

RANYERE SOUSA SILVA

**Aprendizado de máquina aplicado ao planejamento de lavra de curto prazo  
para o aumento do desempenho operacional de equipamentos de mina**

São Paulo  
2021

RANYERE SOUSA SILVA

**Aprendizado de máquina aplicado ao planejamento de lavra de curto prazo  
para o aumento do desempenho operacional de equipamentos de mina**

**Versão Corrigida**

Dissertação apresentada à Escola  
Politécnica da Universidade de São Paulo  
para obtenção do título de Mestre em  
Ciências.

Área de Concentração:  
Engenharia Mineral

Orientador:  
Prof. Dr. Giorgio de Tomi

São Paulo  
2021

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 05 de agosto de 2021

Assinatura do autor: \_\_\_\_\_

Assinatura do orientador: \_\_\_\_\_

#### Catálogo-na-publicação

Silva, Ranyere Sousa

Aprendizado de máquina aplicado ao planejamento de lavra de curto prazo para o aumento do desempenho operacional de equipamentos de mina / R. S. Silva -- versão corr. -- São Paulo, 2021. 112 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Minas e de Petróleo.

1.Análise preditiva 2.Predição da produção 3.Aprendizado de máquina no planejamento de lavra 4.Apoio à tomada de decisão 5.Alocação de escavadeiras I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Minas e de Petróleo II.t.

Nome: SILVA, Ranyere Sousa

Título: Aprendizado de máquina aplicado ao planejamento de lavra de curto prazo para aumento do desempenho operacional de equipamentos de mina

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Aprovado em:

Banca Examinadora

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

A mão de Deus abençoa tudo que o homem faz, sabendo que ele é capaz de ter atitude boa. Até os erros perdoa, lhe colocando na trilha. Por isso meu mundo brilha, pois eu sei de onde venho: toda vitória que tenho, eu dedico à minha família.

Marquinhos da Serrinha.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, por todo o caminho que preparou para minha vida e que jamais me abandonou. A Ele, toda Glória.

Agradeço aos meus pais, Ramos e Fátima, por cada batalha realizada em meu favor, por toda a educação que me foi dada e, principalmente, por serem modelo de caráter e honestidade, sempre me ensinando que não há nada no mundo que supere o valor de ser uma boa pessoa.

Agradeço a Rondinelli, Ramon e Rodrigo por todo apoio, cuidados e ensinamentos. Obrigado pelo zelo constante. Obrigado pela ajuda, por vezes financeira, por outras vezes encorajadora, mas sempre de forma carinhosa e fraternal.

Meu muito obrigado a Deise Raquel, que jamais me deixou desanimar, mesmo diante de tantas dificuldades. Obrigado por ter sido uma base firme em todo esse caminho e por não desistir de me fazer alguém melhor. Minha insistência e perseverança são frutos de sua força, incentivo e resiliência.

Agradeço ao Prof. Dr. Giorgio de Tomi, que me deu a oportunidade de crescer e aprender todos os dias. Serei eternamente grato pela confiança, pelos ensinamentos e, principalmente, pela amizade.

Obrigado a Drauzio Marques e Antônia Marques pela acolhida e por terem sido peças fundamentais para o início de minha jornada.

Agradeço a todos os amigos e colegas do NAP.Mineração/USP. Especialmente, meu muito obrigado a Erbertt Barros, Roger Gonçalves e Brena Barros, que com muito companheirismo, compartilharam comigo vitórias, lutas, dores e alegrias. Desejo-lhes muita felicidade.

Ao Instituto Tecnológico Vale, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, à Tatiane Marin, Carla Campelo e Sandro Freitas, pela oportunidade.

No mais, agradeço a todos que contribuíram direta e indiretamente com essa conquista. Muito obrigado.



## RESUMO

Os processos, sistemas e pessoas envolvidos na cadeia de produção mineral são responsáveis pelo aumento considerável na geração de dados operacionais. Neste cenário, torna-se necessária a utilização de ambientes com capacidades de armazenamento e processamento de dados cada vez mais robustos para que métodos de análises inteligentes possam ser aplicados, a fim de realizar previsões do desempenho produtivo por meio de algoritmos de aprendizado de máquina. O objetivo deste estudo é aumentar a previsibilidade do desempenho operacional de equipamentos de mina, diretamente relacionado com a tomada de decisão nas atividades de planejamento de lavra de curto prazo. Para alcançar este objetivo foi proposta uma metodologia de aprendizado de máquina para o planejamento de lavra que abrange tarefas desde a preparação dos dados envolvidos no processo produtivo, até a tomada de decisão e replicação da rotina de análise para os próximos períodos que serão planejados. A metodologia foi aplicada no planejamento de lavra dos turnos operacionais em uma mina de cobre localizada na região norte do Brasil com o objetivo de prever o desempenho de produção de escavadeiras considerando variáveis de planejamento de lavra, geológicas, geográficas, de manutenção, da programação de horas, das turmas de trabalho escaladas para o turno e dos turnos de trabalho a serem planejados. A partir de diversos algoritmos de aprendizado de máquina, foram criados 175 modelos preditivos, treinados e testados por meio de validação cruzada, para que fosse definido uma modelagem que melhor se ajustasse ao padrão dos dados operacionais. Estes resultados comprovaram que sistemas preditivos podem ser inseridos na rotina operacional do planejamento de lavra de curto prazo, auxiliando na tomada de decisão para a alocação de equipamentos de maneira realística, eficiente e ágil.

**Palavras-Chave:** Análise preditiva. Predição da produção. Aprendizado de máquina no planejamento de lavra. Apoio à tomada de decisão. Alocação de escavadeiras.

## **ABSTRACT**

The complex combination of controls, systems and human aspects presented in the mine value chain are today responsible for an increasing amount of digital data in the mining industry. In this scenario, it is imperative the use of reliable and intelligent systems that can store and process the data to predict mine equipment performance. The objective of this study is to improve the prediction of operational performance in mine equipment for the short-term planning. For this purpose, it is proposed a machine learning (ML) methodology to map the production process from data collection until planning, with replication of the generated routines for subsequent short-term period analysis. The methodology was applied to predict the operational performance of excavators during working shifts in an open pit copper mine located in Northern Brazil, considering a series of variables such as operational, geological, geographic, maintenance. 175 predictive models were generated during the study, which were tested through cross-validation to improve the model adjustment to the collected data. The results obtained using this methodology confirmed that the use of ML predictive models provides a better understanding of the operation and allocation of mine loading fleet through the use of fast and realistic predictive routines.

**Key words:** Predictive analysis. Production process. Machine Learning in short-term planning. Decision making support. Allocation of mine loading equipment.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Mapa de calor das oportunidades em <i>Big Data</i> por indústria.....	18
Figura 2 - Fluxograma dos estágios de planejamento da mina em relação aos modelos econômicos de mineração.....	21
Figura 3 - Horas de atuação dos equipamentos de mina.....	28
Figura 4 - Ciclo de percepção da IA.....	31
Figura 5 - A interdisciplinariedade da Ciência de Dados para projetos de mineração.....	33
Figura 6 - Uma visão geral das etapas que compõem o processo do KDD.....	34
Figura 7- <i>CRISP-DM</i> é um processo adaptativo e iterativo.....	39
Figura 8 - Subajuste, sobreajuste e ajuste ideal.....	43
Figura 9 - Validação cruzada em <i>k-folds</i> .....	44
Figura 10 - Representação de uma aplicação para o algoritmo <i>kNN</i> .....	47
Figura 11 - Esquema para algoritmo de árvores de decisão.....	49
Figura 12 - Representação do algoritmo Floresta Aleatória.....	50
Figura 13 - Representação de uma Rede Neural <i>MLP</i> .....	52
Figura 14 - Representação do hiperplano de algoritmo <i>SVR</i> .....	53
Figura 15 - Comparação entre Busca Exaustiva em Grids e Otimização Aleatória.....	56
Figura 16 - Localização da mina do Sossego.....	60
Figura 17 - Processo padrão de ciência de dados para Planejamento de Lavra.....	62
Figura 18 - Exemplo de transformação pela técnica 1-de-c.....	69
Figura 19 - Registros de alocação da ES3001 conforme litologias da mina (2016) .....	73
Figura 20 - Frota de transporte utilizada nos períodos analisados.....	74

Figura 21 - Registros de horas das escavadeiras.....	75
Figura 22 - Registros das cotas em que as escavadeiras atuaram.....	76
Figura 23 - Registros das horas destinadas para manutenção corretiva.....	77
Figura 24 - Massa produzida (t).....	79
Figura 25 - Correlação entre as variáveis da base de dados.....	80
Figura 26 - Boxplot das variáveis numéricas normalizadas por reescala.....	83
Figura 27 - Boxplot das variáveis numéricas normalizadas por padronização.....	83
Figura 28 - Boxplot das variáveis numéricas normalizadas por reescala, após limpeza.....	85
Figura 29 - Boxplot das variáveis numéricas normalizadas por padronização, após limpeza.....	85
Figura 30 - Resultados preditivos para 75 turnos de operação.....	94
Figura 31 - Resultados para cenários modificados para os turnos planejados.....	98

## LISTA DE TABELAS

Tabela 1 – Sumário de frotas da Mina do Sossego.....	61
Tabela 2 - Discretização do atributo MASSA.....	67
Tabela 3 - Classes do atributo QUALIDADE.....	68
Tabela 4 - Resumo da base de dados após pré-processamento.....	81
Tabela 5 - Limites superiores e inferiores para variáveis das bases normalizadas.....	85
Tabela 6 - Seleção de <i>features</i> para a base normalizada por reescala.....	87
Tabela 7 - Seleção de <i>features</i> para a base normalizada por padronização.....	88
Tabela 8 - Hiperparâmetros avaliados para algoritmo <i>kNN</i> .....	89
Tabela 9 - Hiperparâmetros avaliados para algoritmo Árvore de Decisão.....	90
Tabela 10 - Hiperparâmetros avaliados para algoritmo <i>SVR</i> .....	90
Tabela 11 - Hiperparâmetros avaliados para algoritmo <i>SVR</i> .....	91
Tabela 12 - Hiperparâmetros avaliados para algoritmo <i>MLP</i> .....	92
Tabela 13 - Hiperparâmetros avaliados para algoritmo Regressão de <i>Ridge</i> .....	92
Tabela 14 - Ranking por MAE dos melhores parâmetros para cada algoritmo....	93
Tabela 15 - Cenário para o 1º turno previsto com baixa produção.....	96
Tabela 16 - Cenário para o 2º turno previsto com baixa produção.....	96
Tabela 17 - Cenário modificado para o 1º turno previsto com baixa produção....	97
Tabela 18 - Cenário modificado para o 2º turno previsto com baixa produção....	97



