



MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Especialização em Ciência de Dados



Utilização de Modelos de Regressão para Predição do Teor de Enxofre do Ferro Gusa Produzido em Alto-Forno

Bruno Hermógenes das Graças
Elioenai Dutra da Silva
Helio Bechara de Almeida

Ipatinga, MG
2023

Bruno Hermógenes das Graças
Elioenai Dutra da Silva
Helio Bechara de Almeida

Utilização de Modelos de Regressão para Predição do Teor de Enxofre do Ferro Gusa Produzido em Alto-Forno

Trabalho de conclusão de curso apresentado ao curso de Ciência de Dados do Instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto, como parte dos requisitos necessários para a obtenção do título de Especialista em Ciência de Dados.

Orientador: Prof. Dr. Thiago Augusto de Oliveira Silva

Ipatinga, MG

2023

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S586u Silva, Elioenai Dutra da.
Utilização de modelos de regressão para predição do teor de enxofre do ferro gusa produzido em alto-forno. [manuscrito] / Elioenai Dutra da Silva. Bruno Hermogenes Das Graças. Helio Bechara de Almeida. - 2023. 32 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Thiago Augusto de Oliveira Silva.
Produção Científica (Especialização). Universidade Federal de Ouro Preto. Departamento de Engenharia de Produção.

1. Análise de regressão. 2. Enxofre. 3. Ferro fundido. 4. Predição. 5. Redes neurais (Computação). I. Bechara de Almeida, Helio. II. Hermogenes Das Graças, Bruno. III. Silva, Thiago Augusto de Oliveira. IV. Universidade Federal de Ouro Preto. V. Título.

CDU 519.24:004.8

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431



FOLHA DE APROVAÇÃO

Bruno Hermógenes das Graças
Elioenai Dutra da Silva
Helio Bechara de Almeida

Utilização de Modelos de Regressão para Predição do Teor de Enxofre do Ferro Gusa Produzido em Alto-Forno

Trabalho de conclusão de curso apresentado ao Curso de Especialização em Ciência de Dados da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Especialista em Ciência de Dados.

Aprovada em 30 de junho de 2023

Membros da banca

Prof. Dr. - Thiago Augusto de Oliveira Silva - Orientador(a) (Deenp/Icea)
Prof. Dr. - Alexandre Xavier Martins - (Deenp/Icea)
Prof. Dr. - Paganini Barcellos de Oliveira - (Deenp/Icea)

Thiago Augusto de Oliveira Silva, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 17/07/2023



Documento assinado eletronicamente por **Thiago Augusto de Oliveira Silva, PROFESSOR DE MAGISTERIO SUPERIOR**, em 17/07/2023, às 14:11, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0558745** e o código CRC **CD872181**.

Agradecimentos

Agradecer, em primeiro lugar, a Deus, que nos conduziu e nos deu forças para chegarmos até aqui e concluir com êxito esse trabalho. Agradecer aos nossos pais e familiares pelo apoio e incentivo em todos os momentos que precisamos, dando suporte para a realização desse sonho, apesar de todas as dificuldades. Gostaríamos de agradecer especialmente ao nosso orientador, professor Thiago Augusto de Oliveira Silva pelos conhecimentos compartilhados que nos ajudaram a desenvolver nossa pesquisa e aperfeiçoar nossas habilidades acadêmicas. Agradecemos à Usiminas por investir e acreditar nos seus profissionais dando a oportunidade de realizarmos um excelente curso e posteriormente colocarmos em prática todo o aprendizado adquirido. Agradecer também a equipe técnica da Redução. Agradecemos também a nossa instituição de ensino por nos proporcionar a oportunidade de realizar este trabalho de conclusão de curso somando conhecimentos necessários para enfrentarmos os desafios profissionais. Por fim, agradecer a todos que direta ou indiretamente fizeram parte do processo de conclusão desse trabalho tornando esta jornada educacional tão gratificante e inesquecível.

Resumo

O mercado siderúrgico mundial exige aços de alta qualidade para atender às demandas crescentes por materiais duráveis, resistentes e com desempenho superior em diversas aplicações industriais. Sendo o enxofre uma impureza prejudicial para a obtenção destes aços, a dessulfuração do ferro gusa é de extrema importância. Neste processo, o conhecimento do teor de enxofre do ferro gusa é necessário para o dimensionamento dos agentes dessulfurantes. Na empresa estudada tal variável é obtida através de análises laboratoriais. Eventualmente, essa informação não é disponibilizada em tempo hábil. Neste contexto, esse trabalho tem como propósito o desenvolvimento de um modelo de rede neurais como alternativa para predição do teor de enxofre no gusa a ser dessulfurado em panela pelicano. A metodologia empregada neste estudo é caracterizada como empírico-descritiva, exploratória e explicativa. Inicialmente, realizou-se uma revisão bibliográfica sobre o tema, seguida pela coleta e tratamento de uma base de dados. Em seguida, aplicaram-se modelos de redes neurais para a predição da variável de interesse. O modelo de rede neural desenvolvido demonstrou um erro médio absoluto (MAE) dentro da faixa aceitável estabelecida pelos especialistas do processo estudado. Aproximadamente 77,3% dos valores preditos estavam dentro da faixa de aceitação. No entanto, devido à criticidade do processo, desvios superiores aos limites de aceitação indicam a necessidade de aprimoramento do modelo.

Palavras-chaves: Enxofre. Predição. Redes Neurais.

Abstract

The global steel market demands high-quality steels to meet the growing demands for durable, strong, and high-performance materials in various industrial applications. Since sulfur is a detrimental impurity for the production of these steels, the desulfurization of pig iron is of utmost importance. In this process, knowledge of the sulfur content in pig iron is necessary for the sizing of desulfurizing agents. In the studied company, this variable is obtained through laboratory analysis. However, occasionally, this information is not made available in a timely manner. In this context, the purpose of this work is to develop a neural network model as an alternative for predicting the sulfur content in pig iron to be desulfurized in a ladle furnace. The methodology employed in this study is characterized as empirical-descriptive, exploratory, and explanatory. Initially, a literature review on the topic was conducted, followed by data collection and treatment. Subsequently, neural network models were applied to predict the variable of interest. The developed neural network model demonstrated a mean absolute error (MAE) within the acceptable range established by the experts in the studied process. Approximately 77.3% of the predicted values fell within the acceptance range. However, due to the criticality of the process, deviations beyond the acceptance limits indicate the need for model improvement.

Keywords: *Sulfur. Prediction. Neural Networks*

Lista de ilustrações

| | |
|---|----|
| Figura 1 – Subdivisão interna do alto-forno e suas várias zonas | 3 |
| Figura 2 – Dessulfuração de gusa em carro torpedo (Mansur, 2008) | 4 |
| Figura 3 – Métodos de dessulfuração em panela pelicano (Kirmse, 2006 apud Takano) . | 5 |
| Figura 4 – Fluxograma das etapas do Estudo | 8 |
| Figura 5 – Histograma das variáveis em escala original | 12 |
| Figura 6 – Histograma das variáveis em escala normalizada [-1,1] | 13 |
| Figura 7 – À esquerda regressão linear simples, à direita regressão múltipla | 17 |
| Figura 8 – Correlação positiva 2- Correlação negativa 3- Sem correlação | 17 |
| Figura 9 – Representação esquemática da rede do tipo <i>perceptron</i> | 18 |
| Figura 10 – <i>Bootstrap</i> - Método de reamostragem | 20 |
| Figura 11 – Seleção das variáveis para cada nó | 21 |
| Figura 12 – Representação esquemática da rede do tipo perceptron | 23 |
| Figura 13 – Representação esquemática da rede do tipo perceptron | 24 |
| Figura 14 – Arquitetura da rede neural | 25 |
| Figura 15 – Resultado do treinamento e validação da RNA de predição de teor de enxofre no gusa | 26 |
| Figura 16 – Comparativo de resultados entre modelos de predição | 27 |
| Figura 17 – Validação cruzada dos modelos de predição | 29 |

Lista de quadros

| | |
|---|----|
| Quadro 3.1 – Dicionário das variáveis do modelo de rede neural. | 11 |
| Quadro 3.2 – Limites para identificação de <i>outliers</i> | 11 |
| Quadro 3.3 – Bibliotecas <i>Python</i> Utilizadas. | 14 |
| Quadro 3.4 – Configuração de parâmetros do <i>Keras Tuner</i> | 14 |
| Quadro 4.1 – Configuração obtida pelo <i>Keras Tuner</i> | 25 |
| Quadro 4.2 – Configuração obtida pelo <i>Keras Tuner</i> | 27 |
| Quadro 4.3 – Modelos de Predição | 28 |
| Quadro 4.4 – Resultados ANOVA | 28 |

Sumário

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 1 |
| 1.1 | Objetivo geral | 2 |
| 1.1.1 | Objetivos específicos | 2 |
| 1.2 | Organização do Trabalho | 2 |
| 2 | REVISÃO DA LITERATURA | 3 |
| 2.1 | Processo de Altos-Fornos | 3 |
| 2.2 | Dessulfuração de Gusa | 4 |
| 2.2.1 | Dessulfuração em Carro Torpedo (CT) | 4 |
| 2.2.2 | Dessulfuração em Panela Pelicano | 5 |
| 2.2.3 | Enxofre no Gusa | 6 |
| 2.2.4 | Variáveis que Impactam o Teor de Enxofre no Gusa | 7 |
| 3 | METODOLOGIA | 8 |
| 3.1 | Classificação da Pesquisa | 8 |
| 3.2 | Etapas da Pesquisa | 8 |
| 3.3 | Origem e Coleta dos Dados | 9 |
| 3.4 | Análise Teórica das Variáveis | 9 |
| 3.5 | Tratamento dos Dados | 11 |
| 3.5.1 | Tratamento de <i>Outliers</i> | 11 |
| 3.5.2 | Normalização | 12 |
| 3.5.3 | Valores nulos, ausentes e iguais a zero | 13 |
| 3.6 | Modelos de Predição | 13 |
| 3.6.1 | Arquitetura da RNA | 14 |
| 3.6.2 | Desenvolvimento, treinamento e implementação da RNA | 15 |
| 3.6.3 | Outros modelos de predição e validação cruzada | 15 |
| 3.7 | Modelos de Regressão | 16 |
| 3.7.1 | Regressão Linear | 16 |
| 3.7.1.1 | Conceitos importantes | 16 |
| 3.7.2 | Árvores de Decisão | 18 |
| 3.7.3 | <i>Random Forest</i> | 19 |
| 3.7.3.1 | Métodos <i>Ensemble</i> | 19 |
| 3.7.3.2 | Como funciona o algoritmo <i>Random Forest</i> | 20 |
| 3.7.3.3 | Seleção de amostras | 20 |
| 3.7.3.4 | Seleção das variáveis para cada nó | 20 |
| 3.7.4 | Redes Neurais | 21 |

| | | |
|------------|--|-----------|
| 4 | RESULTADOS E DISCUSSÃO | 25 |
| 4.1 | Resultados da RNA | 25 |
| 4.2 | Modelos Alternativos para Análise Comparativa | 26 |
| 4.3 | Validação Cruzada | 28 |
| 5 | CONCLUSÃO | 30 |
| 5.1 | Sugestão de Trabalhos Futuros | 31 |
| 6 | REFERÊNCIAS BIBLIOGRÁFICAS | 32 |

1 Introdução

O processo siderúrgico pode ser resumido na transformação do minério de ferro em aço. É iniciada na etapa de redução de óxidos do minério de ferro em altos-fornos, o transformando em ferro gusa. Este por sua vez passa pelo processo de refino em aciarias, onde ocorre a oxidação de grande parte do carbono e de impurezas como fósforo, enxofre e outras. Dentre essas impurezas, o enxofre é considerado um dos elementos mais prejudiciais ao aço. A presença deste elemento no aço, reduz sua resistência ao impacto, ductilidade transversal e soldabilidade, podendo afetar tanto a qualidade interna quanto a superficial no produto. Além disso, também pode causar problemas de fragilidade à quente em aços com baixo teor de manganês (Keskinliç, 2007). Portanto, o enxofre torna-se altamente indesejável em aplicações onde essas propriedades são importantes, como por exemplo, chapas para oleodutos, chapas de navio, rodas, vasos de pressão, etc.

