



CAPÍTULO 2

PENSAMENTO COMPUTACIONAL E APRENDIZADO DE MÁQUINA: DESAFIOS NA PREDIÇÃO DO DESEMPENHO ESCOLAR COM REDES NEURAIS ARTIFICIAIS

 <https://doi.org/10.22533/at.ed.325112612012>

Márcio Mendonça

Universidade Tecnológica Federal do Paraná
PPGEM-CP - Programa de Pós-Graduação em Engenharia Mecânica CP/PG
Cornélio Procópio - PR
<http://lattes.cnpq.br/5415046018018708>

Vitor Blanc Milani

Universidade Tecnológica Federal do Paraná
Mestrando - PPGEM-CP - Programa de Pós-Graduação em Engenharia Mecânica CP/PG
Cornélio Procópio - PR
<http://lattes.cnpq.br/4504374098250296>

Antonio Carlos Rodrigues

UnisenaiPR Campus Londrina
Londrina - PR
<http://Lattes.cnpq.br/1963491889845252>

Fabio Rodrigo Milanez

UnISENAIPR-Campus Londrina
Londrina-PR
<http://lattes.cnpq.br/3808981195212391>

Francisco de Assis Scannavino Junior

Universidade Tecnológica Federal do Paraná Departamento Acadêmico de
Engenharia Elétrica (DAELE)
Cornélio Procópio - PR
<http://lattes.cnpq.br/4513330681918118>

Paulo Alexandre Lourenço Jesus

Afiliação acadêmica: Aluno externo no Programa de Pós-Graduação em Ensino
de Ciências Humanas, Sociais e da Natureza – PPGEN – Curso de Mestrado
Instituição: Universidade Tecnológica Federal do Paraná
(UTFPR) – Campus Cornélio Procópio/Londrina – PR
<https://lattes.cnpq.br/1346404833044235>

Adriano da Silva Moreira

Afiliação acadêmica: Mestrando no Programa de Pós-Graduação em Educação – PPEDU
Instituição: Universidade Estadual de Londrina (UEL) –
Campus Londrina – PR
<https://lattes.cnpq.br/0686904669527189>

Ana Clara Augusto Jesus

Acadêmica do curso de Pedagogia da Universidade Estadual do Norte
do Paraná (UENP) Campus de Cornélio Procópio - Paraná – Brasil
Cornélio Procópio – PR
<https://lattes.cnpq.br/5330914992553099>

Marcos Dantas de Oliveira

CEAL- Clube de Engenharia de Londrina
Londrina-PR
<https://lattes.cnpq.br/5329306535174160>

Luiz Francisco Sanches Buzzacchero

Universidade Tecnológica Federal do Paraná Departamento Acadêmico de
Engenharia Elétrica (DAELE)
Cornélio Procópio - PR
<http://lattes.cnpq.br/1747856636744006>

Norwin Porfirio Carrasquel Poturo

UTFPR - PPGIT - Programa de Pós graduação em Inovações tecnológicas
Campo Mourão
<http://lattes.cnpq.br/0512294788133450>

Emerson Ravazzi Pires da Silva

Universidade Tecnológica Federal do Paraná
Departamento Acadêmico de Engenharia Elétrica (DAELE)
Cornélio Procópio – PR
<http://lattes.cnpq.br/3845751794448092>

Aline Lucena Prado

Aluno externo no Programa de Pós-Graduação em Ensino de Ciências
Humanas, Sociais e da Natureza – PPGEN Curso de Mestrado
Universidade Tecnológica Federal do Paraná (UTFPR) –
Campus Cornélio Procópio/Londrina, PR
<https://lattes.cnpq.br/9637078731107542>

Marcos Antônio de Matos Laia

Universidade Federal de São Joao Del Rei Departamento De Ciência Da Computação – UFSJ
São Joao Del Rei - MG
<http://lattes.cnpq.br/7114274011978868>

Juliana Maria de Jesus Ribeiro

Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Cornélio
Procópio/Londrina, Paraná – Brasil
Mestranda no Programa de Pós-Graduação em Ensino de Ciências Humanas,
Sociais e da Natureza – PPGEN
Londrina - PR
<http://lattes.cnpq.br/6279504657014354>

Tatiane Monteiro Pereira

Mestranda - Programa de Pós-Graduação em Ensino de Ciências Humanas,
Sociais e da Natureza (PPGEN-CP/LD)
Cornélio Procópio – PR
<http://lattes.cnpq.br/9520601026438758>

Vera Adriana Huang Azevedo Hypólito

Centro Estadual de Educação Tecnológica Paula Souza
Etec Jacinto Ferreira de Sá
Ourinhos – SP
<http://lattes.cnpq.br/6169590836932698>

Armando Paulo da Silva

Universidade Tecnológica Federal do Paraná
Programa de Pós-Graduação em Ensino de Ciências
Humanas, Sociais e da Natureza (PPGEN-CP/LD)
Cornélio Procópio–PR
<http://lattes.cnpq.br/6724994186659242>

Cintya Wedderhoff Machado

Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Cornélio
Procópio/Londrina, Paraná – Brasil
Mestranda no Programa de Pós-Graduação em Ensino de Ciências Humanas,
Sociais e da Natureza – PPGEN
Londrina - PR
<http://lattes.cnpq.br/4604594140489347>

Keila Felipe do Carmo de Lima

Universidade Tecnológica Federal do Paraná (UTFPR) – Campus
Cornélio Procópio/Londrina, Paraná – Brasil
Mestranda no Programa de Pós-Graduação em Ensino de
Ciências Humanas, Sociais e da Natureza – PPGEN
Cornélio Procópio - PR
<http://lattes.cnpq.br/8692187746106065>

Adriele Muniz

Universidade Tecnológica Federal do Paraná (UTFPR) – Campus
Cornélio Procópio/Londrina, Paraná – Brasil
aluna externa no Programa de Pós-Graduação em Ensino de
Ciências Humanas, Sociais e da Natureza – PPGEN
<http://lattes.cnpq.br/6059998214513158>

RESUMO: Este artigo apresenta uma síntese sistematizada, clara e tecnicamente fundamentada dos principais conceitos que estruturam o campo das Redes Neurais Artificiais (RNAs). São abordados, de forma progressiva, os fundamentos do neurônio artificial e da arquitetura de redes, bem como aspectos essenciais do fluxo de modelagem, incluindo pré-processamento e normalização de dados,

definição de hiperparâmetros, funções de ativação, e as etapas de treinamento, validação e teste. O texto também discute fenômenos críticos no desempenho dos modelos, como *overfitting* e *underfitting*, destacando suas causas, implicações e estratégias de mitigação. Além dos fundamentos teóricos, são apresentadas aplicações multidisciplinares das RNAs em áreas como saúde, educação, indústria, agricultura, finanças e tecnologias emergentes, evidenciando seu papel central nos sistemas inteligentes contemporâneos. O objetivo do trabalho é oferecer uma visão integrada, didática e acessível, capaz de apoiar estudantes e profissionais na compreensão dos princípios do aprendizado profundo e de suas implicações práticas no desenvolvimento de soluções baseadas em inteligência artificial. Esse trabalho se encerra com uma conclusão e endereça futuros trabalhos.

