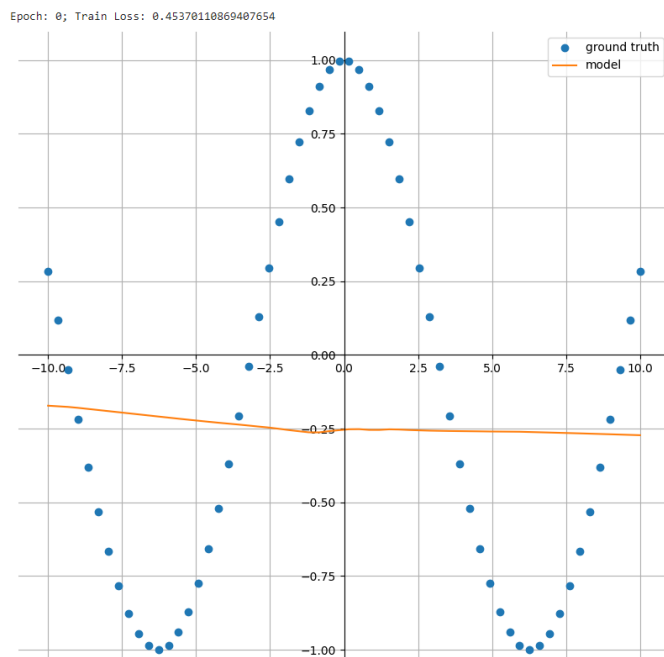
A estruturação da rede deste exemplo é a mesma usada no exemplo anterior, acrescidas de mais neurônios dispostos em mais camadas.

A função MSE é usada como função de custo e há o cálculo do gradiente descendente na retropropagação também para esse exemplo.

Como se trata de uma rede mais complexa que vai trabalhar em um problema mais difícil, serão necessárias mais épocas de treinamento para que a rede aprenda. Neste exemplo há 50001 épocas.

Implementando o código do exemplo anterior com adaptações para esse exemplo, a rede começa a fazer suas tentativas e ajuste para encontrar a solução ideal para o problema.

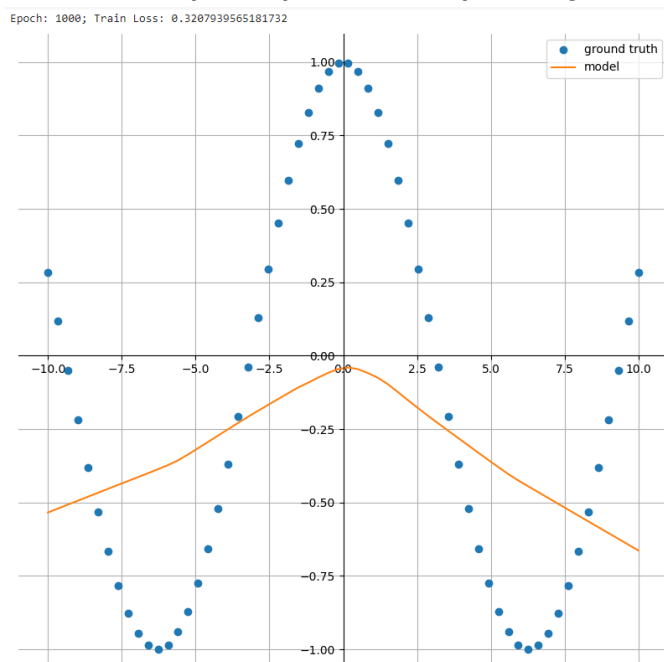
Na época 0 (**Gráfico 17**), o modelo proposto pela rede (em laranja) ainda está distante dos dados de treinamento (pontos em azul), mas assim que calcula o erro, a rede faz a retropropagação e ajusta seus parâmetros e faz novas tentativas.

Gráfico 17 – Exemplo 2: Época 0 da aprendizagem da rede

Fonte: elaborado pelo autor

Na época 1000 (**Gráfico 18**), o modelo proposto já toma uma forma mais curva, mas ainda longe da curva em desejada.

