



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



# **Utilização de técnicas de aprendizado de máquina no processo de produção de síter para os altos-fornos**

**Leandro Ferreira Machado**

João Monlevade, MG  
2023

Leandro Ferreira Machado

## **Utilização de técnicas de aprendizado de máquina no processo de produção de síter para os altos-fornos**

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. Harlei Miguel de Arruda Leite

João Monlevade, MG

2023

## SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

M149u Machado, Leandro Ferreira.

Utilização de técnicas de aprendizado de máquina no processo de produção de sínter para os altos-fornos. [manuscrito] / Leandro Ferreira Machado. - 2023.

45 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Harlei Miguel de Arruda Leite.

Produção Científica (Especialização). Universidade Federal de Ouro Preto. Departamento de Engenharia de Produção.

1. Altos-fornos. 2. Aprendizado do computador. 3. Sinterização. 4. Usinas siderúrgicas. I. Leite, Harlei Miguel de Arruda. II. Universidade Federal de Ouro Preto. III. Título.

CDU 622.78:004.8

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



## FOLHA DE APROVAÇÃO

**Leandro Ferreira Machado**

### Utilização de técnicas de aprendizado de máquina no processo de produção de sínter para os altos-fornos

Monografia apresentada 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 março de 2023

#### Membros da banca

Prof. Dr. Harlei Miguel de Arruda Leite - Orientador (Universidade Federal de Ouro Preto)  
Prof. Dr. George Henrique Godim da Fonseca - (Universidade Federal de Ouro Preto)  
Profa. Dra. Sarah Negreiros de Carvalho Leite - (Universidade Federal de Ouro Preto)

Prof. Dr. Harlei Miguel de Arruda Leite, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 10/07/2023



Documento assinado eletronicamente por **Harlei Miguel de Arruda Leite**, PROFESSOR DE MAGISTERIO SUPERIOR, em 10/07/2023, às 21:16, 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](http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0555984** e o código CRC **6C17321D**.

# Resumo

Este trabalho descreve uma estrutura metodológica para o desenvolvimento de modelos baseados em técnicas de aprendizado de máquina para aplicação no processo de sinterização de uma usina siderúrgica de grande porte. A estrutura proposta compreende modelos analíticos capazes de extrair padrões e tendências nos dados das áreas de sinterização e matérias-primas, a fim de gerar previsões relacionadas à qualidade do sinter produzido, produtividade do processo, consumo de combustíveis e emissão de particulados. O objetivo destes modelos é suportar as decisões operacionais de ajuste de processo e contribuir para um melhor desempenho do sinter nos altos-fornos. Inicialmente, foi proposto um modelo piloto para prever o principal índice de qualidade do sinter, representado pelo *shatter test*. Diversos algoritmos de aprendizado de máquina foram testados, incluindo Regressão Linear, Regressão Polinomial, Árvores de Decisão, *Ensemble Learning* e Redes Neurais Artificiais (RNA). O desempenho dos modelos foi comparado utilizando a métrica de *Mean Absolute Error (MAE)*. Entre os algoritmos testados para previsão do *shatter*, o *Random Forests* apresentou o melhor desempenho. O modelo obteve um *MAE* de 0,31 e exibiu um comportamento gráfico aceitável ao comparar as previsões com os valores reais do conjunto de testes, indicando sua aplicabilidade prática. Os resultados obtidos neste trabalho demonstram que a estrutura proposta foi capaz de aprender com os dados das sinterizações e matérias-primas. Além disso, esta estrutura permite a integração com os demais modelos sugeridos, sendo uma ferramenta flexível e aplicável para as sinterizações.

**Palavras-chaves:** Sinterização. Aprendizado de Máquina. Qualidade do Sinter.

# Abstract

This work describes a methodological structure for the development of models based on machine learning techniques for application in the sintering process of a large steel plant. The proposed structure comprises analytical models capable of extracting patterns and trends in data from the sintering and raw materials areas, in order to generate forecasts related to the quality of the sinter produced, process productivity, fuel consumption and particulate emissions. The objective of these models is to support the operational decisions of process adjustment and contribute to a better performance of the sinter in the blast furnaces. Initially, a pilot model was proposed to predict the main sinter quality index, represented by the shatter test. Several machine learning algorithms were tested, including Linear Regression, Polynomial Regression, Decision Trees, Ensemble Learning and Artificial Neural Networks (ANN). The performance of the models was compared using the Mean Absolute Error (MAE) metric. Among the algorithms tested for shatter prediction, Random Forests presented the best performance. The model obtained a MAE of 0.31 and exhibited an acceptable graphical behavior when comparing the predictions with the actual values of the test set, indicating its practical applicability. The results obtained in this work demonstrate that the proposed structure was able to learn from the sintering and raw material data. In addition, this structure allows integration with the other suggested models, being a flexible and applicable tool for sintering.

**Keywords:** Sinter Plant. Machine Learning. Sinter Quality.

# Lista de ilustrações

Figura 1 – Layout típico de uma usina integrada. . . . .	4
Figura 2 – Fluxograma de sinterização. . . . .	5
Figura 3 – Principais cargas metálicas destinadas aos altos-fornos. . . . .	6
Figura 4 – Principais características do sínter para consumo nos altos-fornos. . . . .	7
Figura 5 – Estrutura de variáveis da sinterização e proposta de aplicação de aprendizado de máquina para realizar previsões. . . . .	8
Figura 6 – Circuito de retorno de uma máquina de sínter. . . . .	9
Figura 7 – Pressão de sucção no processo de sinterização. . . . .	10
Figura 8 – Treinamento, validação e teste: divisão da base de dados. . . . .	17
Figura 9 – Proposta de abordagem do aprendizado de máquina aplicado ao processo de sinterização. . . . .	23
Figura 10 – Proposta de estrutura de tomada de decisão com suporte de algoritmos de previsão da qualidade física do sínter. . . . .	24
Figura 11 – Fluxograma das etapas do trabalho para desenvolvimento dos modelos. . . . .	25
Figura 12 – Informações do banco de dados em formato <i>DataFrame</i> . . . . .	26
Figura 13 – <i>DataFrame</i> após tratamento inicial dos dados. . . . .	27
Figura 14 – Histograma de variáveis do processo de sinterização. . . . .	28
Figura 15 – Tabela de filtro de dados aplicado ao banco de dados do modelo. . . . .	29
Figura 16 – Base de dados tratada após visualização e limpeza dos dados. . . . .	29
Figura 17 – Análise de correlação do <i>shatter</i> com as variáveis de processo e das matérias-primas. . . . .	30
Figura 18 – Configuração do modelo piloto de aprendizado de máquina para previsão do <i>shatter</i> . . . . .	31
Figura 19 – Modelo de rede neural artificial do tipo <i>MLP</i> para previsão do <i>shatter</i> . . . . .	35
Figura 20 – Tabela comparativa da performance dos algoritmos de aprendizado de máquina. . . . .	36
Figura 21 – <i>Random forests</i> - comparativo do valor predito <i>vs</i> valores reais da base de testes. . . . .	37
Figura 22 – Tabela de importância das variáveis na predição do <i>shatter</i> utilizando a técnica de <i>random forests</i> . . . . .	38
Figura 23 – Regressão linear - comparativo do valor predito <i>vs</i> valores reais da base de testes. . . . .	39
Figura 24 – Regressão polinomial - comparativo do valor predito <i>vs</i> valores reais da base de testes. . . . .	39
Figura 25 – Árvores de decisão - comparativo do valor predito <i>vs</i> valores reais da base de testes. . . . .	40
Figura 26 – <i>Ensemble adaboost</i> - comparativo do valor predito <i>vs</i> valores reais da base de testes. . . . .	40

Figura 27 – <i>Gradient boosting</i> - comparativo do valor predito vs valores reais da base de testes. . . . .	41
Figura 28 – Redes neurais artificiais - comparativo do valor predito vs valores reais da base de testes. . . . .	41

# Lista de símbolos

$A$	área da máquina
$h$	altura da camada
$\mu$	média aritmética das variáveis de entrada
$n$	parâmetro da mistura de sínter
$\Delta P_{bed}$	queda de pressão do gás
$\psi$	permeabilidade online média do leito na esteira
$\sigma$	desvio padrão das variáveis de entrada
$\dot{V}$	vazão volumétrica do gás
$X$	valor da variável na base de dados
$Z$	valor da variável $X$ padronizado

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>1.1</b>	<b>Objetivo geral</b>	<b>2</b>
1.1.1	Objetivos específicos	2
<b>1.2</b>	<b>Contribuições</b>	<b>3</b>
<b>1.3</b>	<b>Organização do trabalho</b>	<b>3</b>
<b>2</b>	<b>CONTEXTUALIZAÇÃO TEÓRICA</b>	<b>4</b>
<b>2.1</b>	<b>Contexto da sinterização na siderurgia</b>	<b>4</b>
<b>2.2</b>	<b>O processo de sinterização</b>	<b>5</b>
<b>2.3</b>	<b>Principais cargas metálicas para altos-fornos</b>	<b>6</b>
<b>2.4</b>	<b>Estrutura de variáveis da sinterização</b>	<b>7</b>
2.4.1	Indicadores operacionais	8
2.4.2	Controladores operacionais	11
2.4.3	Características da mistura	13
<b>2.5</b>	<b>Aprendizado de máquina</b>	<b>14</b>
2.5.1	O cenário do aprendizado de máquina	14
2.5.2	Tipos de sistemas do aprendizado de máquina	14
2.5.2.1	Aprendizado supervisionado e não supervisionado	15
2.5.2.2	Aprendizado online e em lote	15
2.5.2.3	Aprendizado baseado em instância e em modelo	15
2.5.3	Preparação da base de dados - coleta, limpeza e visualização	16
2.5.4	Divisão da base de dados para treinamento, validação e teste	16
2.5.5	Medidas de desempenho para avaliar modelos	18
2.5.6	Modelos de aprendizado de máquina	18
2.5.6.1	Regressão linear	18
2.5.6.2	Regressão polinomial	19
2.5.6.3	Árvores de decisão	19
2.5.6.4	Métodos de <i>ensemble learning</i>	20
2.5.6.4.1	<i>Bagging</i>	20
2.5.6.4.2	<i>Boosting</i>	21
2.5.6.5	Redes neurais artificiais	21
<b>3</b>	<b>METODOLOGIA</b>	<b>22</b>
<b>3.1</b>	<b>Abordagem do aprendizado de máquina para aplicação no processo de sinterização</b>	<b>22</b>
<b>3.2</b>	<b>Fluxo de trabalho para desenvolvimento dos modelos</b>	<b>24</b>

<b>3.3</b>	<b>Definição de variáveis do processo de sinterização e coleta de dados . . .</b>	<b>25</b>
<b>3.4</b>	<b>Análise inicial dos dados . . . . .</b>	<b>27</b>
<b>3.5</b>	<b>Visualização dos dados . . . . .</b>	<b>28</b>
<b>3.6</b>	<b>Modelo piloto para previsão do shatter . . . . .</b>	<b>30</b>
3.6.1	Análise de correlação . . . . .	30
3.6.2	Variáveis de entrada e saída selecionadas . . . . .	31
3.6.3	Padronização dos dados . . . . .	32
3.6.4	Divisão do conjunto de dados para treinamento, validação e teste . . . . .	32
3.6.5	Desempenho dos modelos . . . . .	33
3.6.6	Aplicação dos algoritmos de aprendizado de máquina . . . . .	33
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>36</b>
<b>4.1</b>	<b>Avaliação dos modelos de aprendizado de máquina para previsão do <i>shatter</i></b>	<b>36</b>
4.1.0.1	<i>Random forests</i> . . . . .	37
4.1.0.2	Regressão linear e regressão polinomial . . . . .	38
4.1.0.3	Árvores de decisão . . . . .	39
4.1.0.4	<i>Ensembles - adaboost e gradient boosting</i> . . . . .	40
4.1.0.5	Redes neurais artificiais . . . . .	41
<b>5</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>42</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>44</b>

# 1 Introdução

Com as novas tecnologias computacionais e a explosão do registro de dados da era do *Big Data*, as técnicas de aprendizado de máquina vem ganhando um novo e grande impulso, com a aplicação destes métodos em uma ampla variedade de campos, incluindo processamento de imagens e vídeos, reconhecimento de fala, análise de sentimentos, previsão de demanda, recomendação de produtos e muitos outros. Embora a maior parte dos algoritmos de aprendizado de máquina existam há muito tempo, a capacidade de aplicar cálculos matemáticos complexos ao *Big Data* automaticamente é um desenvolvimento recente. Esta nova realidade traz um ambiente de grande oportunidade para as empresas gerarem valor nos seus processos extraindo padrões e tendências nesse ambiente de larga escala (OLIVEIRA, 2020a).

Diante deste contexto que associa o rápido avanço da capacidade computacional e o grande volume de dados sendo gerado nas organizações provenientes das mais diversas fontes, foi realizado um estudo da literatura a fim de avaliar o nível de aplicação destas técnicas no processo produtivo das sinterizações de indústrias siderúrgicas. Esta avaliação partiu de revistas científicas indexadas em bases reconhecidas da área de metalurgia. Foram encontrados diversos artigos relacionados ao tema, principalmente associados ao desenvolvimento de modelos preditivos utilizando redes neurais. Estes estudos foram realizados na sua maior parte por pesquisadores de países que são grandes produtores de aço, como China e Japão. No que tange as sinterizações da siderurgia brasileira, a aplicação destes métodos e a geração de conteúdo científico ainda é escassa.

Alguns dos trabalhos avaliados durante o estudo da literatura apresentaram bons resultados de acurácia dos modelos e demonstraram ser bastante promissores para aplicações práticas. Laitinen e Saxén (2007) utilizaram dados reais de uma determinada planta de sinterização, considerando um período de cinco anos, para construção de um modelo baseado em redes neurais visando a predição de parâmetros de desempenho da sinterização. Neste estudo, o modelo criado foi capaz de prever resultados com uma acurácia entre 65-85%, além de demonstrar as relações complexas e não lineares entre as variáveis de entrada e saídas deste processo. Fan, Li e Chen (2012) partiram de dados de testes de uma sinterização piloto e exploraram os métodos de regressão, séries temporais e redes neurais para criação de modelos preditivos com o objetivo de prever características do sínter como a umidade da mistura, o consumo de combustível, a velocidade da sinterização e a resistência mecânica do sínter com uma acurácia entre 87 e 97%. Shao et al. (2020) desenvolveram um modelo baseado em redes neurais variando as taxas de aprendizado e de momentum para solução do problema de não linearidade, forte agrupamento e atraso na disponibilidade de informações do processo de sinterização. Foi atingida uma acurácia de 81%. Para criar o modelo, foram utilizados 1000 grupos de dados periodicamente atualizados de uma planta de sinterização por um período de 2 anos.

A partir da revisão bibliográfica, foi identificado um grande potencial para a aplicação do aprendizado de máquina no processo de sinterização. Com este foco, foi proposta a criação de um modelo piloto visando a predição do principal índice de qualidade do sínter, representando pelo *shatter test*. Este índice é de grande importância no desempenho e no custo de produção do alto-forno. Para simplificação do problema, foram considerados os valores de uma máquina de sinterizar. Foram testados diversos algoritmos de aprendizado de máquina, como a regressão linear, regressão polinomial, árvores de decisão, *ensemble learning* e redes neurais artificiais. Estes modelos foram comparados utilizando a métrica de *mean absolute error (MAE)* como medida de desempenho.

## 1.1 Objetivo geral

O objetivo geral deste trabalho é promover uma maior estabilidade operacional e ambiental nas áreas de sinterização e matérias-primas, por meio da implementação de modelos de aprendizado de máquina. Estes modelos têm a finalidade de fornecer previsões de indicadores fundamentais do processo, suportando a tomada de decisões operacionais. Atualmente, muitas dessas variáveis críticas não são monitoradas em tempo real, o que resulta em atrasos nas ações operacionais de ajuste do processo. A utilização destes modelos possibilitará um monitoramento contínuo das variáveis, resultando em uma resposta mais ágil e proativa para aprimorar a estabilidade operacional e ambiental das máquinas de sinterização.