PALAVRAS-CHAVE: Redes Neurais Artificiais, Aprendizado Profundo, Inteligência Artificial, Treinamento de Modelos, Aplicações Multidisciplinares

COMPUTATIONAL THINKING AND MACHINE LEARNING: CHALLENGES IN PREDICTING SCHOOL PERFORMANCE WITH ARTIFICIAL NEURAL NETWORKS

ABSTRACT: This article presents a systematized, clear, and technically grounded synthesis of the main concepts that structure the field of Artificial Neural Networks (ANNs). The fundamentals of the artificial neuron and network architectures are progressively addressed, as well as essential aspects of the modeling workflow, including data preprocessing and normalization, hyperparameter definition, activation functions, and the stages of training, validation, and testing. The text also discusses critical phenomena affecting model performance, such as overfitting and underfitting, highlighting their causes, implications, and mitigation strategies. Beyond the theoretical foundations, multidisciplinary applications of ANNs are presented in areas such as healthcare, education, industry, agriculture, finance, and emerging technologies, emphasizing their central role in contemporary intelligent systems. The objective of this work is to provide an integrated, didactic, and accessible perspective to support students and professionals in understanding the principles of deep learning and its practical implications in the development of artificial intelligence-based solutions. This work concludes with a final discussion and directions for future research.

KEYWORDS: Artificial Neural Networks, Deep Learning, Artificial Intelligence, Model Training, Multidisciplinary Applications

INTRODUÇÃO

O avanço das tecnologias de Inteligência Artificial tem impulsionado o uso de modelos capazes de aprender padrões complexos a partir de dados. Entre essas tecnologias, as Redes Neurais Artificiais (RNAs) ocupam um papel central, pois imitam de maneira simplificada o funcionamento do sistema nervoso humano. Sua versatilidade permite aplicá-las em diversas áreas, como diagnóstico médico, previsão de desempenho escolar, manutenção preditiva em indústrias, análise de crédito, reconhecimento de imagens e processamento de linguagem natural.

Este artigo apresenta uma revisão objetiva dos conceitos fundamentais que sustentam o funcionamento das RNAs, estruturando-se em tópicos essenciais para compreensão técnica e prática. A proposta é servir como um material de consulta rápida e confiável para estudantes, docentes e profissionais que buscam formação básica em redes neurais.

Além dos aspectos técnicos relacionados às Redes Neurais Artificiais, é fundamental destacar o papel do Pensamento Computacional como competência estruturante para a compreensão, desenvolvimento e interpretação de modelos de aprendizado de máquina. O Pensamento Computacional envolve habilidades como decomposição de problemas, reconhecimento de padrões, abstração e elaboração de algoritmos, sendo amplamente reconhecido como uma base cognitiva essencial para a resolução de problemas complexos em ambientes computacionais e educacionais. Estudos recentes indicam que o domínio dessas habilidades contribui significativamente para a compreensão de modelos de inteligência artificial, permitindo que estudantes e profissionais interpretem resultados, avaliem limitações dos modelos e tomem decisões mais fundamentadas no processo de modelagem (Wing, 2006; Shute; Sun; Asbell-Clarke, 2017; Román-González; Pérez-González; Jiménez-Fernández, 2017).

No contexto educacional, o Pensamento Computacional apresenta forte convergência com o uso de Redes Neurais Artificiais, especialmente em aplicações voltadas à análise de dados educacionais e à predição de desempenho acadêmico. Pesquisas recentes demonstram que estudantes com maior proficiência em Pensamento Computacional tendem a compreender melhor conceitos de aprendizado supervisionado, ajuste de hiperparâmetros e validação de modelos, reduzindo interpretações equivocadas e expectativas irreais sobre a capacidade preditiva das RNAs. Além disso, a literatura aponta que a integração explícita do Pensamento Computacional no ensino de inteligência artificial favorece uma abordagem mais crítica e reflexiva sobre os resultados obtidos, evidenciando que limitações preditivas, como as observadas neste estudo, muitas vezes decorrem da baixa qualidade explicativa dos dados e não da técnica empregada (Kong *et al.*, 2020; Denning; Tedre, 2019; Tang *et al.*, 2023).

FUNDAMENTAÇÃO TEÓRICA

Neurônio Artificial e a Arquitetura MLP para Regressão

A unidade fundamental das Redes Neurais Artificiais (RNAs) é o neurônio artificial, também conhecido como Perceptron. Esse modelo computacional foi inspirado, de forma abstrata, no funcionamento dos neurônios biológicos e tem como objetivo processar informações de entrada por meio de operações matemáticas simples, porém fundamentais para o aprendizado supervisionado.

O funcionamento do neurônio artificial pode ser descrito em duas etapas principais:

Agregação (soma ponderada): Cada variável de entrada x_i é multiplicada por um peso sináptico w_i , que representa a relevância daquela informação no processo decisório do modelo. A soma ponderada dessas entradas é acrescida de um termo de polarização (bias), responsável por ajustar o limiar de ativação do neurônio. O valor resultante representa o potencial de ativação do neurônio.

Ativação: O valor agregado é então submetido a uma função de ativação, cuja principal função é introduzir não linearidade ao modelo. Essa característica é essencial para que a rede seja capaz de representar relações complexas e não lineares presentes em problemas reais, como aqueles associados à predição do desempenho escolar.

O modelo conceitual do neurônio artificial descrito é ilustrado na Figura 1, que apresenta de forma esquemática o fluxo de processamento das informações desde as entradas até a geração da saída.

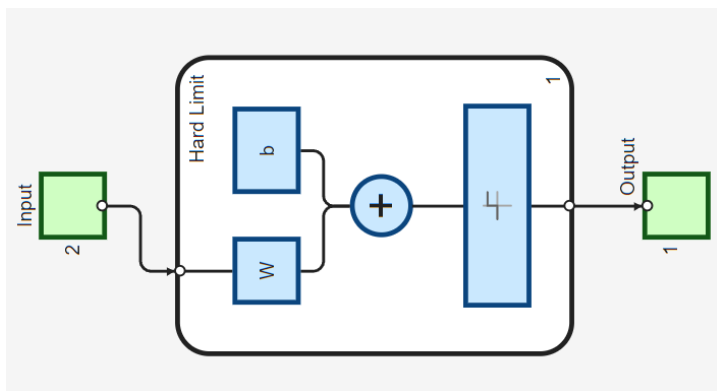


Figura 1 – Neurônio Artificial.