## SUMÁRIO

1.	INTRODUÇÃO .....	14
1.1.	OBJETIVO.....	16
1.2.	JUSTIFICATIVA .....	16
1.3.	QUESTÕES DE PESQUISA .....	19
2.	REVISÃO DA LITERATURA .....	20
2.1.	PLANEJAMENTO DE LAVRA.....	20
2.2.	PLANEJAMENTO DE CURTO DE PRAZO.....	23
2.2.1.	Alocação de equipamentos de mina .....	25
2.3.	INDICADORES CHAVE DE DESEMPENHO .....	26
2.3.1.	Indicadores operacionais de lavra.....	27
2.3.2.	Aderência de lavra .....	29
2.4.	A INTELIGÊNCIA ARTIFICIAL (IA) .....	30
2.5.	A CIÊNCIA DE DADOS ( <i>Data Science</i> ) .....	32
2.5.1.	Descoberta de Conhecimento em Banco de Dados ( <i>KDD</i> ).....	34
2.5.2.	Processo padrão entre indústrias para Mineração de Dados ( <i>CRISP-DM</i> )..	39
2.5.3.	Aprendizado de máquina .....	40
2.6.	VISÃO GERAL DA MINA DO SOSSEGO .....	59
3.	MATERIAIS E MÉTODOS.....	61
3.1.	SELEÇÃO DOS DADOS .....	63
3.2.	PRÉ-PROCESSAMENTO E TRANSFORMAÇÃO DOS DADOS.....	67
3.3.	MODELAGEM E AVALIAÇÃO DE DESEMPENHO DOS ALGORITMOS.....	70
3.4.	ANÁLISE PREDITIVA .....	71
3.5.	TOMADA DE DECISÃO NO CICLO DE PLANEJAMENTO DE LAVRA .....	72
4.	RESULTADOS .....	73
4.1.	RESULTADOS PARA SELEÇÃO DOS DADOS E ANÁLISE DESCRITIVA ...	73
4.1.1.	Litologia da jazida .....	73
4.1.2.	Frota de transporte.....	74
4.1.3.	Registros de horas .....	75
4.1.4.	Local de carregamento.....	75
4.1.5.	Manutenções corretivas (t/h).....	77
4.1.6.	Massa (t) .....	78
4.1.7.	Correlação das variáveis.....	80
4.2.	RESULTADOS PARA PRÉ-PROCESSAMENTO E TRANSFORMAÇÃO .....	81
4.3.	RESULTADOS PARA MODELAGEM E AVALIAÇÃO.....	86
4.3.1.	Análise de importâncias (Seleção de Recursos).....	86
4.3.2.	k-Vizinhos Próximos ( <i>kNN</i> ) .....	88
4.3.3.	Árvores de Decisão.....	89
4.3.4.	Regressão de Vetor de Suporte ( <i>SVR</i> ) .....	89
4.3.5.	Florestas aleatórias.....	90
4.3.6.	Rede Neural - <i>Perceptron</i> de multi camadas ( <i>MLP</i> ).....	90
4.3.7.	Regressão de <i>Ridge</i> .....	91

4.4. RESULTADOS PARA ANÁLISE PREDITIVA.....	91
4.5. RESULTADOS PARA TOMADA DE DECISÃO NO CICLO DE PLANEJAMENTO DE LAVRA .....	94
5. DISCUSSÃO .....	99
6. CONCLUSÕES.....	102
REFERÊNCIAS.....	103

## 1. INTRODUÇÃO

A indústria mineral brasileira participa de forma relevante na economia do país, considerando que a sua ampla produção atende os mais diversos setores da sociedade, fazendo parte tanto da estratégia global de crescimento, quanto do abastecimento de centros de produção locais e regionais.

O setor de mineração demanda de alto investimento de capital, destinado a diferentes necessidades exigidas para a implementação de empreendimentos de diferentes portes, como por exemplo, a compra de equipamentos e instalação de usinas industriais robustas (SACHS; NADER, 2005).

A exigência de altos investimentos, associada ao importante papel do setor para a economia do país torna imprescindível a gestão eficiente das operações, de modo que os empreendimentos sejam impulsionados a se tornarem cada vez mais competitivos.

Sendo assim, os gestores de operações de mina buscam constantemente adaptar os processos, sistemas e pessoas a fim de capacitá-los na elaboração de planos de execução de atividades que visam, principalmente, o melhor aproveitamento das jazidas minerais, possibilitando redução de custos, aumento de eficiência das operações e maior capacidade de gerir riscos.

Johansson *et al.* (2010) propôs que as empresas de mineração devem comprometer-se em criar ambientes de trabalho mais atraentes e motivacionais para as pessoas, de modo que metas sejam alcançadas, geridas e aprimoradas continuamente. Além disso, para assegurar que os roteiros das operações serão seguidos de modo a alcançar os objetivos dos empreendimentos, as equipes de planejadores devem estar empenhadas em compreender os fatores modificadores do negócio, adaptando e aperfeiçoando os meios necessários que alcançarão o sucesso das empresas.

Portanto, o Planejamento de Lavra exerce grande participação na busca por maior eficiência e competitividade na indústria mineral. Assim, Bazante (2004) definiu o planejamento de lavra como o esquema de desenvolvimento da mina que abrange a previsão dos meios e a determinação dos custos essenciais a esta evolução, e caracterizou o planejamento como dinâmico, pelo fato de que novas informações que surgem à medida em que a mina é lavrada, possibilitam a adequação do plano original às novas condições. Assim, Kwiri e Genc (2017), afirmam que as operações que desejam alcançar sucesso dependem de planejamento de lavra e gerenciamento de recursos disponíveis eficientes.

Os trabalhos de planejamento de lavra devem compreender todos os horizontes da mina ao longo dos anos, e por isso, é classificado de forma temporal: longo, médio e curto prazo (OSANLOO; RAHMANPOUR, 2017). O primeiro define a sequência de lavra em anos durante a vida útil da mina. O segundo, fundamentado nos objetivos do longo prazo, elabora as metas dos planos anuais e suas ações. Por fim, o planejamento de curto prazo, foco deste trabalho, descreve como serão executadas as atividades diárias, semanais, mensais, e até trimestrais.

Contudo, para que as decisões tomadas no horizonte de curto prazo sejam mais realísticas, é primordial a utilização de técnicas matemáticas, estatísticas, probabilísticas e computacionais, que promovam a melhoria dos processos minerais de modo que sejam reduzidos os custos e maximizados os benefícios. Cremonese (2014) afirma que é impossível que a cadeia de suprimentos minerais seja otimizada, eficiente e efetiva com a utilização de sistemas isolados e planilhas estáticas, e por isso faz-se necessário o uso de soluções sofisticadas que atendam os objetivos do setor mineral.

A utilização dessas soluções sofisticadas é ainda mais necessária para o atual cenário da mineração, onde os frequentes e constantes avanços tecnológicos, associados ao uso crescente de sistemas da informação, demandam de ambientes de armazenamento de dados com capacidades cada vez maiores (*Big Data*) e que possibilitam a realização de análises inteligentes (SILVA *et al*, 2018).

Por isso, é primordial que sejam utilizados mecanismos de análises de dados que possibilitem a descoberta de conhecimento em um universo robusto e gigantesco de informações. Assim, a Ciência de Dados surge como meio consistente que permite a descrição, classificação, estimativa, predição e agrupamento de grandes quantidades de dados, por meio da aplicação de diversas técnicas avançadas, como por exemplo, o chamado Aprendizado de Máquina (*Machine Learning*).

Sendo assim, esta dissertação propõe um processo padrão de análise preditiva para a operação de equipamentos de carga de uma mina, baseada no Aprendizado de Máquina, a fim de apoiar a tomada de decisão na dinâmica de planejamento de lavra, aumentando a aderência entre a execução operacional e os planos de curto prazo na Mina do Sossego, localizada no Sudeste do Pará.

Esta proposta está vinculada ao Projeto Min\_AO<sup>2</sup>: Gestão Sistêmica do Planejamento e Operação de Lavra na Mina do Futuro (Processo USP: 14.1.02753.03.6), realizado pelo Núcleo de Apoio e Pesquisa para a Mineração Responsável (NAP.Mineração), em parceria com a Companhia Vale e o Instituto Tecnológico Vale.

Os dados coletados nesta pesquisa são oriundos do Sistema de Despacho Eletrônico, registrados ao longo de uma série histórica. Por isso, foi primordial a utilização de técnicas e ferramentas suficientemente capazes para a predição dos padrões operacionais de lavra.

### 1.1. OBJETIVO

O objetivo geral desta pesquisa é aumentar a previsibilidade associada ao desempenho operacional de equipamentos de mina, apoiando a tomada de decisão no planejamento de curto prazo. Assim, os objetivos específicos são os seguintes:

- Aumentar a aderência entre a tomada de decisão do planejamento de curto prazo de uma mina e o que é realizado pela sua operação;
- Predizer o desempenho de produção de equipamentos no horizonte de curto prazo;
- Avaliar as variáveis que mais impactam a operação de carregamento;
- Reduzir incertezas operacionais;
- Melhorar a colaboração entre as áreas de planejamento, operação de lavra e manutenção por meio da integração de informações.