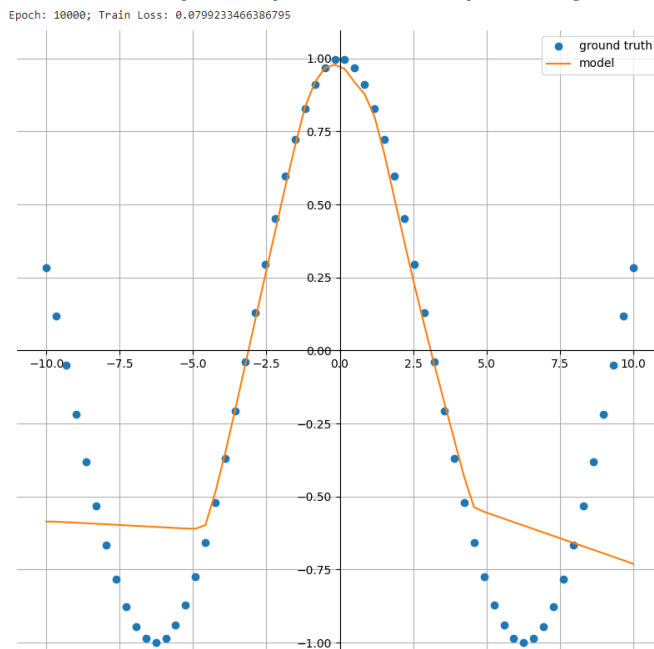
Este é um dos motivos desse tipo de problema necessitar de muitas épocas para aprender e propor um modelo adequado, já que os ajustes feitos são mais finos e sensíveis.

Gráfico 18 – Exemplo 2: Época 1000 da aprendizagem da rede

Fonte: elaborado pelo autor

Na época 10000 (**Gráfico 19**) o modelo proposto pela rede está se moldando aos dados de treinamento, mas ainda não sendo a solução ideal.

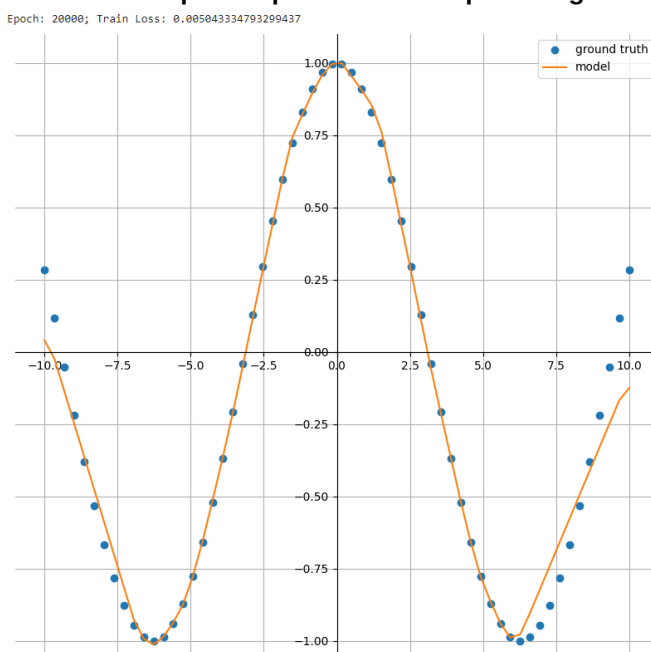
Gráfico 19 – Exemplo 2: Época 10000 da aprendizagem da rede



Fonte: elaborado pelo autor

Na época 20000 (**Gráfico 20**), a rede está perto de encontrar uma solução ideal para o problema, mas ainda seriam necessárias mais épocas para ela moldar melhor seu modelo, o que demanda mais custo computacional.