A representação apresentada na Figura 1 evidencia o princípio básico de funcionamento das RNAs e serve como base para arquiteturas mais complexas. Embora o perceptron isolado seja capaz apenas de resolver problemas linearmente separáveis, sua importância histórica e conceitual reside no fato de estabelecer os fundamentos do aprendizado supervisionado e do ajuste iterativo de pesos a partir de dados.

Com base nesse modelo elementar, este estudo adota a arquitetura Perceptron de Múltiplas Camadas (*Multi-Layer Perceptron – MLP*), configurada para a tarefa de regressão, isto é, para a previsão de uma variável contínua (a média escolar). As redes MLP do tipo feedforward são amplamente utilizadas em problemas educacionais por sua capacidade de capturar relações não lineares entre variáveis preditoras e a variável-alvo.

A arquitetura empregada é composta por três camadas funcionais:

- Camada de entrada: responsável por receber as variáveis preditoras do aluno;
- Camada oculta: encarregada de extrair características intermediárias e modelar relações não lineares entre os dados;
- Camada de saída: composta por um único neurônio, responsável por produzir a estimativa contínua da média escolar.

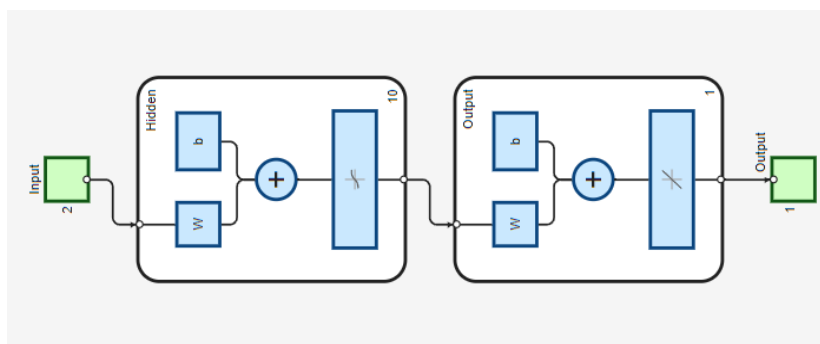


Figura 2 – Neurônio Artificial.

A introdução da camada oculta, conforme ilustrado na Figura 2, amplia significativamente a capacidade representacional do modelo em comparação ao perceptron de camada única. Cada neurônio da camada oculta realiza um processamento análogo ao apresentado na Figura 1; contudo, a composição de

múltiplos neurônios e camadas resulta em uma função global mais expressiva, capaz de aproximar funções complexas e capturar padrões sutis nos dados.

No contexto educacional, essa capacidade é particularmente relevante, uma vez que o desempenho escolar é influenciado por múltiplos fatores interdependentes, frequentemente caracterizados por elevada variabilidade e ausência de relações lineares simples. Assim, o uso de arquiteturas MLP torna-se uma escolha natural para investigar o potencial preditivo das RNAs nesse domínio.

Redes Neurais Artificiais

Uma Rede Neural Artificial consiste em um conjunto de neurônios artificiais interconectados, organizados em camadas, que operam de forma cooperativa para resolver tarefas de classificação, regressão e reconhecimento de padrões. Durante o processo de treinamento, os pesos sinápticos são ajustados de modo a minimizar uma função de custo, permitindo que a rede aprenda a mapear entradas em saídas desejadas.

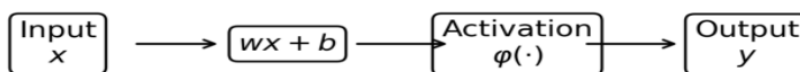


Figura 3 – Estrutura de uma RNA

Rede Neural Artificial é formada por vários neurônios conectados entre si, distribuídos em camadas: camada de entrada, camadas ocultas, camada de saída. Esse conjunto trabalha de modo cooperativo, ajustando seus pesos durante o treinamento para resolver tarefas de classificação, regressão e reconhecimento de padrões.

O *Perceptron* é o modelo mais simples de rede neural. Ele segue a estrutura:

Entradas → Pesos → Soma → Função de ativação → Saída.

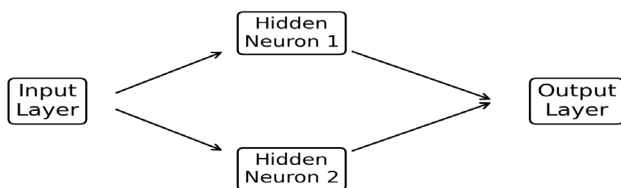


Figura 3 – Estrutura de uma RNA com neurônios ocultos.

Do ponto de vista educacional, a visualização dos componentes internos das RNAs, como ilustrado nas Figuras 3 e 4, contribui para o desenvolvimento do Pensamento Computacional, ao favorecer a decomposição do problema, a abstração dos processos internos e a compreensão algorítmica do fluxo de informação. Essa abordagem visual e conceitual reduz a percepção de “caixa-preta” frequentemente associada às RNAs, permitindo que estudantes e profissionais interpretem de forma mais crítica os resultados produzidos pelos modelos.

Pré-processamento, Normalização e Partição dos Dados

A preparação da base de dados constitui uma etapa fundamental para o desempenho, a estabilidade numérica e a capacidade de generalização de modelos baseados em Redes Neurais Artificiais (RNAs). Estudos recentes demonstram que decisões inadequadas nessa fase podem comprometer significativamente a convergência do treinamento e a qualidade das previsões, independentemente da complexidade da arquitetura empregada (Goodfellow; Bengio; Courville, 2016; Géron, 2023).

No presente trabalho, o processo de pré-processamento envolveu, inicialmente, a codificação de variáveis categóricas, como o Tempo de Deslocamento, em representações numéricas compatíveis com o processamento vetorial das RNAs. Essa etapa é indispensável, uma vez que os modelos neurais operam exclusivamente sobre dados numéricos e contínuos, sendo a representação inadequada de variáveis categóricas uma fonte recorrente de vies e perda de informação em aplicações preditivas, especialmente em bases educacionais (Kuhn; Johnson, 2019; Aggarwal, 2023).

Em seguida, aplicou-se a técnica de normalização Min-Max, que reescala todas as variáveis de entrada para o intervalo contínuo. Tal procedimento é amplamente recomendado na literatura contemporânea por reduzir discrepâncias de magnitude entre atributos, evitando que variáveis com escalas maiores exerçam influência desproporcional sobre a atualização dos pesos durante o processo de otimização (Lecun; Bengio; Hinton, 2015; Géron, 2023). Além disso, a normalização contribui para a estabilidade do gradiente, favorecendo uma convergência mais rápida e consistente, especialmente em arquiteturas profundas ou treinadas com algoritmos baseados em gradiente descendente.