Considerando-se a limitada capacidade de dessulfuração no refino primário do aço, são necessários processos anteriores visando a retirada do enxofre, que usualmente, são realizados em carros torpedos ou em painéis pelicanos. O processo de dessulfuração consiste, na injeção de mistura de óxido de cal e magnésio que reagem com o enxofre formando sulfetos de cal e de magnésio. Para dimensionar a quantidade de óxidos necessários no processo, é necessário conhecer o teor de enxofre no ferro gusa e o teor visado após a dessulfuração.

A análise química do ferro gusa é geralmente realizada em laboratórios a partir de amostras coletadas nos canais de gusa do alto-forno. Não raramente, os resultados das análises de enxofre no gusa não estão disponíveis em tempo hábil, o que pode resultar em erros na dosagem de agentes dessulfurantes. As possíveis consequências desse cenário incluem a desclassificação do aço produzido, a necessidade de correções no teor de enxofre em etapas posteriores do processo com custos mais altos e menor produtividade. Nesse contexto, a implementação de métodos alternativos para a predição do teor de enxofre torna-se extremamente importante. Esses métodos podem se basear em técnicas analíticas rápidas e não laboratoriais.

O uso de redes neurais para a previsão de variáveis tem crescido significativamente na indústria devido aos bons resultados obtidos. A aplicação dessa técnica na modelagem de processos não lineares com características complexas é altamente favorável em comparação com os métodos determinísticos tradicionais.

Os modelos de redes neurais podem ter grande diversidade de configurações. É necessário testar várias possibilidades buscando obter um modelo com eficiência e confiabilidade satisfatória. Essa etapa é fundamental para garantir a qualidade do produto final e a otimização dos recursos utilizados.

1.1 Objetivo geral

O objetivo geral do trabalho é desenvolver um modelo de redes neurais para predição do teor de enxofre em ferro gusa a ser dessulfurado em panela pelicano.

1.1.1 Objetivos específicos

- Identificar variáveis que influenciam no teor de enxofre do ferro gusa a ser dessulfurado;
- Levantar dados históricos das variáveis identificadas que sejam mensuráveis;
- Desenvolver modelos de redes neurais artificiais para predição do teor de enxofre do ferro gusa a ser dessulfurado em panela pelicano.
- Analisar os resultados obtidos pelo modelo de redes neurais.
- Comparar o desempenho do modelo de redes neurais com outros modelos de regressão.

1.2 Organização do Trabalho

No Capítulo 1 será apresentada uma introdução sobre o trabalho desenvolvido, conceituando sua área de aplicação, problema de pesquisa, proposta de desenvolvimento e objetivos. No Capítulo 2 apresenta-se uma revisão bibliográfica contemplando o processo de produção do aço, com foco nas variáveis com influência sobre o teor de enxofre no ferro gusa, além de explorar técnicas e ferramentas utilizadas no desenvolvimento deste trabalho. O Capítulo 3 transcorre sobre a metodologia do trabalho, indicando o método de pesquisa utilizado, o fluxo de trabalho desenvolvido ao longo da pesquisa, as técnicas e parâmetros utilizados para o desenvolvimento do modelo de redes neurais. No Capítulo 4 serão apresentados e discutidos os resultados alcançados. No quinto e último capítulo será apresentada a conclusão do trabalho e discutidas ações futuras sobre o tema.

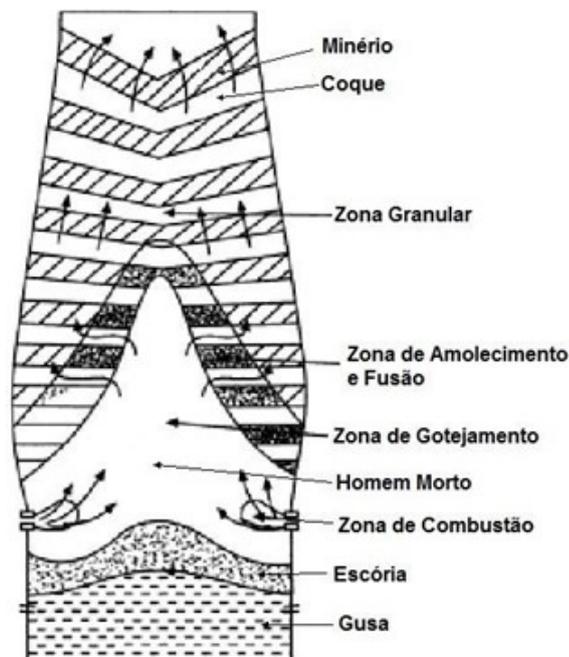
2 Revisão da Literatura

2.1 Processo de Altos-Fornos

A unidade de redução em uma usina siderúrgica é responsável pela conversão da matéria-prima sólida em gusa líquido. Em uma usina integrada, a unidade de redução é composta por três áreas principais: coqueria, sinterização e alto-forno. A coqueria é responsável pela produção de coque através da coqueificação do carvão mineral. A sinterização produz o sinter, um aglomerado metálico poroso rico em ferro, através do processo de aglomeração a quente. O alto-forno é o reator onde o minério de ferro é fundido para se transformar em gusa.

No alto-forno, o oxigênio presente no ar quente injetado nas ventaneiras reage quimicamente com o coque carregado, formando monóxido de carbono (CO), que por sua vez reduz o óxido de ferro presente na carga metálica. A estrutura interna do alto-forno e suas respectivas camadas foram descobertas através de estudos de dissecação realizados inicialmente pelos russos e posteriormente pelos japoneses. A Figura 1 ilustra a subdivisão dessas camadas e zonas.

Figura 1 – Subdivisão interna do alto-forno e suas várias zonas



Fonte: Castro et al., 1997

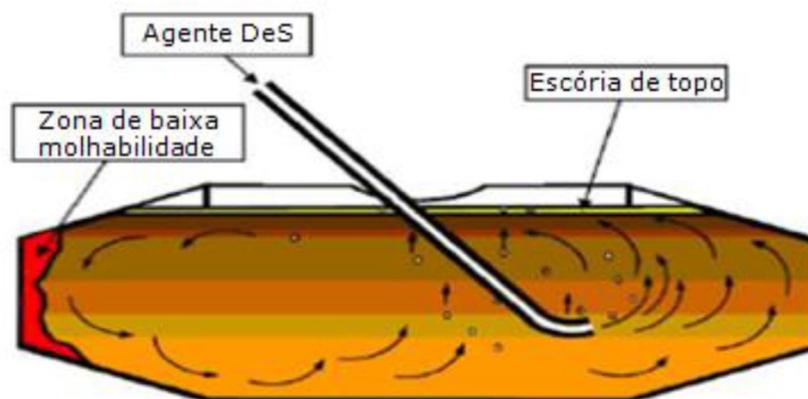
2.2 Dessulfuração de Gusa

A presença indesejável de enxofre no processo de fabricação do aço requer a sua remoção, a fim de atender às exigências do mercado. O enxofre é proveniente de minérios, carvões e cales. A dessulfuração do gusa pode ser realizada em carros torpedos (CT), painelas pelicanos ou transferência.

2.2.1 Dessulfuração em Carro Torpedo (CT)

De acordo com Kirmse (2006), a dessulfuração em CT é amplamente utilizada em todo o mundo. No entanto, o processo apresenta dificuldades devido à forma geométrica do carro torpedo, concebido com o objetivo de preservar a temperatura do gusa. Em decorrência dessa geometria, zonas mortas são criadas durante a dessulfuração, o que afeta sua eficiência (Costa, 1996). A Figura 2 ilustra o tratamento de injeção de pó no carro torpedo.

Figura 2 – Dessulfuração de gusa em carro torpedo (Mansur, 2008)



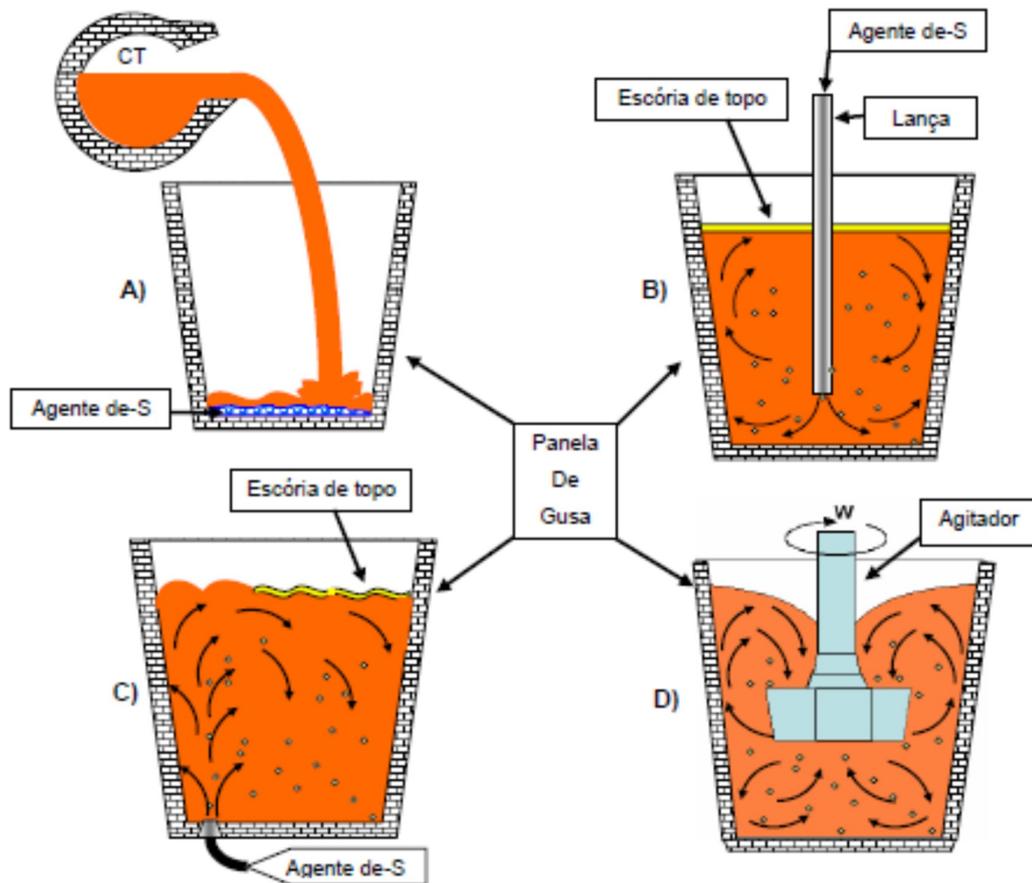
Fonte: Mansur, 2008

Após a dessulfuração realizada no carro torpedo, o ferro gusa é direcionado para a Aciaria. No entanto, dependendo dos níveis de enxofre exigidos, é necessário submetê-lo novamente a um processo de dessulfuração em panela pelicano.

2.2.2 Dessulfuração em Panela Pelicano

Os principais métodos de dessulfuração do gusa na panela pelicano são explicitados por Kirmse (2006) apud Takano (1997), conforme ilustrado na Figura 3.