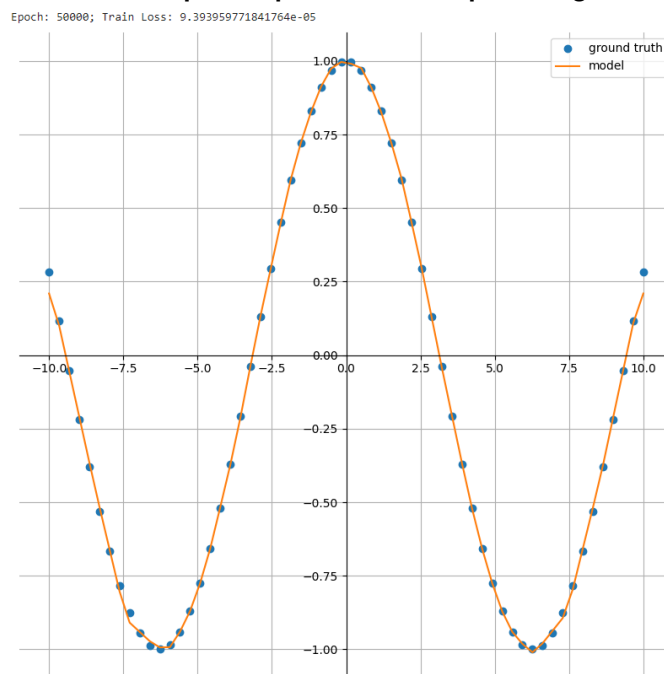
Gráfico 20 – Exemplo 2: Época 20000 da aprendizagem da rede



Fonte: elaborado pelo autor

Finalmente, após 50000 épocas (**Gráfico 21**), a rede consegue uma alta precisão em seu modelo, moldando a curva quase que perfeitamente nos dados de treinamento.

Gráfico 21 – Exemplo 2: Época 50000 da aprendizagem da rede



Fonte: elaborado pelo autor

Problemas mais complexos demandam mais esforço computacional e essa é a tendência entre os dois exemplos abordados. Além do terceiro exemplo ter mais neurônios e camadas escondidas em sua rede, também demanda muito mais épocas de treinamento para alcançar um resultado satisfatório, apesar de não ser 100% fiel à curva da cossenoide.

Apesar dessa diferença de complexidade entre as duas redes criadas nos exemplos, elas partem da mesma arquitetura de construção e treinamento.

6 CONSIDERAÇÕES FINAIS

A integração da Matemática com as tecnologias emergentes, especialmente no contexto das RNAs, revela a importância de repensar e reformular o Ensino da Matemática no Brasil. A análise das práticas pedagógicas atuais demonstra que, apesar dos avanços tecnológicos e das novas demandas do mercado de trabalho, o Ensino da Matemática ainda enfrenta desafios significativos, como a baixa proficiência dos alunos e a resistência à adoção de novas Metodologias.

É preciso reconhecer que a Matemática é uma disciplina central na Formação Educacional, não apenas por sua importância intrínseca, mas também por seu papel como base para diversas áreas do conhecimento. No entanto, as avaliações em larga escala, como o Programa Internacional de Avaliação de Estudantes (PISA), mostram que a Proficiência Matemática dos estudantes brasileiros é insatisfatória. Essa realidade exige uma reflexão profunda sobre as práticas pedagógicas e a necessidade de uma abordagem mais contextualizada e significativa no Ensino da Matemática.

A inclusão de tópicos e novas abordagens, como o uso de RNAs, no currículo do Ensino Médio não apenas enriquece o conhecimento matemático dos alunos, mas também os prepara para um mercado de trabalho cada vez mais orientado por tecnologias digitais. A capacidade de entender e aplicar conceitos matemáticos em contextos práticos, como a análise de dados e a programação, é uma habilidade essencial que deve ser desenvolvida desde os primeiros anos de escolaridade.

Além disso, há a necessidade de uma formação continuada e adequada para os professores de Matemática. A resistência à adoção de novas tecnologias e metodologias de ensino muitas vezes está ligada à falta de preparo e suporte para os educadores. Portanto, se faz necessário que as instituições de ensino e secretarias de educação ofereçam programas de formação que capacitem os professores a integrar tecnologias educacionais em suas práticas pedagógicas. Isso não apenas melhora a qualidade do ensino, mas também aumenta a confiança dos educadores em utilizar recursos tecnológicos de forma eficaz.