### 1.1.1 Objetivos específicos

Para cumprimento do objetivo geral é necessário atender aos seguintes objetivos específicos:

- Construir um base de dados representativa, por meio das atuais fontes de dados existentes;
- Desenvolver um *framework* de aprendizado de máquina aplicado ao processo de sinterização com uma estrutura capaz de aprender com os dados das sinterizações e matérias-primas;
- Desenvolver um modelo piloto de predição da qualidade do sínter de uma máquina de sinterizar e avaliar seu desempenho;
- A partir do *framework* desenvolvido, propor a criação de modelos preditivos alternativos considerando mais variáveis de saída, como produtividade do processo, consumo de combustíveis e emissão de particulados.

## 1.2 Contribuições

Este trabalho visa promover a adoção de técnicas de aprendizado de máquina na análise de dados e inteligência de processo nas sinterizações da siderurgia brasileira, hoje pouco exploradas. Ao propor estas novas formas de análise, espera-se impulsionar a performance das sinterizações, contribuindo para a eficiência e a competitividade do setor.

## 1.3 Organização do trabalho

O trabalho foi dividido em quatro capítulos. O Capítulo 1 apresenta a introdução do presente estudo, no qual é definido o objetivo geral e objetivos específicos, além das principais contribuições no desenvolvimento de métodos de inteligência de negócio para o processo de sinterização. O Capítulo 2 apresenta a revisão bibliográfica que serviu de embasamento para este trabalho. O Capítulo 3 apresenta os métodos e materiais utilizados. O Capítulo 4 apresenta e discute os resultados obtidos e o Capítulo 5 apresenta a conclusão do trabalho.

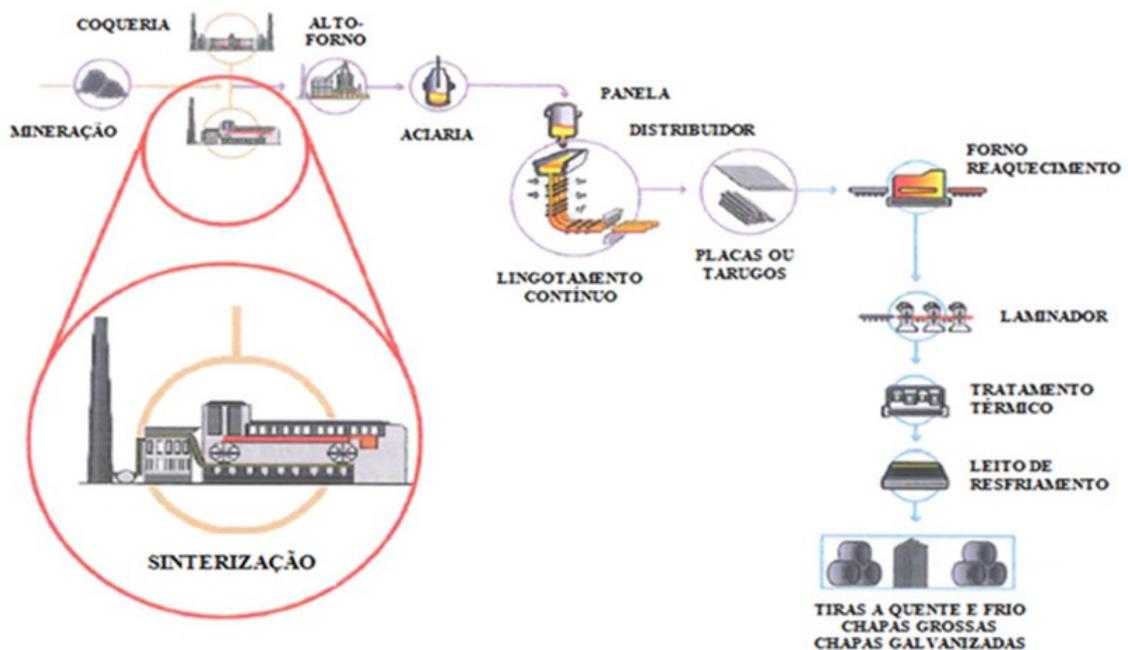
## 2 Contextualização teórica

Este Capítulo busca dar uma visão geral sobre a área de sinterização e seus principais indicadores e controladores operacionais de processo. Por fim, é abordado o tema de aprendizado de máquina contemplando os tipos de sistemas, as etapas de preparação da base de dados e as principais técnicas utilizadas neste trabalho.

### 2.1 Contexto da sinterização na siderurgia

A sinterização é um processo utilizado na siderurgia para aglomeração de misturas de finos de minério de ferro, coque ou carvão, resíduos contendo ferro, finos de retorno e fundentes (LV et al., 2010). O desenvolvimento deste processo de aglomeração possibilitou que parte dos finos de minério de ferro, antes considerados rejeitos de lavra, passassem a ter valor econômico quando agregados em partículas com tamanho e qualidade que possibilitassem sua aplicação direta nos fornos siderúrgicos na forma de sinter (SILVA, 2011). Este contexto demonstra a importância da sinterização na cadeia produtiva do aço. A Figura 1 apresenta um *layout* típico de uma siderúrgica integrada, destacando o papel da sinterização neste fluxo.

Figura 1 – Layout típico de uma usina integrada.



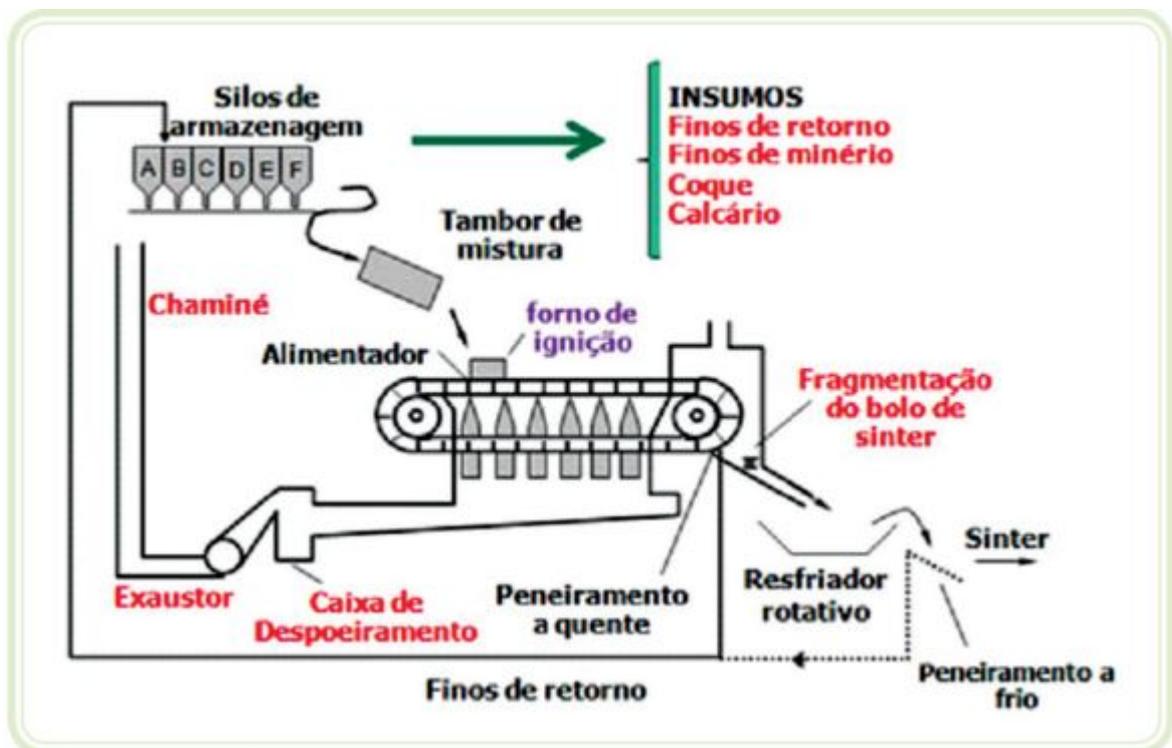
Fonte: Apostila do treinamento: Utilização do Minério de Ferro na Siderurgia - Usiminas (MAGALHÃES; NATAL; SÉRGIO, 2014).

## 2.2 O processo de sinterização

O fluxo da sinterização se inicia no abastecimento dos silos de carregamento com finos de minério, combustíveis (coque e antracito), calcário, reciclados industriais e finos de retorno. Após esta etapa, estes materiais são dosados e transportados para o tambor misturador, responsável pela etapa de aglomeração a frio, onde é adicionado umidade durante o carregamento e ocorre a homogeneização e nucleação da mistura, promovendo a permeabilidade adequada para a subsequente operação de queima. A mistura é distribuída ao longo da esteira sobre uma camada de *bedding* (falsa grelha), que tem a função de proteger as grelhas da máquina.

Após a distribuição na esteira, o material carregado no leito de fusão passa pela etapa de ignição, onde ocorre a queima do combustível contido na parte superior da mistura. A combustão irá se propagar de cima para baixo através da sucção de ar, que esfria o sinter e aquece o ar. Com a movimentação horizontal da esteira e com este fluxo de ar que passa através do material, a frente de queima superficial é transportada até atingir as zonas inferiores do leito, produzindo assim o sinter. Na zona de combustão, a temperatura atinge valores da ordem de 1.000 a 1.300°C, pouco abaixo do ponto de fusão dos minerais. Terminada a queima, o sinter é submetido ao processo de britagem, resfriamento e peneiramento para adequação de granulometria e atendimento à especificação dos altos-fornos, como mostra o fluxograma da Figura 2 (SILVA, 2011).

Figura 2 – Fluxograma de sinterização.



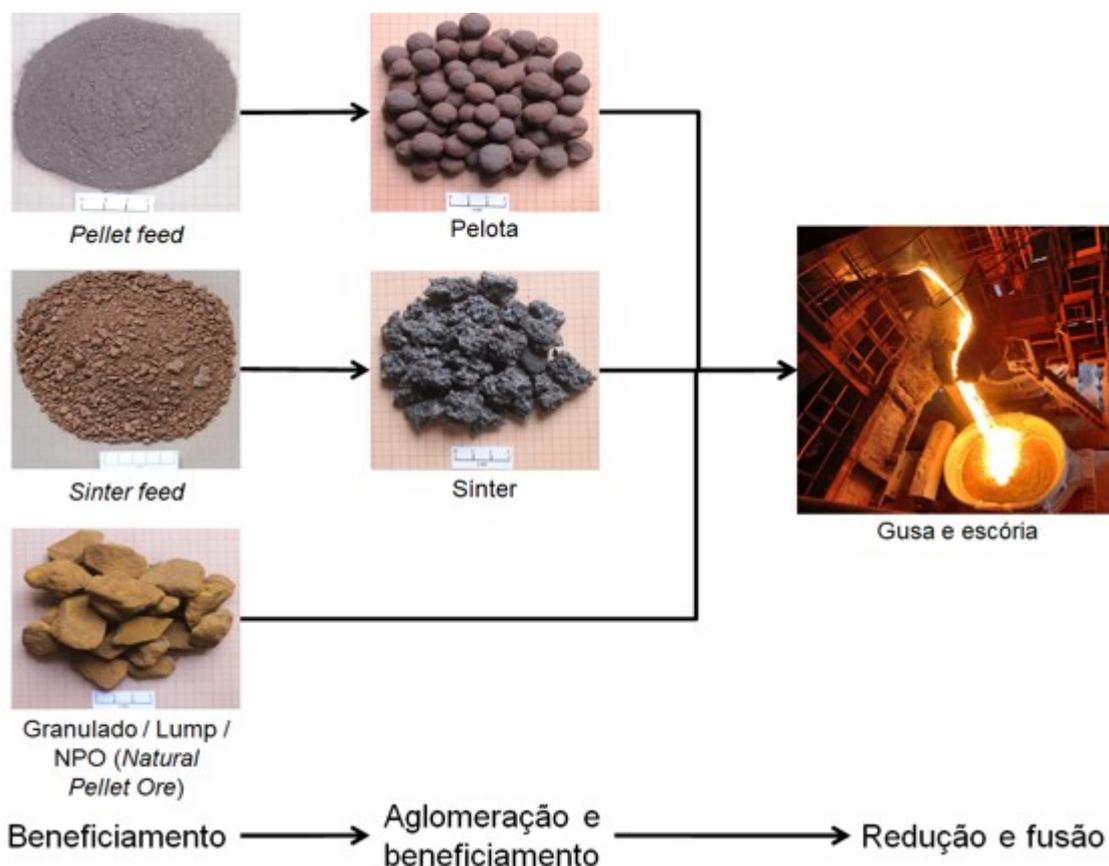
Fonte: [www.dema.puc-rio.br/download/iemaaulaaco.ppt](http://www.dema.puc-rio.br/download/iemaaulaaco.ppt) (SILVA, 2011).

## 2.3 Principais cargas metálicas para altos-fornos

Os minérios de ferro estão cada vez mais sendo preparados antes do uso, através dos processos de sinterização e pelletização. As plantas de sinterizações estão presentes na área de redução das siderúrgicas, tendo como principal objetivo, reaproveitar as partículas finas geradas no beneficiamento do minério de ferro nas unidades de mineração. Estes finos são separados e aglomerados em massas maiores durante o processo de sinterização. O sinter, material resultante deste processo, contribui para melhorar a permeabilidade da carga metálica do alto-forno, uma vez que melhora o contato da carga com o gás redutor e assim, reduz o *coke-rate* e aumenta a produtividade deste reator (SILVA, 2011).

A Figura 3 apresenta as principais cargas metálicas destinadas aos altos-fornos.

Figura 3 – Principais cargas metálicas destinadas aos altos-fornos.

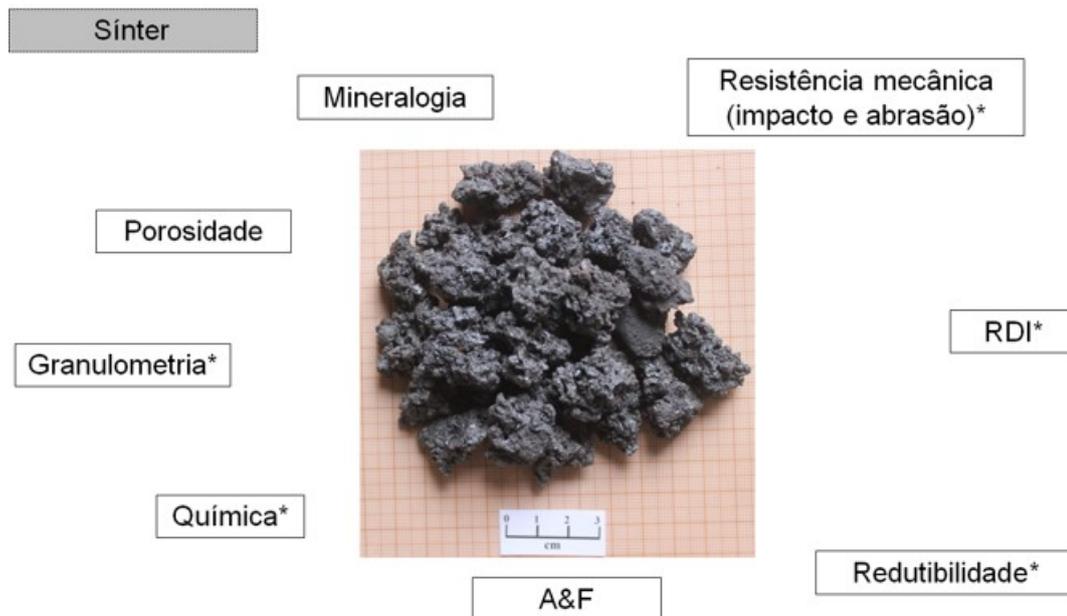


Fonte: Apostila do Treinamento: Utilização do Minério de Ferro na Siderurgia - Usiminas (MAGALHÃES; NATAL; SÉRGIO, 2014).

Este sinter representa a principal carga alimentada na maioria dos fornos, com sua participação em cerca de 60 a 90% do carregamento. A composição da carga metálica total carregada que envolve além do sinter, os minérios granulados e pelota, deve estar dentro das especificações exigidas de uniformidade, isenção de finos, granulometria adequada e resistência mecânica suficiente para assegurar uma boa permeabilidade da coluna de carga.