### 1.2. JUSTIFICATIVA

A execução e melhor desempenho de todas as atividades realizadas em operações de mina serão resultantes das decisões tomadas na dinâmica de planejamento de lavra. Por isso, é essencial que os planejadores utilizem técnicas que agilizem suas tarefas, de tal maneira que sejam eliminadas as falhas provenientes de decisões tomadas sem critério técnico. Assim, na rotina de planejamento, objetivando uma maior aproximação da realidade, é comum a criação de modelos dos mais diversos tipos, aplicações, objetivos e níveis de sofisticação.

Silva (2006) afirma que todas as etapas envolvidas na produção de uma mina podem passar por processos de modelagem e gerenciamento de forma mais eficiente, desde que exista um fluxo contínuo de informações, a fim de possibilitar melhorias significativas no planejamento e operação das minas.

A vasta quantidade de variáveis envolvidas nas operações de lavra não pode ser expressa nem modelada de forma simples, em virtude de suas aleatoriedades. Além

disso, essas variáveis podem ser tanto tangíveis quanto intangíveis, ou seja, as primeiras podem ser mensuradas, diferentemente das segundas.

Outro fator que dificulta a modelagem das operações é que as bases de dados operacionais necessitam de um tratamento adequado e, principalmente, que sejam adotadas metodologias capazes de construir um conhecimento acerca dos dados analisados, a fim de que tais análises possam auxiliar posteriormente em tomadas de decisão benéficas para o empreendimento mineral.

Além da necessidade de construir modelos eficientes, uma boa prática adotada por planejadores é definir indicadores de desempenho relevantes, a fim de se transformar os dados em informações úteis e utilizá-las de forma rotineira na elaboração dos planos de dimensionamento de produção de equipamentos, possibilitando tanto medir, quanto estimar a produção de mina. A prática de se realizar estimativas com baixo erro associado tem como consequência principal o aumento da aderência de lavra aos planos elaborados, ou seja, a redução das incertezas acerca do futuro.

Entretanto, nem sempre a previsão é realizada por meios suficientemente eficazes. Campelo *et al* (2017) explica que, em geral, as previsões do indicador de disponibilidade física são feitas considerando-se inspeções recentes dos equipamentos e o seu plano de revitalização. Já para os casos de dimensionamento de produtividade e utilização física, leva-se em conta dados obtidos na realização dos planos mais recentes que antecedem o período que se quer planejar. Porém, além do método explicado por Campelo *et al* (2017), outros métodos de análises também são aplicados para realizar previsões de indicadores, como por exemplo, a simulação dinâmica de eventos aleatórios.

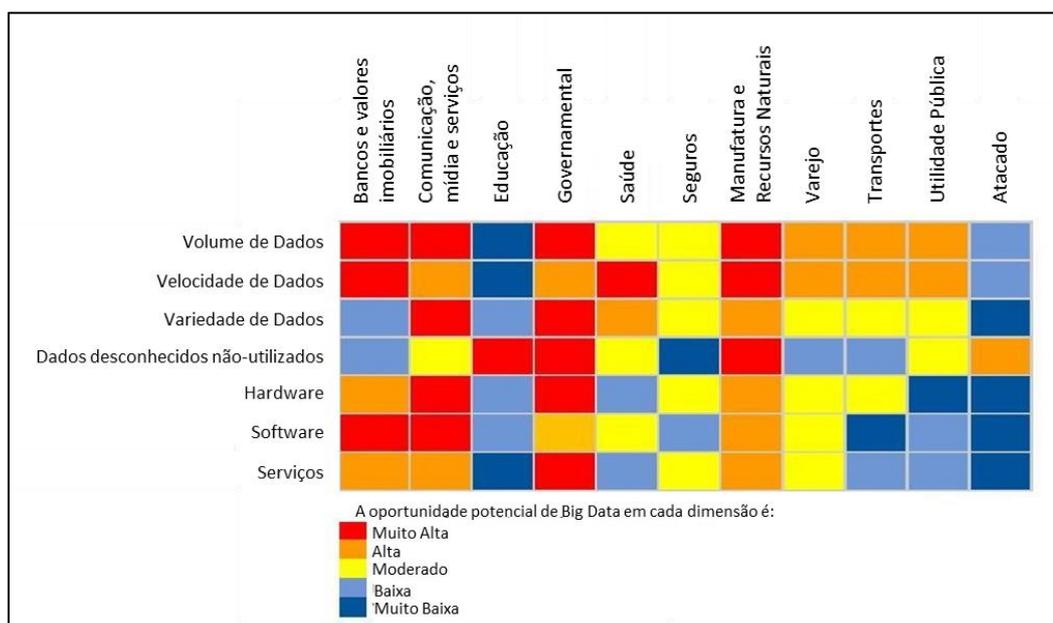
Assim, Campelo *et al* (2017) enfatiza que num cenário de curto prazo, a abordagem de indicadores como produtividade, disponibilidade física e utilização física não deve levar em conta os valores médios dos dados analisados, e por isso, é necessário considerar informações que representem o cenário atual da operação.

Desse modo, dados médios não são suficientes para se prever indicadores operacionais para o curto prazo com maior eficiência, pelo fato desses permitirem desvios distantes da média estimada. Por isso, a equipe de planejamento de lavra, normalmente, obtém dois tipos de informação: (i) as previsões de paradas cedidas pela Manutenção, baseadas em vistorias e previsões; (ii) dados de séries históricas das condições de trabalho de cada equipamento, registrados no sistema de despacho da mina. Assim, considerando este cenário, torna-se relevante a realização de previsões por métodos estocásticos ou de interpretação de padrões por meio da Ciência de Dados.

Além disso, a Ciência de Dados surge como ótima alternativa para melhoria no alcance de resultados em situações em que exista vasta quantidade de dados. Esse contexto, denominado como *Big Data* considera banco dados que possuam, além de alto volume, grandes variedades, alta velocidade, alta veracidade e alto valor de dados.

Desse modo, considerando os 5 V's do *Big Data* (volume, velocidade, variedade, versatilidade e valor), a Figura 1 apresenta um mapa de calor para as oportunidades do uso de tecnologias de *Big Data* em diversas indústrias.

Figura 1 - Mapa de calor das oportunidades em *Big Data* por indústria.



Fonte: Kart (2012).

Como apresentado na Figura 1, o campo das indústrias de manufatura e recursos naturais, no qual se enquadra a mineração, está entre as que apresentam maiores oportunidades para a implementação de *Big Data* em suas rotinas de trabalho. Esta realidade apresenta uma necessidade urgente para a indústria mineral: avançar as condições de seu ambiente de trabalho, tornando-as propícias para se alcançar o estado de Mina Inteligente. Para isso, os processos do ciclo de planejamento de lavra precisam ser submetidos a algumas mudanças, adaptadas a métodos de planejamento modernos.

Por isso, técnicas sofisticadas como o Aprendizado de Máquina, capaz de apoiar a tomada de decisão operacional a partir do universo *Big Data*, contribuem diretamente com o alcance de planos cada vez mais modernos e robustos.

Sendo assim, a relevância desta dissertação se evidencia pela intenção de prover para o problema em questão uma proposta de solução mais apurada que as utilizadas atualmente nas minerações, de forma que a inclusão desta solução na rotina de

atividades de Planejamento de Lavra torne a previsão de desempenho operacional mais realística e semi-automatizada. A sistematização das tarefas, alcançada pela metodologia criada e que é um dos vetores do Projeto Min\_AO<sup>2</sup>, bem como o aumento da previsibilidade operacional na mineração por meio da Ciência de Dados, são importantes inovações, tanto para a indústria quanto para a academia.

### 1.3. QUESTÕES DE PESQUISA

O presente trabalho busca responder as seguintes questões:

1. Qual o impacto da análise preditiva no planejamento de lavra de curto prazo?
2. Quais os desafios e benefícios para a implantação de um sistema preditivo no planejamento de curto prazo?

## 2. REVISÃO DA LITERATURA

Como foi definido por Amankwah (2011), a lavra a céu aberto pode ser entendida como uma operação de mineração realizada na superfície, em que é feita a escavação de minério e estéril.

Blom *et al* (2018) destaca o fato de que a lavra é mais adequada para corpos geológicos localizados perto da superfície, e a define como um conjunto de cavas, nas quais as camadas horizontais de material, chamadas de bancadas, são extraídas das cotas mais superiores para as inferiores.

O volume total de material extraído é determinado pelos limites finais da cava, que obedecem aos parâmetros espaciais, geomecânicos, legais e econômicos. Assim, uma mina é caracterizada pelo conjunto de operações (unitárias e auxiliares) que contribuem direta ou indiretamente com a extração do minério, podendo estas serem cíclicas ou contínuas. Para a melhor execução dessas atividades, faz-se necessário a elaboração de planos, que indicam os objetivos a serem alcançados e os meios disponíveis para atingi-los.

Sendo assim, é de extrema importância que sejam tomadas boas decisões que resultem nas melhores práticas de extração, visando sempre o melhor aproveitamento da jazida de forma segura e sustentável. Portanto, o planejamento de lavra se apresenta como ferramenta essencial para se atingir resultados satisfatórios nas operações de mineração.

### 2.1. PLANEJAMENTO DE LAVRA

O planejamento de lavra possibilita, com antecedência necessária, a previsão dos meios indispensáveis para alcançar a realização dos planos de produção, determinando os recursos imprescindíveis para que os objetivos sejam satisfeitos. Assim, o planejamento é apresentado como o roteiro de operações que se realizarão ao longo da vida útil de uma mina, compreendendo as etapas desde a preparação para o início da produção, até a exaustão da mina e recuperação da área lavrada (BAZANTE, 2004).

Blom *et al* (2018) afirma que as atividades de planejamento seguem uma hierarquia durante a vida útil de uma mina, indo desde a tomada de decisões estratégicas e de longo prazo, até decisões cotidianas e operacionais.

