

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL – UFMS
CAMPUS DE NOVA ANDRADINA - CPNA
CURSO DE ENGENHARIA DE PRODUÇÃO

FERNANDA GOMES MACIEL

**APLICANDO REDES NEURAIS ARTIFICIAIS PARA ESTUDO DE EVASÃO
NO ENSINO SUPERIOR**

NOVA ANDRADINA - MS
2023

FERNANDA GOMES MACIEL

**APLICANDO REDES NEURAIS ARTIFICIAIS PARA ESTUDO DE EVASÃO
NO ENSINO SUPERIOR**

Trabalho de Conclusão de Curso
apresentado como requisito para obtenção
do título de Bacharel em Engenharia de
Produção pela Universidade Federal de
Mato Grosso do Sul - UFMS.

Orientador: Prof^a Dr^a Laiane Andrade.

Coorientador: Prof^o Dr^o Marcelo Furlan.

NOVA ANDRADINA - MS

2023

MACIEL, Fernanda Gomes.

Aplicando Redes Neurais Artificiais para estudo de evasão de Ensino Superior, 2023.

43 f.

Orientador: Prof^a Dr^a Laiane Andrade.

Coorientador: Prof^o Dr^o Marcelo Furlan.

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal de Mato Grosso do Sul, Nova Andradina, 2023.

1. Engenharia de Produção.
 2. Redes Neurais Artificiais.
 3. Pesquisa Operacional.
 4. Evasão.
- I. Universidade Federal de Mato Grosso do Sul, Nova Andradina. II. Aplicando Redes Neurais Artificiais para estudo de evasão de Ensino Superior.

FOLHA DE APROVAÇÃO

APLICANDO REDES NEURAIS ARTIFICIAIS PARA ESTUDO DE EVASÃO NO ENSINO SUPERIOR

FERNANDA GOMES MACIEL

Trabalho de Conclusão de Curso
apresentado como requisito para obtenção
do título de Bacharel em Engenharia de
Produção pela Universidade Federal de
Mato Grosso do Sul – UFMS

APROVADO EM: 20/11/2023

BANCA EXAMINADORA

ORIENTADORA
LAIANE ANDRADE

COORIENTADOR
MARCELO FURLAN

MEMBRO DA BANCA
LUCAS HERMANN NEGRI

MEMBRO DA BANCA
AMÉLIA DE LORENA STANZANI

RESUMO

A evasão no ensino superior é um problema crítico que gera prejuízos socioeconômicos para a sociedade. Assim, o presente trabalho teve como objetivo buscar identificar os fatores que influenciaram a evasão de estudantes do curso de Engenharia de Produção da Universidade Federal do Mato Grosso do Sul do campus de Nova Andradina. O perfil de evasão foi investigado utilizando Redes Neurais Artificiais (RNA). A pesquisa foi realizada por meio da aplicação de questionários em estudantes que permaneceram e que evadiram do curso. Com os dados tratados, os mesmos foram inseridos em uma rede neural do tipo *Perceptron Multicamadas* (PMC). A fim de buscar o melhor modelo, a rede foi executada 20 vezes e os indicadores de desempenho da rede foram analisados. Como resultado, foi observado que o principal fator para a tomada de decisão por parte dos estudantes é a Escolha do Curso, com 100% de importância, e em segundo lugar, a mudança de cidade para frequentar a faculdade, com 78,6%. Outros fatores críticos, com grau de importância superior a 40%, foram a satisfação com o curso escolhido, a localização da faculdade em relação à residência na cidade da faculdade e a falta de base no ensino médio. Os resultados desse trabalho indicam fatores multidimensionais que podem influenciar na decisão de evasão e podem servir como ferramenta para gestores tomarem decisões mais assertivas acerca dos fatores críticos, assim reduzindo a taxa de evasão. Apesar do modelo ter sido aplicado especificamente ao curso de Engenharia de Produção do campus de Nova Andradina, a metodologia utilizada neste trabalho pode ser aplicada a outros cursos de graduação.

PALAVRAS-CHAVES: Engenharia de Produção; Redes Neurais Artificiais; Pesquisa Operacional; Evasão.

ABSTRACT

Dropping out from higher education is a critical issue that generates socioeconomic losses for society. Thus, this study aimed to identify the factors influencing dropout rates. Dropout profiles were investigated using Artificial Neural Networks (ANN) with students from the industrial engineering program at the Nova Andradina/MS campus. The research involved administering questionnaires to both students who remained in the program and those who dropped out. After processing the data, it was fed into a Multilayer Perceptron (MLP) neural network. To find the best model, the network was run 20 times, and performance indicators were analyzed. The results revealed that the primary decision-making factor for students is the Choice of Course, with 100% importance. When combined with other factors, the most significant weight is attributed to the choice of course. Secondarily, living arrangements in the city, including students who had to relocate to attend college, accounted for 78.6%. Other critical factors, with importance levels exceeding 40%, included satisfaction with the chosen course, the location of the college in relation to the student's residence in the city, and a lack of foundation in high school. The findings suggest multidimensional factors influencing dropout decisions and can serve as a tool for managers to make more informed decisions about critical factors, thus reducing dropout rates. Although the model was specifically applied to the industrial engineering program at the Nova Andradina campus, the methodology used in this study can be applied to other undergraduate programs.

KEYWORDS: Production Engineering; Artificial neural networks; Operational Research; Evasion.

SUMÁRIO

1 INTRODUÇÃO	9
2 FUNDAMENTAÇÃO TEÓRICA	11
2.1 REDES NEURAIS ARTIFICIAIS	11
2.2 RNA NA ENGENHARIA	18
2.3 EVASÃO NO ENSINO SUPERIOR	20
2.4. ANÁLISE DA EVASÃO POR MEIO DAS RNAs	22
3 MÉTODO DE PESQUISA	24
3.1 PROCEDIMENTOS DE COLETA DE DADOS	24
3.2 APLICAÇÃO DOS QUESTIONÁRIOS	25
3.3 PROGRAMAÇÃO DA RNA	27
4 RESULTADOS E DISCUSSÕES	29
4.1 SELEÇÃO DO MODELO	29
4.2 PREVISÃO DO PERFIL DE EVASÃO	33
5 CONCLUSÃO	39
REFERÊNCIAS	41

1 INTRODUÇÃO

A busca pelo desenvolvimento de novas tecnologias vem sendo há muito tempo objeto de estudo de pesquisadores e cientistas. As publicações pioneiras mais relevantes em redes neurais datam da década de 40 desenvolvidas por: McCulloch e Pitts (1943), Hebb (1949). Estas publicações introduziram o primeiro modelo de redes neurais simulando “máquinas”, o modelo básico de rede de auto-organização, e o modelo Perceptron de aprendizado supervisionado, respectivamente. Em meados de 1990, houve uma grande ampliação de pesquisas sobre as Redes Neurais Artificiais (RNA), com aplicações variadas entre às quais pode-se citar: avaliação de imagens de satélites, identificação de anomalias em imagens médicas, classificações de padrões, reconhecimento facial, controle de trens de grande velocidade, controle de aparelhos eletrônicos, além de permitir resolução de problemas das mais diversas áreas do conhecimento (SILVA, p.21, 2010).

A utilização das RNAs no contexto das engenharias têm se tornado cada vez mais comuns. Isso porque, quando a mesma é aplicada para a resolução de problemas tem, via de regra, tornado esse processo de busca de soluções mais eficientes (SILVA, et al., 2022). Uma característica potencialmente interessante dessa ferramenta é sua capacidade de auxiliar em processos de pesquisa, permitindo: adaptar por experiência, ampliar a capacidade de aprendizado, promover generalizações, organizar dados, entre outros. Assim, as redes neurais artificiais podem ser utilizadas para os mais diversos fins, com o objetivo de encontrar soluções para problemas complexos.

Um problema complexo e crucial enfrentado pelas universidades é o da evasão escolar, o qual causa perdas sociais, econômicas, políticas, acadêmicas e financeiras (MARTINHO; NUNES; MINUSSI, 2013a). O problema da evasão escolar ocorre em vários países, por isso, a busca por compreender o problema raiz da questão é muito importante. Principalmente, nas Instituições de Ensino Superior (IES), pois muitos cursos podem ser fechados, caso se tenha um índice de evasão elevado, o que gera prejuízos e complicações (MARTINHO; NUNES; MINUSSI, 2013a). Assim, é importante buscar compreender e prever os fatores que levam à evasão, para que se

possa identificar suas causas na tentativa de propor soluções em prol de promover a permanência dos estudantes.

A utilização de métodos e técnicas para prever a evasão entre os estudantes antes que ela ocorra pode ser um método preventivo interessante. Nesse sentido, as RNAs podem ser utilizadas na identificação de fatores que caracterizam a evasão dos cursos de graduação. Os autores Martinho, Nunes e Minussi (2013b) estudaram o perfil de estudantes em Risco de Abandono Escolar utilizando como ferramenta as RNAs. A análise foi realizada com estudantes do Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, das Escolas Superiores de Tecnologia em Automação e Controle Industrial, Controle de Obras, Sistemas de Internet, Redes de Computadores e Secretaria Executiva. O modelo obtido por tais autores se mostrou satisfatório com acurácia global superior a 76%, possibilitando a identificação precoce de discentes passíveis de evasão. Baseando-se no fato de que um dos fatores que pode contribuir para a evasão é a reprovação, Oliveira (2022) construiu utilizando as RNAs um modelo de previsão do desempenho dos alunos na disciplina de programação, no curso de Bacharelado em Sistemas de Informação da UFRN. Seu estudo contribui para prever o rendimento dos estudantes na referida disciplina com uma acurácia de até 79%.