De forma geral, o sinter deve ter mais de 56% de ferro, baixos teores de enxofre e fósforo e boa resistência mecânica. Deve resistir a temperaturas elevadas e aos esforços no interior do forno sem amolecer. As principais características do sinter que influenciam na performance nos altos-fornos estão demonstradas na Figura 4. Elas estão diretamente associadas à estabilidade do processo de sinterização e às qualidades das matérias-primas utilizadas.

Figura 4 – Principais características do sinter para consumo nos altos-fornos.

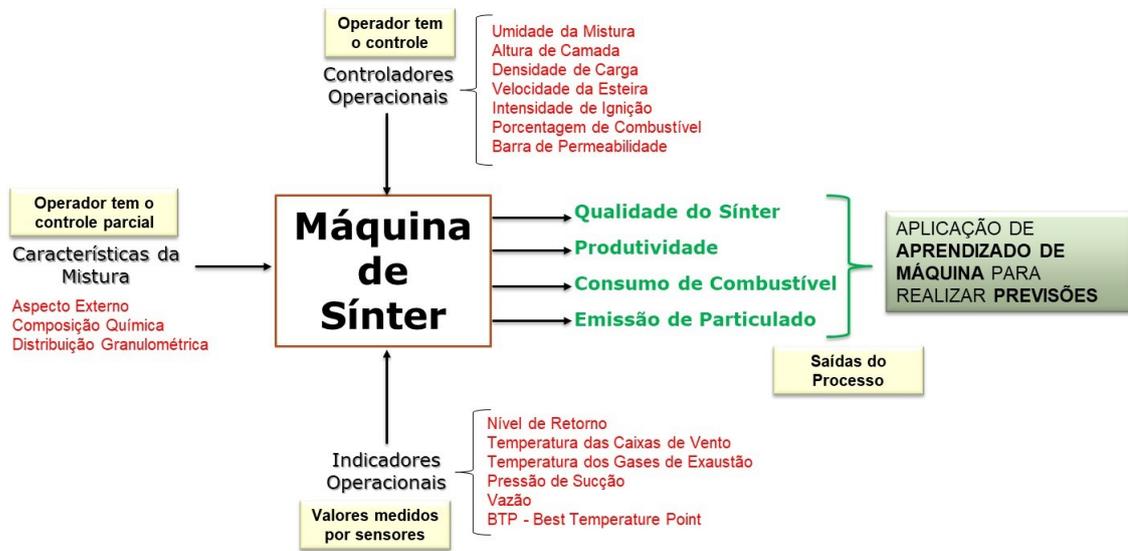


Fonte: Apostila do treinamento: Utilização do Minério de Ferro na Siderurgia - Usiminas (MAGALHÃES; NATAL; SÉRGIO, 2014).

## 2.4 Estrutura de variáveis da sinterização

Os ajustes do processo de sinterização são realizados através de indicadores e controladores operacionais, que são medidos por instrumentos e equipamentos auxiliares. O objetivo é adequar o processo para obter o sinter com a qualidade dentro das metas estabelecidas, com o menor custo e maior produtividade. A Figura 5 demonstra esta estrutura com uma proposta de variáveis de saída que podem ser previstas utilizando técnicas de aprendizado supervisionado.

Figura 5 – Estrutura de variáveis da sinterização e proposta de aplicação de aprendizado de máquina para realizar previsões.



Fonte: Autor.

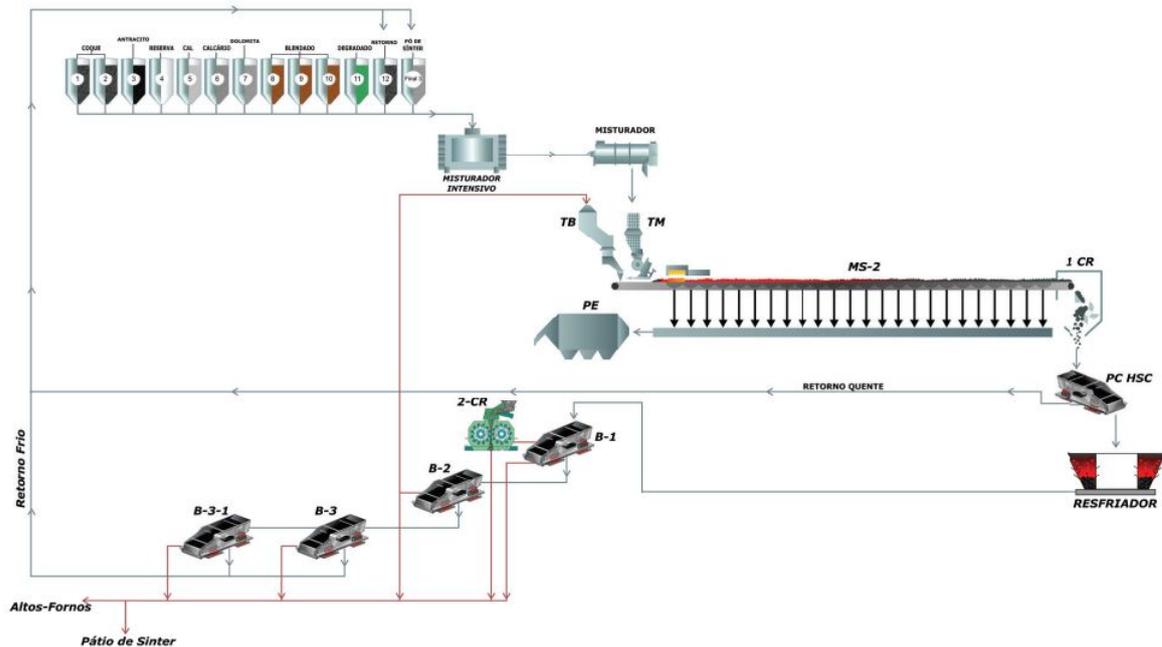
### 2.4.1 Indicadores operacionais

Os valores de variáveis de operação da máquina de sínter registrados por instrumentos são chamados de indicadores operacionais, no qual, através de suas informações, permitem o acompanhamento do processo e auxiliam nas decisões voltadas para obter o melhor desempenho de qualidade, produtividade, custo e o meio ambiente das sinterizações (MAGALHÃES; NATAL; SÉRGIO, 2014). Os indicadores operacionais avaliados estão detalhados a seguir:

#### A. Nível de retorno

É o principal indicador do processo de sinterização. Sua geração está associada diretamente às condições de processo e ao tipo de matéria-prima consumida, assim como o tratamento mecânico no qual o sínter produto é submetido. O nível de recirculação vai depender do pior ou melhor controle operacional e da mistura utilizada. A Figura 6 demonstra um exemplo do circuito de retorno de uma máquina de sínter.

Figura 6 – Circuito de retorno de uma máquina de sinter.



Fonte: Apostila de treinamento: Processo de Sinterização - Módulo IV - Usiminas (MAGALHÃES; NATAL; SÉRGIO, 2014).

Em termos de processo, os principais fatores que influenciam na geração de retorno são a quantidade e granulometria do combustível, umidade da mistura, compactação da camada e qualidade da queima superficial. Para o peneiramento, os impactos se relacionam à condição de desgaste ou obstrução de peneiras, tais como grelhas quebradas ou chapas perfuradas, abertura do britador secundário e anormalidades da balança de retorno.

## B. Temperatura das caixas de vento

Em conjunto com a temperatura dos gases de exaustão, as temperaturas das caixas de vento são importantes indicadores do processo de sinterização. As medições ocorrem através de termopares que ficam localizados na tubulação que sai da caixa de vento antes do anel de dilatação. Em condições normais de processo, as caixas de vento devem estar com temperaturas da ordem de 300°C, o que contribui para a uma boa qualidade do sinter produzido.

## C. Temperatura dos gases de exaustão

Para a temperatura dos gases de exaustão, o seu valor deve se situar na faixa de 110°C. Esse índice representa a média dos gases gerados em todas as caixas de vento. Esse número fornece uma boa indicação do processo e a sua análise para o controle operacional deve ser feita em conjunto com outros indicadores de processo.

### D. Pressão de sucção

A pressão de sucção é o principal indicador de permeabilidade do leito. Ele representa a facilidade ou dificuldade com que o ar passa pela camada de mistura na sinterização e registra também o histórico térmico do processo.

O gráfico da Figura 7 demonstra a pressão de sucção desde o início até o fim do processo. Pode ser observado uma maior pressão entre as caixas 3 e 5. Este fato ocorre pois no início do processo o leito é praticamente uma mistura crua, dificultando a passagem de ar. A medida que a frente de combustão se desloca para a parte inferior do leito, a pressão de sucção é reduzida em função do sínter formado, estabilizando-se nas últimas caixas.

Figura 7 – Pressão de sucção no processo de sinterização.



Pressão de sucção no processo de sinterização

Fonte: Apostila Processo de Sinterização - Módulo IV - Usiminas (MAGALHÃES; NATAL; SÉRGIO, 2014).

Uma melhor referência para avaliar a permeabilidade do leito seria utilizando a *equação de Ergun* que descreve a variação de pressão por unidade de comprimento do leito fluidizado, porém, as variáveis necessárias para alimentar esta equação são de difícil extração no processo de sinterização, devido a grandes variações de temperatura do processo, do tamanho das partículas e da porosidade do leito. Neste sentido, foi desenvolvida uma equação para a permeabilidade do sínter adaptada de acordo com os fundamentos da *Lei de Darcy*, que descreve o fluxo de um fluido em um meio poroso, conforme a equação (2.1) (LAITINEN; SAXÉN, 2007).

$$\left(\frac{\dot{V}}{A}\right) = \psi \left(\frac{\Delta P_{bed}}{h}\right)^n \quad (2.1)$$

onde  $\psi$  é a permeabilidade online média do leito na esteira;  $\dot{V}$  é a vazão volumétrica do gás;  $A$  é a área da máquina;  $\Delta P_{bed}$  é a queda de pressão do gás;  $h$  é a altura da camada;  $n$  é um parâmetro da mistura de sínter na esteira representado por valores entre 0,5 a 1,0.

### **E. Vazão**

Representa a quantidade de ar que passa pela camada de mistura ao longo do leito de fusão. Nas sinterizações em estudo, os medidores estão instalados na entrada do precipitador eletrostático do sistema de exaustão, após o tubulão e canal de fumaça.

### **F. *BTP - Burning Through Point* (Temperatura Final de Queima)**

O *BTP* é uma metodologia de controle que indica o término da combustão. É um importante indicador para suporte na tomada de ação diante das condições do processo. O cálculo da curva para determinar o *BTP* ocorre a partir de um modelo matemático definido pelas temperaturas medidas em termopares posicionados nas caixas de vento. Este modelo pode variar dependendo das características de processo e equipamento de cada empresa.

## **2.4.2 Controladores operacionais**

O processo de sinterização é influenciado por fatores que afetam, em menor ou maior peso, a qualidade do sínter e a produtividade. Para garantir a estabilidade operacional, os controladores operacionais devem ser ajustados para buscar as melhores condições de processo. As principais intervenções são realizadas na umidade da mistura, altura da camada, densidade de carga, velocidade de máquina, porcentagem de combustível, intensidade de ignição e na barra de permeabilidade (MAGALHÃES; NATAL; SÉRGIO, 2014).

### **A. Umidade da mistura**

O ajuste adequado da umidade da mistura influencia diretamente no processo de aglomeração a frio e na condição de permeabilidade da mistura, com impactos na qualidade do sínter e na produtividade da máquina.

### **B. Altura da camada**

Uma maior altura de camada está relacionada ao aumento do rendimento do sínter. Neste sentido, trabalhar com a maior elevação possível da camada em equilíbrio com o rendimento térmico do processo irá contribuir para a produtividade, qualidade e custo do produto. O ajuste da altura da camada é realizado por um sistema manual.

### **C. Densidade de carga**

A densidade de carga é a quantidade de mistura total alimentada na esteira por unidade de volume. Este controlador indica o nível de compactação da camada, sendo que, quanto maior a densidade de carga, mais compacta fica a mistura, com prejuízos na permeabilidade e consequentemente na produtividade.

**D. Velocidade da esteira**

É o principal controlador de operação. Influencia diretamente na produção e reflete o andamento do processo de sinterização. Em anormalidades de processo, a velocidade de máquina é a resposta imediata para corrigir e estabilizar a operação. A velocidade de máquina tem ligação direta com a resistência do síter, o teor de óxido de ferro (FeO), o nível de retorno e a produtividade.

**E. Intensidade de ignição**

É a quantidade de gás consumida por área da máquina, estando relacionada com a qualidade da queima superficial. O controle é realizado visualmente e por indicadores do forno de ignição. A intensidade de ignição é afetada basicamente pela deficiência dos queimadores, sujeira, irregularidade da superfície e altura da camada de mistura. Se a superfície da camada não estiver bem lisa, a ignição da mistura será afetada e ocorrerá elevação do nível de geração de retorno.

**F. Porcentagem de combustível**

O consumo de combustível está diretamente ligado ao nível térmico do processo, refletindo na qualidade e na produção. Se o mesmo estiver em excesso, ocorrerá uma queda da resistência do produto e aumento na geração de retorno. Caso contrário, haverá diminuição da redutibilidade do síter e aumento do FeO.

**G. Barra de permeabilidade**

A barra de permeabilidade tem como principal objetivo melhorar a permeabilidade do leito de sinterização com o controle manual de entrada das barras na mistura. Estas barras tem a função de diminuir a velocidade dos grãos de minério que chegarão à região inferior do leito durante o carregamento da esteira. A diminuição de velocidade dos minérios ao longo do leito promove uma maior permeabilidade da mistura, principalmente na região inferior da camada. Esta região possui maior dificuldade da passagem dos gases por estar sendo compactada pela ação do peso da parte superior do leito já sinterizado e pelo efeito do aumento do volume de gases que estão passando por ela em função da alta temperatura relacionada ao alargamento da frente de combustão.

### 2.4.3 Características da mistura

As características da mistura a ser utilizada nas sinterizações é dependente dos processos relacionados ao plano de recebimento e manuseio de matérias-primas nos pátios das sinterizações. O plano de recebimento vai determinar a disponibilidade, qualidade e granulometria dos minérios, fundentes e combustíveis, onde, em uma etapa posterior, serão considerados para a programação de manuseio e formação das pilhas de material blendado. As principais variáveis da mistura formada estão relacionadas as características dos minérios utilizados na mistura (aspecto externo, química, tamanho, mineralogia, porosidade e termogravimetria), a participação de minérios por tipo e ao consumo de materiais de geração interna.

## 2.5 Aprendizado de máquina

### 2.5.1 O cenário do aprendizado de máquina

O aprendizado de máquina é um campo interdisciplinar que envolve matemática, estatística, ciência da computação e outras áreas. Este conceito vem sendo discutido desde a década de 1950 e sua evolução foi impulsionada, principalmente, por avanços em tecnologia, algoritmos e quantidade de dados disponíveis. Desde então, o aprendizado de máquina vem avançando rapidamente com o surgimento de algoritmos mais complexos, como redes neurais profundas e algoritmos de aprendizado por reforço, com o suporte do aumento exponencial na quantidade de dados disponíveis e na capacidade de processamento computacional. Hoje em dia, sua aplicação envolve uma ampla variedade de campos, incluindo processamento de imagens e vídeos, reconhecimento de fala, análise de sentimentos, previsão de demanda, recomendação de produtos e muitos outros.

As técnicas de aprendizado de máquina passaram por uma grande evolução nas últimas duas décadas, principalmente em desenvolvimentos voltados para aplicações práticas de larga escala. Muitos desenvolvedores de sistemas de inteligência artificial tem utilizado estas técnicas na criação de suas aplicações, de modo que pode ser muito mais fácil treinar um sistema através de exemplos do que programá-los manualmente. O efeito do aprendizado de máquina foi sentido amplamente na ciência da computação e nos vários setores com grande geração de dados. Houve também um grande efeito nas ciências empíricas, com os métodos de aprendizado de máquina apresentando um alto rendimento na análise de dados (JORDAN; MITCHELL, 2015).