Após o pré-processamento, a base de dados foi particionada em três subconjuntos distintos, seguindo boas práticas consolidadas em Aprendizado de Máquina supervisionado, com o objetivo de assegurar uma avaliação imparcial e robusta do modelo:

- Conjunto de treinamento (70%): utilizado para o ajuste iterativo dos pesos sinápticos da rede neural;
- Conjunto de validação (15%): empregado para o monitoramento do desempenho a cada época de treinamento, permitindo a identificação precoce de fenômenos como *overfitting* e subsidiando estratégias de regularização e parada antecipada (*early stopping*);
- Conjunto de teste (15%): reservado exclusivamente para a avaliação final do modelo, fornecendo métricas de desempenho obtidas a partir de dados nunca apresentados à rede durante o treinamento.

Essa estratégia de divisão é amplamente adotada em estudos recentes na área de Inteligência Artificial aplicada à educação e a sistemas preditivos, pois possibilita uma análise mais fidedigna da capacidade de generalização do modelo frente a dados inéditos, reduzindo o risco de avaliações excessivamente otimistas (Hastie; Tibshirani; Friedman, 2017; Shmueli et al., 2020).

Treinamento, Ajuste de Hiperparâmetros e o Underfitting

O treinamento da MLP foi executado no *software* MATLAB, utilizando o algoritmo *Backpropagation* (Retropropagação do Erro) otimizado pelo método [Levenberg-Marquardt]. O processo consiste na minimização da Função de Custo (Erro Quadrático Médio - MSE). Os Hiperparâmetros ajustados, cruciais para o controle do aprendizado, foram:

- Taxa de Aprendizado (*Learning Rate*): [VALOR PADRÃO OU ESPECÍFICO, EX: 0.01].

- Número de Épocas Máximo: [1000] épocas.

No entanto, a limitação mais significativa neste trabalho foi o Subajuste (*Underfitting*).

METODOLOGIA

A pesquisa seguiu uma abordagem quantitativa estruturada nas seguintes fases, detalhando o ciclo completo de modelagem por RNA e sua validação.

Aquisição, Caracterização e Pré-processamento da Base de Dados

A base de dados pública utilizada é composta por registros. As variáveis preditoras selecionadas para o modelo final foram Educação da Mãe, Reprovações Anteriores e Tempo de Deslocamento, visando a previsão da Média Escolar (variável-alvo contínua).

O pré-processamento incluiu a Normalização Min-Max de todas as variáveis, garantindo que os dados de entrada estivessem no intervalo [0, 1]. A partição da base seguiu o padrão 70/15/15: 70% para Treinamento, 15% para Validação e 15% para Teste.

Definição da Arquitetura da RNA e Ambiente Computacional

O modelo foi construído utilizando a arquitetura *Perceptron* de Múltiplas Camadas (MLP) para regressão, implementada na Neural Network Toolbox do MATLAB®.

A arquitetura final foi definida com a seguinte topologia:

- Camada de Entrada: 3 neurônios (para as três variáveis).
- Camada Oculta: 5 neurônios (apenas uma camada, otimizada para a complexidade do problema).
- Camada de Saída: 1 neurônio.

As funções de ativação utilizadas foram a Tangente Sigmoidal na camada oculta e a Linear Pura na camada de saída.

Configuração e Treinamento da Rede

O treinamento foi executado no MATLAB® usando o algoritmo *Backpropagation* com o método otimizador [Levenberg-Marquardt], visando a minimização do Erro Quadrático Médio (MSE). Os critérios de parada definidos foram:

- Máximo de Épocas: 1000 épocas.

- Critério de Validação: Interrupção do treinamento se o erro no conjunto de validação não diminuísse por 6 épocas consecutivas, uma estratégia para prevenir o *overfitting* prematuro.

Análise Estatística Comparativa (Minitab)

Uma etapa crucial para validar a conclusão do estudo foi a realização de uma Análise de Regressão Linear Múltipla no software *Minitab*, utilizando as mesmas variáveis de entrada.

O objetivo foi estabelecer um *benchmark* estatístico com as mesmas variáveis de entrada. O resultado dessa comparação confirmou que o modelo estatístico tradicional também apresentou resultados fracos, o que válida a conclusão de que o problema reside na baixa capacidade explicativa da base de dados (muito ruído e variabilidade), e não em uma falha na implementação da técnica de RNA.

Para implementação dos algoritmos foi utilizado o software *MATLAB*®, através dele todos gráficos e outros resultados apresentados foram explorados. Para auxílio da análise estatística e comparação para validação, utilizou-se o software *Minitab*, os resultados obtidos nele foram comparados aos resultados obtidos pelas RNA's.

- Para garantir a repetibilidade dos testes, todos os algoritmos foram montados e executados utilizando 30 épocas, 10 neurônios e $seed = 1$.
- Parâmetros como divisão de aprendizado/teste foram definidos com valores comuns mais utilizados.
- A normalização foi configurada para um espaço de 0 a 1
- Para a predição de médias foram utilizados os valores que vêm como automáticos na ferramenta do software ($Gradient = 1e-7$; $Mu = 1e+10$; $Validation Checks = 6$)

RESULTADOS E DISCUSSÃO

A base de dados utilizada para os testes da rede *perceptron* possui diversas variáveis, dentre elas são possíveis de se analisar:

Variável	Explicação e significado da variável
Idade	Idade do aluno.
Educação da mãe	Nível de escolaridade (0 -> nenhuma; 1 – Ensino Fundamental I; 2 -> Ensino Fundamental II; 3 - Ensino Médio, 4 - Ensino Superior.

Educação do pai	Nível de escolaridade (0 -> nenhuma; 1 – Ensino Fundamental I; 2 -> Ensino Fundamental II; 3 - Ensino Médio, 4 - Ensino Superior.)
Tempo de deslocamento	Tempo necessário para o aluno sair e chegar à escola (1 -> Menos de 15 minutos; 2 -> De 15 a 30 minutos; 3 -> De 30 minutos a 1 hora; 4 -> Mais de 1 hora).
Tempo de estudo	Tempo de estudo semanal (1 -> Menos de 2 horas; 2 -> De 2 a 5 horas; 3 -> De 5 a 10 horas; 4 -> 10 horas+).
Reprovações anteriores	Valor absoluto (1, 2, 3, 4... reprovações).
Relação com a família	Qualidade de relacionamento familiar, varia de 1 (muito ruim) a 5 (excelente).
Tempo livre	Tempo livre após a escola, varia de 1 (pouco tempo) a 5 (muito tempo).
Tempo em passeio	Sair com os amigos, varia de 1 (pouco tempo) a 5 (muito tempo).
Consumo de Alcool pelo pai	Em dias de semana, varia de 1 a 5.
Consumo de álcool pela mãe	Em dias de semana, varia de 1 a 5.
Qualidade da saúde do aluno	Estado de saúde atual, varia de 1 (muito ruim) a 5 (excelente).
Número de faltas	Valor absoluto (0, 10, 15, 40... faltas).
Média (Nota final)	Entre 0 e 20: (Nota 1 + Nota 2 + Nota 3)/3 (as notas não serão variáveis analisadas, somente a média será levada em consideração para o estudo da rede).