Figura 3 – Métodos de dessulfuração em panela pelicano (Kirmse, 2006 apud Takano)

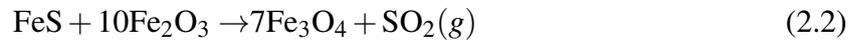


Fonte: Takano, 1997

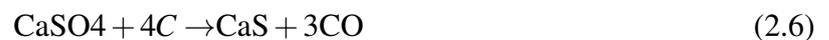
Conforme ilustrado na Figura 3, o método (a), conhecido como carbocal, envolve a adição de carbonato de sódio Na_2CO_3 e calcário no fundo da panela. Em seguida, o gusa líquido é transferido para a panela, promovendo a dessulfuração. A Figura 3 (b) ilustra a injeção de pós dessulfurantes no interior do metal líquido, utilizando lanças refratárias com movimento vertical e um gás de arraste (nitrogênio ou argônio). Outro método utilizado para a dessulfuração em panela de gusa é a injeção de pós dessulfurantes no fundo da panela, conforme mostra a Figura 4 (c). O último método apresentado é o Reator Kambara (KR), que consiste em uma panela agitada mecanicamente por uma pá rotativa inserida em sua parte superior. Nesse sistema, o agente dessulfurante é adicionado por cima, como ilustrado na Figura 3 (d).

2.2.3 Enxofre no Gusa

Aproximadamente 80% do enxofre presente em um alto-forno é proveniente do combustível sólido, coque. O restante provém dos outros materiais adicionados (carga metálica, carvão mineral injetado através das ventaveiras). Cerca de 80% a 90% do enxofre é incorporado na escória, enquanto 10% a 15% é arrastado pelos gases que saem do topo do alto-forno, e aproximadamente 2% a 5% é incorporado no gusa (GHOSH e CHATTERJEE, 2010). O enxofre é considerado uma impureza indesejada no gusa, uma vez que prejudica as propriedades físicas necessárias no aço, tornando, portanto, sua remoção imperativa, a qual ocorre principalmente durante o processo no alto-forno. Geralmente, o enxofre é dissolvido na escória, enquanto no gusa é dissolvido na forma de FeS ou S. As Equações 2.1 a 2.4 descrevem como ocorre essa dissolução, sendo que a porção presente nos gases do topo encontra-se na forma de SO₂ e H₂S (GHOSH e CHATTERJEE, 2008). Os sulfetos presentes na carga são convertidos em sulfetos no gusa por meio da dissolução direta do FeS, conforme a Equação 2.1. Além disso, também pode ocorrer a redução de SO₂ a S₂ em um primeiro momento, conforme as Equações 2.2 e 2.3 seguida pela sulfetização do Fe, formando FeS, conforme a Equação 2.4 (YANG, RAIPALA e HOLAPPA, 2013).

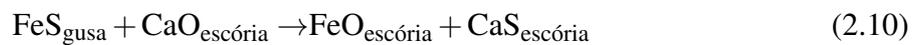


As Equações 2.5 a 2.9 descrevem o processo de transformação de sulfatos em sulfetos na presença de carbono, ferro ou óxido de ferro, bem como na presença de silício gasoso (YANG, RAIPALA e HOLAPPA, 2013). Na Equação 2.5 o sulfato de cálcio reage com o ferro e o carbono, transformando-se em sulfeto de ferro. Já na Equação 2.6, o sulfato de cálcio reage com o carbono, transformando-se em sulfeto de cálcio, o qual, posteriormente, reage com o óxido de ferro na Equação 2.7, resultando no sulfeto de ferro. Por outro lado, na Equação 2.8 o sulfeto de cálcio, na presença de silício gasoso, produz sulfeto de silício gasoso, que, posteriormente, reage com o ferro, conforme a Equação 2.9, resultando no sulfeto de ferro.

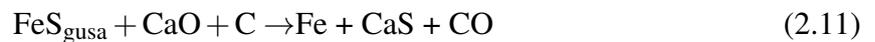


A dessulfuração é baseada no princípio da afinidade. O enxofre segue uma ordem crescente de afinidade pelos metais: FeS, MnS, MgS, Na₂S, CaS, tornando o MgO e o CaO as opções mais eficientes como elementos dessulfurantes. O óxido de cálcio (CaO) é o principal agente dessulfurante, pois o CaS formado é solúvel na escória (WAKELUN E RICKETTS, 1999; SCHRAMA et al., 2017).

A reação de dessulfuração através do óxido de cálcio pode ser expressa pela Equação 2.10, em que o sulfeto de ferro reage com o óxido de cálcio presente na escória, formando óxido de ferro e sulfeto de cálcio na escória.



No entanto, o processo de dessulfuração descrito pela Equação 2.10 é afetado pela formação de óxido de ferro, o qual diminui a velocidade da reação. O método mais eficiente de dessulfuração é governado pela Equação 2.11, na qual o sulfeto de ferro, na presença de óxido de cálcio e carbono, reage para produzir ferro e sulfeto de cálcio, sem a formação de óxido de ferro.



2.2.4 Variáveis que Impactam o Teor de Enxofre no Gusa

Os principais fatores operacionais que influenciam o teor de enxofre no ferro-gusa são: a pressão do forno, a quantidade total de enxofre carregado, a temperatura do ferro-gusa, a taxa de formação de escória, a composição da escória (basicidade, teor de MgO e Al₂O₃) e a composição do ferro-gusa (teor de C e Si) (TAMURA, ONO E NISHIDA, 1981). Observa-se que o enxofre adicionado e a velocidade de produção de gusa exercem uma influência positiva sobre o teor de enxofre no ferro-gusa, enquanto as concentrações de CaO e MgO são variáveis inversamente proporcionais (DAVID, 2003). O processo de dessulfuração do enxofre no gusa é diretamente afetado pelas condições operacionais estáveis do alto-forno e pelas propriedades da escória. Os fatores determinantes são: i) quantidade de enxofre carregada no alto-forno; ii) maior basicidade da escória (através da adição de CaO); e iii) temperatura do alto-forno, especialmente na zona de gotejamento, pois a reação de dessulfuração é altamente endotérmica (YANG, RAIPALA e HOLAPPA, 2013)

3 Metodologia

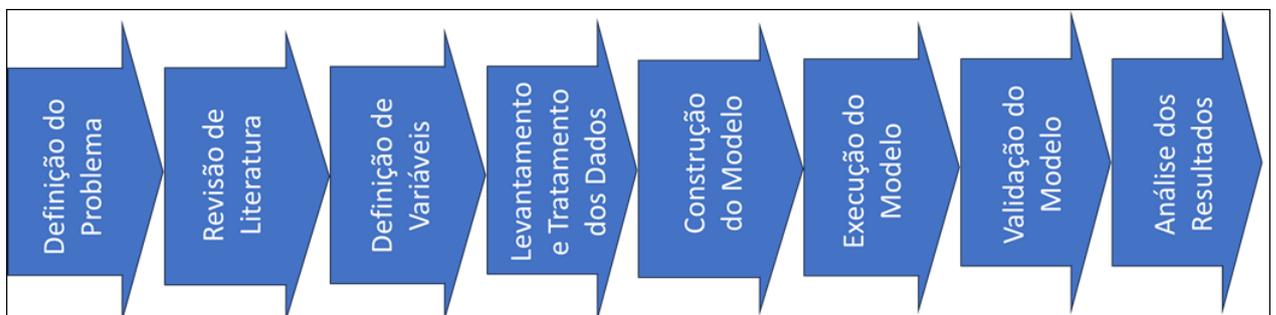
3.1 Classificação da Pesquisa

A pesquisa apresenta características descritivas, ao descrever o modelo de redes neurais e as correlações entre as variáveis, assim como características explicativas, ao buscar compreender as relações de causa e efeito relacionadas à predição do teor de enxofre no ferro gusa. Há também um viés exploratório pois busca através de estudo de caso explorar as causas, consequências e outros elementos envolvidos no problema. Além disso, é considerada empírica, pois envolve a coleta de dados reais, a aplicação de técnicas de redes neurais e a realização de experimentos práticos para desenvolver e validar o modelo de predição.

3.2 Etapas da Pesquisa

Conforme apresentado na Figura 4, a presente pesquisa seguirá as seguintes etapas: primeiramente, será realizada uma abordagem abrangente do tema em questão, onde o problema será definido e os objetivos gerais e específicos serão estabelecidos. Em seguida, será conduzido um estudo da literatura com o propósito de fundamentar a metodologia da pesquisa. Posteriormente, as variáveis que influenciam o teor de enxofre no gusa serão determinadas com a orientação de um especialista. Na sequência, a base de dados contendo as variáveis selecionadas será coletada. A partir dessa base de dados, um modelo de redes neurais será desenvolvido, treinado e aplicado, gerando valores previstos para o teor de enxofre no gusa. Em seguida, os resultados obtidos pelo modelo serão avaliados quanto à sua aderência aos valores reais.

Figura 4 – Fluxograma das etapas do Estudo



Fonte: Elaboração própria, 2023

3.3 Origem e Coleta dos Dados

As amostras foram obtidas de uma fonte sistêmica em um banco de dados estruturado da empresa em estudo, no período de 01/01/2019 a 27/06/2022. Para a extração dos dados, utilizou-se o sistema de gerenciamento de informações de planta (*PIMS - Plant Information Management System*), por meio da criação e consulta de TAGs relacionados aos sensores de processo, relatórios diários e consulta ao *software* de gerenciamento de processos SAP. A granulometria foi definida como a média diária de cada variável do processo, uma vez que o alto-forno leva aproximadamente 14 a 16 horas para responder a variações de parâmetros. Essa abordagem foi adotada para abstrair as variáveis do processo, com foco em compreender e justificar desvios específicos do processo. Portanto, trata-se de uma ferramenta confiável e que pode ser efetivamente utilizada para entender o comportamento das variações do processo.

Inicialmente, foram consideradas 38 variáveis quantitativas para análise, incluindo características do carvão PCI (% H, % O, % N, % S, % Umidade), produtividade do alto-forno, produção de gusa, taxa de escória, % de sinter na carga, consumo de sinter, % de pelota, % de minério de ferro, % de quartzo, *coke rate*, *fuel rate*, enxofre carregado, cinza do coque, umidade do coque, rendimento de CO, vazão de ar, vazão de oxigênio, enriquecimento do ar com oxigênio, umidade do ar, temperatura do ar, pressão do ar, velocidade do ar, fator K de permeabilidade, temperatura de chama, características do gusa (temperatura, Si, S, Mn, C), basicidade binária B2, Al₂O₃, MgO e basicidade quaternária B4. No entanto, após consulta a especialistas e revisão de trabalhos científicos, essas variáveis foram reduzidas para 8, que são: % de enxofre no gusa, produtividade do alto-forno, taxa de escória, % de enxofre carregado, pressão do ar, temperatura do gusa, basicidade binária B2 e basicidade quaternária B4.

3.4 Análise Teórica das Variáveis

Com o auxílio de especialistas e estudos conduzidos sobre o tema, foi descrito o mecanismo teórico de como cada variável independente se relaciona com a variável dependente, que é o teor de enxofre no gusa:

- Produtividade do alto-forno: essa variável determina a taxa de produção do alto-forno e seu estado térmico. Uma maior produtividade só é possível se houver estabilidade térmica. Portanto, quanto maior a produtividade do alto-forno, menor será o teor de enxofre no gusa, devido a um melhor esgotamento do cadinho e a um menor tempo de contato do gusa com a escória, o que dificulta a reincorporação do enxofre da escória no gusa.

- Taxa de escória (*slag rate*): essa variável indica a quantidade de escória gerada no processo (em kg/t de gusa). A escória é onde grande parte das impurezas das matérias-primas carregadas no alto-forno é removida, incluindo o enxofre. Quanto maior a taxa de escória, menor será o enxofre remanescente no gusa.