### 2.5.2 Tipos de sistemas do aprendizado de máquina

Um algoritmo de aprendizado de máquina pode ser:

- supervisionado, não supervisionado, semi-supervisionado e aprendizado por reforço: estão relacionados a serem ou não treinados com supervisão humana;
- aprendizado online vs aprendizado por lotes: se podem ou não aprender rapidamente, de forma incremental;
- aprendizado baseado em instâncias vs aprendizado baseado em modelo: se irão funcionar apenas comparando novos pontos de dados com pontos de dados conhecidos, ou se irão detectar padrões em dados de treinamento permitindo a criação de modelos preditivos.

Esses critérios não são exclusivos e podem ser combinados para que o sistema alcance seu objetivo. A seguir, as abordagens de aprendizado são apresentadas com maior detalhe.

### 2.5.2.1 Aprendizado supervisionado e não supervisionado

No aprendizado supervisionado, um conjunto de dados com respostas conhecidas é fornecido ao algoritmo. O objetivo do algoritmo é aprender a prever as respostas corretas para novos dados que não estão no conjunto de treinamento. Ele envolve o desenvolvimento de modelos estatísticos que busca prever ou estimar uma saída com base em uma ou mais entradas. Ou seja, o modelo é treinado por uma base de dados de entradas e suas respectivas saídas.

No aprendizado não supervisionado, não há respostas conhecidas fornecidas ao algoritmo. Em vez disso, o objetivo é encontrar padrões e estruturas ocultas nos dados. Este formato busca propriedades, relações e estruturas a partir dos dados de entrada. Sendo assim, não existe modelo rígido para a saída a partir dos dados de entrada, porque não existe saída e sim uma estrutura a ser descoberta. A ferramenta, por si só, faz emergir as propriedades dos dados.

### 2.5.2.2 Aprendizado online e em lote

O aprendizado de máquina em lote, também conhecido como aprendizado *offline*, é um método de treinamento em que o modelo é treinado usando um conjunto de dados fixo e completo. Esse tipo de treinamento é mais útil quando todos os dados necessários já estão disponíveis e não há necessidade de atualizar o modelo em tempo real.

Já o aprendizado *online*, também conhecido como aprendizado incremental, é um método de treinamento em que o modelo é atualizado em tempo real, conforme novos dados são coletados. Esse tipo de treinamento é útil quando os dados estão chegando constantemente e é necessário que o modelo se adapte às mudanças rapidamente.

### 2.5.2.3 Aprendizado baseado em instância e em modelo

O aprendizado baseado em instância, é característico de abordagens quando o número de exemplos de treinamento é pequeno e a distribuição dos dados é desconhecida, sendo que o sistema aprende os exemplos por meio da memorização e, em seguida, generaliza para novos casos utilizando medidas de similaridade. Ou seja, os dados de treinamento são armazenados e utilizados diretamente para fazer previsões para novos dados. Cada exemplo de treinamento é tratado como uma instância de dados e a semelhança entre a instância de teste e cada uma das instâncias de treinamento é usada para fazer uma previsão.

Para o aprendizado baseado em modelo, são característicos os problemas que envolvem grandes quantidades de dados disponíveis para treinamento, com suas distribuições conhecidas. Neste sentido, o modelo é construído a partir dos dados de treinamento, com o sistema de aprendizado criado sendo usado para fazer previsões de novos dados. O modelo será representado por uma equação ou conjunto de regras derivadas dos dados de treinamento.

### 2.5.3 Preparação da base de dados - coleta, limpeza e visualização

O aprendizado de máquina apresenta uma série de desafios, sendo um dos principais a coleta e organização dos dados. A qualidade e disponibilidade destes dados são fundamentais para o desempenho dos modelos de aprendizado de máquina. Esta etapa pode ser considerada a mais complexa e a que demanda maior esforço dos cientistas de dados, além de ser essencial para atingir os objetivos do projeto.

Para o tratamento da base de dados, diversos fatores devem ser considerados, como a origem dos dados, se são extraídos de variadas fontes, se a quantidade de dados de treinamento é suficiente e se os dados são representativos para o objetivo do modelo. Sendo assim, diversas técnicas de limpeza deverão ser utilizadas para a preparação da base de treinamento. Algumas das técnicas usadas envolvem a remoção de dados ausentes e *outliers*, detecção de valores inconsistentes, tratamento de dados duplicados, dentre outras.

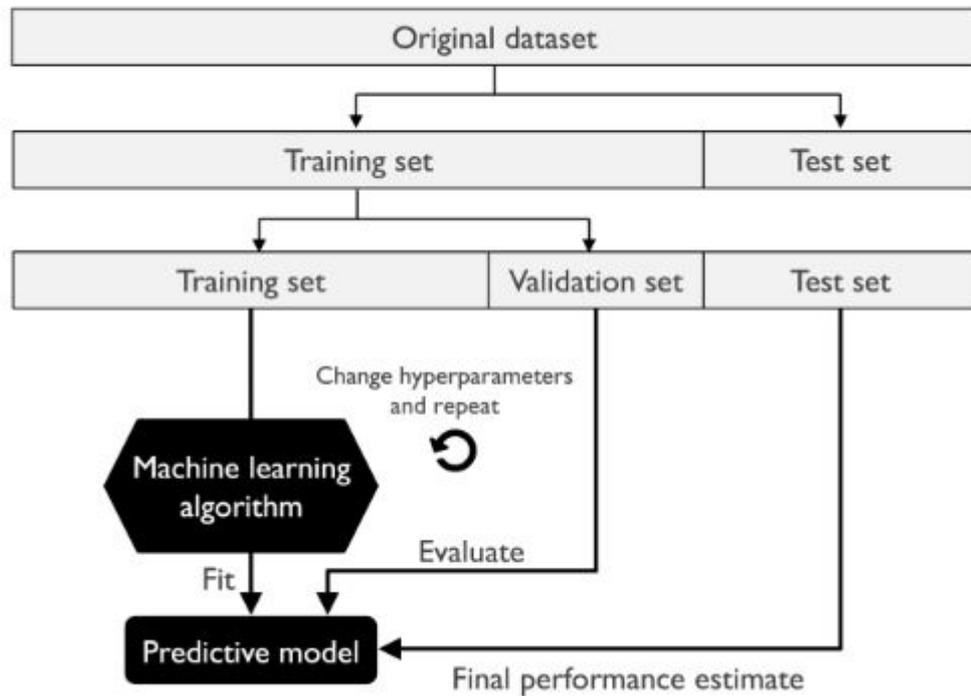
Na sequência é realizada uma análise exploratória dos dados com a função de entender melhor os padrões, tendências e relacionamentos nos dados que estão sendo usados. A visualização de dados ajuda a identificar valores atípicos e verificar como as diferentes variáveis se relacionam umas com as outras. Essas informações são importantes para selecionar as melhores variáveis para alimentar o modelo de aprendizado de máquina.

### 2.5.4 Divisão da base de dados para treinamento, validação e teste

Para avaliar o desempenho dos modelos de aprendizado de máquina de forma mais precisa e evitar o *overfitting* é adotada a prática de divisão da base de dados em treinamento e teste. Na divisão da base em treinamento e teste, a ideia é separar uma parte dos dados para treinar o modelo e outra parte para avaliar o desempenho do modelo. A forma usual de fazer isso é dividindo a base de dados original em dois conjuntos independentes e sem sobreposição. Já para a validação não é recomendado que a comparação de modelos seja feita utilizando a mesma base de testes, uma vez que esse procedimento pode introduzir *bias*. Portanto, a melhor prática é dividir a base de treinamento em duas: treinamento e validação e, após a validação, um novo treinamento deve ser realizado com a base de treinamento completa.

A Figura 8 esquematiza a divisão da base em três partes e define o processo de treinamento, validação e teste do modelo.

Figura 8 – Treinamento, validação e teste: divisão da base de dados.



Fonte: Python Machine Learning: Machine Learning and Deep Learning with Python (RAS-CHKA; MIRJALILI, 2019).

Considerando o conjunto de treinamento e de validação grandes o suficiente, os algoritmos de aprendizado de máquina ou parametrizações de modelos que tiverem os melhores desempenhos nos conjuntos de validação podem ser assumidos como os de melhor capacidade de generalização.

Outro ponto importante a ser considerado é a limitação do conjunto de dados, onde uma nova divisão da base de dados não seria desejada. Desta forma, existem métodos alternativos de separação da base de treinamento para avaliação dos modelos. As técnicas mais comuns para esta prática são o *bootstrap* e o *cross-validation*. O *bootstrap* consiste em criar uma base de dados de mesmo tamanho através de uma amostragem aleatória com reposição da base original para realizar a validação. Já o método de *cross-validation*, consiste em estimar o erro de generalização através de um procedimento iterativo, no qual o conjunto de treinamento é particionado em  $k$  subconjuntos para serem utilizados como conjunto de validação dos modelos treinados nos outros  $k$  menos 1 subconjuntos (OLIVEIRA, 2020b).

### 2.5.5 Medidas de desempenho para avaliar modelos

As principais medidas adotadas para avaliar os modelos de regressão são o *root mean squared error (RMSE)* e o *mean absolute error (MAE)*.

O *RMSE* é uma medida de desvio padrão dos resíduos, ou seja, é uma medida para avaliar se as previsões do modelo estão próximas aos valores reais. Quanto menor o *RMSE*, melhor é o desempenho do modelo. Sua principal vantagem é que ele penaliza mais os erros maiores do que os erros menores. Isso é importante em muitas aplicações de aprendizado de máquina, pois erros maiores podem ter um impacto mais significativo no resultado final. A equação (2.2) apresenta a formulação desta medida de desempenho.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (2.2)$$

Embora a *RMSE* seja geralmente a medida de desempenho preferencial para tarefas de regressão, em alguns contextos outras funções podem gerar melhores avaliações. Neste trabalho, considerando a grande quantidade de *outliers* e a necessidade de alto nível de pré-processamento da base de dados, optou-se por avaliar o modelo utilizando a métrica do *MAE*. O desvio médio absoluto é a média das diferenças absolutas entre as previsões e os valores reais, conforme a equação (2.3).

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (2.3)$$

### 2.5.6 Modelos de aprendizado de máquina

Neste tópico são abordados os diferentes métodos utilizados neste trabalho para o desenvolvimento dos modelos de aprendizado de máquina aplicados ao processo de sinterização.

#### 2.5.6.1 Regressão linear

A Regressão linear é um método estatístico que busca encontrar a relação entre duas ou mais variáveis. Quando aplicado ao aprendizado de máquina, é utilizado para prever uma variável a partir de outras variáveis. A fórmula para previsão do modelo de regressão linear é mostrada na equação (2.4).

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2.4)$$

onde  $\hat{Y}$  é o valor previsto,  $n$  é o número de variáveis de entrada,  $X_i$  é o valor da  $i$ -ésima variável e  $\beta_j$  é o parâmetro  $j$  do modelo.

Existem duas formas bem distintas de treino para estes modelos. A primeira é utilizando uma equação direta de “forma fechada”, onde os parâmetros do modelo são calculados para valores que melhor se ajustam ao conjunto de treinamento, através da minimização da função de custo, utilizando o *mean squared error (MSE)*. A outra é utilizando uma forma iterativa de otimização através do gradiente descendente, que gradualmente busca os parâmetros do modelo de minimização da função custo no conjunto de treinamento, podendo estes valores convergirem para os mesmos do primeiro método (MALLICK; DHARA; RATH, 2021).

A regressão linear é um dos métodos mais simples de aprendizado de máquina, mas pode ser menos apropriado para conjuntos de dados complexos ou com relações não-lineares. Em tais casos, outros métodos de aprendizado de máquina, como árvores de decisão ou redes neurais, podem ser mais adequados.

#### 2.5.6.2 Regressão polinomial

O modelo de regressão supõe que X e Y possuem uma relação linear. Porém, esta relação pode não ser linear e da mesma forma X e Y podem estar correlacionadas. Para estes casos, outros modelos podem ser utilizados, como a regressão polinomial (RIBEIRO, 2021). A principal diferença deste modelo está no fato de poder incluir termos de grau superior. Enquanto a regressão linear tenta ajustar uma linha reta aos dados, a regressão polinomial tenta ajustar uma curva polinomial aos dados.

No entanto, a regressão polinomial pode ser mais propensa ao sobreajuste dos dados de treinamento, especialmente quando o modelo possui muitos parâmetros. Isso significa que o modelo pode ser muito flexível e se ajustar muito bem aos dados de treinamento, mas não ser capaz de generalizar bem para novos dados.

#### 2.5.6.3 Árvores de decisão

As árvores de decisão são um tipo de algoritmo de aprendizagem de máquina que utiliza uma estrutura em forma de árvore para tomar decisões a partir de dados de entrada. Este tipo de modelo apresenta um bom desempenho para extração e interpretação de conhecimentos a partir de uma base de dados. Estas árvores se caracterizam por ser um método hierárquico, não paramétrico e supervisionado. A maior vantagem desta abordagem é a criação de modelos visuais, de fácil compreensão e interpretação. Por outro lado, o método possui muitos parâmetros que podem ser ajustados, gerando diferentes modelos. Pela natureza da construção da árvore de decisão o ajuste pode levar facilmente ao *overfitting* (OLIVEIRA, 2020b).

Neste trabalho são exploradas as árvores de decisão voltadas para aplicações de regressão. A principal diferença em relação as árvores de classificação é que a medida de impureza é substituída pelo erro quadrático médio. Além disto, os nós terminais apresentam valores esperados para as observações e não uma classe estimada.

#### 2.5.6.4 Métodos de *ensemble learning*

Os métodos de *ensemble* no aprendizado de máquina são técnicas que combinam várias modelagens, agrupando as previsões de diversos previsores buscando obter um melhor resultado geral em relação ao modelo individual. Estes métodos geralmente funcionam melhor porque compensam as limitações e as fraquezas dos modelos individuais.

Os *ensembles* são normalmente utilizados em etapas finais dos projetos, quando bons previsores individuais já foram treinados. Sendo assim, estes modelos podem ser combinados para buscar um ainda melhor.

Existem outros vários métodos de *ensemble*, dentre eles, os mais utilizados são o *bagging*, *boosting* e *stacking*. Para este trabalho foram desenvolvidos modelos baseados em *ensembles*, onde foram utilizados os algoritmo *random forests* através dos métodos de *bagging* e os algoritmos *adaboost* e *gradient boosting* através do método de *boosting*.

Um modelo de *ensemble* bastante popular é o *random forests*. Estes modelos são *ensembles* do tipo *bagging* que ajustam diversos modelos de árvores de classificação sem limitação de profundidade.

##### 2.5.6.4.1 *Bagging*

O método de *bagging* utiliza uma abordagem onde um mesmo algoritmo de treinamento é utilizado para cada preditor, porém sendo treinados em diferentes subconjuntos aleatórios do conjunto de treinamento. Sua principal função é tentar evitar o *overfitting*. Normalmente o resultado é um modelo com um viés semelhante ao preditor de modelo único, mas com uma variância menor.

O *bagging* pode ser considerada uma das técnicas de *ensemble* mais utilizadas, principalmente através do algoritmo de *random forests*, que consiste no agrupamento de diversas árvores de decisão visando aumentar a resistência de alta variância característica de uma árvore de decisão individual (OLIVEIRA, 2020b).

#### 2.5.6.4.2 *Boosting*

O *boosting* se refere a qualquer método de *ensemble* que treina vários aprendizes fracos um após o outro. A ideia deste algoritmo é treinar sequencialmente os preditores buscando a correção do anterior. Existem muitos métodos *boosting* disponíveis. Neste trabalho são utilizados os métodos *adaboost* e *gradient boosting*. O algoritmo *adaboost* considera o novo preditor corrigindo seu antecessor, prestando um pouco mais de atenção às instâncias de treinamento que seu antecessor subajustou. Isto resulta em novos preditores focando mais em *cases* difíceis. Já o *gradient boosting* adiciona preditores sequencialmente, cada um corrigindo seu antecessor. No entanto, em vez de ajustar os pesos da instância a cada iteração, como o *adaboost* faz, este método tenta ajustar o novo preditor aos erros residuais feitos pelo preditor anterior (OLIVEIRA, 2020b).