Ao relacionar conceitos matemáticos a situações do cotidiano e a aplicações práticas, como as RNAs, os alunos podem perceber a relevância do que estão aprendendo. Essa abordagem não apenas aumenta o engajamento dos estudantes, mas também os motiva a explorar e aprofundar seus conhecimentos. A Matemática, muitas vezes vista como uma disciplina abstrata e distante da realidade, pode se tornar mais acessível e interessante quando apresentada de forma contextualizada.

Este trabalho também destaca os desafios que as Escolas enfrentam na implementação de um currículo mais abrangente e tecnológico. A falta de infraestrutura adequada, recursos financeiros e acesso a tecnologias educacionais são barreiras que precisam ser superadas. É necessária a adoção de políticas públicas a serem desenvolvidas para garantir

que todas as Escolas, independentemente de sua localização ou contexto socioeconômico, tenham acesso a recursos tecnológicos e formação de qualidade para seus professores.

Por fim, é reafirmada a importância de preparar os alunos para o futuro. Em um mundo em constante evolução, onde as habilidades em ciência de dados e inteligência artificial são cada vez mais valorizadas, é crucial que o Sistema Educacional Brasileiro se adapte a essa realidade., contribuindo, assim, para a formação de uma geração de alunos com uma base sólida em Matemática e Tecnologia não apenas os capacitando para enfrentar os desafios do mercado de trabalho, mas também os preparando para serem cidadãos críticos e informados, capazes de contribuir para a sociedade de maneira significativa.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, Mário de Souza. *Elaboração de Projeto, TCC, Dissertação e Tese. Uma abordagem simples, prática e objetiva*. 2a Edição. Atlas. 2014.

ARRUDA, Eucidio Pimenta. Educação remota emergencial: elementos para políticas públicas na educação brasileira em tempos de Covid-19. *EmRede-Revista de Educação a Distância*, v. 7, n. 1, p. 257-275, 2020.

AVELINO, Wagner Feitosa; MENDES, Jessica Guimarães. A realidade da educação brasileira a partir da COVID-19. *Boletim de Conjuntura (BOCA)*, v. 2, n. 5, p. 56-62, 2020.

BARANWAL, Tanish. *A Hands On Introduction to Deep Learning from a High School Perspective*. Towards Data Science. 2017.

BARONI, Itanimá; PADILHA, Eliandro José. *A MATEMÁTICA POR TRÁS DAS REDES NEURAIS*. 2021.

BARRETO, Andreia Cristina Freitas; ROCHA, Daniele Santos. Covid 19 e educação: resistências, desafios e (im) possibilidades. *Revista encantar*, v. 2, p. 01-11, 2020.

BRASIL. Base Nacional Comum Curricular. Brasília, DF. 2019.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Notas sobre o Brasil no Pisa 2022*. Brasília, DF: Inep, 2023

BRASIL. Ministério da Educação. CNE aprova diretrizes para escolas durante a pandemia. Brasília: junho, 2020. Disponível em: < <http://portal.mec.gov.br/busca-geral/12-noticias/acoes-programas-e-projetos-637152388/89051-cne-aprova-diretrizes-para-escolas-durante-a-pandemia> >

BRASIL. *NOTA SOBRE O BRASIL NO PISA 2022*. MINISTÉRIO DA EDUCAÇÃO. DIRETORIA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA, DAEB. 2023.

BATISTA, Marcio; MARTINS, André Oliveira. *Redes Neurais no Ensino Básico*. PMO v.10, n.4. 2022.

CAFARDO, Renata. Oito em cada dez professores não se sentem preparados para ensinar online. *O Estado de São Paulo*, 2020.

CANSIAN, Ana Bárbara Moulin; COSTA, Andréa Oliveira Souza da; JUNIOR, Esly Ferreira da Costa. *UTILIZAÇÃO DE REDES NEURAIS ARTIFICIAIS COMO ESTÍMULO AO APRENDIZADO DE MATEMÁTICA*. ENCICLOPÉDIA BIOSFERA, Centro Científico Conhecer, V.10, N.19; p. 2567. 2014.