Tendo em vista o exposto acima, o objetivo deste estudo é utilizar Redes Neurais Artificiais para compreender quais são as características e fatores que influenciam a evasão escolar dos estudantes de graduação em Engenharia de Produção do Campus de Nova Andradina - MS. Espera-se que o resultado desse estudo auxilie gestores na tomada de decisões a fim de melhorar os indicadores relacionados à retenção acadêmica por meio de políticas educacionais em suas instituições.

De forma a cumprir o objetivo proposto o texto do presente trabalho foi estruturado da seguinte forma: a próxima seção abordará as fundamentações teóricas sobre os conceitos de redes neurais artificiais. Na terceira seção são apresentados os métodos utilizados na elaboração deste trabalho. Na seção quatro serão apresentados os resultados obtidos e as discussões perante os dados coletados. Por fim, na quinta seção é apresentado a conclusão do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 REDES NEURAIS ARTIFICIAIS

As RNAs foram construídas baseadas em modelos do sistema nervoso biológico e também do cérebro humano (SILVA, p.33, 2010). Nessa arquitetura, se considera um neurônio artificial o qual é composto por sete itens, sendo eles: os sinais de entrada; os pesos sinápticos; o combinador linear; o limiar de ativação; o potencial de ativação; a função de ativação; é o sinal de saída, conforme mostra o Quadro 1.

Quadro 1 - Itens do Neurônio Artificial

Item	Descrição
Sinais de Entrada	São os dados obtidos no meio externo e que são representados por meio de valores.
Pesos Sinápticos	São os valores que tem como objetivo ponderar cada uma das variáveis de entrada, possibilitando quantificar sua relevância em relação a função do neurônio
Combinador Linear	É a junção dos sinais de entrada que foram ponderados pelos seus pesos sinápticos, e resultará em um valor do potencial de ativação
Limiar de Ativação	É uma variável que determina qual será o valor apropriado para o resultado obtido pelo combinador linear, gerando um disparo sentido a saída do neurônio
Potencial de Ativação	É o valor resultante da diferença do combinador linear e o limiar de ativação
Função de Ativação	É responsável pela limitação da saída do neurônio, sendo entre um intervalo de valores que são definidos pela própria imagem funcional
Sinal de Saída	É o valor final resultante do neurônio em contato com um determinado conjunto de sinais de entrada

Fonte: SILVA (p. 34, 2010).

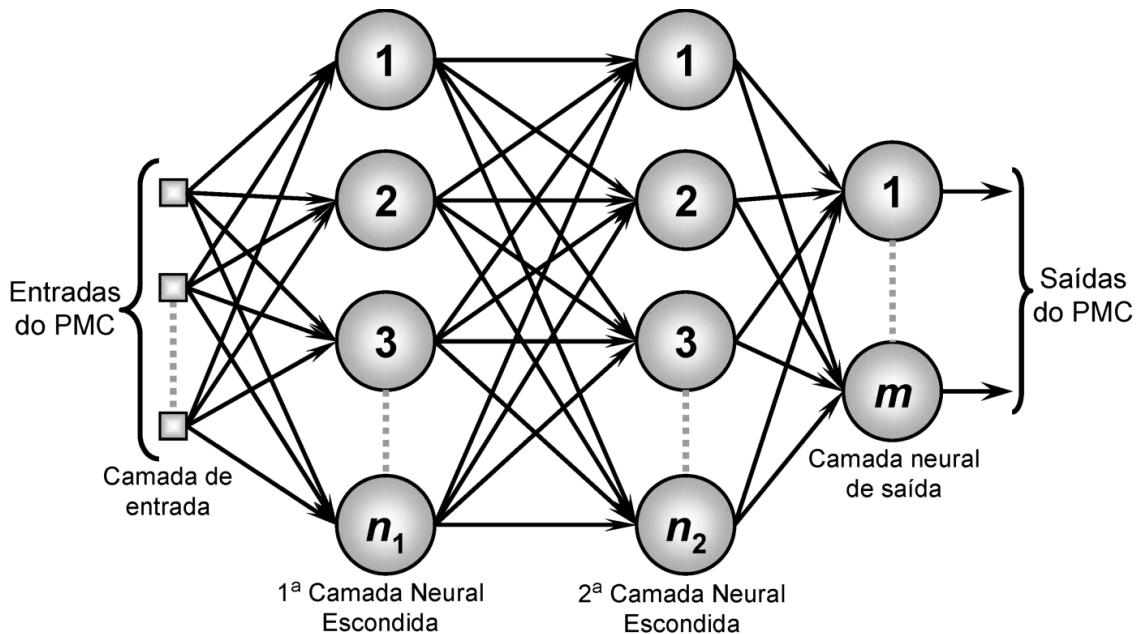
Segundo Matsunaga *et al.* (2012), as RNAs são compostas por conjuntos de neurônios artificiais que são parecidos com os biológicos, tendo a

capacidade de utilizar, calcular e armazenar dados por meio de aprendizagens (repetições) do processo, que são chamados de testes.

As RNAs são compostas por modelos matemáticos que se assemelham com o sistema nervoso central do ser humano, esses modelos são compostos por camadas internas, e incluso nessas camadas, há os neurônios artificiais, que normalmente variam sua quantidade conforme a necessidade da pesquisa. O tipo de RNA mais utilizado é a *Multiplayer Perceptron* (MLP) que é um *Perceptron Multicamadas* (PMC), ou seja, é uma rede que apresenta uma ou várias camadas (SILVA, p. 91, 2010).

Uma rede multicamadas, tal qual a PMC, é composta por uma camada de entrada, uma camada de saída e uma ou mais camadas intermediárias. A camada intermediária é também denominada como camada oculta (SILVA, p.91, 2010). Na Figura 1 é apresentado um modelo da PMC, o qual utiliza a arquitetura *feedforward* de multicamadas.

Figura 1 - Modelo de Rede *Perceptron Multicamadas* (PMC)



Fonte: Silva (p.94, 2010).

Para que ocorra a aprendizagem da rede (conjunto das camadas) é necessário que sejam definidas algumas funções, pois para cada necessidade há funções que trazem melhores resultados. Nesse sentido, é necessário definir um método de reescalonamento para a camada de entrada, uma função de ativação para a camada oculta/intermediária, uma função de erro para a

camada de saída, uma função de ativação para a camada de saída, e por último uma função de treinamento.

Para o método de reescalonamento das variáveis na camada de entrada, quatro opções são possíveis, sendo elas: Padronizado; Normalizado; Normalizado Ajustado; Nenhum. O método padronizado utiliza a média e o desvio padrão para o reescalonamento das variáveis de entrada. Enquanto que o Normalizado padroniza as variáveis de forma que todas fiquem entre o valor 0 e 1; Já a função do tipo Normalizado Ajustado ajusta as variáveis de modo que às mesmas apresentem valores entre -1 e 1 (SILVA, 2010).

A função de ativação desempenha um papel fundamental na determinação da saída de um neurônio ou unidade de processamento. Duas funções de ativação comumente usadas na camada oculta de uma rede neural artificial são a tangente hiperbólica e a curva sigmóide.

A função tangente hiperbólica é uma função matemática que mapeia um valor de entrada em um intervalo entre -1 e 1. Sua forma gráfica possui uma faixa de saída estendida, abrangendo valores negativos (FLECK *et al.*, 2016). A função é simétrica em relação ao ponto (0, 0), o que significa que quando a entrada é próxima de zero, a saída será próxima de zero. À medida que a entrada se afasta de zero, a saída aumenta em magnitude, saturando em -1 para entradas muito negativas e saturando em +1 para entradas muito positivas (FLECK *et al.*, 2016). A função tangente hiperbólica é uma função não linear que pode ajudar a introduzir não linearidades nas camadas ocultas das redes neurais, permitindo a modelagem de relações mais complexas entre os dados (SILVA, 2010). Por outro lado, a curva sigmóide mapeia um valor de entrada em um intervalo entre 0 e 1. Sua forma gráfica é uma curva em forma de "S", que aumenta suavemente do mínimo (0) até o máximo (1).

A função sigmóide é caracterizada por uma região próxima a 0,5, onde a função tem uma resposta mais sensível a pequenas mudanças nos valores de entrada (FLECK *et al.*, 2016). A função sigmóide é amplamente usada em redes neurais devido à sua capacidade de mapear valores de entrada em probabilidades, o que é útil em problemas de classificação binária. Tanto a função tangente hiperbólica quanto a curva sigmóide são funções não lineares e diferenciáveis, o que as torna adequadas para uso como funções de ativação na camada oculta das redes neurais artificiais. Essas funções permitem que as

redes neurais aprendam relações não lineares complexas entre os dados e ajustem os pesos durante o processo de treinamento, permitindo a realização de tarefas de aprendizado mais sofisticadas (SILVA, 2010; FLECK *et al.*, 2016).

Na camada de saída existem a função de ativação e a função de erro. A função de ativação é uma parte essencial de uma RNA, responsável por introduzir a não linearidade nas saídas dos neurônios. Ela determina a saída do neurônio com base na soma ponderada dos sinais de entrada. Há dois modelos comuns de função de ativação: Identidade e Softmax (SILVA, 2010).