#### 2.5.6.5 Redes neurais artificiais

As redes neurais buscam implementar modelos matemáticos que tenham características semelhantes as estruturas neurais biológicas. Apesar de o conhecimento sobre os mecanismos cerebrais serem limitados, tentar modelar os neurônios biológicos e algumas estruturas cerebrais foi o caminho buscado para levar a inteligência nas máquinas.

Diante deste contexto, em 1943, McCulloch e Pitts (1943) desenvolveram o primeiro modelo de neurônio artificial, sendo este trabalho reconhecido como a origem das redes neurais artificiais.

Da forma como ocorrem com as redes neurais biológicas, os neurônios das redes neurais artificiais possuem diversas entradas para receber um valor que simula a atividade de um dentrito. Estas entradas são ponderadas por pesos sinápticos e somadas. Na sequência, uma função matemática conhecida como função de ativação recebe estes valores e simula o processo sináptico biológico determinando se o neurônio artificial deve disparar ou não, gerando desta forma, a resposta do neurônio para o conjunto de entradas (LEITE, 2021).

Para este trabalho foi utilizado o método de Redes Neurais do tipo *multi layer perceptron* (*MLP*) com o algoritmo do *backpropagation*. Estas redes também se caracterizam por ser do tipo *feedforward*, onde o o fluxo de informação é unidirecional, partindo dos neurônios da camada de entrada para os neurônios da camada de saída.

A rede neural do tipo *MLP* é uma extensão imediata da rede de camada única conhecida como *perceptron*. Ela é capaz de tratar com melhor desempenho problemas com mais variáveis e também pode ser capaz de propor superfícies não lineares de separação, justamente por ter camadas de neurônios intermediárias (ocultas). A rede *MLP* possui capacidade de aproximação universal, ou seja, ela é apta a encontrar qualquer função necessária para a resolução de um problema proposto.

## 3 Metodologia

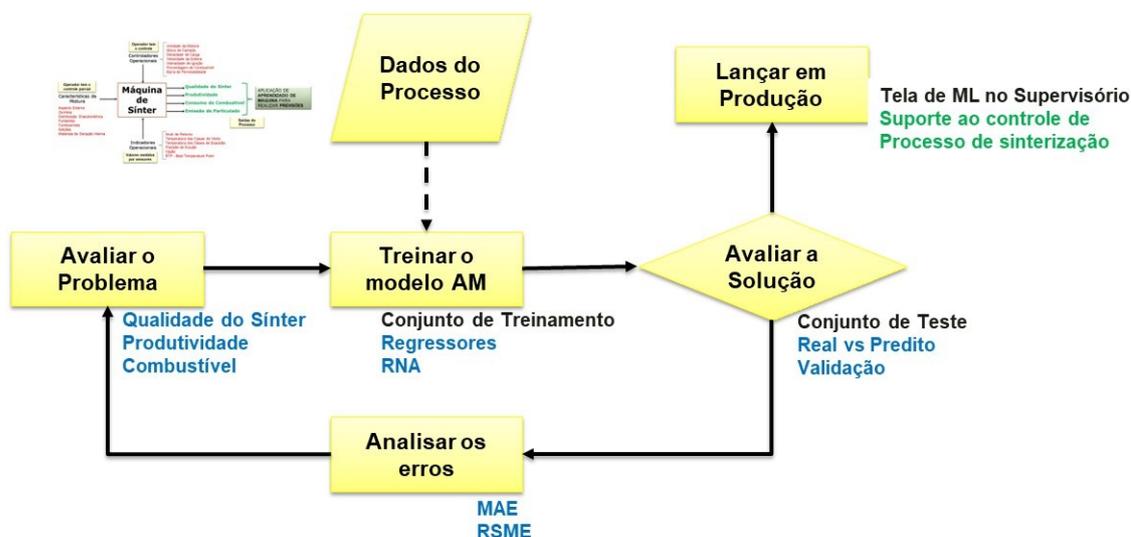
No presente capítulo, será abordado a metodologia utilizada para aplicar a técnica de aprendizado de máquina no processo de sinterização. Será apresentado o fluxo de trabalho adotado para o desenvolvimento dos modelos, desde o processo de definição das variáveis, até a aplicação dos algoritmos de aprendizado de máquina.

Primeiramente, serão discutidos os processos de seleção das variáveis das sinterizações e de coleta de dados. Em seguida, será proposto um modelo piloto para prever um índice que representa a resistência física do sinter. Para este modelo piloto proposto, serão realizadas análises das variáveis, visando compreender suas características e possíveis relações, além de suportar a seleção de suas variáveis das entrada. Após a análise, será aplicado a técnica de padronização nos dados das variáveis deste modelo, com o objetivo de assegurar a comparabilidade entre elas. Em seguida, será abordada o método de medição de desempenho utilizado para avaliar a capacidade do modelo em prever o indicador de interesse. Por fim, serão apresentados os algoritmos de aprendizado de máquina utilizados para esta aplicação prática. Este capítulo fornecerá uma visão abrangente do processo metodológico adotado para a aplicação do aprendizado de máquina nesse contexto específico.

### 3.1 Abordagem do aprendizado de máquina para aplicação no processo de sinterização

As variáveis envolvidas no processo de sinterização apresentam uma complexidade significativa, caracterizada por não linearidade, agrupamento intenso e atraso na disponibilidade das informações. Diante desse cenário, as técnicas tradicionais de programação podem ser bastante limitadas para o desenvolvimento de sistemas capazes de realizar previsões e apoiar as decisões operacionais. A abordagem convencional exigiria uma extensa lista de regras complexas, de difícil manutenção. Por outro lado, um modelo de previsão baseado em técnicas de aprendizado de máquina é capaz de aprender automaticamente as variações do processo de sinterização. Com base nessa abordagem, foi elaborada uma proposta de arquitetura para a implementação dos modelos de aprendizado de máquina no processo de sinterização, conforme ilustrado na Figura 9.

Figura 9 – Proposta de abordagem do aprendizado de máquina aplicado ao processo de sinterização.

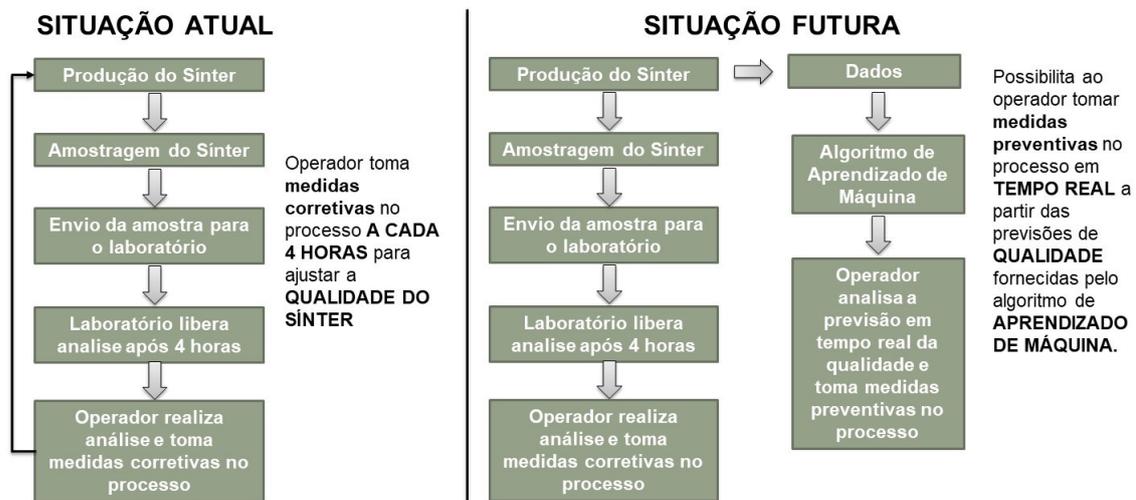


Fonte: Autor.

Em relação ao sistema de aprendizado de máquina utilizado neste estudo, a abordagem principal adotada foi o aprendizado supervisionado, que é amplamente aplicável para resolver esse tipo de problema. Esta abordagem apresenta uma típica tarefa para este tipo de sistema, sendo que envolve dados contínuos característicos de modelos de regressão. No entanto, há também a possibilidade de explorar técnicas de aprendizado não supervisionado em etapas intermediárias do desenvolvimento. Entre as técnicas de aprendizado não supervisionado, destaca-se a mineração de dados (OLIVEIRA, 2020a).

Outro critério utilizado para classificar os sistemas de aprendizado de máquina é se o sistema pode ou não aprender de forma incremental a partir de um fluxo de dados recebido. Existem duas abordagens principais: aprendizado online e aprendizado em lote. Considerando a grande quantidade de dados gerados nas sinterizações, a utilização de dados históricos e a necessidade de adaptação aos novos dados, a abordagem deste trabalho é mais característica do aprendizado em lote. Isso significa que o sistema é inicialmente treinado com todas as instâncias disponíveis no banco de dados extraídas do processo (aprendizado *offline*). Em seguida, para melhorar o desempenho dos modelos, são realizadas atualizações periódicas no treinamento dos modelos, utilizando os dados mais recentes para o sistema se familiarizar com os novos dados. Após cada ciclo de treinamento, uma nova versão do modelo é criada utilizando o conjunto completo de dados, que inclui tanto os dados antigos como os novos. Esta nova versão substitui o modelo anterior e é lançada em produção, onde recebe dados operacionais em tempo real. Essa integração permite a tomada de ações preventivas no processo com base nas previsões geradas pelos algoritmos do modelo atualizado. A Figura 10 ilustra uma proposta de aplicação, com foco na previsão da qualidade física do sinter.

Figura 10 – Proposta de estrutura de tomada de decisão com suporte de algoritmos de previsão da qualidade física do sinter.



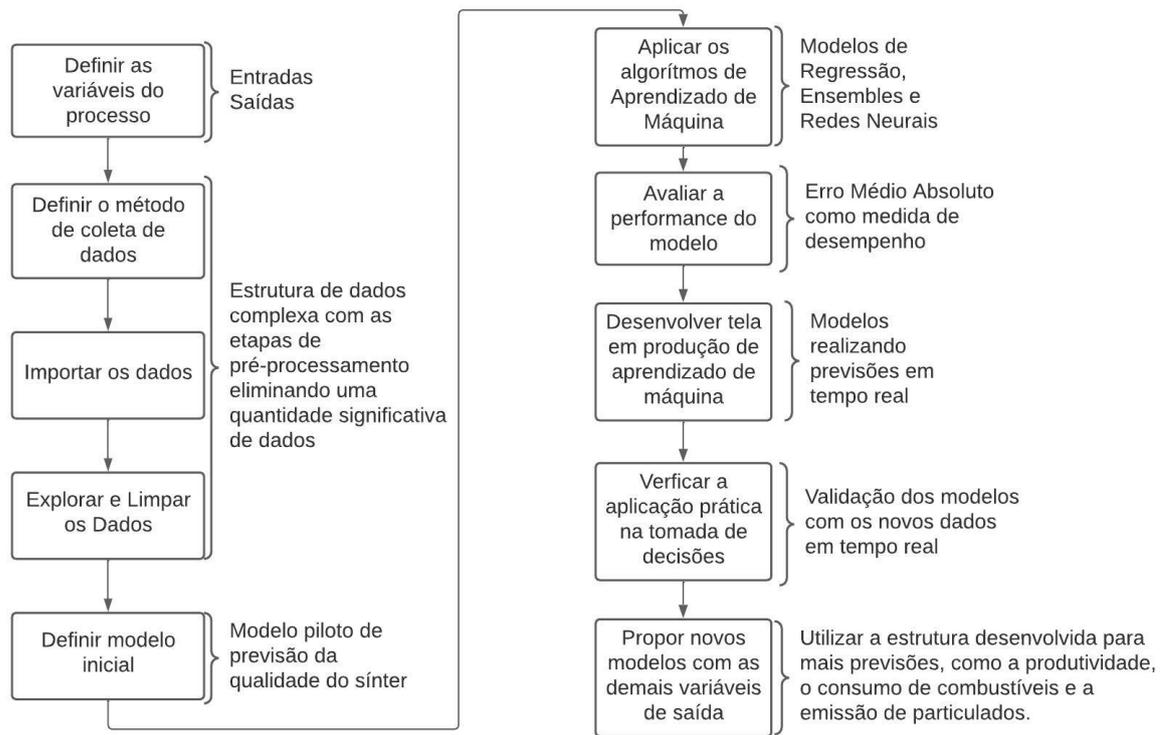
Fonte: Autor.

## 3.2 Fluxo de trabalho para desenvolvimento dos modelos

O sistema em estudo utiliza dados industriais coletados a partir de sensores e amostras de materiais de uma planta de sinterização. Pelos tipos de variáveis coletadas (temperatura, umidade, resistência, velocidade, entre outros) e a finalidade do modelo (previsão de valores), o sistema apresenta uma típica tarefa de aprendizado supervisionado com dados contínuos, característicos de modelos de regressão. Neste sentido, estes dados foram alimentados em algoritmos de aprendizado de máquina para este tipo de aplicação. Foram utilizadas as técnicas de regressão linear, regressão polinomial, árvores de decisão, florestas aleatórias, *ensemble learning* e redes neurais artificiais do tipo *MLP*. Para a avaliação do desempenho dos modelos foi considerado o *mean absolute error (MAE)*.

O Fluxograma da Figura 11 apresenta as etapas para o desenvolvimento dos modelos.

Figura 11 – Fluxograma das etapas do trabalho para desenvolvimento dos modelos.



Fonte: Autor.

### 3.3 Definição de variáveis do processo de sinterização e coleta de dados

As variáveis para o desenvolvimento dos modelos de aprendizado de máquina das sinterizações foram definidas a partir de discussões com os responsáveis técnicos das áreas de sinterizações e matérias-primas juntamente com o estudo de artigos relacionados ao tema (SHAO et al., 2020; LAITINEN; SAXÉN, 2007; FERNANDEZ-GONZALEZ et al., 2017; HONORATO, 2005). Estas discussões foram pautadas na seleção de variáveis para o desenvolvimento de modelos para a predição da qualidade do sínter.

Para simplificar o problema, a análise envolveu os dados dos últimos seis anos de operação de uma máquina de sinterização. Um conjunto de 1.841 amostras foi utilizado, com as variáveis de saída representadas pelos índices de 0 a 6, e as variáveis de entrada representadas pelos índices de 7 a 21. Os dados estão ilustrados na Figura 12.