- % de enxofre carregado: essa variável determina a quantidade de enxofre presente nas matérias-primas carregadas no alto-forno. Quanto maior o % de enxofre carregado, maior será a possibilidade de incorporação desse enxofre no gusa.
- Pressão do ar na base: refere-se à resistência encontrada pelo ar ao ser injetado na região das ventaneiras. Essa variável pode indicar instabilidades e variações na operação, o que pode levar a uma maior incorporação de enxofre. Por outro lado, pode indicar um aumento na produção, o que pode diminuir a concentração de enxofre no gusa.
- Temperatura do gusa: como a reação de dessulfuração que ocorre dentro do alto-forno é endotérmica e ocorre em torno de 1.500 °C, quanto maior a temperatura do gusa, mais favorável será a reação de dessulfuração, reduzindo o teor de enxofre no mesmo.
- Basicidade binária B2: essa variável determina a relação CaO/SiO_2 . O CaO é um dos agentes dessulfurantes do gusa. Portanto, quanto maior a basicidade B2, mais favorável será a dessulfuração e menor será o teor de enxofre no gusa.
- Basicidade quaternária B4: essa variável determina a relação $(\text{CaO}+\text{MgO})/(\text{SiO}_2+\text{Al}_2\text{O}_3)$. Assim como o CaO, o MgO também é um agente dessulfurante do gusa. Portanto, quanto maior a basicidade B4, menor será o teor de enxofre no gusa.

Quadro 3.1 – Dicionário das variáveis do modelo de rede neural.

| Nº | Variável | Descrição | Tipo | Definição |
|----|------------|-----------------------------|----------|---|
| 1 | S Gusa | Enxofre do Gusa | Numérica | Determina o Teor de Enxofre do Gusa(%) |
| 2 | Prod | Produtividade do alto-forno | Numérica | Determina o ritmo de produção do alto-forno (t/dxm ³) |
| 3 | Slag | Slag rate | Numérica | Indica a taxa de escória gerada no processo (kg/t de gusa) |
| 4 | S Carreg | Enxofre carregado | Numérica | Determina a quantidade de enxofre contido nas matérias primas carregadas no Alto-Forno (kg/t de gusa) |
| 5 | Pressao Ar | Pressão de ar na base | Numérica | Determina a resistência do ar ao ser injetado na região das ventaneiras (g/cm ²) |
| 6 | Temp Gusa | Temperatura do gusa | Numérica | Mede a temperatura do gusa para reação de dessulfuração do gusa (°C) |
| 7 | B2 | Basicidade binária B2 | Numérica | Determina a relação CaO / SiO ₂ (adimensional) |
| 8 | B4 | Basicidade quaternária B4 | Numérica | Determina a relação (CaO+MGO) / (SiO ₂ +Al ₂ O ₃) (adimensional) |

Fonte: Elaboração própria, 2023

3.5 Tratamento dos Dados

A base de dados utilizada consistiu em 1.271 observações. Esses dados foram submetidos a processos de tratamento, adotando a remoção de valores nulos, *outliers* e iguais a zero.

3.5.1 Tratamento de *Outliers*

Os limites (inferior e superior) utilizados para identificação de *outliers* foram determinados por meio de análises gráficas realizadas na base de dados e com o auxílio de especialistas no processo de alto-forno. As observações que apresentaram *outliers* foram removidas. O Quadro 3.2 a seguir apresenta os limites inferior e superior estabelecidos para a identificação dos *outliers*.

Quadro 3.2 – Limites para identificação de *outliers*.

| Variável | S_Gusa | Prod | Slag | S_carreg | Pressao_Ar | Temp_Gusa | B2 | B4 |
|-----------------|--------|------|--------|----------|------------|-----------|------|------|
| Limite Inferior | 0,01 | 0,10 | 200,00 | 1,00 | 1.000,00 | 1.350,00 | 1,00 | 0,80 |
| Limite Superior | 0,15 | 2,50 | 400,00 | 10,00 | 3.500,00 | 1.600,00 | 1,30 | 1,50 |

Fonte: Elaboração própria, 2023

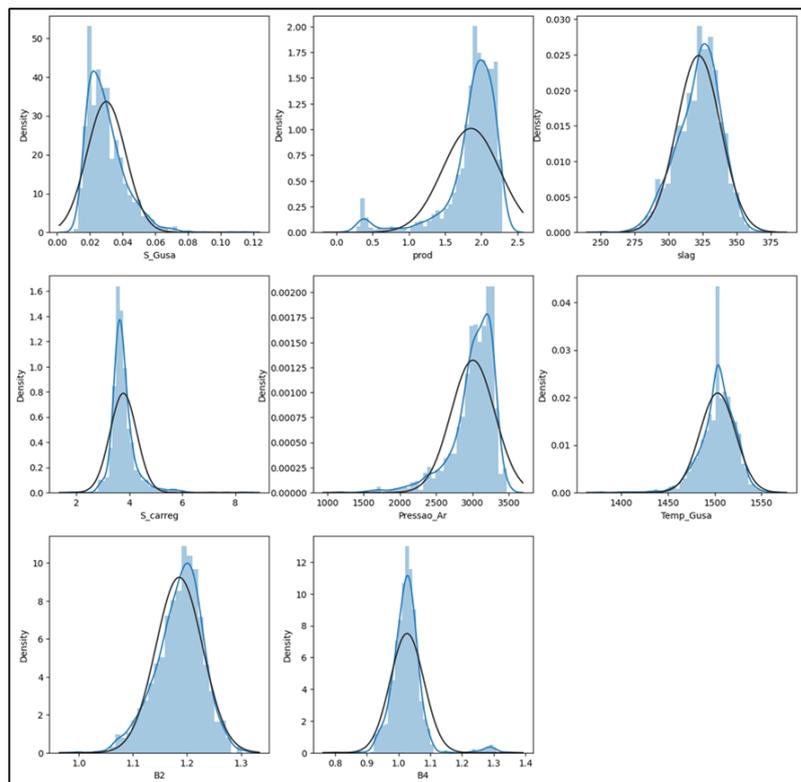
3.5.2 Normalização

As variáveis de entrada consideradas possuem ordens de grandeza significativamente diferentes entre si. Nesse caso, é recomendável realizar uma transformação para uma escala comum, a fim de evitar que certas características possam prevalecer sobre as outras durante o processo de aprendizado.

Foram testados dois modelos para lidar com esse comportamento. O primeiro modelo consistiu em normalizar os dados de todas as variáveis, colocando-os em intervalos de $[-1, 1]$. O segundo modelo atribuiu valores zero para as médias e um para os desvios-padrão de cada variável. Optou-se pelo primeiro modelo, pois ele apresentou resultados mais interessantes.

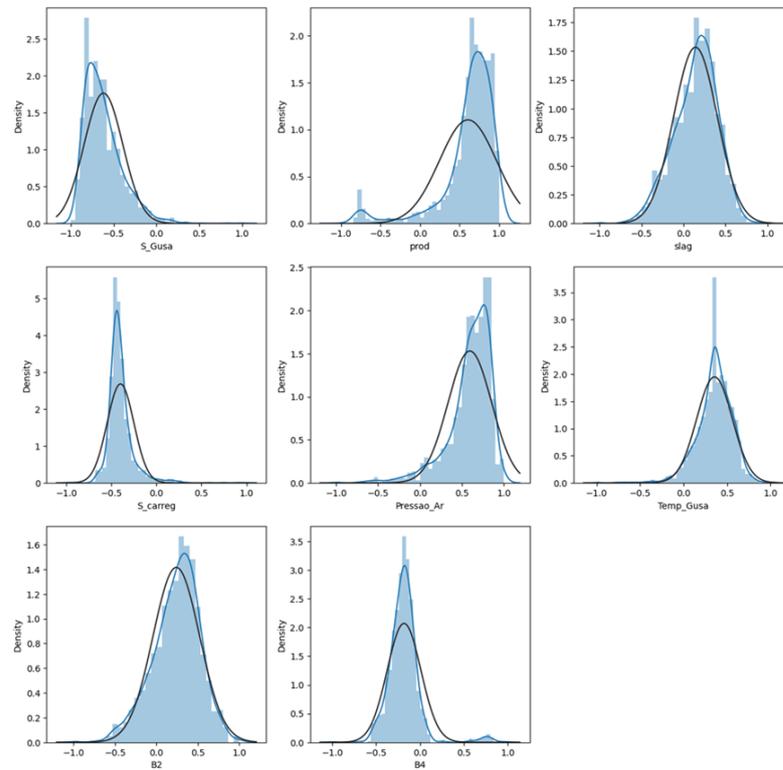
A Figura 5 e a Figura 6 a seguir mostram a distribuição dos dados de cada variável nas escalas original e normalizada, respectivamente.

Figura 5 – Histograma das variáveis em escala original



Fonte: Elaboração própria, 2023

Figura 6 – Histograma das variáveis em escala normalizada [-1,1]



Fonte: Elaboração própria, 2023

3.5.3 Valores nulos, ausentes e iguais a zero

Os valores nulos, ausentes e iguais a zero foram substituídos pela média das demais observações da variável em questão. Esse método é conhecido como Método de Imputação Simples ou Única (IU), conforme Engels e Diehr (2003), por ser realizado apenas uma vez. Tal método foi adotado com o objetivo de minimizar a redução no tamanho da base de dados.

Após processar tais tratamentos estatísticos, a base de dados passou a conter 1.206 observações.

3.6 Modelos de Predição

Foi desenvolvido uma RNAs para predição do teor de enxofre no gusa a partir das sete variáveis de entrada pré-selecionadas conforme já apresentado no estudo. Várias combinações de configurações de RNA's foram realizadas, buscado alcançar uma melhor assertividade. Visando obter valores de referência para avaliar tal assertividade, testamos a base de dados em outros modelos de predição como: Regressão Linear, Árvore de Decisão e *Random Forest*.

Todos os modelos testados foram desenvolvidos usando a linguagem *Python* com a implementação de diversas bibliotecas e funções conforme apresentado no Quadro 3.3 a seguir:

Quadro 3.3 – Bibliotecas *Python* Utilizadas.

| Bibliotecas | Funcionalidades Utilizadas |
|--------------------|--|
| Numpy | Armazenamento e processamento de dados. |
| Scipy | Processamento de dados. |
| Pandas | Análise exploratória de dados. |
| Keras | Otimização de hiperparâmetros de redes. |
| SealBorn | Análise exploratória de dados. |
| Matplotlib | Representação gráfica de dados. |
| Tensor Flow | Desenvolvimento, treinamento e implementação de modelos de Machine Learning. |
| Scikit-learn | Desenvolvimento, treinamento e implementação de modelos de Machine Learning. |

Fonte: Elaboração própria, 2023

3.6.1 Arquitetura da RNA

Para selecionar a configuração do modelo desenvolvido neste estudo adotou-se a ferramenta *Keras Tuner*, disponível na biblioteca *Keras*. Com o *Keras Tuner* é possível definir intervalos para cada hiperparâmetro de uma rede neural, e as possíveis combinações destes hiperparâmetros são testadas e seus resultados são registrados e comparado, retornando ao final a configuração que apresentou o melhor resultado.

A quantidade de camadas intermediárias a serem utilizadas não está contemplada nas interações do *Keras Tuner*, sendo necessário sua pré definição. Durante a rotina do *Keras Tuner*, a quantidade de camadas intermediárias foi constante em todas as configurações testadas. Diante disso, efetuou-se testes variando de duas a cinco camadas, para avaliar quais apresentaram melhores resultados.

O Quadro 3.4 a seguir apresenta os intervalos de quantidade de neurônios (valores mínimo e máximo) e o *step* (valor de salto em cada iteração do algoritmo) para cada camada do modelo.