A função de ativação Identidade, retorna a entrada exatamente como ela é, sem alterações. Isso significa que a saída do neurônio será igual à soma ponderada dos sinais de entrada. A função Identidade é frequentemente usada em problemas de regressão, nos quais a saída desejada é um valor contínuo. Como a função Identidade não introduz não linearidade, ela é geralmente utilizada na camada de saída de redes neurais de regressão (SILVA, 2010; FLECK *et al.*, 2016).

A função de ativação Softmax é amplamente utilizada em problemas de classificação multiclasse. Ela é aplicada à camada de saída da rede neural e produz uma distribuição de probabilidades sobre as diferentes classes. A função Softmax normaliza as saídas dos neurônios, garantindo que a soma das probabilidades seja igual a 1 (FLECK *et al.*, 2016). Ela transforma a saída linear em uma probabilidade relativa entre as classes, permitindo que a rede neural faça previsões sobre a classe mais provável para uma determinada entrada. A função é calculada aplicando a exponencial a cada saída e, em seguida, normalizando pela soma exponencial de todas as saídas (FLECK *et al.*, 2016).

As duas funções de erro comumente usadas nas RNAs são: Entropia Cruzada e Soma dos Quadrados. A função de erro de Entropia Cruzada é amplamente usada em problemas de classificação, especialmente quando se utiliza a função de ativação Softmax na camada de saída. Ela mede a diferença entre a distribuição de probabilidade prevista pela rede neural e a distribuição de probabilidade real dos rótulos de treinamento (SILVA, 2010; FLECK *et al.*, 2016). A entropia cruzada penaliza mais fortemente as previsões erradas e, portanto, é útil para treinar a rede neural a produzir probabilidades mais precisas para cada classe (FLECK *et al.*, 2016). A fórmula da entropia cruzada varia dependendo do número de classes e do formato dos rótulos, mas em

essência, ela envolve o cálculo do logaritmo da probabilidade prevista para a classe correta (FLECK *et al.*, 2016).

A função de erro da Soma dos Quadrados, é comumente usada em problemas de regressão, nos quais a saída desejada é um valor contínuo. Ela mede a média dos quadrados das diferenças entre as saídas previstas pela rede neural e os valores reais dos rótulos (FLECK *et al.*, 2016). O objetivo da função é minimizar a discrepância entre as previsões da rede e os valores reais. A fórmula envolve o cálculo do quadrado da diferença entre a saída prevista e o valor real para cada exemplo de treinamento seguido pelo cálculo da média desses quadrados (FLECK *et al.*, 2016).

Em resumo, a função de ativação determina a saída de um neurônio, introduzindo não linearidades nas redes neurais, enquanto a função de erro mede a discrepância entre as previsões da rede e os valores reais. A escolha da função de ativação e da função de erro depende do tipo de problema a ser resolvido e do formato dos dados de entrada e saída.

E por fim, a escolha da função de treinamento em redes neurais, refere-se ao processo de ajuste dos pesos e dos neurônios da rede para minimizar o erro e melhorar o desempenho do modelo (SILVA, 2010). É durante o treinamento que a rede neural aprende a mapear corretamente os padrões nos dados de entrada para as saídas desejadas. Uma das abordagens mais comuns para treinar as RNAs é o método de otimização chamado Descida de Gradiente. Nesse método, a rede neural é atualizada interativamente, ajustando os pesos e nós na direção oposta ao gradiente da função de erro em relação aos parâmetros. Em outras palavras, a descida de gradiente procura encontrar o mínimo da função de erro movendo-se em pequenos incrementos ao longo da direção do gradiente negativo (SILVA, 2010).

Outra opção para otimizar o treinamento de redes neurais é o método de Gradiente Conjugado em Escala (GCE). O Gradiente Conjugado é um algoritmo iterativo usado para encontrar mínimos em funções de múltiplas variáveis. No contexto de redes neurais, o GCE é uma variação do Gradiente Conjugado projetada especificamente para problemas de aprendizado de máquina (SILVA, 2010). O GCE utiliza informações sobre as derivadas da função de erro em relação aos parâmetros da rede neural para ajustar os pesos e nós. Ele combina características do Gradiente Conjugado com uma

técnica de escalonamento para determinar tamanhos de passo otimizados. O GCE é conhecido por sua eficiência computacional e pela capacidade de lidar bem com redes neurais com muitos parâmetros (SILVA, 2010).

A escolha entre eles depende do problema específico, tamanho do conjunto de dados, complexidade da rede neural e requisitos de desempenho computacional (SILVA, 2010). Cada método tem suas vantagens e desvantagens, e a seleção adequada pode impactar a velocidade e eficiência do treinamento, bem como a qualidade das soluções alcançadas. Dentre as duas funções, a do tipo da Descida de Gradiente é a mais usual. Isso porque, ela auxilia na minimização do erro, e assim colabora com as adaptações dos pesos dos indicadores de entrada (RUMELHART *et. al.*, 1986; SILVA, 2010).

Após definidos as funções e o método que serão utilizados é necessário definir a porcentagem de partições dos casos observados, ou seja, definir a porcentagem de treinamento e de testes, normalmente é utilizado 70% de treinamento e 30% de testes (SILVA, 2010). A principal razão para essa divisão é avaliar o desempenho e a capacidade de generalização do modelo treinado. Durante o treinamento, a rede neural ajusta seus pesos e nós com base nos dados do conjunto de treinamento para aprender os padrões presentes nesses dados (SILVA, 2010). No entanto, a capacidade de generalização do modelo não pode ser avaliada apenas com base no desempenho nos dados de treinamento, pois isso pode levar a um superajuste (*overfitting*) (BEJANI; GHATEE, 2021).

O *overfitting* ocorre quando o modelo se ajusta excessivamente aos dados de treinamento, aprendendo padrões específicos dos dados de treinamento em vez de capturar relações mais gerais que se aplicariam a novos dados (BEJANI; GHATEE, 2021). Para verificar se o modelo é capaz de generalizar adequadamente, é necessário avaliá-lo em um conjunto de dados não visto anteriormente, chamado conjunto de teste. Com isso, utiliza-se a porcentagem mais adequada para cada conjunto de dados, sendo como já mencionado o mais amplamente usado o de 70%-30% (SILVA, 2010).

Com todos os passos definidos, passa-se para a execução do treinamento da rede. Tipicamente, a etapa de treinamento pode ocorrer de forma supervisionada ou não supervisionada, ficando a critério do executor

(SILVA, 2010). Usualmente, durante o processo de treinamento não há realimentação, assim esse processo de treino deve ser executado algumas vezes, sendo importante a não alteração de dados e de sua configuração durante a execução do treinamento. O processo de realimentação é possível, porém exige mais complexidade computacional (FLECK *et al.*, 2016).

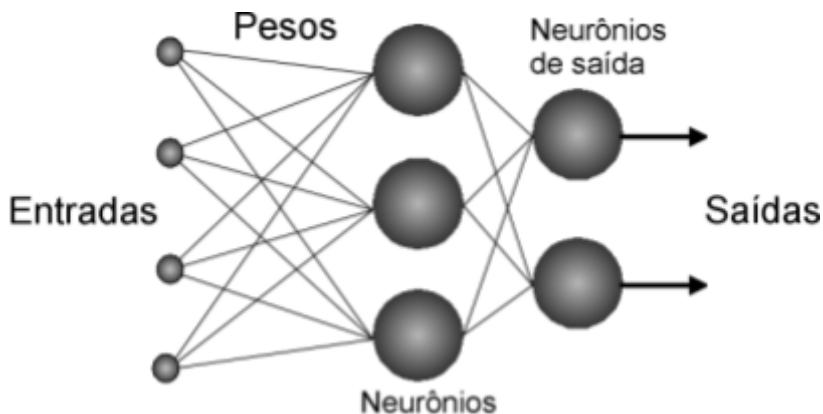
Após a construção da configuração e rodada sua primeira execução, é importante analisar a quantidade de neurônios que foi fornecida pela própria rede e comparar com a expressão matemática que permite estimar o número de neurônios na camada intermediária como mostra a Equação 1 (ACHEAMPONG; BOATENG, 2019).

$$N = (\text{quant. de indicador entrada} + \text{quant. de indicador saída}) / 2 \quad (\text{eq.1})$$

Segundo Lee, Lee, Yoon (2019), é de extrema importância ter a definição correta da quantidade de neurônios na camada intermediária, pois esse está diretamente relacionado ao desempenho da RNA. A determinação correta da quantidade de neurônios na camada intermediária de uma rede neural é crucial para o desempenho do modelo. Um número insuficiente pode levar a um modelo subajustado, enquanto um número excessivo pode resultar em superajuste. Encontrar o equilíbrio adequado requer experimentação e análise empírica para encontrar a quantidade ideal de neurônios que capture a complexidade dos dados sem comprometer a generalização do modelo (LEE; LEE; YOON, 2019).

Na execução da rede os indicadores de entrada representam as informações/dados, usualmente levantados na fase de pesquisa. Esses indicadores ganham pesos e estão conectados a cada um dos neurônios presentes na Rede. De forma subsequente cada neurônio, presente na camada intermediária, se relaciona com um neurônio de saída, como apresentado na Figura 2.

Figura 2 - Apresentação dos Neurônio na RNA



Fonte: Silva (2010).

A definição dos pesos, os quais são colocados em cada indicador de entrada, é realizada de maneira aleatória e independente pelo próprio *software* utilizado para a execução da RNA.