Figura 12 – Informações do banco de dados em formato *DataFrame*.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1841 entries, 0 to 1840
Data columns (total 22 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   FeTMedia                                   1841 non-null   float64
1   FeoMedia                                   1841 non-null   float64
2   BasMedia                                   1841 non-null   float64
3   ShatterMedia                               1841 non-null   float64
4   TumblerMedia                               1841 non-null   float64
5   RDIMedia                                   373 non-null    float64
6   RIMedia                                    373 non-null    float64
7   Permeabilidade (JPU)                      1841 non-null   float64
8   Velocidade esteira (m/min)                1841 non-null   float64
9   Temp. FI (°C)                             1841 non-null   float64
10  % Umidade mistura 1                       1841 non-null   float64
11  % Umidade mistura 2                       1841 non-null   float64
12  Densidade de Carga                        1841 non-null   float64
13  Temp Entrada PE1                          1841 non-null   float64
14  Temp Gas Exausto                          1841 non-null   float64
15  >1,00mm                                   1841 non-null   float64
16  REL                                         1841 non-null   float64
17  <0,106mm                                  1841 non-null   float64
18  >6,30mm                                   1841 non-null   float64
19  PPC                                         1841 non-null   float64
20  U.T. MS3                                   1841 non-null   float64
21  PROD MS3                                   1841 non-null   float64
dtypes: float64(22)
memory usage: 316.5 KB

```

Fonte: Autor.

Os dados de entrada e saída foram extraídos principalmente do repositório geral de dados de processo, *Plant Information Management Systems (PIMS)*, e do sistema de gestão integrado da empresa, *Enterprise Resource Planning (SAP ERP)*, sendo posteriormente integrados em formato *excel*. A estrutura de coleta dos dados estão listados abaixo:

**Processo de Sinterização:** coletados a partir do repositório *PIMS*, onde são armazenados os dados dos sensores das máquinas de sinterização. Estes dados estão identificados no *DataFrame* como variáveis de entrada, pelos índices 7 a 14.

**Mistura de Minérios:** obtidos a partir de registros relacionados à amostragem da mistura no sistema *SAP*. O processo de amostragem ocorre 24 horas após a formação da pilha de material blendado, e as características da mistura são analisadas em laboratório. Estes dados representam variáveis de entrada do processo, identificadas pelos índices de 15 a 21.

**Qualidade do Sínter:** os dados foram coletados a partir dos registros relacionados à qualidade do sínter armazenados no sistema *SAP*. Para registrar os dados representados pelos índices de 0 a 4, são realizadas seis amostragens diárias do sínter, nas quais as características de qualidade deste produto são analisadas em um laboratório. Para os índices 5 e 6, é feita uma amostragem do sínter por pilha de material blendado formada, que também passa por análise laboratorial. Devido ao fato de uma pilha de blendado ser consumida a cada 5 dias, a disponibilidade de dados

de RDI (Índice de Degradação sob Redução) e RI (Índice de Redutibilidade) é menor, conforme pode ser observado no DataFrame apresentado na Figura 12. Estas características representam as variáveis de saída do processo.

As sinterizações em estudo não possuem uma estrutura de engenharia de dados preparada para aplicação direta de aprendizado de máquina. Portanto, para a aplicação prática deste trabalho, será necessário desenvolver, futuramente, uma arquitetura de coleta de dados que possa integrar os modelos de aprendizado de máquina realizando as predições em tempo real e suportando a análise do processo e a tomada de decisão.

Outro ponto considerado foi a necessidade de realizar um pré-processamento da base de dados antes de utilizá-la no desenvolvimento dos modelos. Isso se deve pela presença de dados faltantes, erros de medição, formatos diferentes e falta de integração dos dados. Através desta etapa foi possível identificar oportunidades de melhorias nos sistemas e instrumentos de medição, bem como no controle e manutenção destes equipamentos.

### 3.4 Análise inicial dos dados

Após o processo de coleta e análise da qualidade dos dados, foi necessário realizar uma exclusão manual de um número significativo de instâncias. Como resultado, o banco de dados foi reduzido de 1.841 amostras para 748, visando garantir um melhor desempenho do sistema. Todas as instâncias que continham dados faltantes foram diretamente removidas na planilha. Em seguida, o banco de dados foi importado e convertido para o formato *DataFrame* do pandas. A Figura 13 apresenta as informações detalhadas do banco de dados.

Figura 13 – *DataFrame* após tratamento inicial dos dados.

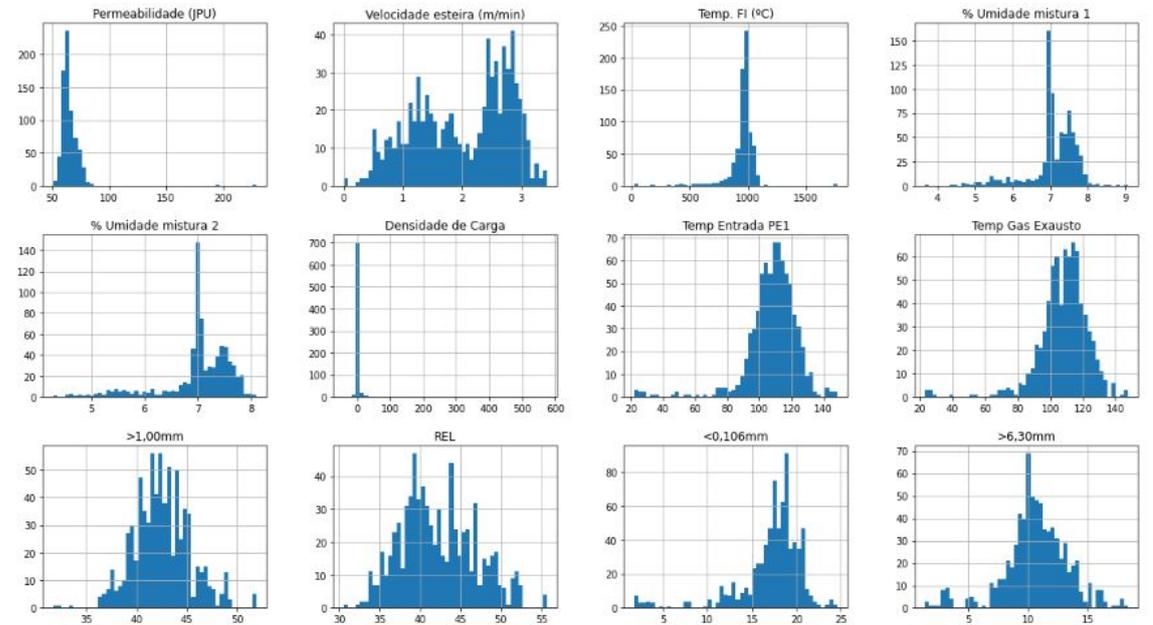
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 748 entries, 0 to 747
Data columns (total 22 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   FeTMedia                              734 non-null   float64
1   FeoMedia                              734 non-null   float64
2   BasMedia                              734 non-null   float64
3   ShatterMedia                          734 non-null   float64
4   TumblerMedia                          734 non-null   float64
5   RDIMedia                              140 non-null   float64
6   RIMedia                               140 non-null   float64
7   Permeabilidade (JPU)                  742 non-null   float64
8   Velocidade esteira (m/min)           748 non-null   float64
9   Temp. FI (°C)                        748 non-null   float64
10  % Umidade mistura 1                   748 non-null   float64
11  % Umidade mistura 2                   748 non-null   float64
12  Densidade de Carga                    748 non-null   float64
13  Temp Entrada PE1                      748 non-null   float64
14  Temp Gas Exausto                      748 non-null   float64
15  >1,00mm                               740 non-null   float64
16  REL                                    740 non-null   float64
17  <0,106mm                              748 non-null   float64
18  >6,30mm                                748 non-null   float64
19  PPC                                    740 non-null   float64
20  U.T. MS3                              748 non-null   float64
21  PROD MS3                              748 non-null   float64
dtypes: float64(22)
memory usage: 128.7 KB
```

Fonte: Autor.

### 3.5 Visualização dos dados

Com o objetivo de perceber os tipos de dados utilizados no estudo, foram plotados histogramas para os atributos do processo de sinterização, de acordo com a Figura 14.

Figura 14 – Histograma de variáveis do processo de sinterização.



Fonte: Autor.

A análise gráfica demonstra a existência de ruídos consideráveis nos dados, os quais podem estar associados a períodos de paradas e reinício de operações das máquinas, bem como a possíveis contaminações no abastecimento. Além disso, há uma observação importante em relação a possíveis erros de medição. Com o intuito de mitigar o impacto desses valores atípicos no desempenho dos modelos, foi realizado um segundo estágio de limpeza. Nesse estágio, foi aplicada uma filtragem por faixa de valores em cada atributo, levando em consideração a sua relevância ou representatividade. A faixa de filtros utilizadas está representada na Figura 15.

Figura 15 – Tabela de filtro de dados aplicado ao banco de dados do modelo.

	Faixa de Valores		
	Unidade	Mínimo	Máximo
BasMedia	-	1,50	1,90
ShatterMedia	%	88,0	91,0
Permeabilidade	JPU	50	90
Velocidade esteira	m/s	1,0	3,0
Temp. FI	°C	920	1050
% Umidade mistura	%	6,8	8,0
Densidade de Carga	-	1,80	2,05
Temp Gas Exausto	°C	75	350
>1,00mm	%	36,0	50,0
REL	%	32,0	55,0
PPC	%	< 10,0	
U.T. MS3	%	> 80,0	
PROD MS3	t/h	> 250	

Fonte: Autor.

A nova configuração do *DataFrame*, após a filtragem, está apresentada na Figura 16, onde o total de dados foi reduzido de 748 para 293.

Figura 16 – Base de dados tratada após visualização e limpeza dos dados.

```
Data columns (total 22 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   FeTMedia                                   293 non-null    float64
1   FeoMedia                                   293 non-null    float64
2   BasMedia                                   293 non-null    float64
3   ShatterMedia                               293 non-null    float64
4   TumblerMedia                               293 non-null    float64
5   RDIMedia                                   58 non-null     float64
6   RIMedia                                    58 non-null     float64
7   Permeabilidade (JPU)                       293 non-null    float64
8   Velocidade esteira (m/min)                 293 non-null    float64
9   Temp. FI (°C)                              293 non-null    float64
10  % Umidade mistura 1                         293 non-null    float64
11  % Umidade mistura 2                         293 non-null    float64
12  Densidade de Carga                          293 non-null    float64
13  Temp Entrada PE1                           293 non-null    float64
14  Temp Gas Exausto                            293 non-null    float64
15  >1,00mm                                     293 non-null    float64
16  REL                                          293 non-null    float64
17  <0,106mm                                    293 non-null    float64
18  >6,30mm                                     293 non-null    float64
19  PPC                                          293 non-null    float64
20  U.T. MS3                                    293 non-null    float64
21  PROD MS3                                    293 non-null    float64
dtypes: float64(22)
memory usage: 52.6 KB
```

Fonte: Autor.

## 3.6 Modelo piloto para previsão do shatter

### 3.6.1 Análise de correlação

Para avaliar as variáveis a serem consideradas neste modelo foi realizada uma análise dos fatores de correlação da variável de saída, o *shatter*, com as variáveis de entrada de processo inicialmente selecionadas de acordo com a Figura 17.

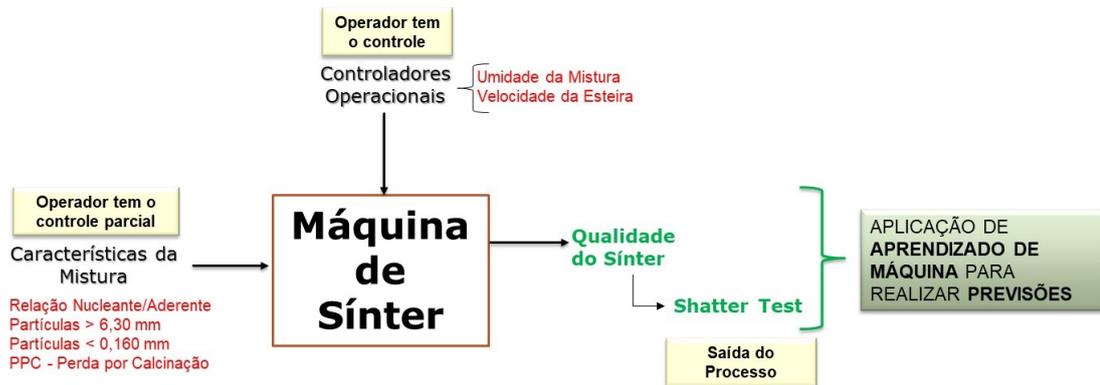
Figura 17 – Análise de correlação do *shatter* com as variáveis de processo e das matérias-primas.

ShatterMedia	1.000000
REL	0.523834
Velocidade esteira (m/min)	0.303777
Densidade de Carga	0.047761
BasMedia	-0.013593
Permeabilidade (JPU)	-0.017837
U.T. MS3	-0.040894
FeoMedia	-0.068781
Temp Gas Exausto	-0.130062
PROD MS3	-0.145250
Temp. FI (°C)	-0.158337
Temp Entrada PE1	-0.159815
>1,00mm	-0.160736
RDIMedia	-0.221802
RIIMedia	-0.261857
>6,30mm	-0.282514
% Umidade mistura 2	-0.335897
% Umidade mistura 1	-0.367127
<0,106mm	-0.403846
FeTMedia	-0.440597
PPC	-0.450288

Name: ShatterMedia, dtype: float64

Fonte: Autor.

Com base na análise dos fatores de correlação e levando em consideração a experiência técnica do processo, foi estabelecida uma estrutura para o modelo de previsão do *shatter*. Esta estrutura utiliza as seguintes variáveis: resistência mecânica do sínter (ShatterMedia), relação nucleante/aderente (REL), % umidade mistura 1, velocidade esteira, percentual de partículas supergrossas na mistura (>6,30mm), percentual de partículas superfinais na mistura (<0,106mm) e perda por calcinação (PPC). Esta configuração é ilustrada na Figura 18.

Figura 18 – Configuração do modelo piloto de aprendizado de máquina para previsão do *shatter*.

Fonte: Autor.

### 3.6.2 Variáveis de entrada e saída selecionadas

A seguir são apresentados os significados das variáveis selecionadas através da análise de regressão.

#### • Variável de Saída

1. A variável de saída descrita como **ShatterMedia** representa a resistência do sínter e consiste no teste de impacto que simula o percentual de finos gerados durante o envio deste produto para o alto-forno.

#### • Variáveis de Entrada

1. A primeira variável de entrada selecionada para o modelo foi a relação nucleante/aderente, representada como **REL** no *DataFrame*. O coeficiente de correlação associado a essa variável foi de +0,52. Esta relação refere-se à proporção entre as partículas nucleantes e aderentes presentes no material blendado, e possui um impacto direto na capacidade de aglomeração a frio da mistura. É essencial manter uma relação nucleante/aderente adequada para assegurar uma permeabilidade adequada da carga durante o processo de sinterização. A capacidade de aglomeração a frio é altamente desejável, uma vez que ela aumenta a produtividade da máquina e melhora a qualidade do sínter.
2. A variável **% Umidade mistura 1** (correlação de -0,36) representa o percentual de umidade presente na mistura após a etapa de homogeneização e umidificação no misturador. A utilização adequada de umidade é essencial para garantir a eficiência do processo de micro pelotização durante a etapa de aglomeração a frio, contribuindo para uma maior produtividade e qualidade do sínter.

3. A variável **Velocidade da esteira**, com correlação de +0,30, desempenha um papel fundamental no controle operacional. Ela tem um impacto direto na produção e deve refletir o progresso do processo de sinterização. A velocidade da esteira, quando bem controlada, é um reflexo da velocidade de sinterização. A velocidade da máquina deve ser ajustada de forma que a queima termine na região inferior ao atingir a penúltima caixa de vento. Isso assegurará que todo o leito tenha sido sinterizado, contribuindo para uma maior resistência do sinter.
4. A variável **>6,30mm**, com correlação de -0,28, representa o percentual de partículas supergrossas na mistura. Estas partículas não contribuem para o processo de aglomeração a frio da mistura, uma vez que as partículas mais finas ao seu redor não conseguem aderir à sua superfície
5. A variável **<0,106mm**, com correlação de -0,40, considera o percentual de partículas superfina (*pellet feed*). Essas partículas, devido à sua baixa granulometria, exercem uma forte influência na permeabilidade do leito durante o processo de sinterização.
6. A medida de perda por calcinação, representada como **PPC**, com correlação de -0,45, é uma variável significativa que desempenha um importante papel no comportamento do sinter no alto-forno. O PPC é uma indicação da quantidade de material que é eliminada durante o aquecimento da mistura. Esta medida é importante para entender as transformações químicas e físicas que ocorrem durante o aquecimento do material, e tem impacto direto na qualidade do sinter obtido.

### 3.6.3 Padronização dos dados

Antes da formulação dos modelos de aprendizado de máquina, as variáveis de entrada foram padronizadas para uma média igual a 0 e desvio padrão igual a 1. A padronização foi realizada utilizando a fórmula do *z-score*, de acordo com a equação (3.1):

$$Z = \frac{X - \mu}{\sigma} \quad (3.1)$$

onde:  $Z$  é o valor da variável  $X$  padronizado;  $X$  é o valor da variável na base de dados;  $\mu$  é a média aritmética das variáveis de entrada;  $\sigma$  é o desvio padrão das variáveis de entrada.

Esta transformação foi definida para evitar que os modelos fiquem enviesados para as variáveis com maior ordem de grandeza.

### 3.6.4 Divisão do conjunto de dados para treinamento, validação e teste

A divisão do conjunto de dados para criação dos modelos foi feita com 25% dos dados distribuídos para a base de testes e o restante para a base de treinamento. Para a validação dos modelos foi utilizada a técnica de *cross-validation*, com um número de partições  $k = 10$ .

### 3.6.5 Desempenho dos modelos

Para medir e comparar o desempenho dos modelos foram utilizadas as métricas de *mean absolute error (MAE)* e *root mean squared error (RMSE)* aplicados a base de testes. O *MAE* tem menos influência de valores discrepantes, enquanto o *RMSE* considera mais o impacto dos *outliers* na performance do modelo. O *MAE* foi a principal métrica avaliada para escolha do melhor modelo.