CENSO EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2018 [livro eletrônico]/[organização] ABED – Associação Brasileira de Educação a Distância. Camila Rosa (tradutora). Curitiba: InterSaberes, 2019.

COAN, L. G. W; VISEU, F; MORETTI, M. T. As TIC no ensino de Matemática: a formação dos professores em debate. *REVEMAT*. Florianópolis (SC), V. 08, n. 2, p. 222-244, 2013. Disponível em: <https://periodicos.ufsc.br/index.php/revemat/issue/view/2057>.

DELGADO, Jorge; FRENSEL, Katia; CRISSAFF, Lhaylla. Geometria Analítica. Coleção PROFMAT. Sociedade Brasileira de Matemática. 2017.

GÉRON, Aurélien. Hands-on Machine Learning with Scikit-Learn, Keras, and Tensor Flow Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media. 2019.

GIL, A. C. Métodos e Técnicas de Pesquisa Social. São Paulo, SP: Atlas, 2008.

GOMES, Suzana; ZEN, Giovana Cristina; D'ÁVILA, Cristina. Desafios e perspectivas do ensino do componente curricular didática em meio a pandemia da Covid-19. Revista Ibero-Americana de Estudos em Educação, p. 1453-1475, 2022.

HASENBALG, Carlos Alfredo; DO VALLE SILVA, Nelson. Origens e destinos: desigualdades sociais ao longo da vida. Iuperj/Ucam, 2003.

IEZZI, Gelson; MURAKAMI, Carlos. Fundamentos de matemática elementar, 1: conjuntos, funções. SARAIVA S. A. Livreiros Editores 9. ed. 2013.

IEZZI, Gelson; DOLCE, Osvaldo; MURAKAMI, Carlos. Fundamentos de matemática elementar, 2: logaritmos. SARAIVA S. A. Livreiros Editores 10. ed. 2013.

IEZZI, Gelson. Fundamentos de matemática elementar, 3: trigonometria. SARAIVA S. A. Livreiros Editores 9. ed. 2013.

IEZZI, Gelson; HAZZAN, Samuel. Fundamentos de matemática elementar, 4: sequência, matrizes, determinantes e sistemas. SARAIVA S. A. Livreiros Editores 8. ed. 2013.

LAKATOS, E. M.; MARCONI, M. A. Fundamentos de metodologia científica. São Paulo, SP: Atlas, 2003. p.44-73.

MAIA, Berta Rodrigues; DIAS, Paulo César. Ansiedade, depressão e estresse em estudantes universitários: o impacto da COVID-19. Estudos de psicologia (Campinas), v. 37, p. e200067, 2020.

MOREIRA, J. António; HENRIQUES, Susana; BARROS, Daniela Melaré Vieira. Transitando de um ensino remoto emergencial para uma educação digital em rede, em tempos de pandemia. Dialogia, p. 351-364, 2020.

MOTTA, Marcelo Souza. Formação inicial do professor de matemática no contexto das tecnologias digitais. Contexto & Educação, v. 32, n. 102, p. 170-204, 2017.

NELSON, Hala. Essential Math for AI Next-Level Mathematics for Efficient and Successful AI Systems. O'Reilly Media. 2023.

ORNELL, FELIPE et al. Pandemia de medo e Covid-19: impacto na saúde mental e possíveis estratégias. Debates em psiquiatria, v. 10, n. 2, p. 12-16, 2020.

PONTES, Ana Paula Florêncio Ferreira; BARBOZA, Pedro Lucio. O professor de matemática frente às tecnologias e as dificuldades em integrá-las na sala de aula. Ensino em Foco, v. 3, n. 8, p. 33-47, 2020.

RECH, Gislaine Zanoto; PESCADOR, Cristina Maria. Ensino remoto em tempos de pandemia: COVID-19 suas implicações na interação professor-estudante-uma perspectiva freireana. Revista Ibero-Americana de Estudos em Educação, p. 1264-1278, 2022.

RIOS, Mirivan Carneiro. O gestor escolar e as novas tecnologias. Revista Eletrônica

Educação Em Foco, Amparo, v. 4, 2011.