2.2 RNA NA ENGENHARIA

No que tange a enfoque na engenharia o uso das RNAs, frequentemente, tem como objetivo a melhoria nas condições das empresas e indústrias, seja nas áreas de: qualidade, segurança, meio ambiente, pesquisa operacional, educação, produção e outros.

Segundo Silva (p.120, 2010) as RNAs do tipo *Perceptron Multicamadas* são as mais utilizadas. Isso porque, as mesmas são abrangentes e conseguem se adaptar para solucionar problemas nas mais diversas áreas, fato que demonstra sua versatilidade. A seguir alguns exemplos da utilização dessa ferramenta na área da engenharia serão destacados.

Segundo Ishida (2021), a utilização da RNA na engenharia química ainda é pouca discutida em âmbito escolar, entretanto essa ferramenta possui potencial para auxiliar na resolução de diversos problemas relacionados a essa área. O autor Azarpour *et al.* (2017) estudou a aplicação de RNA's para prever o mecanismo de desativação de catalisadores em processos químicos. Vale destacar que o catalisador é uma substância de grande relevância em processos catalíticos, a qual auxilia a manter a alta produtividade e qualidade do produto. Gonzaga *et al.* (2009) desenvolveu um sensor virtual para previsão da viscosidade intrínseca durante o processo produtivo de Polietileno

tereftalato (PET). Apesar da viscosidade ser uma variável passível de ser medida diretamente problemas como fluxo do polímero não desejáveis, obstrução capilar e mau funcionamento da bomba, frequentemente ocorrem, e podem levar a medida incorreta dessa variável, que é de suma importância para qualidade do produto final. Na engenharia química utilização das RNAs não se limita apenas a ambientes acadêmicos, empresas como a Nouryon Industrial Chemicals utilizam redes neurais artificiais para identificar as previsões para as manutenções e para as substituições dentro do processo produtivo (ISHIDA, 2021).

Outra área de utilização é a da engenharia geotécnica, os estudos discutidos no artigo de Bicalho *et al.* (2022), destacam o potencial das redes neurais artificiais (RNAs) como uma alternativa às análises empíricas de regressão na estimativa de propriedades geotécnicas de solos em diversas regiões. Enquanto as correlações empíricas tendem a ser mais eficazes no local de origem dos dados, as RNAs demonstraram ser menos influenciadas pela variabilidade natural do ambiente geográfico, geológico e geotécnico de diferentes regiões. Isso se deve à capacidade das RNAs de aprender e se adaptar aos dados de treinamento, proporcionando resultados mais precisos e generalizáveis. Além disso, a técnica das RNAs pode ser continuamente atualizada conforme novos dados experimentais se tornam disponíveis, permitindo melhorias contínuas nos resultados. Esses achados destacam o potencial das RNAs como uma abordagem promissora para a estimativa de propriedades geotécnicas em diferentes contextos geográficos (BICALHO, *et al.*, 2022).

O trabalho de Liu e Chen (2023), destaca a aplicação das RNAs na Engenharia Civil. Com base na existência de incerteza e aleatoriedade nos fatores de risco da engenharia civil e utilizando uma rede neural convolucional, os autores apresentam um modelo de alerta precoce de risco. O modelo dos autores foi capaz de identificar e analisar os pontos de risco em projetos de engenharia por meio de imagens do local de construção, fornecendo alertas precoces para esses pontos de risco dentro de um prazo adequado. Os experimentos realizados comprovam a precisão do método em alertar sobre os riscos do projeto, alcançando uma taxa de prontidão dos alertas de 91,3% (LIU; CHEN, 2023).

2.3 EVASÃO NO ENSINO SUPERIOR

A evasão é o processo do estudante se matricular no curso e pedir a exclusão da matrícula ou transferência por qualquer motivo. Para a universidade é um fator de risco se a taxa de evasão for alta, principalmente porque gera diminuição de verba para investimento no curso, diminuição de auxílios e bolsas entre outros recursos. Esses fatores acabam por agravar ainda mais a situação do curso, pois levam a uma diminuição de discentes concomitante à diminuição de recursos disponíveis.

Dentre os aspectos que podem contribuir com a evasão escolar destaca-se: colocação profissional no mercado, vocação pessoal, dificuldades financeiras, tempo para estudar e tempo para ir às aulas, como os que mais podem influenciar na decisão de abandonar a graduação (TONTINI; WALTER, 2014).

Os autores Santos, Júnior e Ribeiro (2015), identificaram que um dos principais fatores para os estudantes continuarem no curso é a identificação com o curso escolhido e a estrutura do curso na universidade. Nesse sentido, os resultados do estudo dos autores demonstraram que os estudantes que evadiram tiveram como principais fatores a falta de conhecimento sobre o curso e ausência de orientação na escolha de carreira e profissão. Com relação aos fatores institucionais que levaram à evasão os autores apontam, como os mais relevantes: a falta de estrutura de apoio à graduação como os laboratórios, a má didática e desinteresse dos professores e a rotina cansativa e estressante do curso. Nesse sentido, foi pontuado a importância de uma boa estruturação do curso, para que o peso que tais fatores exercem na evasão seja diminuído. No que diz respeito aos fatores externos que culminam na evasão, os autores Santos, Júnior e Ribeiro (2015) destacam: a elevada exigência de dedicação ao curso; incompatível com outras necessidades do aluno, como a de trabalhar para pagar suas despesas. As ações corretivas propostas com base nos fatores de evasão identificados e nas sugestões dos participantes do estudo de Santos, Júnior e Ribeiro (2015) foram: orientação na escolha profissional, informação para os alunos quanto à disponibilidade de serviços de orientação psicológica, mais monitorias, condensação das aulas

em um período do dia, melhoria da didática dos professores e mais assistência estudantil.

Segundo Capelas (2014), existem diversos fatores que estão diretamente ligados ao processo de decisão do estudante de permanecer no curso de graduação. Entre os fatores internos à Instituição de Ensino Superior (IES) se destacam como os mais relevantes: a relação com o curso, os conteúdos aprendidos, a infraestrutura para o curso e o quadro docente (CAPELAS, 2014).

Segundo Cislaghi (2008), a dimensão que afeta o estudante a optar por abandonar a graduação, está relacionada à interpretação do mesmo sobre a qualidade do curso que ele está matriculado, bem como, a percepção de interesse por parte dos coordenadores pelo curso. Entretanto, Albuquerque (2008), pontua que a evasão está relacionada mais aos estudantes que optam por se matricular em um curso que não seja a sua primeira opção. Ainda acrescenta que, mesmo estes alunos se matriculando em um curso que não seja do seu interesse, há uma possibilidade do discente permanecer, caso se sinta pertencente ao curso, à faculdade e se notar que o curso pode lhe oferecer benefícios na sua carreira profissional.

Os autores Capelas (2014), Moraes e Theóphilo (2006) e Rollo e Pereira (2003), destacam que a percepção dos estudantes sobre o conhecimento e as habilidades técnicas dos docentes influenciam diretamente na satisfação com o curso, e consequentemente em sua permanência na IES. Nesse sentido, é denotada a importância de se ter docentes não apenas dotados de conhecimentos técnicos, mas também capacitados para ensinar.

Outro ponto importante destacado por diversos autores, e que influencia sobremaneira a permanência dos discentes, são as questões relacionadas à renda. Isso porque, muitas vezes é necessário que o estudante trabalhe para arcar com os custos financeiros de se manter fora de sua cidade natal. Tal fato, frequentemente, impacta de forma negativa no desempenho estudantil, o que ,usualmente, leva à evasão (MAYORGA; DE SOUZA, 2012). Assim, as políticas de permanência se mostram essenciais para evitar os processos de evasão. Isso porque, as políticas públicas além de contribuírem com os estudantes, garantem que o enfoque da universidade seja o discente. Tal fato, permite um acesso mais igualitário à educação contribuindo com a

vida financeira do estudante e colaborando com sua permanência (MAYORGA; DE SOUZA, 2012).

Além da parte financeira, a percepção do discente sobre o atendimento e os serviços oferecidos pela IES pode impactar sobre a decisão de permanência (MAYORGA; DE SOUZA, 2012). No que tange os serviços que podem ser oferecidos pelas instituições de ensino, Albuquerque (2008) afirma que programas sociais, bolsas de pesquisa e extensão e auxílios financeiros, permitem que se tenha uma maior possibilidade de permanência na instituição. Adicionalmente, os autores Mayorga e De Souza (2012), apresentam que programas de atendimento psicológicos, podem auxiliar os estudantes a permanecerem no curso. E ainda, acrescentam que essas práticas devem ser ofertadas desde seu início na graduação, pois os primeiros anos do estudante no ensino superior podem ser os mais difíceis.

2.4. ANÁLISE DA EVASÃO POR MEIO DAS RNAS

A aprendizagem de máquina, tal qual proporcionado pelas RNAS, é uma ferramenta promissora que pode auxiliar na construção de modelos preditivos para evasão escolar, fornecendo à gestão e autoridades responsáveis a possibilidade de adotar alternativas antecipadas a fim mitigar o abandono escolar (DEL BONIFRO *et al.*, 2020). Na literatura são reportados estudos de caso para previsão da evasão escolar usando RNAs tanto no cenário internacional como no cenário nacional, o que demonstra a relevância do tema e sua ocorrência de modo global.