A técnica de *cross-validation* também foi utilizada para a comparação dos modelos. Esta comparação foi feita considerando os valores de *scoring* com base nos métodos do *RMSE* e *MAE*. Além disso, o *cross-validation* pode ser considerado para avaliar o efeito de diferentes valores de hiperparâmetros a serem utilizados nas aplicações de aprendizado de máquina abordadas neste trabalho.

### 3.6.6 Aplicação dos algoritmos de aprendizado de máquina

Na sequência foram aplicados diversos algoritmos de aprendizado de máquina considerando as variáveis selecionadas durante a etapa de definição das variáveis do processo de sinterização, sendo testados os métodos de regressão linear, regressão polinomial, árvores de decisão, florestas aleatórias, *ensemble learning* e redes neurais artificiais do tipo *MLP*.

#### **Regressão linear e regressão polinomial**

Para o treinamento dos modelos utilizando regressão linear e regressão polinomial, foi aplicada a função *LinearRegression* da biblioteca do *scikit-learn* mantendo os hiperparâmetros destes métodos no formato padrão e sem a utilização de técnicas de regularização.

Na regressão polinomial, foram testados vários graus para o polinômio, com o grau 2 obtendo um melhor ajuste para o modelo. Um maior grau levou o treinamento para o nível de *overfitting*.

#### **Árvores de decisão**

Para os modelos de árvores de decisão foram utilizadas uma profundidade máxima da árvore (*max\_depth*) de 5 níveis e um número mínimo de amostras em cada nó folha (*min\_samples\_leaf*) de 10. Foram testados outros valores para estes parâmetros e esta configuração demonstrou o melhor desempenho para o modelo. Valores muito altos levaram ao *overfitting* e valores muito baixos, ao *underfitting*.

#### **Ensembles - random forests, adaboost e gradient boosting**

Para o modelo com o método de *random forests*, foi definido o número de árvores de decisão para a floresta aleatória (*n\_estimators*) em 500. O número máximo de folhas que cada árvore de decisão poderá ter (*max\_leaf\_nodes*) foi 8. Esta configuração foi feita buscando restringir a complexidade do modelo e evitar o *overfitting*.

Já para o método de *adaboost*, foi utilizado o valor de 50 estimadores fracos (*n\_estimators*) para o conjunto de árvores de decisão definido. A profundidade máxima das árvores (*max\_depth*) foi de 5 e a taxa de aprendizado que controla a contribuição de cada estimador fraco no *ensemble* (*learning\_rate*) foi de 0,5.

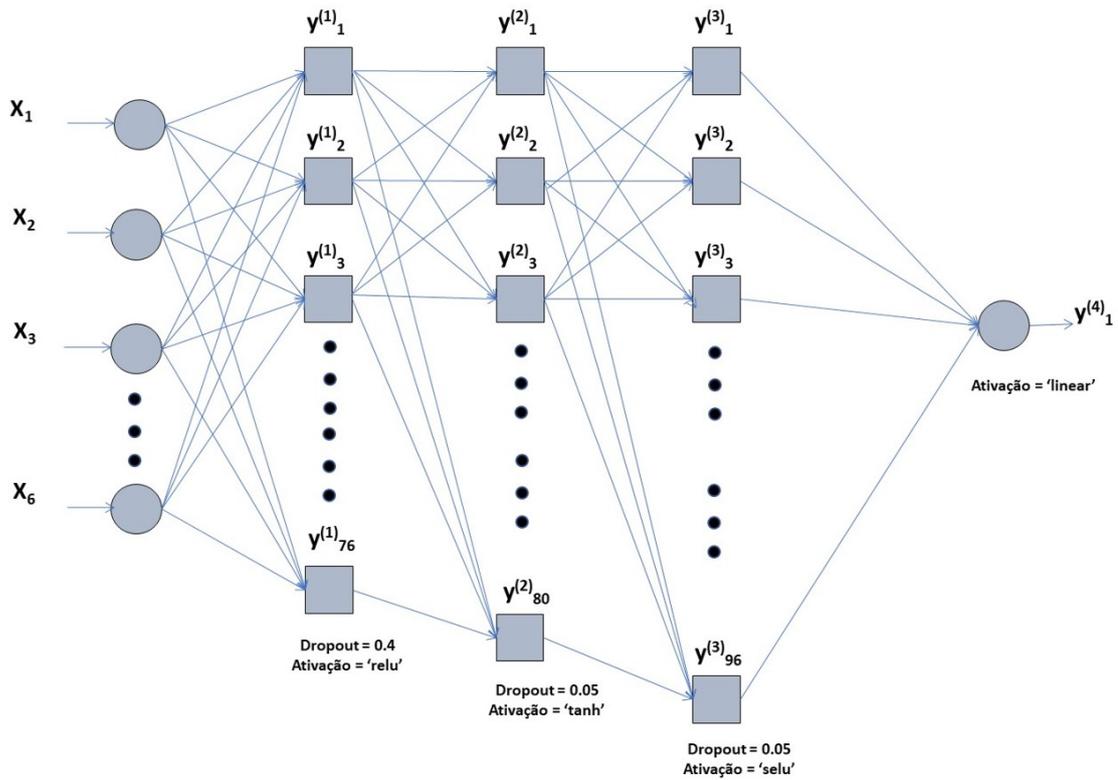
No *gradient boosting* foi considerado o valor de 3000 árvores de decisão (*n\_estimators*). A taxa de aprendizado que controla a contribuição de cada árvore de decisão no *ensemble* (*learning\_rate*) foi de 0,1. A profundidade máxima das árvores de decisão no conjunto (*max\_depth*) foi de 8. O número mínimo de amostra em cada folha (*min\_samples\_leaf*) foi de 3.

### **Redes neurais artificiais**

Inicialmente, foi empregada a biblioteca *kerastuner* para otimizar as redes neurais, possibilitando ajustar automaticamente os hiperparâmetros do modelo e buscar uma configuração de rede mais adequada para o problema em questão. A função ajustada pelo *kerastuner* utilizou três camadas densas. Os hiperparâmetros ajustados foram o número de neurônios na camada densa (variando de 12 a 128 na primeira camada e de 0 a 128 nas demais camadas, ambos com incrementos de 16), a taxa de *dropout* (com variação de 0.0 a 0.5 em incrementos de 0.05) e a função de ativação. Para a busca da melhor configuração, dentro das faixas consideradas, foi definido o *tuner* do *kerastuner* utilizando o método *RandomSearch*, estabelecendo o número máximo de modelos que são testados durante a busca (*max\_trials=15*) e a quantidade de vezes que cada modelo é avaliado com diferentes conjuntos de dados de treinamento e validação (*executions\_per\_trial=3*).

Em seguida, a partir dos resultados do *kerastuner*, foi desenvolvida uma rede neural artificial do tipo *MLP* de acordo com a configuração de rede mostrada na Figura 19.

Figura 19 – Modelo de rede neural artificial do tipo *MLP* para previsão do *shatter*.



Fonte: Autor.

Para compilar o modelo foi utilizado o método *adam* como otimizador da função. O *mean squared error (MSE)* foi definido como função de perda para calcular a diferença média quadrática entre os valores previstos e os valores reais e o *mean absolute error (MAE)* como métrica de avaliação do desempenho do modelo. Foram realizadas 250 épocas de treinamento para ajustar os pesos da rede neural e minimizar a função de perda.

## 4 Resultados

### 4.1 Avaliação dos modelos de aprendizado de máquina para previsão do *shatter*

Entre os algoritmos testados, o método de *ensemble* utilizando a técnica de *random forests* demonstrou o menor erro. O *mean absolute error (MAE)* para esse método foi de 0,31, enquanto o *root mean squared error (RMSE)* foi de 0,40. Ao avaliar os modelos por meio da validação cruzada, os melhores resultados foram alcançados pelos métodos *random forests* e *ensemble adaboost*. Por outro lado, os demais métodos apresentaram um desempenho inferior, com *MAE* variando em torno de 0,35 e *RMSE* na faixa de 0,44. As redes neurais artificiais exibiram um maior erro, com um valor de *MAE* de 0,42, podendo ser característico de *overfitting*, com o *kerasturner* não conseguindo chegar a uma configuração de rede que minimizasse esse erro. Vale ressaltar que há um grande potencial para reduzir o erro dos modelos, melhorando a qualidade e a integração dos dados utilizados.

Considerando a faixa de trabalho da base de testes, que utilizou valores de *shatter* entre 88,5 e 90,5%, o *MAE* de 0,31%, atingido pelo *random forests*, pode ser considerado aceitável para a utilização deste modelo no suporte as decisões operacionais de ajuste dos processos de sinterização e matérias-primas. A Tabela da Figura 20 faz uma comparação do desempenho dos modelos na previsão do *shatter*.

Figura 20 – Tabela comparativa da performance dos algoritmos de aprendizado de máquina.

Modelo Aplicado	Erro Médio	Raiz do Erro Quadrado	Cross-Validation	Cross-Validation	R2 Score
	Absoluto (MAE)	Quadrático Médio (RMSE)	Score - MAE	Score - RMSE	
Regressão Linear	0,35	0,44	0,42	0,51	34,07
Regressão Polinomial - Grau 2	0,36	0,44	0,42	0,51	33,22
Árvores de Decisão	0,33	0,44	0,46	0,54	42,48
<i>Random Forests</i>	0,31	0,40	0,41	0,51	46,22
<i>Ensemble - Adaboost Regressor</i>	0,31	0,42	0,42	0,51	47,38
<i>Ensemble - Gradient Boosting</i>	0,30	0,46	0,42	0,53	49,85
Redes Neurais	0,42	0,52	0,59	0,73	10,44

Fonte: Autor.

O desempenho inferior dos algoritmos mais simples, como regressão linear, polinomial e árvores de decisão, pode estar relacionado à complexidade das variáveis envolvidas nas sinterizações. Essa complexidade inclui características como não linearidade dos dados, agrupamento forte e atraso das informações do processo. Esses modelos mais simples têm a tendência de suba-

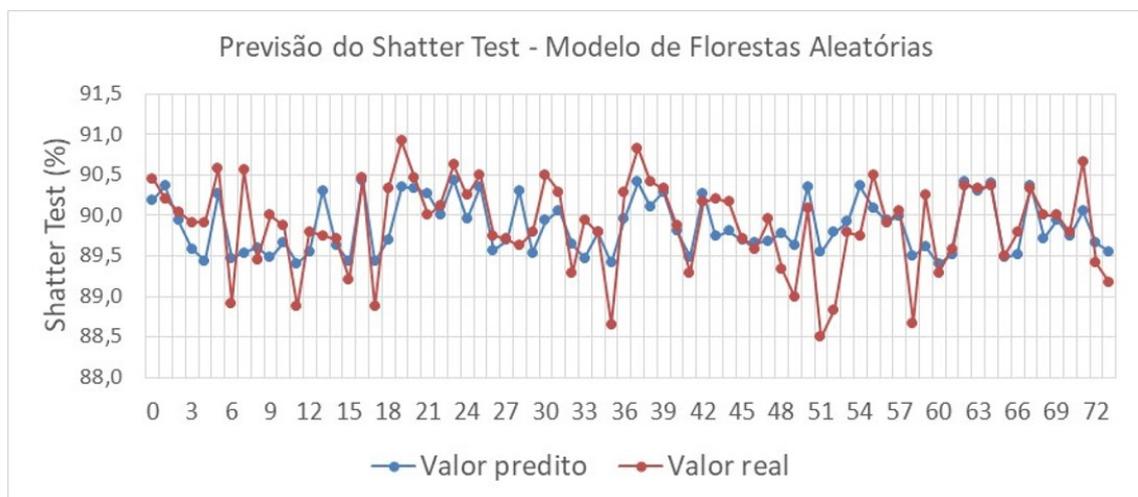
justar aos dados, o que significa que eles não conseguem capturar adequadamente as interações entre as variáveis. Por outro lado, as técnicas de *ensemble*, que combinam vários modelos mais fracos para formar um preditor mais conciso, têm mostrado resultados superiores. Essas técnicas ajudam a mitigar o problema de subajuste dos modelos mais fracos, melhorando a precisão e a capacidade de generalização do modelo. É importante pontuar que os modelos de redes neurais artificiais e o *gradient boosting*, considerados de alto desempenho, apresentaram maiores erros nos resultados. Isso indica a necessidade de explorar novos ajustes para os hiperparâmetros desses modelos, aproveitando ao máximo o seu potencial.

Na sequência são apresentados detalhes do comportamento dos modelos na base de testes.

#### 4.1.0.1 *Random forests*

Como demonstrado na tabela comparativa, a técnica de *random forests* apresentou o menor erro entre os modelos testados, sendo o mais viável para ser lançado em produção. A avaliação gráfica dos resultados de previsão considerando a base de testes vs valor predito também se mostrou razoável, conforme o gráfico da Figura 21.

Figura 21 – *Random forests* - comparativo do valor predito vs valores reais da base de testes.



Fonte: Autor.

O gráfico demonstra uma boa aderência do modelo aos dados reais do processo, o que permite lançar o sistema em produção para realizar as previsões em tempo real e dar suporte ao controle de processo da sinterização. Os resultados de maior erro no gráfico podem estar relacionados a questões pontuais, como contaminações no carregamento, erros de medição de balanças e paradas de equipamento. Para minimizar o erro do modelo, uma proposta é durante a etapa de limpeza dos dados, desenvolver um método para remoção destes dados que podem ser considerados como *outliers* do processo.

Utilizando o método de *random forests* foi verificada também a importância atribuída para as variáveis de entrada nas previsões, conforme mostra a tabela na Figura 22. Com isso, foi possível medir o impacto de cada variável no resultado final.

Figura 22 – Tabela de importância das variáveis na previsão do *shatter* utilizando a técnica de *random forests*

Variável	Importância na Previsão do Modelo
Perda por Calcinação (PPC)	25,4%
Granulometria Mistura >6,30 mm	19,5%
Velocidade da Esteira	17,0%
Relação Nucleante/Aderente	16,2%
% Umidade da Mistura	12,0%
Granulometria Mistura <0,106 mm	10,0%

Fonte: Autor

A tabela mostra a maior relação do *shatter* com as variáveis relacionadas a mistura utilizada, com a perda por calcinação com importância de 25,4% na previsão, seguida pela granulometria de > 6,30 mm, com 19,5%. A variável de processo selecionada pela análise de correlação foi a velocidade da esteira, com importância de 17,0%.

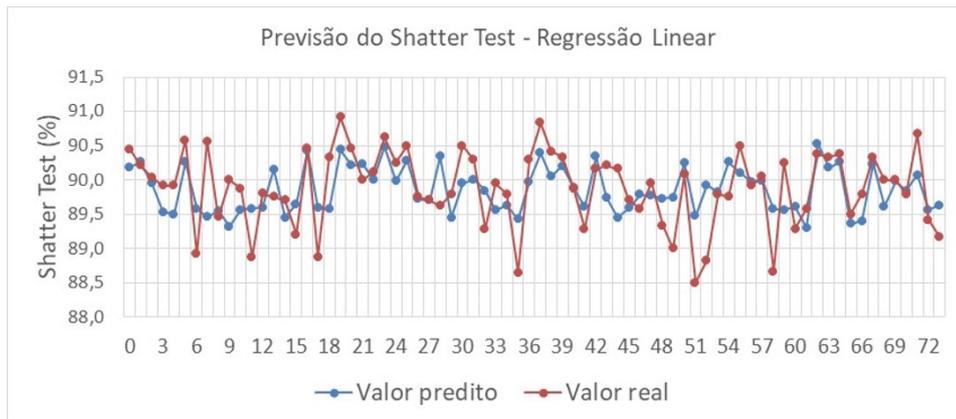
Como todos estes parâmetros podem ser parcialmente ajustados, a depender das matérias-primas disponíveis, estes resultados podem ser avaliados no sentido de buscar as melhores condições para melhoria da qualidade do síter.

#### 4.1.0.2 Regressão linear e regressão polinomial

As Figuras 23 e 24 apresentam o comportamento gráfico de previsão dos algoritmos de regressão linear e regressão polinomial. A análise gráfica demonstra uma melhor aderência do método linear em relação ao polinomial.