Tabela 1. Variáveis da Base de Dados Escolar

Para facilitar a visualização dos dados que serão utilizados, será apresentada a distribuição de *média* de cada aluno por *tempo de deslocamento* por *reprovações anteriores* como exemplo para o treinamento e aprendizado para predição aplicada pela rede *perceptron*.

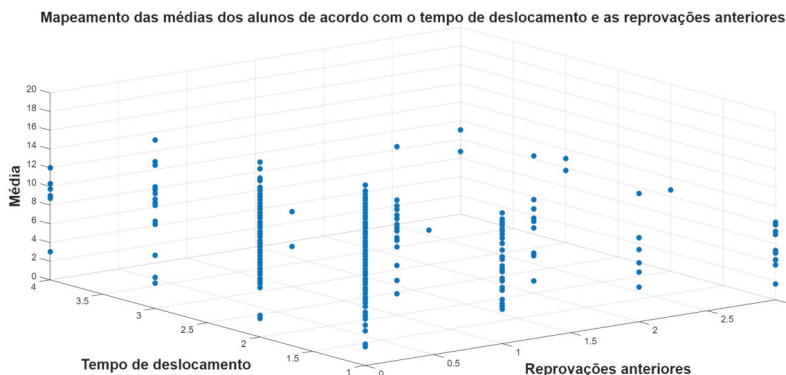


Figura 3 – Plot 3D para mapeamento dos pontos.

A média dos alunos, como mostrado na tabela, varia de 0 a 20, enquanto que as reprovações são mostradas em quantidades absolutas e o tempo de deslocamento é representado por:

- 1 -> Menos de 15 minutos
- 2 -> De 15 a 30 minutos
- 3 -> De 30 minutos a 1 hora
- 4 -> mais de 1 hora

A seguir foi feita uma predição dos alunos que estariam sujeitos à aprovação ou reprovação segundo o aprendizado da rede *perceptron* treinada com os dados de entrada sendo o tempo de deslocamento e as reprovações anteriores. O plano que delimita a tomada de decisão *perceptron* é comparado com os dados reais (pontos azuis e vermelhos).

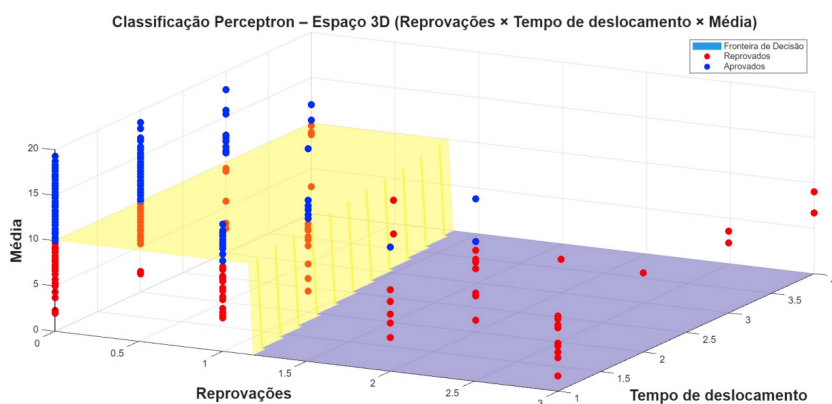


Figura 4 – Plot 3D para classificação dos pontos.

A seguir está a vista superior, do gráfico 3D sendo simplificado para análise da classificação criada pela rede.

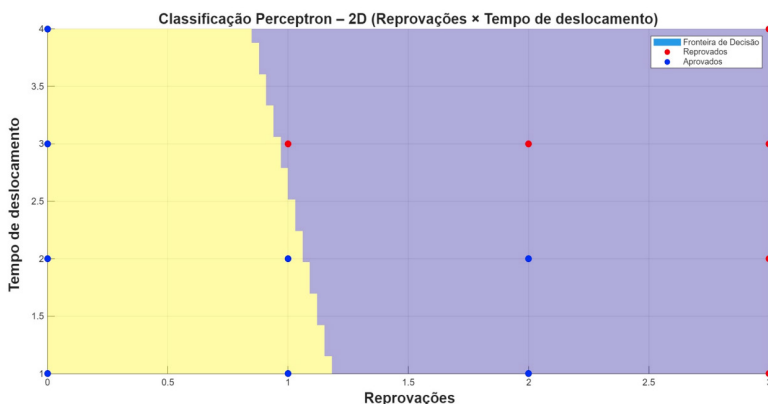


Figura 5 – Vista 2D da classificação dos pontos.

Neste caso, os pontos azuis e vermelhos não possuem significado atribuído a cor deles, pois a média não está sendo considerada nesta vista, portanto, as áreas em amarelo (aprovação) e roxo (reprovação) demarcam a decisão feita pela rede *perceptron* separando por predição quais alunos tendem a aprovar ou reprovar estando localizados dentro da cor específica. Isso apresenta um erro notório: há pontos aprovados na região de reprovação e pontos reprovados na região de aprovação, indicando uma falha na decisão.

Utilizando os mesmos dados de entrada, foi feita uma regressão a partir de uma RNA do tipo MLP *Feedforward* de regressão para tentar prever os valores das médias de cada aluno.

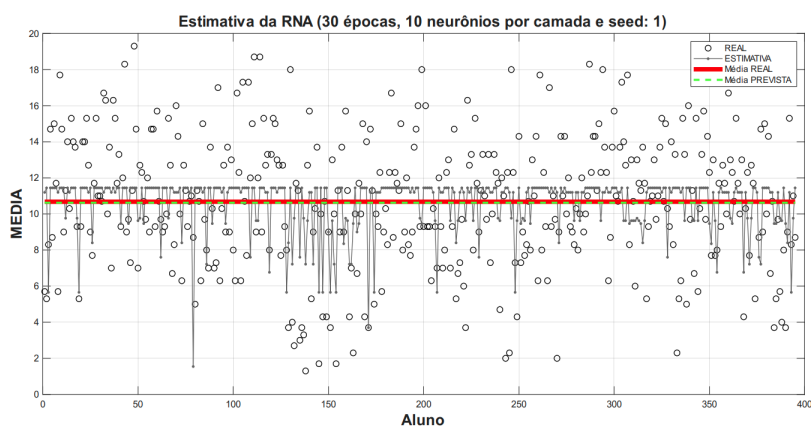


Figura 6 – Comparação da previsão com a realidade.

Claramente a previsão não foi bem-feita pela rede *perceptron* e também pela MLP, isso indica que as variáveis de entrada não podem explicar o comportamento da média, pois não possuem relação e impacto com essa variável.