ROSENBLATT, Frank (1957). "The Perceptron—a perceiving and recognizing automaton" (PDF). Report 85-460-1. Cornell Aeronautical Laboratory.

ROSENBLATT, F. (1958). "The perceptron: A probabilistic model for information storage and organization in the brain". *Psychological Review*. 65 (6): 386–408. doi:10.1037/h0042519. ISSN 1939-1471

SÁ, Adriana Lourenço; MACHADO, Marília Costa. O uso do software GeoGebra no estudo de funções. XIV EVIDOSOL e XI CILTEC online, junho 2017. Disponível em: <https://eventos.textolivre.org/moodle/course/view.php?id=12>.

SANTOS, José Elyton Batista; ROSA, Maria Cristina; DA SILVA SOUZA, Denize. O ensino de matemática online: um cenário de reformulação e superação. *Revista Interacções*, v. 16, n. 55, p. 165-185, 2020.

SANTOS, José Elyton Batista; DOS SANTOS, Iris Grasielle Xavier. O COMPUTADOR INTERATIVO NO ENSINO DE MATEMÁTICA E SUAS IMPLICAÇÕES. RFB Editora.

SANTOS, José Elyton Batista; ROSA, Maria Cristina; DA SILVA SOUZA, Denize. O ensino de matemática online: um cenário de reformulação e superação. *Revista Interacções*, v. 16, n. 55, p. 165-185, 2020.

SANT'ANA, Claudinei de Camargo; AMARAL, Rúbia Barcelos; BORBA, Marcelo de Carvalho. O uso de softwares na prática profissional do professor de matemática. *Ciência & Educação (Bauru)*, v. 18, p. 527-542, 2012.

SOUZA, Janicléia Pereira; FERNANDES, Flávia Emília Cavalcante Valença. Os impactos na saúde dos professores da educação básica durante a pandemia da covid-19: uma revisão de literatura. *Travessias*, v. 17, n. 1, p. 7, 2023.

QUINZANI, Marcia Angela Dahmer. O avanço da pobreza e da desigualdade social como efeitos da crise da covid-19 e o estado de bem-estar social. *Boletim de Conjuntura (BOCA)*, v. 2, n. 6, p. 43-47, 2020.

ZANELLA, Lilian; MACHADO, Simone Raquel Casarin. O USO DE RECURSOS TECNOLÓGICOS PARA O ENSINO DE MATEMÁTICA. 2019.

APÊNDICE

Códigos de programação em linguagem Python dos exemplos abordados no capítulo

5.

Todos os códigos foram implementadas na plataforma Google Colab.