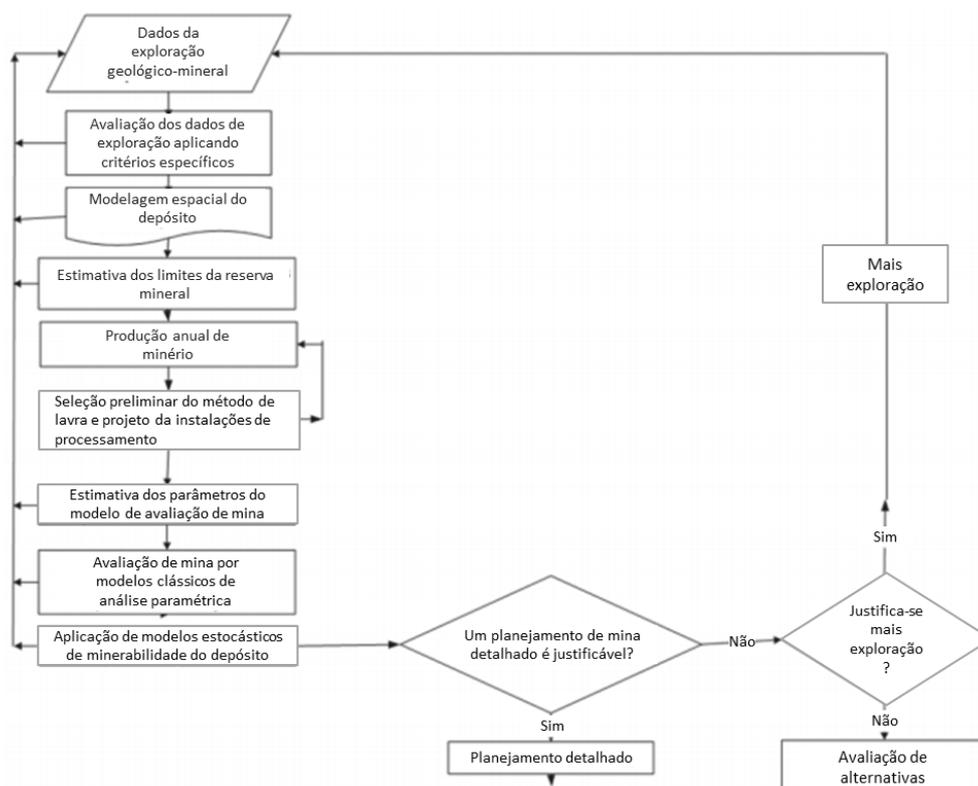
Roumpou *et al* (2016) destaca que o uso de diversas técnicas pelo planejamento de lavra permite a redução de incertezas e contribui com o aumento da vida útil de uma mina.

Essas incertezas, segundo Roumpou e Papakosta (2013), devem ser levadas em conta pelo fato de estarem relacionadas com as características geológicas, técnicas, ambientais, sociais, econômicas, entre outras condições do depósito.

Portanto, o planejamento de lavra deve estar presente em todas as fases do empreendimento mineral, para que seja possível alcançar o aumento da aderência das operações aos objetivos definidos pelo planejamento. Desse modo, o planejamento é classificado de acordo com o horizonte de tempo das ações a serem executadas nas atividades de mineração, ou seja, em longo, médio e curto prazo.

À medida que o nível de informações aumenta, as incertezas associadas à abertura do empreendimento mineral diminuem, possibilitando que o planejamento de lavra progrida para etapas cada vez mais detalhadas. A Figura 2 representa um fluxograma proposto para justificar avanços nos estudos de planejamento, considerando o nível de informações adquiridas ao longo do tempo.

Figura 2 - Fluxograma dos estágios de planejamento da mina em relação aos modelos econômicos de mineração.



Fonte: Roumpou *et al* (2016).

Nesse contexto, o planejamento de longo prazo está vinculado à quantidade lavrada em todo o decorrer da vida de uma mina e o retorno financeiro da atividade mineira. Segundo Blom (2018), nesta etapa é construído um modelo de blocos geológicos, estruturado em um grid de blocos de igual tamanho, que contém valores de estimativas de teores para cada bloco, além da variação litológica do material e outros atributos de qualidade relevantes.

O modelo de blocos, em geral, é submetido a uma modelagem econômica, que busca encontrar o maior valor presente líquido do empreendimento. Para isso, além do modelo de blocos, o modelo econômico utiliza como informação de entrada parâmetros espaciais e, principalmente econômicos, como: custos de mineração, custos de processamento, custos administrativos, preço de venda, diluição e recuperação de lavra, ângulo geral de talude, etc. Estes parâmetros são chamados de fatores modificadores.

O modelo econômico resulta numa cava otimizada, denominada de Cava Final. Campos *et al* (2018), indica que a cava final representa o volume de material que apresenta viabilidade econômica para ser extraído e que retornará o maior lucro possível para o negócio.

Silva (2001) aponta que a cava final representa a maior superfície da mina que possibilita o benefício operacional instantâneo máximo (fluxo de caixa não descontado) de uma jazida, que possui os fatores para ser lavrada a céu aberto satisfazendo os requerimentos operacionais, ambientais e de segurança.

O tamanho e forma da cava final são influenciados por alguns fatores, sendo eles, além dos econômicos, o limite legal de concessão de lavra e das restrições de projeto e produção. Os ângulos de taludes também afetam significativamente o design da cava, determinando a quantidade de remoção de estéril necessária para que o minério possa ser lavrado, sendo comum a variação destes parâmetros para diferentes setores e profundidades da cava (SILVA, 2008).

O planejamento de médio prazo desenvolve planos que determinam quando cada setor da mina será lavrado e quais os efeitos disso nos resultados anuais. Smith & Wicks (2014) acrescentam que o planejamento de médio prazo é útil para operações ativas que precisam lavar as camadas de estéril que estão sobrejacentes ao minério, pois este planejamento pode ser utilizado para definir metas de produção anuais, que serão definidas posteriormente de formas trimestrais e mensais.

Por fim, o planejamento de curto prazo destina-se a definir as ações a serem tomadas no dia-a-dia da mina, de modo que estas ações culminem no cumprimento das

metas estabelecidas pelos planos de médio e longo prazo. Ou seja, os esforços de curto prazo são direcionados sob o ponto de vista operacional de mina, compreendendo as condições e necessidades para que equipamentos, geometria de mina, pessoal e sistemas envolvidos no processo produtivo alcancem adequadamente os objetivos preestabelecidos.

## 2.2. PLANEJAMENTO DE CURTO DE PRAZO

A finalidade primordial do planejamento de curto prazo é definir os passos intermediários para que sejam atendidos os objetivos de longo prazo, ajustando a operação para atender aos requisitos estabelecidos, como: contorno de cava, teor, relação estéril / minério, etc. Estes requisitos precisam antecipar informações básicas sobre os lucros necessários para previsões de produções futuras e necessidade de novos equipamentos (BAZANTE, 2004).

Quevedo (2009) atribui as diferenças entre os planos de lavra de longo e curto prazo ao fato de o segundo exigir um nível de detalhamento muito maior, em virtude de que é necessário que as imprecisões a ele associadas sejam reduzidas.

Jewbali e Dimitrakopoulos (2018) acrescenta uma importante diferenciação entre horizontes de longo e curto prazos: enquanto a programação de produção de longo prazo objetiva a maximização do Valor Presente Líquido (VPL) de um projeto, a programação de produção de curto prazo concentra suas ações no atendimento das demandas de produção de curto prazo dentro do plano de longo prazo, entretanto serão considerados a capacidade de processamento da usina e a gestão de qualidade do minério que foi produzido na lavra e será processado posteriormente.

Segundo Silva (2008), também confirma que o horizonte de planejamento de curto prazo deve estar restringido pelos objetivos de longo prazo, de tal maneira que a visão estratégica global seja traduzida para as bases mensais, semanais e diárias.

Malhotra e List (1989) afirmam que os planos de curto prazo englobam horizontes diários, semanais e até mensal, que são desenvolvidos e implementados para garantir a integração de todos os equipamentos e atividades.

Para Osanloo e Rahmanpour (2017), além de seguir as diretrizes de longo prazo, o planejamento de curto prazo deve buscar meios para se reduzir o custo operacional, sempre que possível.

Nesta etapa de planejamento devem ser consideradas restrições de capacidade e disponibilidade de equipamentos, pessoal, acessos às áreas que serão lavradas, qualidade do material produzido, limitações geotécnicas, entre outras.

Além das restrições mencionadas anteriormente, o horizonte de curto prazo deve observar as dimensões definidas para o modelo de blocos nas etapas de longo prazo, a fim de que seja realizada a programação da produção.

Smith e Wicks (2014), por exemplo, destacam que a frente produtiva de um bloco geológico deve ter dimensões mínimas adequadas para suportar o trabalho de um equipamento de carga.

Esta mesma assertiva pode ser ampliada para todos os equipamentos e equipes de trabalho que estarão atuando nas frentes de lavra, onde a praça de trabalho deve ser dimensionada, de modo que a operação e a segurança não sejam comprometidas.

Já Lipovetzky *et al* (2014), acrescenta que a alocação destes equipamentos de carga deve seguir a sequência de lavra e a capacidade das frentes em atenderem a massa e a especificação química requerida.

Os principais objetivos dos planos de curto prazo consistem em: i) atingir metas operacionais de requisitos de qualidade e quantidade, de modo que sejam atendidas as exigências requeridas pelas plantas de processamento; ii) atingir níveis de utilização máxima de equipamentos e; iii) atender um alto nível de conformidade com os planos estratégicos. A conformidade dos planos de curto prazo durante as operações é essencial para o cumprimento dos planos estratégicos para alcançar os objetivos econômicos da mina (UPADHYAY; NASAB, 2018).

Blom *et al* (2018) confirmam também que dentre os mais variados objetivos que devem ser elevados a curto prazo, estão inclusos: manter o nível de produção dentro dos limites desejados em cada período de tempo, maximizar a utilização de equipamentos para atingir as metas de produção, e acrescentam, minimizar o remanuseio de material e minimizar custos operacionais.