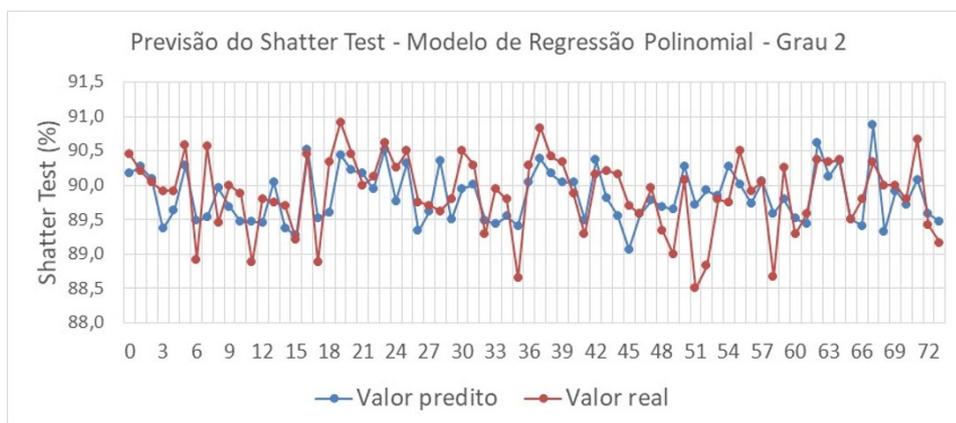
Considerando a simplicidade do algoritmo de regressão linear e o comportamento gráfico das previsões na base de testes, este algoritmo poderia ser utilizado como uma alternativa na previsão da qualidade do síter. No entanto, com o desenvolvimento dos modelos de *ensemble*, que apresentaram maior precisão nas previsões da qualidade do síter, a utilização da regressão linear torna-se menos viável. Como mencionado no início deste Capítulo, o menor desempenho destes algoritmos em relação a métodos mais robustos como os que utilizam as técnicas de *ensemble* se deve as características de complexidade dos dados das sinterizações.

Figura 23 – Regressão linear - comparativo do valor predito vs valores reais da base de testes.



Fonte: Autor.

Figura 24 – Regressão polinomial - comparativo do valor predito vs valores reais da base de testes.

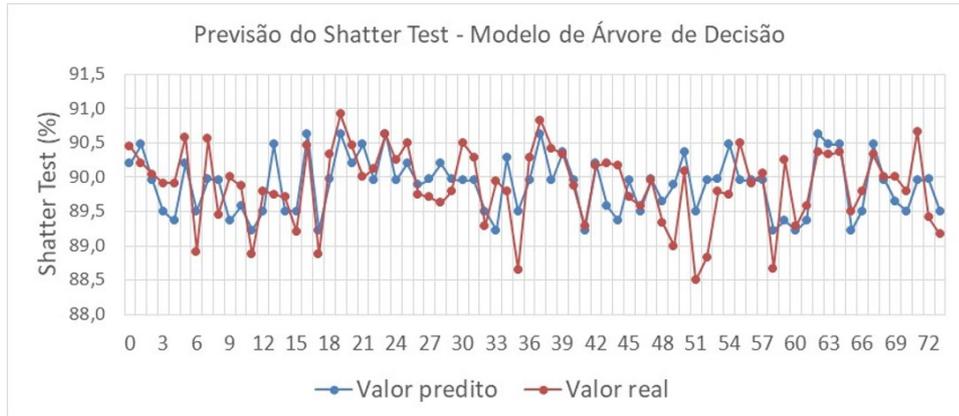


Fonte: Autor.

#### 4.1.0.3 Árvores de decisão

O modelo de árvore de decisão única apresentou *mean absolute error (MAE)* e *root mean squared error (RMSE)* com valores próximos aos da regressão linear e regressão polinomial, além de um comportamento gráfico também razoável para realizar previsões, conforme mostrado na Figura 25.

Figura 25 – Árvores de decisão - comparativo do valor predito vs valores reais da base de testes.

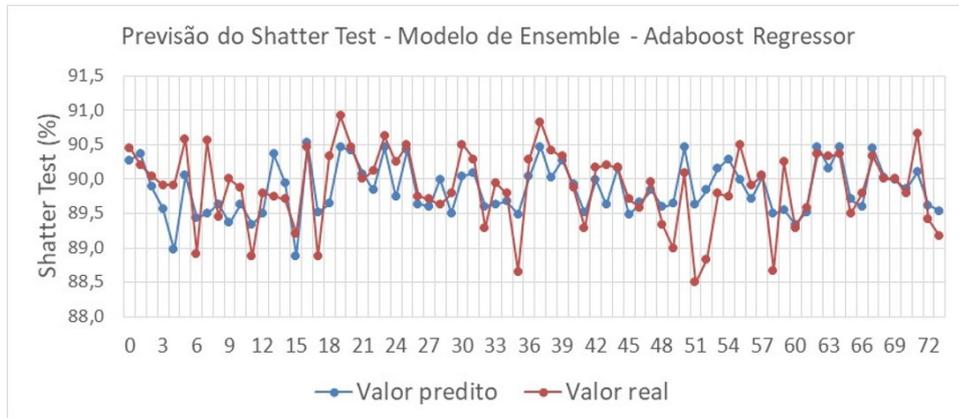


Fonte: Autor.

#### 4.1.0.4 Ensembles - adaboost e gradient boosting

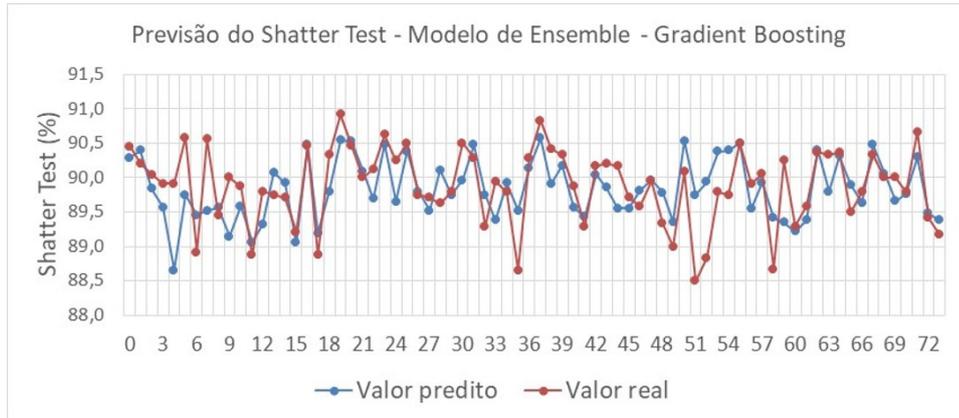
O modelo desenvolvido com o algoritmo *adaboost* apresentou valores de *MAE* entre os mais baixos, podendo ser comparado ao método de *random forests*. Conforme a Figura 26, o comportamento gráfico apresentou boa aderência das previsões ao conjunto de testes, principalmente nos últimos resultados. Já o algoritmo de *gradient boosting* apresentou menor aderência aos dados de testes, além de maior *MAE* como mostrado na Figura 27.

Figura 26 – Ensemble *adaboost* - comparativo do valor predito vs valores reais da base de testes.



Fonte: Autor.

Figura 27 – Gradient boosting - comparativo do valor predito vs valores reais da base de testes.



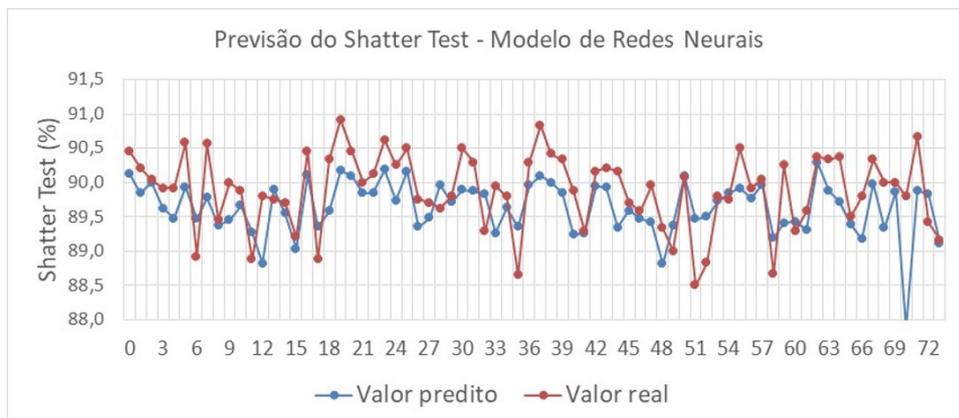
Fonte: Autor.

#### 4.1.0.5 Redes neurais artificiais

O modelo de RNA foi desenvolvido utilizando o algoritmo de busca iterativa de configurações de rede chamado *kerastuner*. A rede sugerida pelo *kerastuner* obteve um *MAE* de 0,42 e um *RMSE* de 0,52. No entanto, há potencial para reduzir o erro desse modelo, explorando outras estratégias de definição dos hiperparâmetros a fim de otimizar seu desempenho. Além disso, é fundamental melhorar a base de dados, aumentando a quantidade de dados disponíveis, integrando-os de forma mais abrangente e aprimorando sua qualidade.

A Figura 28 ilustra o comportamento da curva de predição do *shatter* utilizando redes neurais artificiais.

Figura 28 – Redes neurais artificiais - comparativo do valor predito vs valores reais da base de testes.



Fonte: Autor.

## 5 Considerações Finais

Com base na revisão da literatura sobre a aplicação de modelos preditivos nas sinterizações, foi possível identificar que a maioria dos estudos se concentrou na utilização da técnica de redes neurais. No entanto, este trabalho demonstrou que há um potencial significativo para a utilização de outros algoritmos de aprendizado de máquina nos dados das sinterizações. Este potencial foi amplamente explorado através do modelo piloto desenvolvido, fornecendo uma ferramenta eficaz de análise preditiva para suporte nas decisões operacionais de ajuste do processo. Além disto, o modelo piloto criado poderá contribuir para a melhoria da qualidade do sinter e abrir espaço para a construção de novos modelos preditivos que irão contribuir para a estabilidade operacional da sinterização.

Outro ponto de destaque está relacionado ao processo de construção da base de dados, onde foi possível realizar uma análise mais aprofundada da estrutura de dados existente no processo de sinterização. Nesta etapa, foi identificada a presença de dados faltantes, erros de medição, formatos diferentes e falta de integração, determinando uma necessidade do pré-processamento dos dados antes do treinamento dos modelos. Esta condição apontou para oportunidades de melhorias no controle, manutenção e desenvolvimento dos sistemas e instrumentos de medição. Em uma visão de longo prazo, a aplicação de novas tecnologias de coleta e armazenamento de dados, com foco em precisão e integração, poderá garantir mais qualidade dos dados disponíveis e, assim, contribuir para melhores decisões operacionais. Além disso, a análise dos dados também pode fornecer informações valiosas para a tomada de decisões estratégicas, permitindo uma melhoria contínua no processo e aprimorando a qualidade e a eficiência da produção.

Na etapa de resultados, foram avaliados os principais algoritmos de aprendizado de máquina relacionados a tarefas de regressão. Dentre os algoritmos testados, o método de *ensemble* com a técnica de *random forests* se mostrou o mais promissor com um *MAE* de 0,31. Outro algoritmo de *ensemble* testado, que utiliza a técnica *adaboost regressor* também se mostrou promissor, com um *MAE* de 0,33. Diante deste melhor desempenho do *random forests*, ele é o modelo mais indicado para ser lançado em produção, com grande potencial para suportar as decisões operacionais com foco na melhoria da qualidade do sinter. Em relação aos outros algoritmos testados e considerando a complexidade das variáveis envolvidas no processo de sinterização, tais como a não linearidade dos dados, forte agrupamento e atraso de informações de processo, é possível explicar a menor performance dos algoritmos mais simples, como a regressão linear, polinomial e árvores de decisão. Isso ocorre porque esses modelos tendem a se subajustar aos dados, não conseguindo capturar adequadamente as interações das variáveis. Por

outro lado, técnicas mais robustas de *ensemble*, que utilizam a combinação de diversos modelos mais fracos, tem apresentado resultados superiores, mitigando a subajustagem dos modelos mais fracos e melhorando a precisão e generalização do modelo.

Por fim, o *framework* desenvolvido para aplicação de algoritmos de aprendizado de máquina aos dados da sinterização poderá se estender em aplicações para previsões de variáveis importantes deste processo, como a produtividade, consumo de combustíveis e emissão de particulados, permitindo uma maior estabilidade operacional (qualidade e produtividade) das sinterizações e contribuindo para uma maior participação de sinter na carga e redução de *coke-rate* nos altos-fornos.

# Referências

- FAN, X. hui; LI, Y.; CHEN, X. ling. Prediction of iron ore sintering characters on the basis of regression analysis and artificial neural network. **Energy Procedia**, Elsevier, v. 16, n. Part B, p. 769–776, 2012.
- FERNANDEZ-GONZALEZ, D.; RUIZ-BUSTINZA, I.; MOCHON, J.; GONZALEZ-GASCA, C.; VERDEJA, L. F. Iron ore sintering: Quality indices. **Mineral Processing and Extractive Metallurgy Review**, v. 38, n. 4, p. 254–264, Maio 2017.
- HONORATO, E. P. Adequação granulométrica das matérias-primas e do sistema de segregação contínua (i.s.f), para melhorias na produtividade e qualidade do sínter para os altos-fornos. **Dissertação de Mestrado**, Universidade Federal de Minas Gerais, DEMET - Departamento de Metalurgia, 2005.
- JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. **ISIJ International**, v. 349, p. 255, 2015.
- LAITINEN, P. J.; SAXÉN, H. A neural network based model of sínter quality and sinter plant performance indices. **Ironmaking and Steelmaking**, Maney on behalf of the Institute, v. 34, n. 2, p. 109–114, 2007.
- LEITE, S. N. de C. **Redes Neurais Artificiais e Deep Learning**. Universidade Federal de Ouro Preto: PPGEP - UFOP / ICEA (Campus: João Monlevade), 2021.
- LV, X.; BAI, C.; QIU, G.; ZHANG, S.; HU, M. Moisture capacity: Definition, measurement, and application in determining the optimal water content in granulating. **ISIJ International**, v. 50, n. 5, p. 695–701, Março 2010.
- MAGALHÃES, J. A. D.; NATAL, L. F. G.; SÉRGIO, P. **Processo de Sinterização**. Usiminas Ipatinga: Usiminas, 2014. 102 p.
- MALLICK, A.; DHARA, S.; RATH, S. Application of machine learning algorithms for prediction of sinter machine productivity. **Machine Learning with Applications**, v. 6, December 2021.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics** 5.4, p. 115–133, 1943.
- OLIVEIRA, R. P. B. **Introdução ao Aprendizado de Máquina e Clusterização**. Universidade Federal de Ouro Preto: PPGEP - UFOP / ICEA (Campus: João Monlevade), 2020.
- OLIVEIRA, T. A. de. **Árvores de Regressão e Classificação**. Universidade Federal de Ouro Preto: PPGEP - UFOP / ICEA (Campus: João Monlevade), 2020.
- RASCHKA, S.; MIRJALILI, V. **Python Machine Learning: Machine Learning and Deep Learning with Python**. Birmingham, UK: Packt Publishing, 2019. <<http://www.deeplearningbook.org>>.
- RIBEIRO, M. C. **Regressão Linear**. Universidade Federal de Ouro Preto: PPGEP - UFOP / ICEA (Campus: João Monlevade), 2021. 54 p.

SHAO, H.; YI, Z.; CHEN, Z.; ZHOU, Z.; DENG, Z. Application of artificial neural networks for prediction of sinter quality based on process parameters control. **Transactions of the Institute of Measurement and Control**, v. 42, n. 3, p. 422–429, Novembro 2020.

SILVA, J. N. S. **Siderurgia**. IFPA - Campus Belém: Instituto Federal de Educação, Ciência e Tecnologia do Pará e Universidade Federal de Santa Maria para o Sistema Escola Técnica Aberta do Brasil – e-Tec Brasil, 2011. 110 p.