Quadro 3.4 – Configuração de parâmetros do *Keras Tuner*

| Camada | Intervalo |
|------------------------|--------------------------------|
| Camada Intermediária 1 | Mín: 0 Máx: 120 Step: 10 |
| Dropout 1 | Mín: 0 Máx: 0,5 Step: 0,25 |
| Camada Intermediária 2 | Mín: 0 Máx: 120 Step: 10 |
| Dropout 2 | Mín: 0 Máx: 0,5 Step: 0,25 |
| Camada Intermediária 3 | Mín: 0 Máx: 120 Step: 10 |
| Dropout 3 | Mín: 0 Máx: 0,5 Step: 0,25 |

Fonte: Elaboração própria, 2023

A respeito das funções de ativação a serem testadas no *Keras Tuner*, foram selecionadas as seguintes: Relu, Elu, Selu, Sigmoid e Tangente hiperbólica. Para a camada saída, especificamente adotou-se a função Linear para todas as interações.

Os otimizadores para compilação do modelo de RNA testados no *Keras Tuner* foram: Adam, RMSProp, Adamax, Nadam e SGD.

O algoritmo utilizado para buscar a melhor arquitetura da rede foi o *Random Search*, com uma função objetivo de maximizar o MSE dentre as combinações de parâmetros possíveis dentro das faixas pré-determinadas no *Keras Tuner*.

3.6.2 Desenvolvimento, treinamento e implementação da RNA

Com base nos resultados obtidos com o *Keras Tuner*, foi desenvolvido um modelo de redes neurais utilizando os hiperparâmetros sugeridos. Esse modelo foi treinado usando uma base de treinamento selecionada aleatoriamente. Para as variáveis de entrada da base de validação, o modelo foi aplicado para obter os valores previstos da variável de interesse. Esses valores obtidos foram comparados com os valores reais da base de validação. A métrica utilizada para avaliar o desempenho do modelo foi o MAE. A fim de determinar se o valor de MAE obtido era satisfatório, consultou-se especialistas no processo de dessulfuração da empresa estudada para obter um valor de referência considerado aceitável como desvio. Com o objetivo de realizar uma análise mais detalhada dos resultados, esses valores foram representados em gráficos de correlação. No mesmo gráfico, foram estabelecidos limites inferiores e superiores para a área de aceitação dos erros, com base no desvio aceitável mencionado anteriormente.

3.6.3 Outros modelos de predição e validação cruzada

Ainda com o objetivo de avaliar a qualidade de previsão do modelo de RNA desenvolvido, foram aplicados outros métodos de previsão considerados mais simples sobre a mesma base de dados. Os métodos escolhidos foram regressão linear, árvore de decisão e *random forest*.

Essa análise comparativa incluiu a técnica de validação cruzada (*K-fold Cross Validation*), um processo de amostragem utilizado para avaliar modelos com um número limitado de amostras. Esse procedimento envolve o uso de um parâmetro "K", que determina o número de grupos nos quais uma determinada amostra de dados deve ser dividida (BROWNLEE, 2018). A vantagem dessa técnica é que todos os elementos do conjunto de dados são utilizados tanto para treinamento quanto para teste em algum momento. Normalmente, escolhe-se um valor de K igual a 10 ou 30 (Alpaydin, 2004). Neste trabalho, adotou-se um valor de K igual a 20.

3.7 Modelos de Regressão

3.7.1 Regressão Linear

O termo “Regressão” surgiu em 1885 com o antropólogo, matemático e estatístico Francis Galton. As primeiras aplicações do método surgiram na Antropometria, ou seja, estudo das medidas e da matemática dos corpos humanos.

Ao estudar as estaturas de pais e filhos, Galton observou que filhos de pais com altura baixa em relação à média tendem a ser mais altos que seus pais, e filhos de pais com estatura alta em relação à média tendem a ser mais baixos que seus pais, ou seja, as alturas dos seres humanos em geral tendem a regredir à média.

Regressão linear é um tipo de algoritmo supervisionado, portanto antes de entender como funciona o algoritmo é importante conhecer o que seria aprendizado supervisionado.

3.7.1.1 Conceitos importantes

Antes de prosseguir, há alguns conceitos importantes que precisam estar bem claros na sua mente.

Variável independente ou preditora: é aquela que será passada para o modelo, tendo influência na variável que queremos encontrar. Por exemplo: Se queremos prever as vendas de sorvete, a estação do ano pode interferir nas vendas. Variável alvo ou dependente: é a variável que queremos prever. No exemplo acima seria as vendas de sorvete.

Regressão linear é um algoritmo supervisionado de machine learning usado para estimar o valor de algo baseado em uma série de outros dados históricos, portanto olhando para o passado você pode “prever” o futuro.

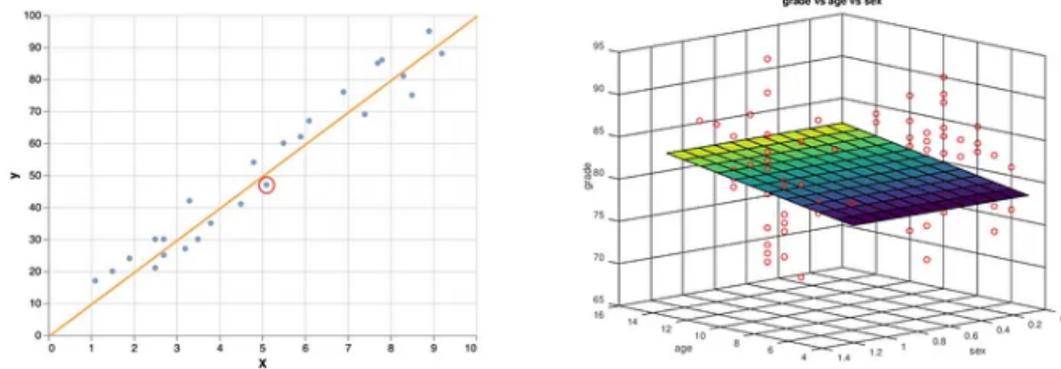
Existem 2 tipos de regressão linear: simples e a múltipla.

Regressão linear simples : refere-se quando temos somente uma variável independente (X) para fazermos a predição.

Regressão linear múltipla: refere-se a várias variáveis independentes (X) usadas para fazer a predição.

E com isso a forma de representação de cada regressão varia conforme o tipo, como por exemplo a representação gráfica em uma regressão linear simples é uma reta em um plano de 2 dimensões, já em uma regressão linear múltipla, sua representação é feita em um plano que pode ser nD.

Figura 7 – À esquerda regressão linear simples, à direita regressão múltipla

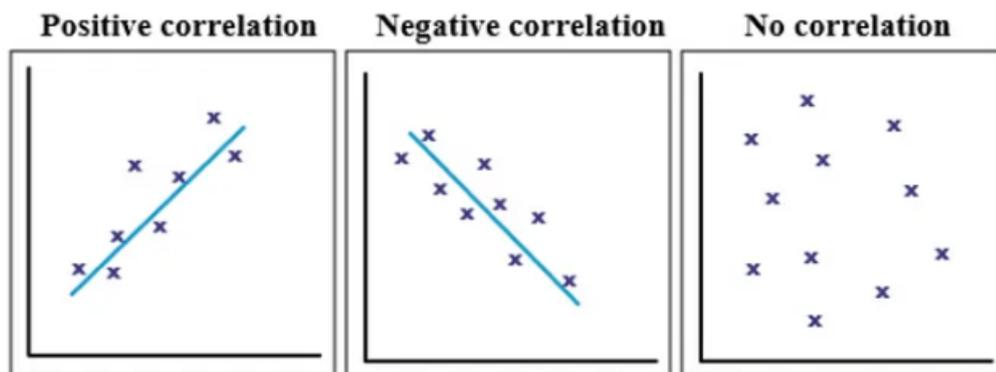


Fonte: <https://medium.com/>, 2023

Correlação Linear de Pearson: mede a correlação linear entre a nuvem de pontos. O resultado varia entre -1 e 1:

- 1: Correlação linear perfeita negativa
- 1: Correlação linear perfeita positiva
- 0: Não tem correlação linear

Figura 8 – Correlação positiva 2- Correlação negativa 3- Sem correlação



Fonte: <https://medium.com/>, 2023

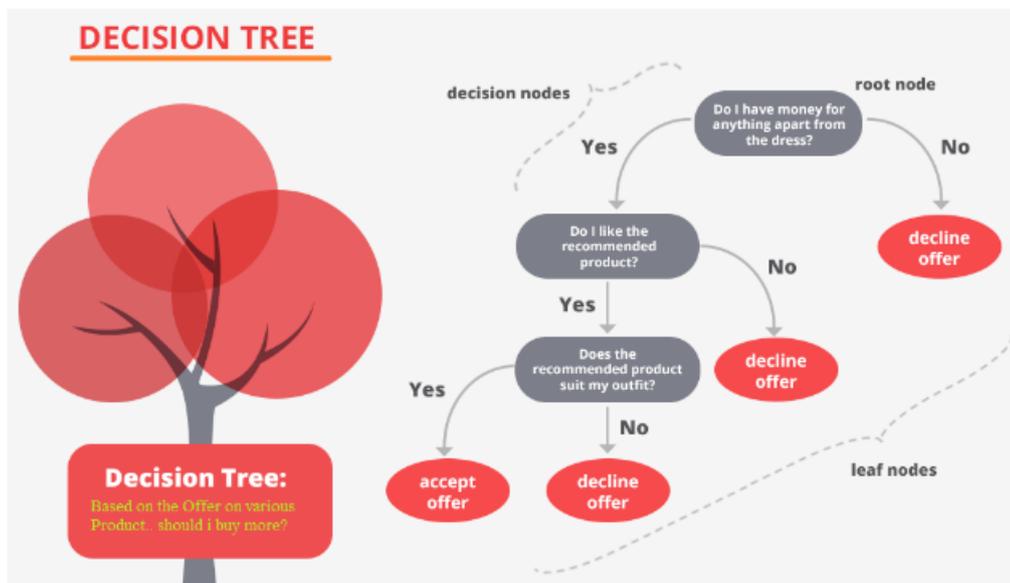
O objetivo da regressão linear é encontrar uma reta que consiga definir bem os dados e minimizar a diferença entre o valor real e a saída calculada pelo modelo.

3.7.2 Árvores de Decisão

Uma árvore de decisão é um algoritmo de aprendizado de máquina supervisionado que é utilizado para classificação e para regressão. Isto é, pode ser usado para prever categorias discretas (sim ou não, por exemplo) e para prever valores numéricos (o valor do lucro em reais).

Assim como um fluxograma, a árvore de decisão estabelece nós (*decision nodes*) que se relacionam entre si por uma hierarquia. Existe o nó-raiz (*root node*), que é o mais importante, e os nós-folha (*leaf nodes*), que são os resultados finais. No contexto de machine learning, o raiz é um dos atributos da base de dados e o nó-folha é a classe ou o valor que será gerado como resposta.

Figura 9 – Representação esquemática da rede do tipo *perceptron*



Fonte: Data Science Foundation, 2023

Na ligação entre nós, temos regras de “se-então”. Ao chegar em um nó A, o algoritmo se pergunta acerca de uma regra, uma condição, como “se a característica X do registro analisado é menor do que 15?”. Se for menor, então ele vai para um lado da árvore; se for maior, então ele vai para outro. No próximo nó, segue a mesma lógica.

É um algoritmo que segue o que chamamos de “recursivo” em computação. Ou seja, ele repete o mesmo padrão sempre na medida em que vai entrando em novos níveis de profundidade. É como se uma função chamasse a ela mesma como uma segunda função para uma execução paralela, da qual a primeira função depende para gerar sua resposta.

O grande trabalho da árvore é justamente encontrar os nós que vão ser encaixados em cada posição. Quem será o nó raiz? Depois, quem será o nó da esquerda? E o da direita?

Para isso, é preciso realizar alguns importantes cálculos. Uma abordagem comum é usar o ganho de informação e a entropia. Essas duas variáveis dizem respeito à desorganização e falta de uniformidade nos dados. Quanto mais alta a entropia, mais caóticos e misturados estão os dados. Quanto menor a entropia, mais uniforme e homogênea está a base.