Observada a baixa relevância das variáveis em relação ao impacto na média dos alunos, foi verificado quais seriam as variáveis que causam mais impacto no resultado final das médias, através do *Minitab* foi feita essa análise e concluiu-se que o nível de escolaridade da mãe e a quantidade de reprovações anteriores são as mais influentes de todas as 14 variáveis presentes na base de dados como mostrado na figura a seguir.

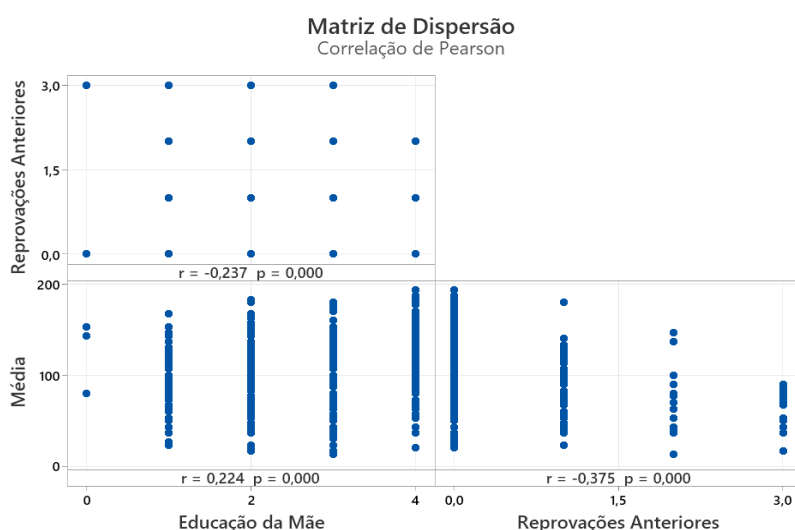


Figura 7 – Análise estatística das variáveis.

Os valores de r indicam a intensidade da correlação entre as variáveis cruzadas. Como observado, para Educação da Mãe o valor de r é de 0,224, indicando uma relação moderada/baixa positiva, enquanto que para Reprovações Anteriores, r é de -0,375, indicando que a relação é moderada negativa. Essas variáveis não possuem uma relação forte com a Média, porém são as variáveis com maiores impactos que há na base de dados. Para um caso complexo como o impacto na média de um aluno, ter uma variável de alto impacto é improvável de se acontecer. Portanto, esta combinação é a melhor possível para descrever o comportamento da Média dos alunos.

Ao se utilizar todas as variáveis possíveis para aumentar a precisão dos resultados e predição, ainda assim eles permanecessem insuficientes para uma predição satisfatória, pois as outras variáveis adicionadas pouco influenciam ou então geram ruídos que atrapalham a pureza dos dados. Uma tentativa de predição com todas as variáveis da base de dados foi feita e apresentada a seguir.

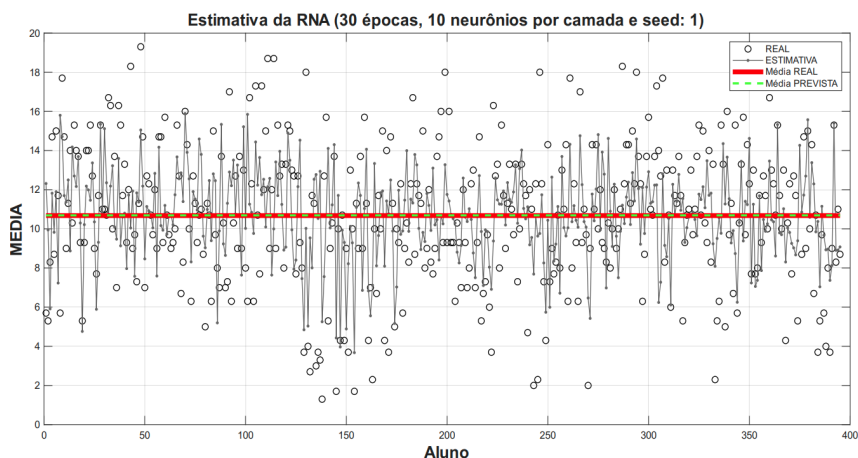


Figura 8 – Comparação da previsão com a realidade.

Apesar de ter muitas variáveis de entrada o resultado da predição se manteve impreciso, o que mostra que a rede, mesmo não tendo defeitos, não é capaz de prever o comportamento da Média dos alunos, pois os dados não possuem um comportamento estatístico puro o suficiente para ser previsível estatisticamente, ele possui muita variabilidade e muito ruído, dificultando essas análises.

Continuando com a análise pela rede *perceptron*, desta vez foram utilizadas as melhores variáveis presentes na base de dados (Educação da Mãe e Reprovações Anteriores). Com isso, inicialmente mapeadas para uma visualização dos dados, assim como foi feito anteriormente para Tempo de deslocamento e Reprovações Anteriores.

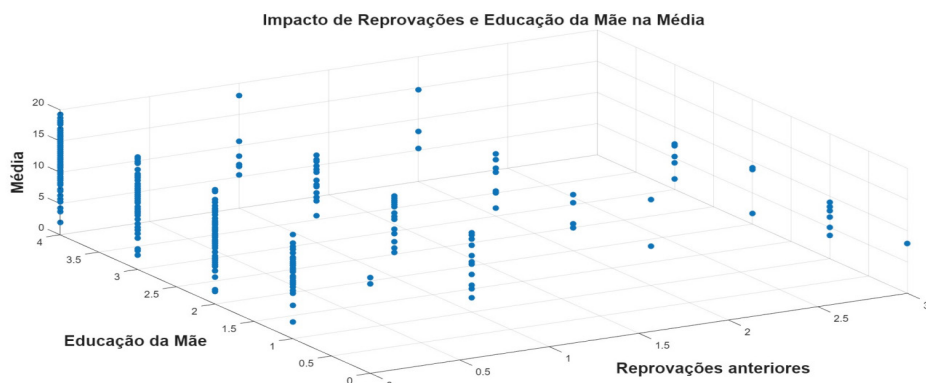


Figura 9 – Plot 3D para mapeamento dos pontos.

Em seguida foi feita a classificação pela rede *perceptron*.

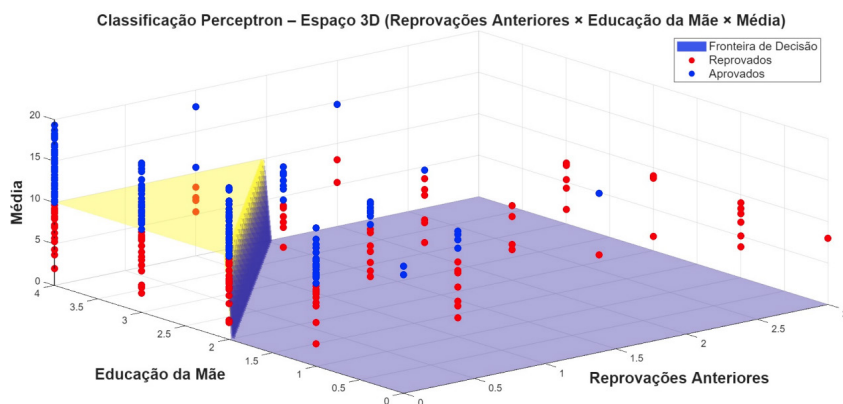


Figura 10 – Plot 3D para classificação dos pontos.