No trabalho de Segura, Mello e Hernández (2022), foi abordado o estudo da evasão universitária, concentrando nas desistências ocorridas no final do primeiro ano de estudos. Foram usados dados de uma grande universidade europeia e considerado estudantes de diversas áreas, o estudo analisou 10 cursos universitários em cinco áreas distintas. Segundo Segura, Mello e Hernández (2022) os estudantes que desistiram (evadiram) obtiveram os piores resultados em ciências e ciências da saúde.. Com relação à importância dos preditores nos diferentes modelos de aprendizado de máquina, como esperado, a taxa de aprovação em matérias do primeiro semestre e as notas do primeiro semestre foram as variáveis mais importantes. Por fim, a opção de admissão, que expressa a preferência do aluno pelo curso que ele

estuda, é também um fator com alto grau de importância (SEGURA; MELLO; HERNÁNDEZ, 2022).

Utilizando uma Rede Neural Fuzzy-ARTMAP, uma rede com possibilidade de aprendizado contínuo Martinho, Nunes, Minussi (2013b) construíram um modelo para identificação precoce de grupos de alunos com maior risco de abandonar os cursos de Tecnologia em Automação e Controle Industrial, Obras de Controle, Sistemas de Internet, Redes de Computadores e Secretariado Executivo do Instituto Federal do Mato Grosso do Sul. A aplicação da RNA se mostrou satisfatória com uma acurácia global maior que 76% é um grau significativo de confiabilidade.

Oliveira (2022) abordou a previsão do desempenho dos alunos na disciplina de programação do curso de Bacharelado em Sistemas de Informação da Universidade Federal do Rio Grande do Norte. Os resultados dos experimentos mostraram que os algoritmos de Aprendizado de Máquina tiveram bons resultados, com destaque para o Naive Bayes, SVM e Redes Neurais Artificiais (RNAs), que apresentaram acurárias acima de 70% na previsão de rendimento dos estudantes. Para as RNAs, os autores observaram que um menor número de neurônios na camada oculta das RNAs resultou em um melhor desempenho na predição do desempenho dos alunos.

3 MÉTODO DE PESQUISA

A partir de dados fornecidos pela secretaria acadêmica do Campus de Nova Andradina, considerando a quantidade de discentes matriculados e excluídos de 2018 a 2022 no curso de engenharia de produção, foi possível constatar uma taxa de evasão média de 46,45% (UFMS, 2022). Esse dado revela-se como alarmante, demonstrando que quase metade dos estudantes que se matricularam evadiram do curso. A taxa elevada de evasão reforça a necessidade de estudos que auxiliem na previsão de fatores que levam a esse problema, visto que os gestores da instituição podem utilizar- se de tais resultados para tomar ações antecipadas a fim de mitigar a evasão. Assim, com a finalidade de traçar o perfil de evasão dos estudantes do curso de engenharia de produção do campus de Nova Andradina, questionários foram aplicados junto aos discentes que evadiram e os que permaneceram no curso. Os dados coletados foram tratados e utilizados para elaboração e treinamento da RNA utilizando o *software IBM SPSS Statistics v.28*.

3.1 PROCEDIMENTOS DE COLETA DE DADOS

A metodologia utilizada para coleta de dados foi baseada na classificação de Oliveira (2022), no qual a aquisição de dados ocorreu por meio da aplicação de formulários. A escolha deste método foi principalmente pela agilidade de compartilhamento. Sendo essa uma característica importante, pois muitos estudantes que evadiram o curso não são mais residentes da cidade de Nova Andradina/MS. Adicionalmente, os questionários são elaborados para serem preenchidos sem necessidade de auxílio, sendo assim, não há a necessidade de experiência ou treinamento, e por esse motivo, é uma das escolhas mais recorrentes no processo de levantamento e coleta de dados (GIL, 2002).

O questionário foi elaborado e criado na plataforma do Google Forms e compartilhado via *link* direto com os estudantes que evadiram e permaneceram no curso. As questões foram separadas em três grandes seções, mencionadas e explicadas a seguir:

- Seção 1: informações pessoais/socioeconômicas: onde foi realizada a coleta características sociodemográficas dos estudantes e julgadas como variáveis independentes;
- Seção 2: motivos da evasão: na qual o usuário elencou os principais motivos que favoreceram a evasão;
- Seção 3: grau de importância: avaliação, sob a visão do respondente, a influência de determinados fatores sob a sua evasão.

3.2 APLICAÇÃO DOS QUESTIONÁRIOS

A divulgação dos questionários foi realizada via *e-mail* e redes sociais. Desse modo, o público-alvo da pesquisa pôde responder ao questionário no horário ou local que achasse mais adequado, não havendo limitação geográfica.

A aplicação do questionário foi dificultada pela política de privacidade imposta pela universidade, onde, não é permitido compartilhar informações de estudantes e ex-estudantes para terceiros. Por esse motivo, foi necessário um trabalho de campo coletando números e *e-mails* dos próprios alunos do campus que tinham contato com os que saíram. Já para os alunos que permaneceram esse processo foi mais fácil, tendo em vista que os mesmos ainda se encontram matriculados no curso.

Com intuito de traçar um perfil mais completo dos estudantes do curso de engenharia de produção, optou-se pela coleta de dados tanto dos estudantes que evadiram como dos estudantes que permaneceram no curso. O período de coleta de dados ocorreu entre 10 de maio de 2022 a 31 de agosto de 2022, sendo obtidos 65 questionários preenchidos e válidos para a análise. Nesse procedimento de coleta dos dados o formulário foi preenchido por 20 estudantes que evadiram o curso e 45 estudantes que permaneceram no curso.

Para aplicação da RNA foram usadas 10 questões do questionário que foram consideradas mais importantes, e por conseguinte utilizadas no tratamento de dados. A escolha das questões relevantes para aplicação da RNA foi baseada na metodologia adotada por Santos, Júnior e Ribeiro (2015), sendo elas: gênero, se possui filho(s), estado civil, renda, se tinha conhecimento sobre as políticas de permanência da IES, escolha do curso,

satisfação com o curso, em qual lugar a pessoa morava antes do ingresso na IES, distância da residência a faculdade e a lacuna acerca de conhecimentos que deveriam ser fundamentados no ensino médio. O Quadro 2 apresenta um resumo relacionado ao perfil das questões selecionadas.

Quadro 2 - Perguntas Utilizadas na RNA

Variável da Pergunta	Medição
Gênero	Feminino, Masculino, Outro
Tem Filhos	Sim e Não
Faixa de Renda	Até um (1) salário mínimo Até dois (2) salários mínimos Entre dois (2) e cinco (5) salários mínimos Entre cinco (5) e dez (10) salários mínimos Mais que dez (10) salários mínimos
Estado Civil	Casado(a)/União Estável/Mora Junto com um(a) companheiro(a) Separado(a)/Divorciado(a) Solteiro(a)
Moradia - cidade	De 1 ao 5, de pouco importância a muita
Localização - Distância	De 1 ao 5, de pouco importância a muita
Escolha do Curso Equivocada	Sim e Não
Satisfação com o Curso	De 1 ao 5, de pouco importância a muita
Base de conhecimentos do Ensino Médio	Sim e Não
Política de Permanência	Sim e Não

Fonte: a autora (2023).

Para aplicação na rede neural os dados nominais foram tratados de forma a representarem valores numéricos. Assim, a variação binária de 0 e 1 foi utilizada para as perguntas cujas opções de respostas eram Sim e Não, respectivamente. Para a pergunta sobre Gênero, os dados foram codificados da seguinte forma: Feminino = 0, Masculino = 1 e Outros = 2. Na pergunta da faixa de renda os dados foram classificados de 1 ao 5, da sendo: um (1) salário mínimo = 1, Até dois (2) salários mínimos = 2, entre dois (2) e cinco (5) salários

mínimos = 3, Entre cinco (5) e dez (10) salários mínimos = 4 e Mais que dez (10) salários mínimos = 5.

As respostas sobre o estado civil foram codificadas de 1 a 3 sendo: Casado(a)/União Estável/Mora Junto com um(a) companheiro(a) = 1, Separado(a)/Divorciado(a) = 2 e Solteiro(a) = 3. Por fim, as perguntas sobre a moradia, localização (distância) e satisfação com o curso por já serem em escala numérica de 1 a 5, onde 1 representava que o fator era pouco importante e 5 muito importante, não precisavam de codificação.

3.3 PROGRAMAÇÃO DA RNA

Os métodos e funções escolhidos para a configuração da rede são apresentados no Quadro 3. Para a elaboração da Rede foi utilizado como apresentado na metodologia: o software IBM SPSS Statistics v. 28.

Quadro 3 - Método e Funções da RNA

Método e funções	Tipo Escolhido
Método de Ajuste de Variável	Normalizado Ajustado
Função Ativação	Tangente Hiperbólica
Função de Erro	Soma dos Quadrados
Função Ativação Camada Saída	Função Identidade
Função de Treinamento	Descida de Gradiente

Fonte: a autora (2023).

No Método de Ajuste de Variáveis, um dos modelos que é utilizado na RNA utiliza de algoritmos que fazem a leitura de variáveis entre -1 a 1, portanto se as variáveis estiverem normalizadas e ajustadas neste padrão, os modelos matemáticos da RNA conseguem ser executados com maior facilidade. Já para a Função de Ativação da camada oculta, a escolhida foi a Tangente Hiperbólica, por essa ser a mais indicada para os dados escolhidos. A função de treinamento foi escolhida com base nos tipos de dados da pesquisa, já a função de erro e a função ativação camada saída, foi escolhida pelo software com base nos dados. Na elaboração da RNA foram contabilizados 10 indicadores de entrada, 5 neurônios na camada oculta e um indicador de saída. Os indicadores de entrada e saída são especificados no Quadro 4.