Em virtude destas variadas e detalhadas atribuições que compõem a atuação da equipe de planejamento de curto prazo, que normalmente, assim como é explicado por Araújo (2008), este grupo de colaboradores deve ser bastante diversificado, englobando profissionais das mais diversas áreas do conhecimento, como por exemplo: geologia, engenharia de minas, qualidade, produção, administração, segurança do trabalho, e especialmente, da área social.

Na intenção de se alcançar planos de curto prazo cada vez mais realísticos e que considerem o elevado número de variáveis envolvidas no processo de produção, diversas técnicas de estimativas foram implementadas ao longo dos anos.

Montiel e Dimitrakopoulos (2017), por exemplo, explicam que diversas abordagens foram aplicadas nas últimas décadas para alcançar a redução da incerteza geológica de depósitos, como teor e tipo de material, tanto no projeto de desenho de mina, quanto na programação das atividades de produção.

Ramazan e Dimitrakopoulos (2012), se utilizam de programação estocástica inteira de dois estágios para solucionar o problema de programação de mina, onde a formulação pode ser considerar também a projeção de estoques, condições do processamento de material e informações de curto prazo.

Upadhyay e Askari-Nasab (2018) apresentaram uma opção de ferramenta para o desenvolvimento de cronogramas de curto-prazo, que utiliza modelagem de simulações de eventos discretos em interação com otimizações operacionais de mina baseadas em programação de objetivos. Esta ferramenta contabiliza incertezas operacionais e propicia tomadas de decisão proativas, ao contrário das decisões reativas que são tomadas forçadamente em ambientes de previsibilidade limitada.

Blom *et al* (2018) descrevem métodos modernos utilizados no planejamento de mina de curto prazo. Dentre as diversas aplicações, apresenta metodologias metaheurísticas baseadas em Programação Inteira Mista que capturam o comportamento dos equipamentos, considera multi-objetivos, etc. Destaca ainda aplicações que utilizam Simulação Condicional de corpos minerais, diferentes abordagens usadas para otimização multi-objetivo no contexto de curto prazo, entre outras.

Jewbali e Dimitrakopoulos (2018) criaram uma nova abordagem para otimizar a programação de produção de curto prazo a partir de uma integração com dados de planejamento de longo prazo. A abordagem foi dividida em duas etapas, utilizando simulação sequencial, pseudo-variogramas de dados de controle e de exploração e simulação condicional por resíduos e gerou modelos com maiores NPV e reserva mineral, quando comparados à abordagem convencional.

### **2.2.1. Alocação de equipamentos de mina**

A ótica da rotina diária se torna bastante evidente nos horizontes de curto e curtíssimo prazo, onde há tomada de decisões operacionais cotidianas como

posicionamento e despacho de equipamentos, controle de qualidade e controle de pessoal (BLOM; PEARCE; STUCKEY, 2018).

Atingir os requisitos de qualidade e quantidade de produção, além da utilização máxima dos equipamentos para cumprimento do plano estratégico são os objetivos centrais do planejamento de curto prazo. A utilização ideal dos equipamentos constitui um ponto essencial na economia de custos de uma mina, em vista que cerca de 60% dos custos operacionais totais são relativos às operações com caminhões e escavadeiras (UPADHYAY; ASKARI-NASAB, 2018).

A otimização da alocação dos ativos móveis de mina é dependente de diversas variáveis que implicarão em variações de tempo e custo associado a utilização de caminhões e escavadeiras como, custo de movimento por metro, horas/homens, configuração e disponibilidade da frota, recursos disponíveis e características das rotas. A aplicação de técnicas de pesquisa operacional e modelos de programação linear têm contribuído para a eficiência da alocação desses equipamentos e, muitas das operações de carregamento e transporte estão sendo otimizadas atualmente por meio do uso de algoritmos (UPADHYAY; ASKARI-NASAB, 2016).

Neste contexto, além de serem aplicados diversos tipos de técnicas em busca do alcance de maior previsibilidade no planejamento de curto prazo no que diz respeito ao melhor uso dos ativos de mina, uma outra abordagem que também se difundiu na indústria mineral é o uso de Indicadores Chave de Desempenho (ICDs), e que serão explorados a seguir.

### 2.3. INDICADORES CHAVE DE DESEMPENHO

Os indicadores chave de desempenho são resultados de medidas que podem ser avaliados de forma a qualificar uma determinada tarefa. Segundo Lima (2005, *apud* Costa 2015), indicador de desempenho é um conjunto de medidas que fornece informações acerca do desempenho de processos e produtos, tornando-se úteis no processo de tomada de decisões.

Segundo Dias (2008), os indicadores são compostos de:

- Índice: representa o padrão de medida ou unidade de medida dos indicadores;

- Referencial comparativo: é um índice utilizado como padrão de comparação;
- Meta: é um índice definido como ponto a ser alcançado num futuro previamente determinado e constituídas por objetivo, valor e prazo. Este componente se torna uma motivação para o desenvolvimento da gestão, à medida que os gestores trabalham desenvolvendo métodos para que as metas sejam alcançadas;
- Fórmula de obtenção do indicador: metodologia utilizada para que o índice seja apurado.

Para Peral *et al* (2017), os indicadores chave de desempenho têm um papel fundamental no monitoramento da execução dos negócios das organizações, fornecendo informações rápidas e precisas, através da comparação entre o desempenho atual e os objetivos indispensáveis para a concretização dos negócios.

Assim, Costa (2015) considera características relevantes de um indicador a medição do que é importante, a facilidade em ser compreendido, fácil identificação e a possibilidade de importar um pacote de indicadores, e não apenas um, particularmente.

Meier *et al* (2013), diz que os indicadores chave de desempenho são bastante eficazes, uma vez que criam transparência para o planejamento e controle, através de informações de suporte, que auxiliam nas tomadas de decisões por parte da administração.

Sendo assim, para que sejam alcançados os objetivos definidos pela equipe de planejamento de lavra, é imprescindível que as operações de mina sejam avaliadas a partir do uso de indicadores, observando os gargalos a serem vencidos e identificando oportunidades de melhorias para o negócio.

### **2.3.1. Indicadores operacionais de lavra**

Como descrito por Brandão e De Tomi (2011), o que irá permitir a quantificação dos impactos dos desvios dos ICDs na operação de lavra, será a escolha da ferramenta de estimativa de produtividade, que resultará numa melhor eficácia do sistema produtivo.

Nader *et al* (2012), comenta sobre a importância de se medir o impacto da cadeia de valor mineral através de indicadores, tornando possível quantificar os ganhos obtidos.

Dentre os diversos tipos de indicadores utilizados na rotina do planejamento, existem os ICDs voltados para quantificar o desempenho das frotas de equipamentos de

lavra. Estes indicadores possuem uma alta relação com as horas programadas para a operação e os eventos associados a elas.

A Disponibilidade Física (DF), Utilização Física (UF) e Produtividade são exemplos de indicadores comumente utilizados para se avaliar o desempenho de equipamentos de mineração. Em geral, são usados como informações importantes que auxiliam a programação de lavra de uma mina, integrando dados da política de manutenção das frotas, bem como controle das causas que resultam em horas não trabalhadas, etc.

A Figura 3 ilustra como as horas programadas podem ser interpretadas no cotidiano operacional dos equipamentos de mina.

Figura 3 - Horas de atuação dos equipamentos de mina.

Horas Calendário							
Horas Disponíveis				Horas de Manutenção			
Horas Trabalhadas			Horas Ociosas		Horas de Manutenção Preventiva		Horas de Manutenção Corretiva
Horas Trabalhadas Produtivas		Horas Trabalhadas Não-produtivas		Horas Ociosas Externas	Horas Ociosas Internas	Horas de Manutenção Preventiva Sistemática	Horas de Manutenção Preventiva Não-Sistemática
Horas Efetivas		Hora de Atraso Operacional	Horas Trabalhadas de Infraestrutura	Hora Trabalhada Diversa		Horas de Manutenção Corretiva	Hora Acidente

Fonte: Elaborado pelo autor.

Desse modo, temos que as Horas Calendário podem ser entendidas como a quantidade de horas programadas para as atividades a serem realizadas nas operações de mina.

A DF é o indicador que informa o quão disponível determinado equipamento estará para ser operado. Ou seja, a DF é um indicador que considera como horas disponíveis aquelas em que os equipamentos não se encontram em manutenção. A DF é dada pela equação (1).

$$DF (\%) = \frac{\text{Horas Programadas} - \text{Horas de Manutenção}}{\text{Horas Programadas}} \times 100 \quad (1)$$

A UF é um importante indicador que avalia, dentre as horas disponíveis, a quantidade de horas que um determinado equipamento trabalhou. Ou seja, a UF indica qual o tempo que o equipamento realizou algum tipo de trabalho (produtivo ou não-

produtivo), de forma a desconsiderar as horas que o equipamento esteve disponível, porém se manteve ocioso. A UF é dada pela equação (2).

$$UF (\%) = \frac{\text{Horas Trabalhadas}}{\text{Horas Programadas} - \text{Horas de manutenção}} \times 100 \quad (2)$$

O indicador de Produtividade (P) relaciona a produção de um dado equipamento com a quantidade de horas que este trabalhou para atingir esta mesma produção. Ou seja, o aumento da produtividade não está simplesmente ligado ao aumento da produção isoladamente, mas sim, ao aumento desta mesma produção em função de um dado período de tempo. A P é expressa pela equação (3).

$$P (t/h) = \frac{\text{Massa produzida}}{\text{Horas trabalhadas Produtivas}} \quad (3)$$

A Taxa de Produção (TP) é o indicador que representa a produtividade efetiva, ou seja, desconsidera das horas trabalhadas produtivas todas as horas de atraso operacional. Entende-se como causa para atraso operacional os eventos de causas internas e naturais que resultam em tempos de paradas momentâneas, ao ponto de interromper o ciclo das atividades produtivas. Por isso, TP pode ser expresso pela equação (4).

$$TP (t/h) = \frac{\text{Massa produzida}}{\text{Horas Efetivas}} \quad (4)$$

Desse modo, cada vez mais as operações de mina aproximarão o desempenho de suas tarefas realizadas com as decisões tomadas pela equipe de planejamento de lavra, aumentando a continuidade de acertos e diminuindo a frequência de desvios na realização de suas tarefas.