5.1. Código do primeiro exemplo

```
import numpy as np
from random import *
def escolheValoresRandomicos():
    x1 = randint(0, 1) #fator nota
    x2 = randint(0, 1) #fator amigos
    x3 = randint(0, 1) #fator distância
    #quanto maior o valor, mais se motiva a ir pela nota
    w1 = randint(0, 8)
    #quanto maior o valor, mais se motiva a ir com amigos
    w2 = randint(0, 8)
    #quanto maior o valor, mais se motiva a ir se for perto
    w3 = randint(0, 8)
    #quanto maior o valor, mais rígido é pra decidir se vai ao evento de matemática
    limiar = randint(0, 11)
    return x1, w1, x2, w2, x3, w3, limiar
def perceptron(x1, w1, x2, w2, x3, w3, limiar):

    #fator nota
    if(w1 >= 0 and w1 < 3): #se cair neste intervalo, significa que Matheus não se
motiva a ir pela nota
        print("Você não se motiva com a nota a ir ao evento de matemática")
    elif(w1 >= 3 and w1 < 6): #se cair neste intervalo, significa que Matheus se motiva
um pouco a ir pela nota
        print("Você se motiva um pouco com a nota a ir ao evento de matemática")
    elif(w1 >= 6 and w1 < 9): #se cair neste intervalo, significa que Matheus se motiva
muito a ir pela nota
        print("Você se motiva fácil com a nota a ir ao evento de matemática")

    #fator amigos
    if(w2 >= 0 and w2 < 3): #se cair neste intervalo, significa que Matheus não se
motiva a ir pelos amigos
```

```
print("Nem seus amigos conseguem te empurrar para ir ao evento de matemática")
elif(w2 >= 3 and w2 < 6): #se cair neste intervalo, significa que Matheus se motiva
um pouco a ir pelos amigos
    print("Seus amigos empurram um pouquinho a ir ao evento de matemática")
    elif(w2 >= 6 and w2 < 9): #se cair neste intervalo, significa que Matheus se motiva
muito a ir pelos amigos
        print("Com os seus amigos topando, é alta a probabilidade de você ir ao evento de
matemática")

#fator distancia
if(w3 >= 0 and w3 < 3): #se cair neste intervalo, significa que Matheus não se
motiva a ir mesmo se o evento for perto
    print("Seja longe ou perto, a distância pouco te motiva a ir ao evento de matemática")
    elif(w3 >= 3 and w3 < 6): #se cair neste intervalo, significa que Matheus se motiva
um pouco a ir mesmo se o evento for perto
        print("Se for perto, isso te dá um empurrãozinho para ir ao evento de matemática")
        elif(w3 >= 6 and w3 < 9): #se cair neste intervalo, significa que Matheus se motiva
muito a ir se o evento for perto
            print("É perto e isso te motiva muito a ir ao evento de matemática")

#imprimindo as variáveis x
if(x1 == 0):
    print("Fator nota está desligado")
elif(x1 == 1):
    print("Fator nota está ligado")
if(x2 == 0):
    print("Fator amigos está desligado")
elif(x2 == 1):
    print("Fator amigos está ligado")
if(x3 == 0):
    print("Fator distância está desligado")
elif(x3 == 1):
    print("Fator distância está ligado")
print("\n")

#imprimindo o limiar
if(limiar >= 0 and limiar < 4):
    print("O limiar escolhido está frouxo")
elif(limiar >= 4 and limiar < 8):
```

```

print("O limiar escolhido está flexível")
elif(limiar >= 8 and limiar < 12):
print("O limiar escolhido está rigoroso")
output = w1 * x1 + w2 * x2 + w3 * x3
return output
def main():
x1, w1, x2, w2, x3, w3, limiar = escolheValoresRandomicos()
output = perceptron(x1, w1, x2, w2, x3, w3, limiar)
if(output >= limiar):
print("Matheus vai ao evento de Matemática! =)")
else:
print("Matheus desistiu de ir ao evento de Matemática! =)")
if __name__ == "__main__":
main()

```

5.2. Código do segundo exemplo

```

from IPython.display import HTML
HTML('')
import numpy as np
class Perceptron(object):
def __init__(self, no_of_inputs):
self.weights = np.zeros(no_of_inputs + 1)
def predict(self, inputs):
summation = np.dot(inputs, self.weights[1:]) + self.weights[0]
if summation > 0:
activation = summation
else:
activation = 0
return activation
import torch
import numpy as np
from torch import nn
class LineNetwork(nn.Module):
# Inicialização
def __init__(self):
super().__init__()
self.layers = nn.Sequential(
nn.Linear(1, 1)

```

```

)
# Como a rede computa
def forward(self, x):
    return self.layers(x)

from torch.utils.data import Dataset, DataLoader
import torch.distributions.uniform as urand

class AlgebraicDataset(Dataset):
    def __init__(self, f, interval, nsamples):
        X = urand.Uniform(interval[0], interval[1]).sample([nsamples])
        self.data = [(x, f(x)) for x in X]
    def __len__(self):
        return len(self.data)
    def __getitem__(self, idx):
        return self.data[idx]

line = lambda x: 2*x + 1
interval = (-10, 10)
train_nsamples = 1000
test_nsamples = 100
train_dataset = AlgebraicDataset(line, interval, train_nsamples)
test_dataset = AlgebraicDataset(line, interval, test_nsamples)
train_dataloader = DataLoader(train_dataset, batch_size=train_nsamples, shuffle=True)
test_dataloader = DataLoader(test_dataset, batch_size=test_nsamples, shuffle=True)

# Função de custo (loss function)
# Erro quadrático médio (Mean Squared Error)
lossfunc = nn.MSELoss()
# Gradiente Descendente Estocástico
# SGD = Stochastic Gradient Descent
optimizer = torch.optim.SGD(model.parameters(), lr=1e-3)
# taxa de aprendizado lr = learning rate

def train(model, dataloader, lossfunc, optimizer):
    model.train()
    cumloss = 0.0
    for X, y in dataloader:
        X = X.unsqueeze(1).float().to(device)