Quadro 4 - Indicadores de Entrada e Saída

Indicadores	Entrada	Saída
Gênero	X	
Tem Filhos	X	
Faixa de Renda	X	
Estado Civil	X	
Moradia	X	
Localização - Distância	X	
Escolha do Curso Equivocada	X	
Satisfação com o Curso	X	
Base Ensino Médio	X	
Política de Permanência	X	
Evadiu ou Não do Curso		X

Fonte: a autora (2023).

A determinação do número de neurônios na camada oculta (5,5 neurônios) foi calculada, pela Equação 1, conforme apresentado na seção 2.1, arredondando para o menor valor inteiro, e assim a rede contou com 5 neurônios na camada oculta. O arredondamento para o menor valor inteiro após o cálculo do número de neurônios é necessário para se evitar problemas de superajuste e promover uma melhor capacidade de generalização.

4 RESULTADOS E DISCUSSÕES

4.1 SELEÇÃO DO MODELO

Para se obter a melhor configuração de rede possível o programa de treinamento da RNA foi executado 20 vezes. Assim, na Tabela 1 cada modelo descrito representa uma execução da rede, conforme será discutido a seguir, baseado nos indicadores de desempenho, a melhor rede para previsão de evasão foi selecionada. Na Tabela 1 é possível observar a comparação dos dados obtidos nas 20 execuções, onde os dados foram separados em colunas de teste e de treinamento.

A Curva Característica de Operação do Receptor, ou, do inglês, *Receiver Operating Characteristic Curve*, ou, simplesmente, Curva ROC (segunda coluna da Tabela 1), é um dos parâmetros que indicam o desempenho de um modelo. A técnica de avaliação ROC utiliza a proporção de duas métricas de avaliação baseadas em coluna única, a saber, a taxa de verdadeiros positivos (sensibilidade) e a taxa de falsos positivos (especificidade). A Curva ROC é útil porque fornece uma representação visual das compensações relativas entre os benefícios (refletidos pelos verdadeiros positivos) e os custos (refletidos pelos falsos positivos) da classificação em relação às distribuições dos dados (HE; GARCIA, 2009). O valor desse parâmetro pode variar de 0 a 1, sendo que o limite inferior indica (0,0) indica 100% de probabilidade de erro na previsão pelo modelo e o limite superior (1,0) indica 100% de probabilidade de previsões corretas pelo modelo. Assim, usualmente, modelos com valores mais próximos a 1,0 do referido parâmetro apresentam previsões com melhor qualidade. Observando a Tabela 1, os três modelos que mais se aproximaram do valor ideal (Curva ROC 1,0) foram os modelos 2, 3 e 15, com os valores: 0,948, 0,973 e 0,946, respectivamente. Porém, a escolha do melhor modelo não se baseia unicamente no valor de Curva ROC, assim outros parâmetros para avaliar o desempenho da rede devem ser analisados.

Tabela 1 - Comparação dos Modelos

Modelo	ROC Curve - <u>Área sob a curva</u>	Sum of Squares - Train	Sensitivity - Train	Specificity - Train	Accuracy (ACC) - Train	G-mean Train	Sum of Squares - Test	Sensitivity - Test	Specificity - Test	Accuracy (ACC) - Test	G-mean Test
1	0,854	6,996	100,000	33,300	79,200	0,577	2,492	83,300	80,000	82,400	0,816
2	0,948	4,604	90,300	85,700	88,900	0,880	2,114	92,900	66,700	85,000	0,787
3	0,973	3,886	94,100	64,300	85,400	0,778	1,851	90,900	83,300	88,200	0,870
4	0,827	7,490	96,700	40,000	77,800	0,622	3,055	93,300	20,000	75,000	0,432
5	0,873	4,734	96,700	54,500	85,400	0,726	3,868	86,700	66,700	79,200	0,760
6	0,832	6,771	96,300	24,000	74,400	0,481	3,118	94,400	62,500	84,600	0,768
7	0,719	8,748	95,800	35,300	70,700	0,582	2,890	100,000	33,300	91,700	0,577
8	0,833	5,727	97,400	33,300	85,100	0,570	3,260	100,000	54,500	72,200	0,738
9	0,922	3,973	91,400	86,700	90,000	0,890	2,812	100,000	40,000	80,000	0,632
10	0,907	4,110	96,200	58,300	84,200	0,749	3,744	89,500	50,000	77,800	0,669
11	0,859	4,336	91,400	83,300	89,400	0,873	3,864	100,000	12,500	61,100	0,354
12	0,853	6,601	84,400	58,300	77,300	0,701	3,298	100,000	37,500	76,200	0,612
13	0,861	5,646	94,300	45,500	82,600	0,655	2,867	100,000	66,700	84,200	0,817
14	0,643	10,069	100,000	0,000	68,800	0,000	3,305	100,000	0,000	70,600	0,000
15	0,946	4,969	97,100	81,300	92,200	0,888	1,519	90,000	75,000	85,700	0,822
16	0,830	8,472	100,000	37,500	79,600	0,612	2,239	100,000	50,000	87,500	0,707
17	0,930	3,666	96,800	84,600	93,200	0,905	3,103	92,900	42,900	76,200	0,631
18	0,877	5,935	96,800	53,300	82,600	0,718	2,555	92,900	60,000	84,200	0,747
19	0,840	5,284	88,900	58,300	79,500	0,720	4,155	94,400	50,000	80,800	0,687
20	0,820	7,290	90,600	46,700	76,600	0,650	2,235	84,600	80,000	83,300	0,823

Fonte: a autora (2023).

As colunas 3 e 8 da Tabela 1 mostram os valores de Soma dos Quadrados (*Sum of square train* e *Sum of square test*, respectivamente). A soma dos quadrados é utilizada no aprendizado da máquina como uma função de erro. A soma dos quadrados de treinamento é calculada no conjunto de dados de treinamento, enquanto que a soma dos quadrados de teste é calculada no conjunto de dados de teste. Sendo que o conjunto de dados de teste são aqueles que o modelo da rede neural não viu durante seu treinamento. Um valor baixo em comparação com os outros modelos, para ambos indicadores é desejável, sendo que um valor baixo para soma dos quadrados de treinamento indica que o modelo está se ajustando de modo satisfatório aos dados de treinamento e um valor baixo para soma dos quadrados de teste indica que o modelo está fazendo a generalização de maneira adequada.

Os elementos da quarta e quinta coluna da tabela representam a sensibilidade e especificidade dos dados do treinamento, enquanto que a nona e décima apresentam os mesmos parâmetros porém para os dados de teste, respectivamente. A sensibilidade em um modelo mede a proporção de casos positivos que foram identificados corretamente, enquanto que a especificidade mede a proporção de casos negativos que foram identificados de forma correta. Esses elementos: sensibilidade e especificidade, são inversamente proporcionais, graficamente ao plotarmos esses dois parâmetros tem-se a Curva ROC. Na prática em um modelo é desejável um equilíbrio entre esses dois indicadores (sensibilidade e especificidade), o ponto onde esse equilíbrio ocorre é aquele em que se tem maiores valores para taxas de verdadeiros positivos e menores valores para taxa de falsos positivos, ou seja, aquele que se encontra mais próximo do canto superior esquerdo do gráfico da Curva ROC. Nesse ponto, o modelo alcança um balanço aceitável entre minimizar falsos negativos (aumentando a sensibilidade) e minimizar falsos positivos (aumentando a especificidade).

Sabe-se que os dados de treinamento são usados para alimentar o algoritmo de treinamento da rede neural. Esses dados consistem em exemplos rotulados, ou seja, pares de entrada e saída esperada. Enquanto que os dados de teste são usados para avaliar o desempenho da rede neural após o treinamento. Esses dados também consistem em exemplos rotulados, mas são

dados que a rede não viu durante o treinamento. Assim, tipicamente os parâmetros de desempenho da rede são sempre avaliados tendo por base os dados de teste e de treinamento de forma conjunta. Porém para um bom ajuste é esperado que os parâmetros de teste sejam melhores, visto que os mesmos medem a capacidade da rede de fazer previsões em situações não vistas anteriormente.

A Acurácia (*Accuracy (ACC) - Train* e *a Accuracy (ACC) - Test*), tem o intuito de mostrar a exatidão do modelo, com valor de 0 a 100, sendo que quanto mais próximo de 100 melhor exatidão dos resultados. Pela Tabela 1 é possível observar que a maioria dos modelos executados pela rede neural teve acurácia acima de 70%.

O índice *G-mean* é um bom indicador de desempenho da RNA, pois diz respeito a análise de desempenho de aprendizado da rede, em relação aos dados, para que seja identificado problemas de desequilíbrio, ou seja, quando um conjunto de dados tem um valor menor do que outro conjunto em comparação (HE; GARCIA, 2009). Esse índice avalia o grau de viés indutivo, analisando como o modelo se comporta tanto em relação às taxas de positivos verdadeiros quanto às taxas de falsos positivos (LU *et al.*, 2016). Tanto para o *G-mean* de teste como para o *G-mean* de treinamento a intenção é maximizá-los. Assim, quanto mais próximo de 1 tiver esse parâmetro, melhor performance do modelo.