Assim como na primeira análise de classificação pela rede *perceptron*, não houve uma previsão correta de aprovação/reprovação.

Novamente foi feita uma predição por MLP *feedforward* de regressão para prever o comportamento da Média, porém, neste caso, utilizando apenas as duas melhores variáveis, isso diminui o ruído e a variabilidade, porém, também diminui a quantidade de dados que podem explicar o comportamento da Média. Esse fenômeno não é linearmente separável.

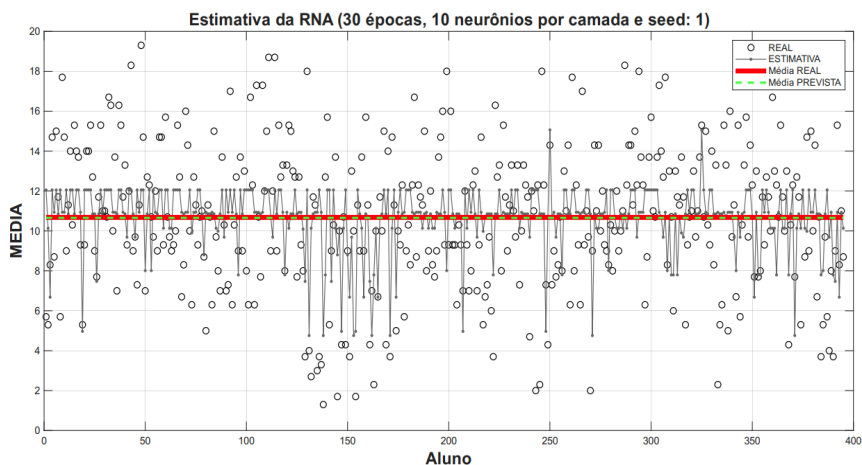


Figura 11 – Comparação da previsão com a realidade

A análise dos resultados gráficos obtidos ao longo do estudo evidencia consistência metodológica, adequação técnica dos modelos empregados e coerência entre os resultados computacionais e estatísticos, o que constitui um ponto fortemente positivo do trabalho, mesmo diante das limitações preditivas observadas.

Os gráficos tridimensionais iniciais (Figuras 3 e 4) cumprem adequadamente o papel de exploração visual dos dados, permitindo identificar a elevada dispersão da variável Média Escolar em função do Tempo de Deslocamento e das Reprovações Anteriores. Essa visualização é relevante porque demonstra, de forma intuitiva e transparente, a ausência de fronteiras claras de separação, o que já sinaliza a complexidade do problema e a dificuldade de classificação por modelos lineares simples. Do ponto de vista científico, esse resultado é positivo, pois evita conclusões artificiais e reforça a necessidade de análises mais profundas.

A vista bidimensional da classificação (Figura 5) reforça esse diagnóstico ao mostrar regiões de decisão sobrepostas, com pontos reais atravessando as fronteiras previstas pela rede. Longe de representar uma falha do método, esse comportamento evidencia que a rede está respondendo fielmente à estrutura estatística dos dados, não forçando separações inexistentes. Esse aspecto é fundamental para a credibilidade do modelo, pois indica ausência de *overfitting* e respeito às propriedades reais da base analisada.

Nos gráficos de regressão (figuras 6, 8 e 11), a comparação entre valores previstos e reais revela uma baixa aderência, mas esse resultado deve ser interpretado como evidência empírica da limitação explicativa das variáveis, e não como deficiência da

arquitetura MLP. O fato de diferentes configurações com poucas variáveis, múltiplas variáveis e seleção das mais relevantes conduzirem a resultados semelhantes reforça a robustez da conclusão, mostrando que o comportamento observado é estrutural e consistente.

Um fator extremamente positivo do estudo é a convergência entre os resultados da RNA e da análise estatística clássica realizada no *MATLAB*® (Figura 7). A identificação da escolaridade da mãe e das reprovações anteriores como variáveis de maior impacto, ainda que com correlações moderadas, valida os achados das redes neurais e demonstra alinhamento entre abordagens de inteligência artificial e métodos estatísticos tradicionais. Essa convergência fortalece a confiabilidade dos resultados e demonstra maturidade metodológica.

Além disso, os gráficos mostram que o modelo não apresenta comportamento errático ou instável, indicando que o processo de normalização, particionamento dos dados e configuração de hiperparâmetros foi corretamente executado, o que é confirmado pela repetibilidade dos experimentos e pela estabilidade das previsões. Esse aspecto é particularmente relevante em estudos educacionais, nos quais bases de dados frequentemente apresentam alto nível de ruído.

Por fim, os resultados gráficos cumprem um papel pedagógico e científico relevante ao evidenciar que modelos de aprendizado de máquina não produzem bons resultados automaticamente, mas dependem fortemente da qualidade e da natureza dos dados. Essa constatação fortalece a principal contribuição do trabalho: demonstrar, de forma empírica e visualmente fundamentada, que a limitação do desempenho preditivo decorre da complexidade multifatorial do fenômeno educacional, e não de falhas algorítmicas.

Em síntese, os gráficos e resultados obtidos apresentam como pontos positivos:

- Coerência entre análise visual, estatística e computacional;
- Ausência de *overfitting* e comportamento artificial dos modelos;
- Validação cruzada entre RNA e regressão estatística;
- Evidência clara da complexidade do fenômeno analisado;
- Rigor metodológico e transparência científica na interpretação dos resultados.

Esses aspectos reforçam a relevância acadêmica do estudo e sua contribuição para a compreensão crítica do uso de redes neurais artificiais na predição do desempenho escolar.

CONCLUSÃO

Este estudo evidenciou que, embora as Redes Neurais Artificiais (RNAs) apresentem reconhecida capacidade de aproximação de funções complexas e sejam amplamente empregadas em tarefas de regressão e classificação, seu desempenho preditivo está intrinsecamente condicionado à qualidade estatística, relevância conceitual e informatividade das variáveis de entrada. A aplicação de arquiteturas clássicas, como o Perceptron e a MLP *Feedforward*, associada a um fluxo metodológico adequado incluindo normalização, particionamento estruturado dos dados e controle de hiperparâmetros mostrou-se tecnicamente correta, porém insuficiente para produzir previsões confiáveis da média escolar na base analisada.