```

```
y = y.unsqueeze(1).float().to(device)
pred = model(X)
loss = lossfunc(pred, y)
# zera os gradientes acumulados
optimizer.zero_grad()
# computa os gradientes
loss.backward()
# anda, de fato, na direção que reduz o erro local
optimizer.step()
# loss é um tensor; item pra obter o float
cumloss += loss.item()
return cumloss / len(dataloader)

def test(model, dataloader, lossfunc):
    model.eval()
    cumloss = 0.0
    with torch.no_grad():
        for X, y in dataloader:
            X = X.unsqueeze(1).float().to(device)
            y = y.unsqueeze(1).float().to(device)
            pred = model(X)
            loss = lossfunc(pred, y)
            cumloss += loss.item()
    return cumloss / len(dataloader)

import matplotlib.pyplot as plt

# Pra visualizar
def plot_comparinson(f, model, interval=(-10, 10), nsamples=10):
    fig, ax = plt.subplots(figsize=(10, 10))
    ax.grid(True, which='both')
    ax.spines['left'].set_position('zero')
    ax.spines['right'].set_color('none')
    ax.spines['bottom'].set_position('zero')
    ax.spines['top'].set_color('none')
    samples = np.linspace(interval[0], interval[1], nsamples)
    model.eval()
    with torch.no_grad():
        pred = model(torch.tensor(samples).unsqueeze(1).float().to(device))
```

```

ax.plot(samples, list(map(f, samples)), "o", label="ground truth")
ax.plot(samples, pred.cpu(), label="model")
plt.legend()
plt.show()

```

```

epochs = 1000
for t in range(epochs):
    train_loss = train(model, train_dataloader, lossfunc, optimizer)
    if t % 10 == 0:
        print(f"Epoch: {t}; Train Loss: {train_loss}")
        plot_comparinon(line, model)
    test_loss = test(model, test_dataloader, lossfunc)
    print(f"Test Loss: {test_loss}")

```

5.3 Código do terceiro exemplo

```

class MultiLayerNetwork (nn.Module)
def __init__(self):
    super().__init__()
    self.layers = nn.Sequential(
        nn.Linear(1, 128),
        nn.ReLU(),
        nn.Linear(128, 64),
        nn.ReLU(),
        nn.Linear(64, 32),
        nn.ReLU(),
        nn.Linear(32, 8),
        nn.ReLU(),
        nn.Linear(8, 1),
    )
def forward(self, x):
    return self.layers(x)

multimodel = MultiLayerNetwork().to(device)
from math import cos
f = lambda x: cos(x/2)

train_dataset = AlgebraicDataset(f, interval, train_nsamples)
test_dataset = AlgebraicDataset(f, interval, test_nsamples)

```

```
train_dataloader = DataLoader(train_dataset, train_nsamples, shuffle=True)
test_dataloader = DataLoader(test_dataset, test_nsamples, shuffle=True)

lossfunc = nn.MSELoss()
optimizer = torch.optim.SGD(multimodel.parameters(), lr=1e-3)

epochs = 50001
for t in range(epochs):
    train_loss = train(multimodel, train_dataloader, lossfunc, optimizer)
    if t % 100 == 0:
        print(f"Epoch: {t}; Train Loss: {train_loss}")
        plot_comparinon(f, multimodel, nsamples=60)
    test_loss = test(multimodel, test_dataloader, lossfunc)
    print(f"Test Loss: {test_loss}")
```