Observando os valores de *G-mean* para os modelos de previsão 2, 3 e 15, é possível observar que os modelos 2 e 15 apresentam uma queda ao se comparar o *G-mean* de treinamento com o *G-mean* de teste, tal fato sugere a ocorrência de um sobreajuste (*overfitting*). Isso significa que o modelo 2 e 15, se adaptam bem há um conjunto de dados mas se tornam ineficazes para a predição de novos dados, ou seja, muito bom na fase de treinamento mas nem tanto na fase de testes. O sobreajuste não foi observado no modelo 3, além disso o mesmo apresentou valores satisfatórios dos demais parâmetros de desempenho da rede neural, como ROC Curve, Soma dos Quadrados e Acurácia. Sendo assim, o modelo 3 foi selecionado para rede neural de previsão de evasão.

4.2 PREVISÃO DO PERFIL DE EVASÃO

Na Tabela 2 é apresentado um resumo do processamento dos casos, sendo indicado a porcentagem de treinamento, de testes e de excluídos, conforme os dados coletados pela aplicação dos questionários, do modelo 3.

Tabela 2 - Resumo do Processamento dos Casos

		N	Porcentagem
Amostra	Treinamento	48	73,8%
	Testes	17	26,2%
Válidos		65	100%
Excluídos		0	
Total		65	

Fonte: a autora (2023).

Os pesos sinápticos das ligações da rede neural do modelo selecionado iniciando nos indicadores de entrada, passando pelos neurônios da rede até o indicador de saída estão representados na Figura 3. As ligações sinápticas destacadas em azul são aquelas que possuem maior peso para o indicador de saída, enquanto que às ligações em cinza representam um menor peso (grau de importância) para o indicador de saída.

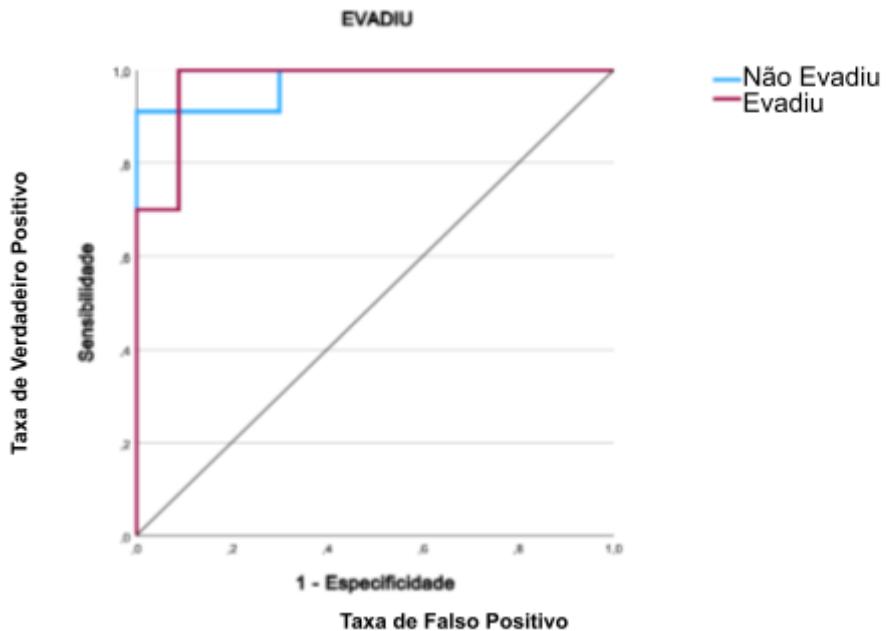
Figura 3 - Ponderações Sinápticas



Fonte: a autora (2023).

Na Figura 4 é apresentado a curva ROC que mostra a relação entre Taxa de Verdadeiro Positivo (sensibilidade) x Taxa de Falso Positivo (especificidade) para o modelo selecionado.

Figura 4 - Verdadeiro Positivo x Falso Positivo



Fonte: a autora (2023).

É possível observar pelo formato das curvas da Figura 4 que o modelo selecionado, tanto para curva de permanência (em azul) como para curva de evasão (vermelho), à medida que o limiar (*threshold*) diminui o mesmo ainda possui boa capacidade de detecção correta da resposta, ou seja, o mesmo tem menor probabilidade de levar a uma falsa identificação de possíveis perfis de evasão.

Os valores dos parâmetros estimados pelo modelo 3 são apresentados no Quadro 5.

Quadro 5 - Estimativas de Parâmetros

	Camada Oculta					
	Neurônio	N 1	N 2	N 3	N 4	N 5
C A M A D E E N T R A D A	(Viés)	-0,482	0,189	0,545	0,574	0,443
	Gênero	-0,015	-0,342	0,662	-0,432	0,402
	Renda	0,435	-0,215	-0,079	-0,156	-0,587
	Estado Civil	-0,358	-0,607	0,416	-0,121	0,472
	Tem filhos?	0,394	0,200	-0,170	-0,183	-0,228
	Falta Base Ensino Médio	-0,466	0,044	-0,437	0,311	0,485
	Escolha do Curso	0,342	-0,777	0,258	0,502	-0,223
	Política de Permanência	-0,016	-0,292	-0,472	-0,252	-0,282
	Satisfação com o Curso	-0,380	0,628	-0,278	0,069	-0,062
	Moradia - cidade	-0,387	0,258	0,688	-0,178	1.160
	Localização - distância	0,355	0,826	0,017	0,635	0,163

Fonte: a autora (2023).

A Tabela 4 apresenta o grau de importância de cada uma das variáveis independentes estudadas sobre a resposta avaliada (evasão) na forma bruta e normalizada. O grau de importância foi calculado pelo método de Sensibilidade do Gradiente (Gradient Sensitivity), o qual calcula a derivada da saída da rede em relação às entradas. As entradas com maiores derivadas têm maior impacto na saída. A análise é conduzida utilizando as amostras combinadas de treinamento e teste de forma automática pelo próprio software. A Tabela apresenta a importância e a importância normalizada de cada variável preditora.

Tabela 4 - Importância das Variáveis

	Importância	Importância Normalizada
Gênero	0,071	32,1%
Renda	0,076	34,4%
Estado Civil	0,059	27,0%
Tem Filhos	0,042	19,0%
Falta Base Escolar	0,092	41,6%
Escolha do Curso	0,220	100,0%
Política de Permanência	0,073	33,2%
Satisfação com o Curso	0,097	44,0%
Moradia - cidade	0,173	78,6%
Localização - distância	0,098	44,7%

Fonte: a autora (2023).

Pela análise da Tabela 4, é possível perceber que a escolha do Curso tem um elevado grau de importância sobre a decisão de evasão de 0,220 ou de 100% na importância normalizada. Tal fato ressalta a importância de projetos de extensão que sejam voltados para possíveis ingressantes tenham conhecimento sobre o curso, desde o mercado de atuação até o itinerário de formação. Assim, os estudantes poderão fazer escolhas mais assertivas. Nesse sentido, apesar de a Universidade Federal do Mato Grosso do Sul ter um programa institucional de visita às escolas (denominado: Vem para a UFMS), o mesmo tem foco na divulgação científica e não tem como público alvo estudantes do ensino médio, o que pode ser um caminho estratégico a ser adotado pelos gestores.

O segundo indicador com maior grau de importância foi o relacionado à moradia, que se refere a cidade de origem do estudante (0,173 ou 78,6% na importância normalizada). O elevado grau de importância desse indicador, aponta que a mudança de cidade para estudar acaba por dificultar a permanência desse estudante no curso. Outros indicadores considerados críticos, que possuíram grau de importância acima de 40% são: a satisfação com o curso escolhido, a localização da faculdade em relação à residência e a falta de base do ensino médio. Apesar de não ter sido analisada a correlação

entre os fatores que levam à evasão, possivelmente os fatores críticos podem apresentar interação entre si, por exemplo, a falta de conhecimentos básicos do ensino médio pode dificultar a aprendizagem, desestimulando o estudante diminuindo, assim, sua satisfação com o curso.

Já os fatores de: gênero (32,1%), estado civil (27,0%) e ter filhos (19,0%) apresentaram grau de importância menor na decisão de evadir. Isso indica que, embora esses fatores possam ter algum impacto, eles não são tão determinantes como outros aspectos mencionados acima.

5 CONCLUSÃO

O objetivo dessa pesquisa foi buscar a identificação de fatores e características que pudessem colaborar com a permanência dos estudantes no curso de Engenharia de Produção no campus de Nova Andradina da UFMS. E com a aplicação da Rede Neural Artificial (RNA), foi possível identificar os fatores que mais influenciam na tomada de decisão do estudante de evadir da Instituição de Ensino Superior.

A RNA foi estruturada, configurada e aplicada com base em um conjunto de dados de 65 questionários, sendo estes respondidos por alunos que evadiram e que permaneceram, a rede distribuiu esses dados em 2 conjuntos, sendo: de treinamento e de teste.