3.7.3 *Random Forest*

Em português, *Random Forest* significa floresta aleatória. Este nome explica muito bem o funcionamento do algoritmo.

3.7.3.1 Métodos *Ensemble*

Para entender o algoritmo *Random Forest*, precisamos primeiramente conhecer os métodos ensemble, dos quais ele faz parte.

Estes métodos são construídos da mesma forma que algoritmos mais básicos, como regressão linear, árvore de decisão ou knn, por exemplo, mas possuem uma característica principal que os diferenciam, a combinação de diferentes modelos para se obter um único resultado.

Essa característica torna esses algoritmos mais robustos e complexos, levando a um maior custo computacional que costuma ser acompanhado de melhores resultados.

Normalmente na criação de um modelo, escolhemos o algoritmo que apresenta o melhor desempenho para os dados em questão. Podemos testar diferentes configurações deste algoritmo escolhido, gerando assim diferentes modelos, mas no fim do processo de *machine learning*, escolhemos apenas um.

Com um método *ensemble* serão criados vários modelos diferentes a partir de um algoritmo, mas não escolheremos apenas um para utilização final, e sim todos.

Com muitos instrumentos para um resultado - Orquestratodologia, poderemos ter um resultado para cada modelo criado. Se criarmos 100 modelos, teremos 100 resultados, que serão agregados em apenas um.

Em problemas de regressão poderá ser utilizada a média dos valores para obtenção do resultado final, e em problemas de classificação o resultado que mais se repete será o escolhido.

Há casos onde o resultado de um modelo será utilizado na criação do próximo, criando uma dependência entre os modelos, e levando a um único resultado final, gerado a partir de vários resultados intermediários.

Muitos métodos *ensemble* dependem do conceito de árvore de decisão, sendo de grande valia o conhecimento deste conceito no aprendizado dos métodos. Inclusive, quem já conhece árvores de decisão aprenderá os métodos ensemble com muita facilidade e rapidez.

3.7.3.2 Como funciona o algoritmo *Random Forest*

Como já mencionamos, no algoritmo *Random Forest* serão criadas várias árvores de decisão, sendo este conhecimento fundamental para o entendimento do algoritmo.

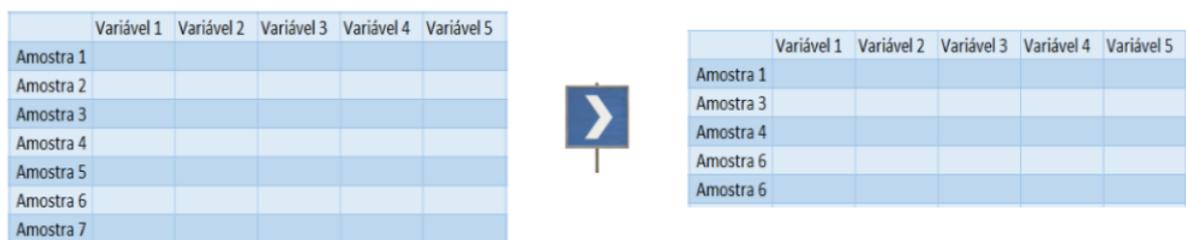
As Árvores de Decisão, ou *Decision Trees*, estabelecem regras para tomada de decisão. O algoritmo criará uma estrutura similar a um fluxograma, com “nós” onde uma condição é verificada, e se atendida o fluxo segue por um ramo, caso contrário, por outro, sempre levando ao próximo nó, até a finalização da árvore.

3.7.3.3 Seleção de amostras

Diferentemente do que acontece na criação de uma árvore de decisão simples, ao utilizar o *Random Forest*, o primeiro passo executado pelo algoritmo será selecionar aleatoriamente algumas amostras dos dados de treino, e não a sua totalidade.

Nesta etapa é utilizado o bootstrap, que é um método de reamostragem onde as amostras selecionadas podem ser repetidas na seleção. Com esta primeira seleção de amostras será construída a primeira árvore de decisão.

Figura 10 – *Bootstrap* - Método de reamostragem



Fonte: <https://didatica.tech/>, 2022

3.7.3.4 Seleção das variáveis para cada nó

Conforme vimos nos detalhes sobre a construção de uma árvore de decisão, para começar é preciso definir o primeiro nó da árvore (nó raiz), que será a primeira condição verificada, dando origem aos dois primeiros ramos.

Utilizando o algoritmo de entropia ou o índice Gini, será escolhida a melhor variável para compor o nó raiz, variando de acordo com o método utilizado.

No *Random Forest* a definição desta variável não acontece com base em todas as variáveis disponíveis. O algoritmo irá escolher de maneira aleatória (*Random*) duas ou mais variáveis, e então realizar os cálculos com base nas amostras selecionadas, para definir qual dessas variáveis será utilizada no primeiro nó.

Para escolha da variável do próximo nó, novamente serão escolhidas duas (ou mais) variáveis, excluindo as já selecionadas anteriormente, e o processo de escolha se repetirá. Desta forma a árvore será construída até o último nó. A quantidade de variáveis a serem escolhidas pode ser definida na criação do modelo.

Figura 11 – Seleção das variáveis para cada nó



Fonte: <https://didatica.tech/>, 2022

3.7.4 Redes Neurais

Redes Neurais Artificiais são modelos computacionais formadas por um conjunto de neurônios artificiais, modelados matematicamente com inspiração no funcionamento de um neurônio biológico. Suas principais características são associadas à possibilidade de aprender, a partir das informações do ambiente em que está operando, a capacidade de processamento paralelamente distribuído e a capacidade de resolver problemas que não são linearmente separáveis. Haykin (2001, p. 28) aponta que “uma rede neural é uma máquina que é projetada para modelar a maneira como o cérebro realiza uma tarefa particular (...); a rede é normalmente implementada utilizando-se componentes eletrônicos ou é simulada por programação”.

O primeiro estudo sobre redes neurais e proposta de modelamento matemático do comportamento do neurônio e da rede de neurônios que formam o cérebro foi discutida por McCulloch e Pitts no artigo “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” publicado originalmente em 1943 no *Bulletin of Mathematical Biophysics* e republicado em 1990 pelo *Bulletin of Mathematical Biology*.

Podem-se destacar duas características como sendo as principais do modelo proposto por McCulloch e Pitts, a saber: i) há uma única saída (resposta) no neurônio, em função da soma de suas entradas; ii) considerando seu caráter de tudo ou nada (*all-or-none*), o neurônio está em estado ativado ou inibido, a depender do somatório de suas entradas.

A partir desse princípio originou-se a função de ativação (*threshold*) como sendo o mecanismo para definição se a soma das entradas nos neurônios resultaria em uma excitação ou inibição, o que propiciará ou não na ativação de um próximo neurônio.

Outro aspecto importante da proposta de McCulloth e Pitts (1943) está relacionada às influências das entradas recebidas por um neurônio. No caso de haver uma entrada de valor positivo, haverá contribuição para a excitação do neurônio, por outro lado, um valor negativo contribuirá para a sua inibição.

A partir dessa proposição originou-se o conceito de pesos sinápticos, amplamente utilizados nas redes neurais artificiais, como forma de atribuir diferentes pesos para os estímulos recebidos pelo neurônio.

De acordo com Eberhart e Dobbins (p. 17, 1990, tradução nossa) foi Hebb “o primeiro a definir um método de atualização dos pesos sinápticos”. Ainda, segundo Eberhart e Dobbins (1990) quatro pontos de contribuição de Hebb são utilizados atualmente nas redes neurais artificiais, a saber:

1. ele afirmou que em uma rede neural, a informação é armazenada no peso das sinapses (conexões);
2. Ele postulou uma taxa de aprendizagem de peso de conexão que é proporcional ao produto dos valores de ativação dos neurônios. Observe que seu postulado assumiu que os valores de ativação eram positivos. Porque não foi fornecido um meio para que os pesos fossem diminuídos, eles poderiam, teoricamente, serem infinitamente altos. Aprendizado que envolve neurônios com valores de ativação negativa também foi rotulado por Hebb. Isso não está incluído na formulação original de Hebb, mas é uma extensão lógica da mesma;
3. Ele assumiu que os pesos são simétricos. Isto é, o peso de uma conexão do neurônio A ao neurônio B é o mesmo que o de B para A. Embora isso possa ou não ser verdade em uma rede neural biológica, é frequentemente aplicado a implementações em ferramentas de redes neurais artificiais;
4. Ele postulou uma teoria de montagem de células, afirmando que, à medida que a aprendizagem ocorre, as forças e os padrões de conexões de sinapse (pesos) mudam, e as montagens de células são criadas por essas mudanças. Dito de outra forma, se a ativação simultânea de um grupo de células fracamente ligadas ocorre repetidamente, estas células tendem a se fundir numa montagem mais fortemente ligada (Eberhart e Dobbins p. 18 e 19, 1990, tradução nossa).

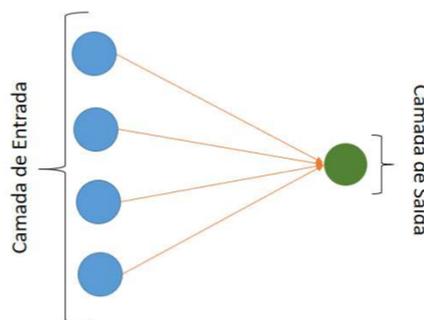
Embora os estudos de Hebb tenham sido originalmente realizados com o objetivo de melhor compreensão da rede neural biológica, sua contribuição para o desenvolvimento das redes artificiais foi essencial, uma vez que a partir da teoria de Hebb que se propõe a alteração dos pesos sinápticos dos neurônios como forma de aprendizado das redes neurais artificiais.

Em 1951, Marvin Minsky desenvolveu o Snark, o primeiro neurocomputador da história que era capaz de realizar os ajustes de seus pesos sinápticos automaticamente utilizando a regra de aprendizagem de Hebb. Contudo, o sistema nunca foi implementado para funções de processamento com alguma aplicação prática (POULTON 2001; ALDRICH, 2002; KRIESEL, 2007).

Rosenblatt (p. 387, 1958, tradução nossa) publicou, em 1958, o artigo *The Perceptron: A Probabilistic Model for Informations Storage and Organization in the Brains* em que foi apresentada “a teoria desenvolvida para um sistema nervoso hipotético, ou máquina, chamada perceptron”.

Foi com base no algoritmo perceptron que Rosenblatt, Wightman e colaboradores desenvolveram o primeiro neurocomputador a ser implementado com sucesso. Denominado de *Mark I Perceptron*, o sistema conseguia reconhecer números com base em um sensor de imagem de 20x20 pixels, além de reconhecer padrões e realizar associações. Rosenblatt desenvolveu ainda outras versões do perceptron com camadas de neurônios, comutadores de limiar e regra de aprendizagem para ajustes dos pesos sinápticos. Destaca-se que as redes perceptrons eram limitadas a duas camadas de neurônios, conforme apresentado na Figura 12, uma de entrada de dados e a outra de saída de dados (POULTON 2001; KRIESEL, 2007).

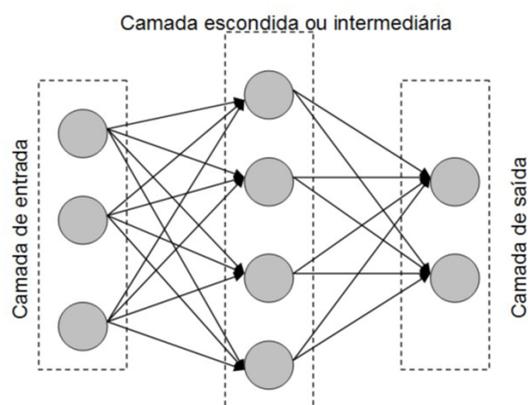
Figura 12 – Representação esquemática da rede do tipo perceptron



Fonte: POULTON 2001; KRIESEL, 2007

Segundo Pulton (2001), em 1986, com a publicação de *Parallel Distributed Processing*, é apresentada as redes perceptrons multicamadas (*multilayer perceptron*) com a capacidade para solução de problemas não linearmente separáveis, sendo a principal característica a adição de pelo menos uma camada intermediária de neurônios entre as camadas de entrada e saída, conforme representado na Figura 13.