### 2.3.2. Aderência de lavra

O papel do planejamento de lavra na dinâmica de produção de mina é buscar constantemente que as suas decisões estejam sempre antecipadas em relação à modificação de fatores que influenciarão no resultado das operações.

Estes fatores, precisam ser constantemente monitorados e controlados, para que a sua imprevisibilidade seja mitigada. Isto, pelo fato de que eventos inesperados podem desviar a execução das tarefas, distanciando-as gradativamente das metas traçadas.

Para se avaliar o quão próximo a operação de lavra está dos planos elaborados é possível observar a Aderência como importante indicador. Este, por sua vez, pode ser

calculado a partir da razão entre o que foi executado pela operação e o que foi planejado anteriormente.

Esta medida gera um fator facilmente compreendido e bastante eficaz para gestores que desejam observar a oportunidade de melhorias e a eficiência de suas operações.

$$Aderência (\%) = \frac{Produção Real}{Produção Planejada} \quad (5)$$

Desse modo, a Aderência se torna um importante indicador para o planejamento de lavra, pois por meio dele será possível medir desvios oriundos da execução operacional, bem como identificar as causas resultantes desses.

Assim, considerando a vasta quantidade de dados gerados nas rotinas operacionais, bem como a necessidade de gerenciá-los por meio de ICDs e/ou prevê-los, é necessário que as equipes de planejamento de curto prazo absorvam técnicas capazes de acelerar o processo de planejar e gerar previsões para as operações, a fim de melhorar as tarefas e aumentar a aderência da lavra aos planos elaborados.

É nesse contexto que a Ciência de Dados surge como uma alternativa capaz de utilizar as metas do planejamento de curto prazo, o padrão das rotinas operacionais revelado pelos dados e os indicadores de desempenho que representem as operações para a geração de previsões realísticas que poderão apoiar eficazmente a tomada de decisão dos planejadores.

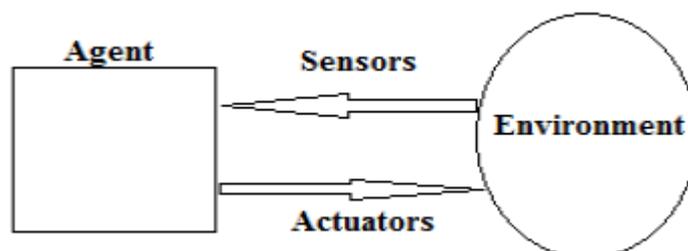
#### 2.4. A INTELIGÊNCIA ARTIFICIAL (IA)

A Inteligência Artificial, apesar de ter se tornado um tema atual e bastante frequente, se trata de um campo da ciência estudado há décadas. Quando este termo foi utilizado pela primeira vez, na década de 1950, vários pesquisadores já desenvolviam estudos que resultavam no desenvolvimento de sistemas cognitivos artificiais. Esta realidade, para a época, impulsionou o cientista da computação John McCarthy, considerado o pai da IA, a realizar uma conferência denominada de *Dartmouth Artificial Intelligence (AI) Conference*.

Como se tratava de um campo científico bastante novo, McCarthy precisou criar o termo *Artificial Intelligence*, e desde então, a IA é a ciência e engenharia de fazer máquinas e programas de computador inteligentes (MCCARTHY, 2007).

Das *et al* (2015) apresenta as máquinas e computadores inseridos na IA como “Agentes Inteligentes”. Um agente inteligente é capaz de interagir com o meio ambiente, identificando o seu estado a partir de sensores, podendo posteriormente afetar esse estado por meio de seus atuadores, como mostrado na Figura 4

Figura 4 - Ciclo de percepção da IA.



Fonte: Das *et al* (2015)

Uma das aplicações da IA é o chamado *Machine Learning* (Aprendizado de Máquina) que em 1959 foi definido por Arthur Samuel como o campo da ciência da computação que possibilita que computadores tenham a capacidade de aprender sem serem explicitamente programados. Outra definição mais formal, dada por Tom Mitchell para Aprendizado de Máquina (AM), é que esta ocorre se uma máquina programada para aprender com a experiência, melhora o seu desempenho a cada vez que realizar determinada tarefa (DAS *et al*, 2015).

Assim, computadores programados para aprender por meio de experiência adquirida devem eliminar significativamente os esforços dos trabalhos de programação (SAMUEL, 1959).

Durante a segunda metade do século XX o avanço da ciência em Inteligência Artificial e Aprendizado de Máquina foi bastante significativo.

McCulloch e Pitts (1943) propuseram uma maneira eficiente para cálculo baseado nas ações nervosas, partindo do princípio de que toda atividade nervosa segue uma lógica proposicional “tudo ou nada” (*all-or-none*), ou seja, para cada reação de qualquer neurônio há uma afirmação correspondente de uma proposição simples. Estes pesquisadores utilizaram o conhecimento da fisiologia básica e da função dos neurônios no cérebro, a lógica proposicional criada por Russell e Whitehead e a teoria da computação de Turing como as bases principais do estudo.

Turing (1950) publicou uma visão completa da Inteligência Artificial no artigo “Computing Machinery and Intelligency”, apresentando temas marcantes e ainda atuais para a área como o teste de Turing, aprendizagem de máquina, algoritmos genéticos e aprendizagem por reforço.

Em 1951, Marvin Minsky construiu a primeira rede neural artificial, que simulava um rato que percorria um labirinto em busca de comida. O neurocomputador, denominado de SNARC (Stochastic Neural Analog Reinforcement Computer), constituído de vários neurônios artificiais, conseguiu aprender conceitos durante o experimento (KELEMEN, 2007).

Em 1957 Frank Rosenblatt criou o primeiro Perceptron. Tratava-se de uma rede neural artificial construída segundo um algoritmo de aprendizagem supervisionada, que consiste num classificador binário (ROSENBLATT, 1957).

Uma das aplicações do aprendizado de máquina bastante útil para a indústria mineral é a chamada *Data Mining* (Mineração de Dados). Das *et al* (2015) explica que essa aplicação se evidenciou após o crescimento da geração de dados, gerados em virtude do aumento do uso da web e sistemas autônomos.

## 2.5. A CIÊNCIA DE DADOS (*Data Science*)

A vasta quantidade de dados agora disponíveis nos mais diversos setores industriais tem resultado em exploração de dados, visando vantagem competitiva. Ao passo que os computadores se tornaram mais poderosos e foram desenvolvidos diversos algoritmos que podem conectar dados, houve significativo aumento do volume e variedade de dados, ao ponto de superar a capacidade de bancos de dados convencionais e das análises manuais. Estes fatores convergiram na difusão da chamada Ciência de Dados – CD (PROVOST; FAWCETT, 2013).

Segundo Provost e Fawcett (2013), as tecnologias de *Big Data* são usadas para muitas tarefas, como por exemplo a engenharia de dados, a implementação de técnicas de Mineração de Dados, processamento de dados, etc.

O contexto *Big Data* também é uma atual realidade na mineração, conforme Figura 1. Segundo Silva *et al* (2018), a mineração tem passado por diversos avanços tecnológicos constantes, que associados ao aumento do uso de sistemas da informação, resultaram em gigantesco aumento de dados gerados a cada instante, necessitando de repositórios com maiores capacidades de armazenamento.

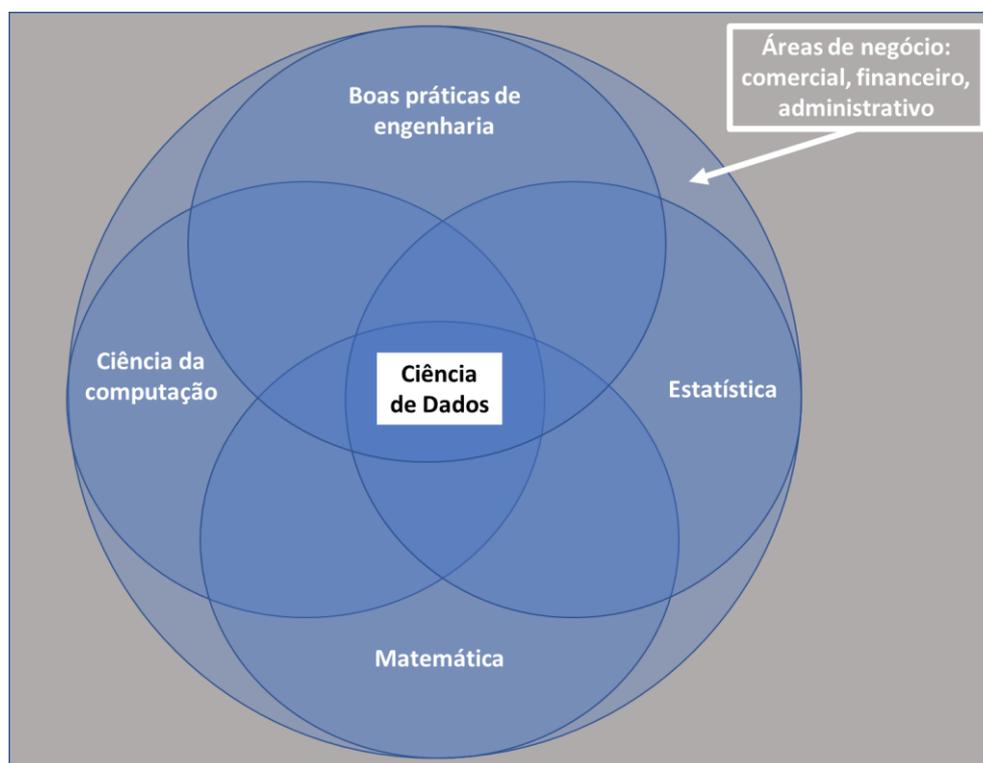
Outro fator que tem contribuído nesse contexto é a utilização cada vez mais frequente de equipamentos, sistemas e processos autônomos. Desse modo, a automação permite, além de redução de custos operacionais, aumentar a possibilidade de melhorar as tarefas.