Para que fosse encontrado o melhor modelo da rede para esses dados, a rede foi executada 20 vezes seguindo a mesma configuração, com base nos indicadores de desempenho curva ROC, acurácia é G-mean o melhor modelo para previsão de evasão foi selecionado. O modelo escolhido apresentou o melhor valor de desempenho da Curva ROC de 0,973, além de ter valores altos de Acurácia, tanto no treinamento quanto no teste. De tal modo, que a sua acurácia no teste foi de 88,2% melhor que no treinamento que foi de 85,4%, demonstrando um melhor resultado em dados novos. Isso se apresentou também nos resultados do G-mean, ambos, no treinamento e no teste, com bons resultados, e também apresentando resultados melhores na fase de teste, refletindo em uma rede eficiente e que tem bons resultados em um conjunto de dados novos.

Com base nesse modelo, foi possível obter informações sobre o grau de importância de cada co-variável (fatores que influenciam na evasão), em que as variáveis com maior impacto na decisão sobre evadir puderam ser identificadas. Dos 10 fatores em questão, a que obteve 100% de importância normalizada foi a: Escolha do Curso, o que significa que estudantes que tiveram problemas e dificuldades no processo de escolha de curso, sendo o maior índice nas respostas dos 65 estudantes, o que demonstrar que mesmo que com dificuldades em outras áreas se o estudante escolheu o curso de forma assertiva ele vai buscar maneiras de continuar e o oposto também, quando o estudante por sua vez não fez uma escolha muito assertiva sobre o

curso de graduação, com ou sem dificuldades em outras áreas pode resultar em uma evasão. Em segundo lugar, vem a localidade de origem do estudante 78,6%, que demonstra que estudantes que optaram por se mudar de cidade para fazer a graduação, tem uma taxa alta no processo de decisão para optar pela evasão. Logo depois, nas porcentagens entre 40% a 50%, tem: a falta de base escolar do ensino médio, a localização entre a moradia e a universidade e a satisfação com o curso.

Como todas essas informações, é possível notar pontos importantes para a universidade analisar para buscar soluções que aumentem a taxa de permanência nos cursos de graduação de Engenharia de Produção do campus de Nova Andradina. Uma sugestão é ampliar programas institucionais com foco em gerar maior conhecimento por parte dos estudantes do ensino médio sobre o curso em questão, suas abordagens, áreas de atuação, o que pode resultar em uma melhoria nas escolhas dos mesmos no momento de ingresso em uma universidade.

Para um entendimento mais global, é necessário um estudo em outros campus e também em outros cursos, para que seja possível traçar projetos que sejam mais efetivos. Um estudo amplo, pode apontar informações importantes para os gestores das universidades e assim, gerar uma mobilização conjunta em prol da universidade e dos estudantes mitigando o risco de evasão.

REFERÊNCIAS

- ACHEAMPONG, A. O.; BOATENG, E. B. *Modelling carbon emission intensity: Application of artificial neural network*. *Journal of Cleaner Production*, v. 225, p. 833-856, 2019.
- ALBUQUERQUE, Teresa. Do abandono à permanência num curso de ensino superior. **Sísifo - Revista de Ciências da Educação**. Lisboa, n. 7, p. 19-28, 2008.
- AZARPOUR, Abbas. *et al.* *A generic hybrid model development for process analysis of industrial fixed-bed catalytic reactors*. *Chemical Engineering Research and Design*, 2017.
- BEJANI, Mohammad Mahdi.; GHATEE, Mehdi. *A systematic review on overfitting control in shallow and deep neural networks*. *Artificial Intelligence Review*, p. 1-48, 2021.
- BICALHO, Katia Vanessa. *et al.* Estudo do uso da Técnica de Redes Neurais Artificiais em Engenharia Geotécnica. **XX Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica**. 2022.
- CAPELAS, Mauricio. **Análise de evasão de discentes em cursos de Engenharia de Produção**. 2014. 119f. Tese de Doutorado apresentada ao Programa de Pós Graduação em Engenharia de Produção da Universidade Paulista. São Paulo-SP, 2014.
- CISLAGHI, Renato. **Um modelo de sistema de gestão do conhecimento em um framework para a promoção da permanência discente no ensino de graduação**. 2008. 253f. Tese (Doutorado em Engenharia e Gestão do Conhecimento) - Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis-SC, 2008.

DEL BONIFRO, F. et al. *Student Dropout Prediction. Lecture Notes in Computer Science*, v. 12163, 2020. Disponível em: https://doi.org/10.1007/978-3-030-52237-7_11. Acesso em: 20 jul. 2023.

FLECK, Leandro. et al. Redes neurais artificiais: Princípios básicos. **Revista Eletrônica Científica Inovação e Tecnologia**, v. 1, n. 13, p. 47-57, 2016.

GIL, A. C. **Como elaborar Projetos de Pesquisa**. Ed. Atlas. São Paulo - SP, 2002.

GONZAGA, J. C. B. et al. *ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process*. **Computers & chemical engineering**, v. 33, n. 1, p. 43-49, 2009.

HE, H.; GARCIA, E. A. *Learning from Imbalanced Data*. **IEEE Transactions on Knowledge and Data Engineering**, vol. 21, n. 9, p. 1263-1284, 2009.

ISHIDA, Denise Miki Tawaraya. **Perspectivas do aprendizado de máquina no ensino da Engenharia Química**. 2021.

LEE, Sanghoon.; LEE, Kang-Kun.; YOON, Heesung. *Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors*. **Hydrogeology Journal**, v. 27, n. 2, p. 567-579, 2019.

LIU, Qixin. CHEN, Ziwei. *Early warning control model and simulation study of engineering safety risk based on a convolutional neural network*. **Neural Computing and Applications**, p. 1-8, 2023.

LU, Y. et al. *GOBoost: G-mean optimized boosting framework for class imbalance learning*. **12th World Congress on Intelligent Control and Automation (WCICA)**, China, p. 3149-3154, 2016.

MARTINHO, Valquíria R. C.; NUNES, Clodoaldo.; MINUSSI, Carlos Roberto. *Prediction of School Dropout Risk Group Using Neural Network*. Anais: **Federated Conference on Computer Science and Information Systems**, p. 111-114, 2013a.

MARTINHO, Valquíria R. C.; NUNES, Clodoaldo.; MINUSSI, Carlos Roberto. *An Intelligent System for Prediction of School Dropout Risk Group in Higher Education Classroom based on Artificial Neural Networks*. Anais: **IEEE 25th International Conference on Tools with Artificial Intelligence**, 2013b.

MATSUNAGA, Yohei. et al. RNA interference-mediated growth control of the southern root-knot nematode *Meloidogyne incognita*. **Bioscience, biotechnology, and biochemistry**, v. 76, n. 2, p. 378-380, 2012.

MAYORGA, Claudia; DE SOUZA, Luciana Maria. Ação afirmativa na universidade: a permanência em foco. **Revista Psicologia Política**, v. 12, n. 24, p. 263-281, 2012.

MORAES, J. O. de.; THEÓPHILO, C. R. **Evasão no ensino superior:** estudo dos fatores causadores da evasão no curso de Ciências Contábeis da Universidade Estadual de Montes Claros & UNIMONTES. Congresso USP de iniciação científica em contabilidade. São Paulo: USP, 2006.

NIYOGISUBIZO, Jovial. et al. *Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization*. **Computers and Education: Artificial Intelligence**, v. 3, 2022.

OLIVEIRA, Artur Carlos Santiago de. **Aplicação de técnicas de aprendizado de máquina para identificação de alunos com tendência à evasão escolar**. 2022. Trabalho de Conclusão de Curso. Universidade Federal do Rio Grande do Norte, 2022.

ROLLO, Lúcia F.; PEREIRA, Anísio C. Análise do processo educacional contábil sob o prisma de seus elementos de maior relevância. **Revista Brasileira de Contabilidade**, Brasília-DF, ed. 31, p. 49-53, 2003.

RUMELHART, D. E. et al. Processamento distribuído paralelo: Explorações na microestrutura da cognição. Vol. 1: Fundações. **Imprensa do MIT**, 1986.

SANTOS, Nicolas Victor Martins dos.; JUNIOR, Muris Lage.; RIBEIRO, Monica Luiz de Lima. Evasão no curso de engenharia de produção da universidade federal de Goiás - regional Catalão. **XXXV Encontro Nacional de Engenharia de Produção (ENEGET)**: Perspectivas Globais para a Engenharia de Produção. Fortaleza-CE, 2015.

SEGURA, M.; MELLO, J.; HERNÁNDEZ, A. *Machine Learning Prediction of University Student Dropout: does preference play a key role?* **Mathematics**, ed. 10, 2022. Disponível em: <https://doi.org/10.3390/math10183359>. Acesso em: 7 ago. 2023.

SILVA, Ivan Nunes da Silva. **Redes Neurais Artificiais Para Engenharia E Ciências Aplicadas**: Curso Prático. Editora Artliber, 2010.

SILVA, Leonária Araújo. et al. Modelagem preditiva de propriedades mecânicas em concretos reforçados com fibra de aço utilizando redes neurais artificiais. **Ambiente Construído**, v. 22, n. 2, p. 215-232, 2022. Disponível em: <https://doi.org/10.1590/s1678-86212022000200602>. Acesso em: 11 ago. 2022.

TONTINI, Gérson.; WALTER, Silvana Anita. Pode-se identificar a propensão e reduzir a evasão de alunos? Ações estratégicas e resultados táticos para instituições de ensino superior. **Avaliação**, Campinas-SP, v. 19, n. 1, p. 89-110, 2014

UFMS. Universidade Federal de Mato Grosso do Sul - câmpus Nova Andradina. **Quantidade de discentes matriculados e excluídos de 2018 a 2022**. 2022.