Figura 13 – Representação esquemática da rede do tipo perceptron



Fonte: Pulton, 2001

4 Resultados e Discussão

Este capítulo apresenta e discute acerca dos resultados obtidos pela rede neural desenvolvida neste trabalho, assim como transcorre sobre a análise comparativa com os resultados obtidos pelos demais métodos de predição descritos na seção 3.6.

4.1 Resultados da RNA

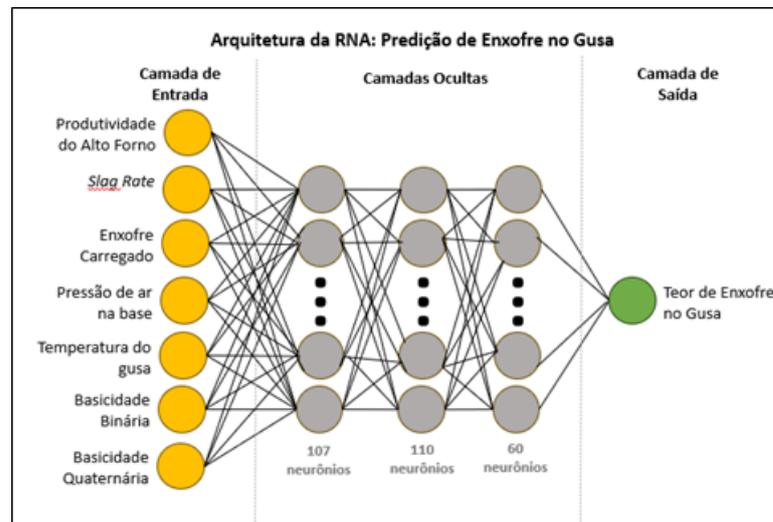
Conforme mencionado na metodologia deste trabalho, a arquitetura adotada na rede foi obtida através do *Keras Tuner*. O Quadro 4.1 apresenta os resultados obtidos por esta ferramenta.

Quadro 4.1 – Configuração obtida pelo *Keras Tuner*

| Qtde. Camadas | Qtde. de Neurônios | Função de Ativação | Taxa de Dropout | Otimizador |
|------------------------|--------------------|--------------------|-----------------|------------|
| Camada de Entrada | - | relu | 0.05 | Adam |
| Camada Intermediária 1 | 107 | elu | 0.05 | |
| Camada Intermediária 2 | 110 | selu | 0.25 | |
| Camada Intermediária 3 | 60 | tanh | 0.15 | |
| Camada de Saída | - | linear | - | |

A Figura 14 a seguir apresenta a arquitetura adotada para a rede neural testada.

Figura 14 – Arquitetura da rede neural



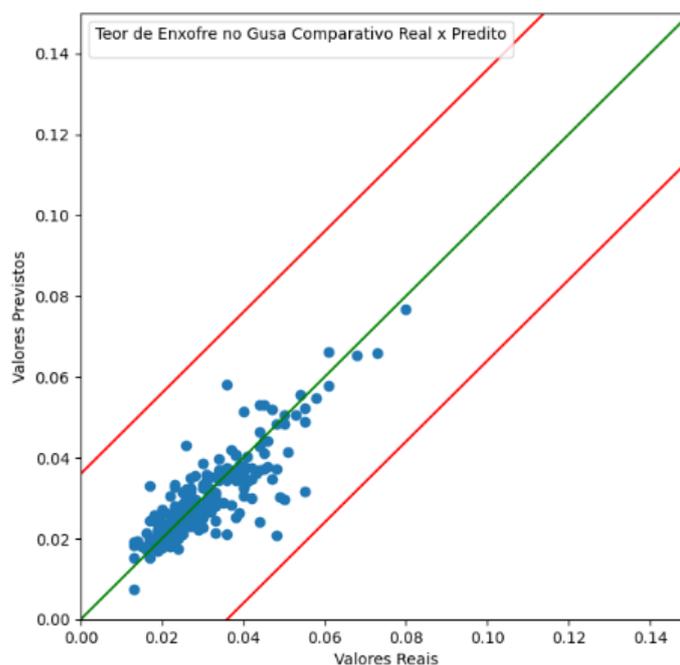
Fonte: Elaboração própria, 2023

A rede neural configurada a partir destes resultados retornou um MAE de 0,0044%. Vale lembrar que este valor já está na base original da variável, e não na base normalizada em escala [-1,1].

Com o intuito de avaliar a adequação do modelo à uma precisão satisfatória do ponto de vista dos usuários do modelo, definiu-se limites superior e inferior com variações de $\pm 0,006\%$.

A Figura 15 a seguir apresenta um gráfico comparativo entre os valores preditos pela rede e os valores reais da base de validação.

Figura 15 – Resultado do treinamento e validação da RNA de predição de teor de enxofre no gusa



Fonte: Elaboração própria, 2023

É possível verificar que a maior parte das predições ficou dentro dos limites superior e inferior definidos. No entanto, existem valores preditos cujo o desvio obtido supera o valor real correspondente.

4.2 Modelos Alternativos para Análise Comparativa

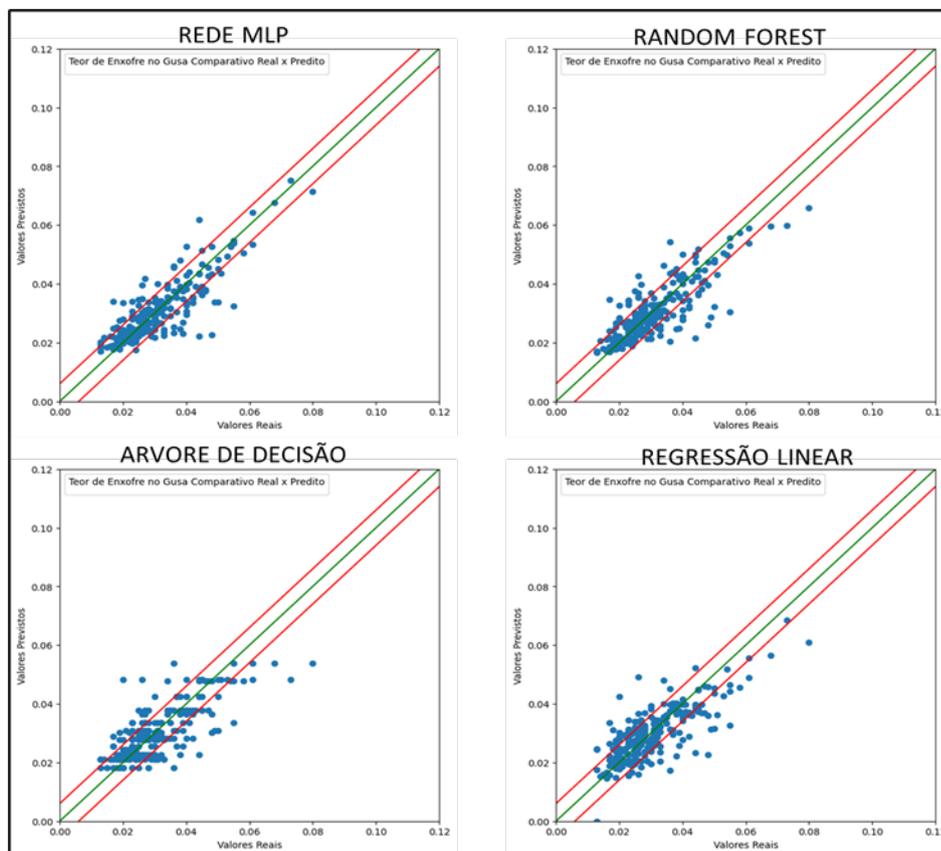
Os métodos de Regressão Linear, Árvore de Decisão e *Random Forest* foram utilizados sobre a mesma base de treinamento e testados na mesma base de validação. Em primeiro momento, calculou-se para cada método o erro absoluto médio (MAE) para efeito de comparação, conforme demonstrado no quadro 4.2 a seguir.

Quadro 4.2 – Configuração obtida pelo *Keras Tuner*

| Modelo de Predição | MLP "RNA " | LR "Regressão Linear" | DT "Árvore de Decisão" | RF "Random Forest" |
|--------------------|---------------|--------------------------|---------------------------|-----------------------|
| MAE | 0.0043 | 0.0053 | 0.0052 | 0.0044 |

Entre os modelos contemplados nessa análise, a MLP apresentou o menor erro absoluto médio (0,0043%), seguida dos modelos *Random Forest* (0,0044%), *Árvore de Decisão* (0,0052%) e *Regressão Linear* (0,0053%). A Figura 16 a seguir apresenta os gráficos com os valores preditos e reais de cada modelo testado.

Figura 16 – Comparativo de resultados entre modelos de predição



Fonte: Elaboração própria, 2023

É possível perceber que os modelos MLP e RF apresentaram maior concentração de pontos dentro da faixa de desvios considerados aceitáveis pelos potenciais usuários dos modelos de predição do enxofre no gusa, conforme Quadro 4.3 a seguir:

Quadro 4.3 – Modelos de Predição

| Modelos de Predição | | MLP | DT | RF | LR |
|--------------------------------------|-------|-------|-------|-------|-------|
| Observações da Base de Validação | | 242 | 242 | 242 | 242 |
| Valores Preditos com erro <0,006% | Qtde. | 187 | 169 | 183 | 166 |
| | % | 77,3% | 69,8% | 75,6% | 68,6% |

Fonte: Elaboração própria, 2023

No entanto, se faz necessário verificar que a diferença entre os erros de predição obtidos em cada método são estatisticamente representativos. Para tal, utilizou-se o teste ANOVA com a hipótese nula de que não há diferença entre os métodos. O Quadro 4.4 expressa os resultados deste teste.