Davis e Newman (2008) sugerem diversas ações para que se alcance na mineração um planejamento estratégico e moderno, sendo a primeira delas, a aplicação de algoritmos de otimização holísticos que sejam apropriados para o negócio, além de técnicas de soluções que oferecem melhorias em todo o projeto de uma só vez, eliminando ações desconexas.

Assim, uma tendência global das empresas de mineração, atualmente, tem sido direcionar esforços para que a exploração de dados faça parte de suas rotinas de gestão, acrescenta Silva *et al* (2018). Sendo assim, a CD é bastante adequada para tratar grandes quantidades de dados. Estes dados são armazenados em ambientes de capacidades gigantescas, denominados *Big Data*.

A CD pode ser denominada como uma área interdisciplinar que direciona esforços para o estudo e análise aprofundados de dados estruturados ou não estruturados, que visa a extração de conhecimento ou indícios de informações para possíveis tomadas de decisão. Ciência de Dados aplica diversas áreas interdisciplinares como estatística, economia, engenharia, administração e ciência da computação para solucionar problemas de diversos tipos, incluindo os inseridos no contexto *Big Data*. A Figura 5 ilustra a multidisciplinariedade da Ciência de Dados.

Figura 5 - A interdisciplinariedade da Ciência de Dados para projetos de mineração.



Fonte: Elaborado pelo autor.

De acordo com a Figura 5, é imprescindível que as áreas de negócio atuem diretamente em um projeto de Ciência de Dados aplicada na mineração. Além disso, os profissionais que atuam em todas as disciplinas envolvidas necessitam atuar de forma colaborativa e conjunta.

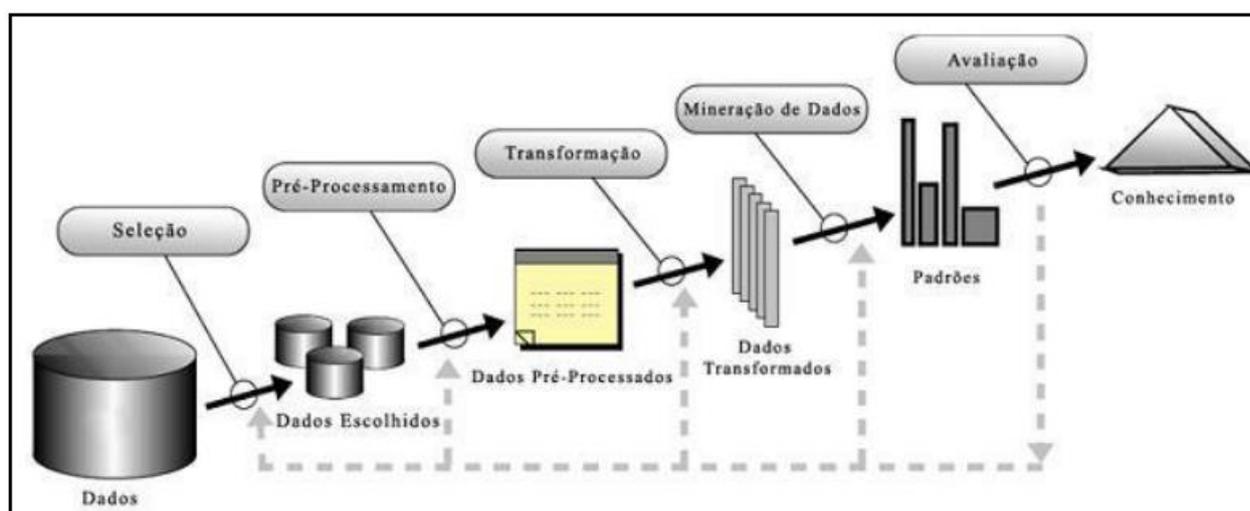
Portanto, ao compreender todas as vantagens competitivas oferecidas pelo uso de CD, certifica-se que extrair conhecimento de banco de dados é uma necessidade real para a indústria mineral, fazendo com que os atores desta indústria busquem se tornar, além de especialistas da atividade de mineração convencional, especialistas também na habilidade de extrair conhecimento útil de bancos de dados operacionais.

Assim, torna-se imprescindível conhecer um componente eficaz que, aliado à tomada de decisão no universo *Big Data*, complementa a rotina da Ciência de Dados: a Mineração de Dados.

### 2.5.1. Descoberta de Conhecimento em Banco de Dados (KDD)

Denominado por Fayyad *et al* (1996) como *Knowledge Discovery in Database – KDD* (Descoberta de Conhecimento em Bases de Dados), o *KDD* pode ser entendido como um procedimento padrão para se aplicar rotinas de descoberta de conhecimento. As etapas deste processo são detalhadas na Figura 6.

Figura 6 - Uma visão geral das etapas que compõem o processo do KDD.



Fonte: Fayyad *et al* (1996).

Desse modo, o fluxo de ações tem como base atividades ordenadas que garantem maior controle durante a busca de conhecimento em bases de dados, conforme será explanado a seguir.

### 2.5.1.1. Seleção dos dados

Nesta etapa, a partir de um repositório robusto de dados, se define quais destes dados serão relevantes para o processo de descoberta, considerando o objetivo global da análise. Em um universo gigantesco de registros, diversos conjuntos de dados podem ter variáveis que não apresentam contribuição ou benefício para o processo de extração de conhecimento. Esses dados podem ser excluídos do conjunto que comporá a análise.

Além disso, em função do custo associado ao processamento, tamanho de memória exigida e esforço computacional adotado no processo de *KDD* em grandes repositórios de dados, algumas técnicas de amostragem estatística podem ser aplicadas a fim de auxiliar na seleção de uma parcela reduzida dos dados, preservando a representatividade da população total.

Dentre as técnicas de amostragem estatística, as mais utilizadas são a amostragem simples e a amostragem estratificada.

A amostragem simples considera a extração de objetos com igual probabilidade da base de dados original. Esta extração pode ser com ou sem reposição, ou seja, no primeiro caso, a probabilidade de extrair cada objeto se mantém a cada rodada de extração. No segundo caso, por não haver reposição, a probabilidade é alterada.

A amostragem estratificada é adotada para casos em que existe desbalanceamento das classes de cada variável. Ou seja, neste cenário, normalmente uma ou outra classe possuirá uma quantidade de objetos consideravelmente maior do que as demais. Assim, a amostragem estratificada pode ser utilizada mantendo a mesma quantidade de objetos para cada classe no novo subconjunto, ou mantendo a mesma proporção de desbalanceamento da base de dados original.

### 2.5.1.2. Pré-processamento de dados

O pré-processamento de dados pode ser compreendido como a fase de preparação dos dados. Camilo e Silva (2009) explicam que esta preparação consiste basicamente de algumas etapas, como:

- Limpeza dos dados: etapa destinada para a verificação de inconsistências como registros incompletos, valores errados, dados inconsistentes, etc. Visa, sobretudo, evitar que inconsistências influenciem o resultado final dos algoritmos utilizados;

- Integração dos dados: consiste na etapa que visa integrar os dados oriundos das mais diversas fontes;
- Redução dos dados: nesta etapa pretende-se reduzir o volume de dados analisados, quando esse volume torna as análises e a própria mineração impraticável. Desse modo, a massa de dados original pode ser reduzida, desde que não se perca a representatividade dos dados originais.

### 2.5.1.3. Transformação dos dados

A transformação é feita quando se é necessário mudar o tipo dos valores observados em função das exigências de determinado algoritmo. Um exemplo comum ocorre pelo fato de que, os dados podem ser classificados em quantitativos (discretos ou contínuos) ou qualitativos (nominais e ordinais). (CAMILO; SILVA, 2009). Portanto, a depender do objetivo da análise, seja necessária a transformação de dados numéricos em dados categóricos ou vice-versa.

Algumas outras técnicas como a normalização também são utilizadas nessa etapa. Esta técnica é utilizada pelo fato de alguns conjuntos de dados reais apresentarem variáveis contínuas, com valores que distribuídos em faixas completamente distintas ou que possuem diferentes variações, em função do comportamento natural dessas variáveis ou devido a escala utilizada em suas medições. Em algumas situações, valores não-normalizados podem afetar o desempenho de algoritmos. Os tipos mais comuns de normalização são os de reescala ou padronização.

A normalização por reescala, consiste em escalar os dados em um intervalo padrão para todos os atributos, a partir da definição de um valor mínimo e um valor máximo. Geralmente, o intervalo escolhido para a nova escala é definido como [0;1]. A normalização por reescala de um atributo  $j$  de um objeto  $x_i$  pode ser calculada como:

$$X_{ij} = \frac{X_{ij} - \min_j}{\max_j - \min_j} \quad (6)$$

Onde os valores mínimo e máximo do atributo  $j$  são dados como  $\min_j$  e  $\max_j$ , respectivamente.

Por sua vez, a normalização por padronização, conforme Faceli *et al* (2011), considera que os diferentes atributos contínuos estarão dispersos em diferentes intervalos, porém, deverão possuir os mesmos valores para alguma medida de posição e de variação. Geralmente são utilizadas a média e desvio-padrão como medida de posição

e variação, respectivamente. A normalização por padronização de um atributo  $j$  em um objeto  $i$  pode ser calculada como:

$$X_{ij} = \frac{X_{ij} - \bar{X}_j}{s_j} \quad (7)$$

Onde  $\bar{X}_j$  e  $s_j$  representam, nesta ordem, a média do atributo  $j$  e o seu desvio-padrão. Fazendo com que cada atributo  $j$  tenha média zero e desvio-padrão 1.