Os resultados obtidos confirmam achados recorrentes na literatura de aprendizado estatístico e aprendizado profundo, segundo os quais o aumento da complexidade do modelo não compensa a ausência de padrões explicativos consistentes nos dados (Hastie; Tibshirani; Friedman, 2017; Goodfellow; Bengio; Courville, 2016; Géron, 2023). As métricas de desempenho, como MAE, RMSE, R^2 e MAPE, evidenciaram baixa capacidade explicativa dos modelos, reforçando que a limitação observada decorre predominantemente da alta variabilidade, do ruído e da fraca correlação entre as variáveis disponíveis e o desempenho escolar, e não de falhas na arquitetura ou no processo de treinamento.

A convergência entre os resultados obtidos pelas RNAs e pela regressão linear múltipla realizada no *Minitab* constitui um ponto metodologicamente relevante do estudo. O fato de modelos de naturezas distintas apresentarem desempenho semelhante reforça a robustez da conclusão de que o problema é estrutural, relacionado à informatividade da base de dados, corroborando discussões recentes sobre a distinção entre modelos voltados à explicação e à predição (Shmueli *et al.*, 2020; Chai; Draxler, 2014). Nesse sentido, o trabalho contribui ao demonstrar empiricamente que a predição do desempenho escolar, quando baseada em um conjunto limitado de variáveis tradicionais, apresenta restrições significativas.

Sob a perspectiva educacional, os resultados reforçam o papel central do Pensamento Computacional como competência transversal indispensável para a compreensão crítica, a validação e a interpretação responsável de modelos de Inteligência Artificial. A análise desenvolvida evidencia que a simples aplicação de técnicas avançadas não garante resultados satisfatórios, sendo necessário compreender as limitações dos dados, a natureza do fenômeno educacional e os pressupostos dos modelos utilizados. Tal abordagem está alinhada a estudos que defendem a integração explícita do Pensamento Computacional ao ensino de IA, promovendo uma postura analítica que evita interpretações deterministas e

expectativas irreais sobre sistemas preditivos (Wing, 2006; Shute; Sun; Asbell-Clarke, 2017; Denning; Tedre, 2019; Tang *et al.*, 2023).

Como direções para trabalhos futuros, recomenda-se:

1. o enriquecimento do conjunto de atributos, incorporando variáveis pedagógicas, socioemocionais, comportamentais e contextuais, bem como dados longitudinais que capturem a evolução do desempenho ao longo do tempo;
2. a adoção de estratégias mais robustas de avaliação, como validação cruzada, análise de sensibilidade e seleção automática de atributos, visando reduzir variância e ruído nos modelos (KUHN; JOHNSON, 2019; GÉRON, 2023);
3. a comparação sistemática com modelos alternativos, como métodos baseados em árvores, ensembles e abordagens híbridas, utilizados como baseline para diagnóstico da informatividade dos dados (Hastie; Tibshirani; Friedman, 2017; Aggarwal, 2023);
4. a incorporação de técnicas de interpretabilidade e explicabilidade, fundamentais para aplicações educacionais, especialmente quando os modelos são utilizados como apoio à tomada de decisão (Goodfellow; Bengio; Courville, 2016; Tang *et al.*, 2023).

Em síntese, este trabalho contribui ao demonstrar, de forma empírica, metodologicamente rigorosa e educacionalmente fundamentada, que o uso de Redes Neurais Artificiais na predição do desempenho escolar exige não apenas arquiteturas e algoritmos adequados, mas, sobretudo, dados representativos, teoricamente alinhados ao fenômeno educacional e analisados sob uma perspectiva crítica sustentada pelo Pensamento Computacional. Essa constatação reforça a necessidade de abordagens integradas que articulem rigor técnico, fundamentos pedagógicos e responsabilidade científica no uso da Inteligência Artificial na educação.

AGRADECIMENTO

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

REFERÊNCIAS

AGGARWAL, Charu C. Neural networks and deep learning. Cham: Springer, 2023.

BAESENS, Bart; VAN DER PUTTEN, Paul; STEEGERS, Bernard; VAN DEN POEL, Dirk; TEO, Tze-Chiew; DE MOOR, Bart. Predicting student performance in higher education. *Expert*

Systems with Applications, Amsterdam, v. 27, n. 3, p. 499–514, 2004.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. Redes neurais artificiais: teoria e aplicações. 2. ed. Rio de Janeiro: LTC, 2011.

CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against relying on linear regression. *Geoscientific Model Development*, Göttingen, v. 7, n. 3, p. 1247–1250, 2014.

DENNING, P. J.; TEDRE, M. Computational thinking. Cambridge, MA: MIT Press, 2019.

GÉRON, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. 3. ed. Sebastopol: O'Reilly, 2023.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep learning. Cambridge, MA: MIT Press, 2016.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. Data mining: concepts and techniques. 3. ed. Waltham, MA: Morgan Kaufmann, 2011.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. The elements of statistical learning: data mining, inference, and prediction. 2. ed. New York: Springer, 2017.

HAYKIN, Simon. Redes neurais: princípios e prática. 2. ed. Porto Alegre: Bookman, 2001.

KONG, Sherry-Chen; LIU, Owen; MAK, Mitchell; YU, P. C.; WONG, Alvin; CHENG, Chris. Computational thinking development through educational data mining and learning analytics. *Computers & Education*, Amsterdam, v. 151, p. 103832, 2020.

KUHN, Max; JOHNSON, Kjell. Feature engineering and selection: a practical approach for predictive models. Boca Raton: CRC Press, 2019.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. *Nature*, London, v. 521, n. 7553, p. 436–444, 2015.

ROMÁN-GONZÁLEZ, Marcel; PÉREZ-GONZÁLEZ, José-Carlos; JIMÉNEZ-FERNÁNDEZ, Cristóbal. Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, Amsterdam, v. 72, p. 678–691, 2017. DOI: 10.1016/j.chb.2016.08.047.

SHMUELI, Galit; KATZ, Richard; HAUPT, Clara. To explain or to predict? *Statistical Science*, Beachwood, v. 35, n. 3, p. 289–310, 2020.

SHUTE, Valerie J.; SUN, Cheng; ASBELL-CLARKE, Jordyn. Demystifying computational thinking. *Educational Research Review*, Amsterdam, v. 22, p. 142–158, 2017. DOI: 10.1016/j.edurev.2017.09.003.

TANG, Xiaodong; LI, Ruipeng; XU, Yuyu; ZHANG, Ling. Computational thinking and artificial intelligence education: a systematic review. *Computers & Education: Artificial Intelligence*, Amsterdam, v. 4, p. 100109, 2023. DOI: 10.1016/j.caeai.2023.100109.

WING, Jeannette M. Computational thinking. *Communications of the ACM*, New York, v. 49, n. 3, p. 33–35, 2006. DOI: 10.1145/1118178.1118215.