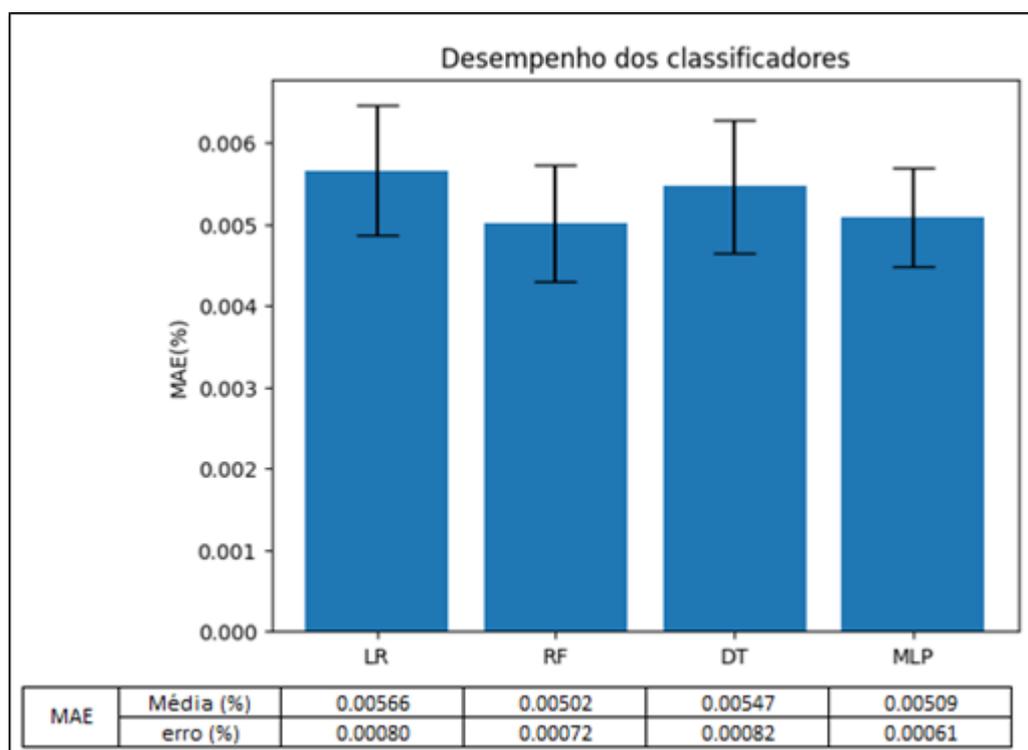
Quadro 4.4 – Resultados ANOVA

| ANOVA | | | | | | |
|-------------------|-----------|-----------|-----------|----------|----------------|------------------|
| Fonte da variação | <i>SQ</i> | <i>gl</i> | <i>MQ</i> | <i>F</i> | <i>valor-P</i> | <i>F crítico</i> |
| Entre grupos | 4,75E-05 | 3 | 1,58E-05 | 0,364147 | 0,778912 | 2,614136 |
| Dentro dos grupos | 0,041909 | 964 | 4,35E-05 | | | |
| Total | 0,041956 | 967 | | | | |

4.3 Validação Cruzada

A técnica de validação cruzada *k-fold* foi utilizada visando minimizar o risco de que a base de treinamento, escolhida aleatoriamente, beneficie um método em relação aos outros. Os resultados obtidos mostraram que o desempenho do modelo de MLP (MAE = 0,00509%), *Random Forest* (MAE = 0,00502%), Árvore de Decisão (MAE = 0,00547%) e Regressão Linear (MAE = 0,00566%) apresentaram valores similares, conforme representado na Figura 17 a seguir.

Figura 17 – Validação cruzada dos modelos de predição



Fonte: Elaboração própria, 2023

5 Conclusão

Este trabalho apresentou o uso de redes neurais artificiais com o objetivo de prever o teor de gusa a ser dessulfurado em aciarias. A assertividade de um modelo de predição com este objetivo tem grande influência no consumo de agentes dessulfurantes, impactando nos custos de produção de placas na aciaria.

A partir de consultas a especialistas em processos de altos-fornos, identificaram-se as variáveis com influência sobre o teor de enxofre no gusa. Esta etapa contemplou o atendimento ao primeiro objetivo específico do trabalho.

O segundo objetivo do trabalho, que contempla a obtenção da base de dados, foi atingido através do processo de extração pela ferramenta PIMS. Técnicas estatísticas como gráficos de *box-plot*, *heatmap* e regressões elucidaram o comportamento das variáveis. Tais análises permitiram reduzir o número de variáveis de entrada dos modelos de regressão.

Com o *Keras Tuner*, foi possível testar várias topologias de redes neurais, obtendo a combinação de parâmetros que apresentou o melhor resultado. A acurácia do modelo, expressa pelo indicador de erro médio absoluto (MAE), foi de 0,0044%. Este resultado é inferior ao desvio considerado aceitável ($\pm 0,006\%$) pelos especialistas no processo de dessulfuração estudado.

A análise gráfica dos valores reais e preditos pela rede sobre a base de validação também mostraram seu bom desempenho, uma vez que 77,3% dos valores preditos apresentaram erros dentro de limites aceitáveis. No entanto, é desejável que esse valor seja ainda maior, sendo necessário buscar aprimoramentos no modelo.

Outro ponto importante foi que algumas observações que se encontraram fora da faixa tiveram desvios com ordens de grandeza superiores ao valor absoluto da variável de interesse. Em processos industriais, os desvios de maior magnitude, ainda que em menor frequência, exigem atenção significativa. Isso se intensifica quando a variável de interesse tem relação de elasticidade representativa sobre aspectos importantes, como custos, segurança e meio ambiente.

Embora a rede neural desenvolvida tenha apresentado bons resultados, foi feita uma análise comparativa com o desempenho de outros modelos de predição. Os modelos testados (Regressão Linear, Árvore de Decisão e *Random Forest*) apresentaram desempenho similar ao obtido pela rede neural. Essa equidade se confirmou com a validação cruzada *k-fold*. Pelo fato dos modelos de predição alternativos à RNA serem de desenvolvimento mais simples, vale a reflexão sobre a escolha da aplicação destes em detrimento da RNA.

Diante dos resultados apresentados, os modelos de regressão se mostraram como ferramentas de grande potencial para o problema abordado neste trabalho. No entanto, devido à criticidade do processo, faz-se necessário desdobramentos acerca dos processos de obtenção da base de dados para treinamento (qualidade e tamanho da base) e de avaliações de novas topologias de redes. Estes desdobramentos poderão oferecer resultados mais assertivos, refletindo em melhorias nos indicadores do processo de dessulfuração.

5.1 Sugestão de Trabalhos Futuros

Diante dos resultados apresentados neste trabalho, foram identificados alguns possíveis desdobramentos, gerando demandas e sugestões para trabalhos futuros:

1. Realizar a atualização da base de dados, testando a inclusão de novas variáveis, com o objetivo de aumentar a precisão das predições em relação a alguns poucos desvios de alta magnitude em relação aos valores reais.
2. Testar novas configurações de redes neurais e outros modelos de regressão, buscando obter uma maior acurácia nas predições.
3. Integrar o modelo como uma ferramenta online para a predição do teor de enxofre no gusa, a fim de determinar a massa de dessulfurantes a ser adicionada ao gusa no processo de dessulfuração na aciaria.
4. Testar a rede neural para a predição de outros elementos presentes no gusa.

6 Referências Bibliográficas

BARCELOS, A. F.; TOSHIMITSU, H. S. Modelomatemático Neural Artificial para Cálculo de Rendimento Metálico Teórico de um Laminador de Tiras A Frio. 2018.

CARVALHO, André Carlos Ponce de Leon Ferreira e BRAGA, Antonio de Pádua e LUDERMIR, Teresa Bernarda. Fundamentos de redes neurais artificiais. 1998. Rio de Janeiro: DCC/IM, COPPE/Sistemas, NCE-UFRJ. . Acesso em: 21 jun. 2023.

CARVALHO, V. P. Previsão de Séries Temporais no Mercado Financeiro de Ações com o uso de Rede Neural Artificial. 2018. CASTRO, L. F. A.; TAVAJRES, R. P.; MORELATO, A. P. Injeção de carvão pulverizado nas ventaneiras do Alto-Forno. 1. ed. Belo Horizonte: UFMG, Departamento de Engenharia Metalúrgica, 1997.

CHOLLET, F. et al. Keras. 2015. Disponível em: <https://keras.io>.

DAVID, S. F. Análise da influência dos fatores operacionais dos altos-fornos sobre o teor do silício e enxofre no ferro gusa por meio de redes neurais artificiais. 2013. 89 f: Dissertação (Mestre em Engenharia Metalúrgica e de Materiais), Instituto Federal do Espírito Santo, Vitória, 2013.

ENGELS JM, DIEHR P. *Imputation of missing longitudinal data: a comparison of methods. Journal of Clinical Epidemiology.* 2003; 56:968-976.

FACELI, K. et al. Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina. 2. ed. Rio de Janeiro, RJ, Brasil: LTC, 2011. GOODFELLOW, I. BENGIO, Y. COURVILLE, A. Deep Learning. [S.l.] MIT Press, 2016.

FURTADO M. I. V., Redes Neurais Artificiais: Uma Abordagem Para Sala de Aula. Ponta Grossa (PR): Atena Editora, 2019. GHOSH, A., e CHATTERJEE, A. (2008). *Ironmaking and Steelmaking: Theory and Practice. New Delhi: PHI Learning Provate Limited.*

GHOSH, A., e CHATTERJEE, A. (2010). *Ironmaking and Steelmaking: Theory and Practice. New Delhi: PHI Learning Private Limited.*

Hinton, G., Srivastava, N., and Swersky, K. (2012). *Neural networks for machine learning. lecture 6e. BIBLIOGRAPHY BIBLIOGRAPHY.*

KIM, D. K. K. J. *Elastic exponential linear units for convolutional neural networks.* 2020.

KINGMA, D. P.; BA, J. Adam: *A Method for Stochastic Optimization.* 2017. Citado na página 44.

KRENKER, A.; KOS, A.; BEŠTER, J. *Introduction to the artificial neural networks. [S.l.]: INTECH Open Access Publisher, 2011.*

JANKOWSKI, D. W. N. *Survey of neural transfer functions. Neural Computing Surveys*, 1999.

JUNIOR, M. A. V. Modelo híbrido associando cálculos termodinâmicos e redes neurais artificiais para a predição da temperatura do aço desde o fim de sopro até a metalurgia secundária, 2018.

MARCON, A. Saiba o que são redes neurais artificiais e de que forma aplicá-las na indústria. *A voz da Indústria*, 2019. Disponível em: <https://avozdaindustria.com.br/inovacao/saiba-o-que-sao-redes-neurais-artificiais-e-de-que-forma-aplica-las-na-industria>. Acesso em: 20 de jun. 2023.

MARIANO, D. C. B. Métricas de avaliação em machine learning. DIEGOMARIANO, 2021. Disponível em: <https://diegomariano.com/metricas-de-avaliacao-em-machine-learning/>. Acesso em: 15, jun. 2023.

NWANKPA, Chigozie et al. *Activation Functions: Comparison of trends in Practice and Research for Deep Learning. [S.l.: s.n.], 2018. eprint: arXiv:1811.03378. Citado nas pp. 19, 20.*

SAMPAIO, P. T. Modelagem Térmica de um Forno Panela utilizando Redes Neurais Artificiais: Uma abordagem para Detecção de *Outliers* em Dados Categóricos. Universidade Federal de Minas Gerais, 2006.

SAYD, D. F.; SILVA, M. B.; NASCIMENTO, R. C.; MACHADO, M. L. P. Uma abordagem para Detecção de *Outliers* em Dados Categóricos. [s.l.] (Dissertação de Mestrado no Intituto de Computação) - Universidade Estadual de Campinas, 2004. SILVA, F. R. Uma abordagem para Detecção de *Outliers* em Dados Categóricos. [s.l.] (Dissertação de Mestrado no Intituto de Computação) - Universidade Estadual de Campinas, 2004.

SRIVASTAVA, Nitish et al. *Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research*, v. 15, n. 1, p. 1929-1958, 2014.

RIZZO, I. V., e CANATO, L. C. C. INTELIGÊNCIA ARTIFICIAL: FUNÇÕES DE ATIVAÇÃO. *Revista Prospectus*, Itapira-SP, v. 2, n. 2, p. 51-65, Fev, 2020. Disponível em: <https://prospectus.fatecitapira.edu.br/index.php/pst/article/view/37>. Acesso em: 20 de jun. 2023.

TAMURA, K.; ONO, K.; NISHIDA, N. *Effects of Operating Factors of Blast Furnaces on the Contents of Silicon and Sulphur in Pig Iron. Tetsu-to-Hagané, Tokyo*, v. 67, p. 2635-2644, Mai 1981.

WAKELUN, D. H., e RICKETTS, J. A. (1999). *The Nature of Ironmaking. Em D. H. Wakelun, e J. A. Ricketts, The Making, Shaping, and Treating of Steel: Ironmaking Volume 2 de The Making, Shaping, and Treating of Steel (11ª ed.). AISE Steel Foundation.*

YANG, Y., RAIPALA, K., e HOLAPPA, L. (2013). *Ironmaking. Em S. Seetharaman, A. McLean, R. Guthrie, e S. Seetharaman, Treatise on Process Metallurgy - Volume 3: Industrial Processes (p. 1810). Stockholm: Elsevier Science.*