#### 2.5.1.4. Mineração de dados

O termo Mineração de Dados (MD) pode ser compreendido como o direcionamento de esforços para se realizar a extração de conhecimento em bancos de dados robustos por meio de diversas ferramentas matemáticas, computacionais, estatísticas, etc.

Fayyad *et al* (1996) definiu a Mineração de Dados como sendo a utilização de diversas ferramentas computacionais a fim de realizar a descoberta de informações valiosas e potencialmente úteis, de modo que o entendimento de um problema possa ser melhorado, visando a melhor tomada de decisão.

Ou seja, a mineração de dados é formada por um conjunto de ferramentas e técnicas que através do uso de algoritmos de aprendizagem, são capazes de explorar um conjunto de dados, extraíndo ou ajudando a evidenciar padrões nestes dados e auxiliando na descoberta de conhecimento.

Como ilustrado na Figura 5, a Mineração de Dados é concretizada a partir da identificação de padrões comportamentais dos dados analisados.

Segundo Amo (2004), a Mineração de Dados é um ramo da computação que se desencadeou no início dos anos 80, quando as empresas e organizações começaram a se preocupar com os grandes volumes de dados informáticos estocados e que até então não eram utilizados. Nesta época, acrescenta Amo (2004), o conceito de Mineração de Dados consistia essencialmente em extrair informação de bases de dados gigantescas, da maneira mais automatizada possível.

Petermann (2006) explica que a MD pode se dar por meio de 3 principais áreas: a estatística, a inteligência artificial e o aprendizado de máquina. A estatística é a base da maioria das tecnologias de Mineração de Dados disponíveis que se baseiam em conceitos clássicos como média, mediana, moda, variância, correlação, etc. A inteligência artificial realiza extração de conhecimento construída a partir de fundamentos de heurística, apoiadas por estatística clássica. A terceira linhagem da MD é o aprendizado

de máquina, que de maneira geral consiste na junção das melhores práticas de estatística clássica e inteligência artificial.

Após realizar a preparação dos dados a serem utilizados são aplicadas tarefas de mineração de dados. A MD é bastante caracterizada pela capacidade em realizar determinadas tarefas como as descritas a seguir.

- Descrição: descreve padrões e tendências revelados pelos dados;
- Classificação: visa identificar a qual classe um determinado registro pertence;
- Estimativa ou Regressão: é semelhante à classificação, porém utilizada quando um registro é identificado por valor numérico, e não categórico;
- Predição: similar às tarefas de estimativa e classificação, porém, visa descobrir o valor futuro de um determinado atributo;
- Agrupamento: visa identificar e aproximar os registros similares em grupos (clusters).

Como indica Côrtes *et al* (2002), após serem definidas as funcionalidades a que se deseja chegar com Mineração de Dados, escolhe-se as técnicas a serem utilizadas a fim de se alcançar maior aderência na obtenção dos resultados e que sejam mais precisas em relação ao objetivo proposto. Cada tarefa de mineração de dados possui uma série de métodos de aplicação direta. Serão descritos alguns destes métodos de Mineração de Dados no decorrer deste trabalho.

#### 2.5.1.5. Avaliação

O processo de descoberta de conhecimento em banco de dados se encerra quando após a mineração de dados é realizada uma profunda avaliação de desempenho dos resultados alcançados pela etapa anterior.

As métricas de avaliação são definidas de acordo com o tipo de tarefa de MD realizada, tendo em vista que cada medida avaliativa foi criada visando um tipo de problema específico que a MD se propõe a solucionar.

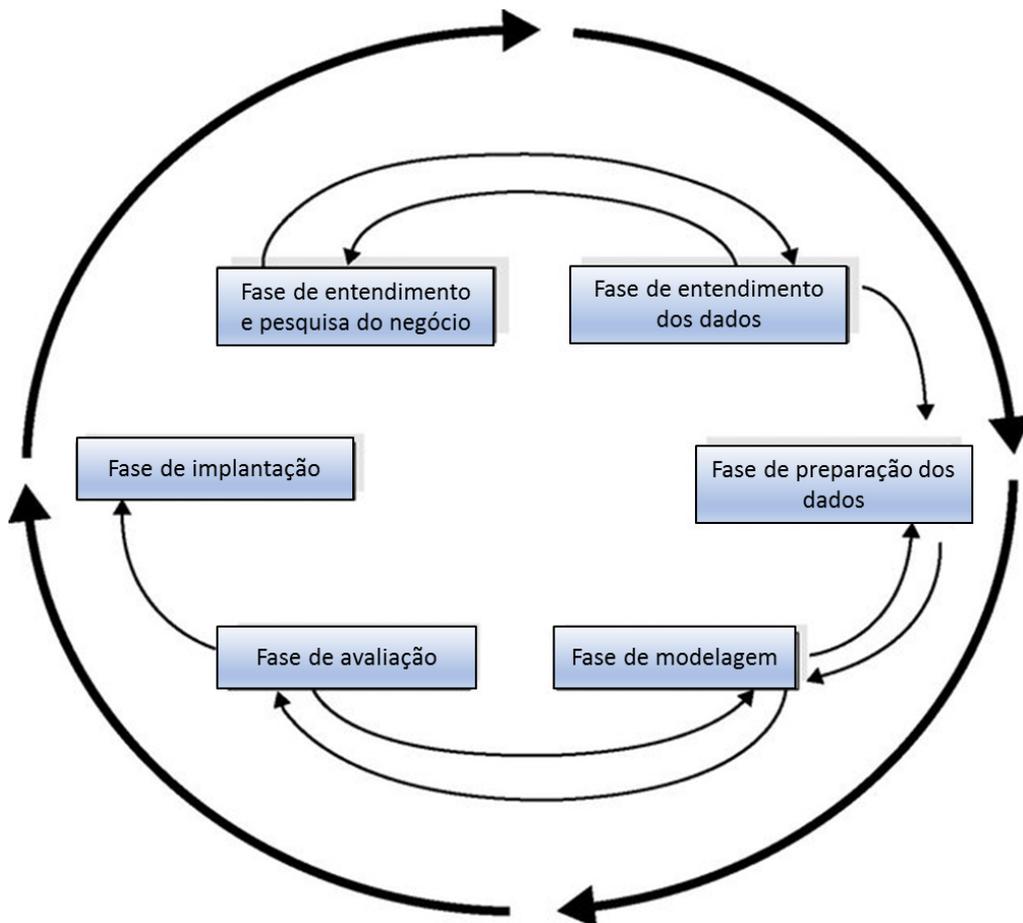
### 2.5.2. Processo padrão entre indústrias para Mineração de Dados (*CRISP-DM*)

Em virtude da dinâmica e rapidez exigidas por empresas na extração de conhecimento em seus bancos de dados, não se torna adequado adotar rotinas de análises que demandem muito tempo. Assim como também, a inércia departamental pode ser um fator que atrase um processo de análise eficiente baseado em Ciência de Dados.

Desse modo, é necessário que os processos de extração de conhecimento em ambiente corporativo sejam suficientemente rápidos, mas ao mesmo tempo dinâmicos, para que sejam gerados resultados que possam auxiliar a tomada de decisão em tempo instantâneo.

Para isso, foi criada a metodologia *CRISP-DM* (*Cross-industry standard process*).

Figura 7- *CRISP-DM* é um processo adaptativo e iterativo



Fonte: Larose (2005)

Larose (2005) apresenta essa metodologia como um processo padrão para ajustar a MD na estratégia geral de solução de problemas de uma unidade de negócios ou de pesquisa. Portanto, o *CRISP-DM* divide o projeto de mineração de dados para empresas em 6 fases, conforme apresentado na Figura 5.

### **2.5.3. Aprendizado de máquina**

Vieira, Lopez e Mechelli (2019) afirmam que não há uma definição universal para o aprendizado de máquina, mas que este é conhecido como uma área da inteligência artificial cujo objetivo é identificar padrões em dados fornecidos e usá-los para fazer previsões. Arthur Samuel, um dos pioneiros da Inteligência Artificial, definiu o Aprendizado de Máquina em 1959, como o “campo de estudo que dá aos computadores a capacidade de aprender sem serem explicitamente programados”. Em outras palavras, o sistema aprende por meio da experiência, adquirida a partir dos dados de treinamento.

As previsões realizadas por um sistema de aprendizado de máquina são fundamentadas pela generalização, que é a capacidade do modelo de fornecer resultados corretos para um conjunto de dados de entrada até então não vistos por ele (COHEN, 2020).

O princípio da Inteligência Artificial ocorreu na primeira metade do século XX, quando entre os anos de 1943 e 1955 foram publicadas a lógica proposicional de Russel e Whitehead, reconhecida como o primeiro trabalho da área, e a teoria da computação de Alan Turing. O ano de 1956 ficou marcado como o ano do nascimento da IA em decorrência das contribuições de John McCarthy, conhecido como o pai da IA, a criação das redes neurais e diversos outros avanços nos estudos de inteligência artificial. As décadas de 1950 e 1960 destacaram-se, em sua maior parte, por um grande entusiasmo e expectativa de crescimento e visibilidade para a área. No entanto, os anos seguintes mostraram que o esperado não foi concretizado, visto que os resultados iniciais aplicados a problemas simples não foram eficazes para problemas mais complexos. Entre 1969 e 1979, surgiu o primeiro sistema de conhecimento intensivo para atender a necessidade de reter conhecimento humano, baseado na ideia de Dendral sobre ser preciso ter respostas prévias para solucionar problemas difíceis. No mesmo período, ocorreu o surgimento da PROLOG e PLANNER, linguagens de representação de conhecimento. A partir de 1980, a Inteligência Artificial passou por novas revoluções, tornando-se uma

ciência com firme método científico e uma nova indústria, com significativos avanços na robótica, visão computacional e representação do conhecimento, progredindo até os dias atuais (BRAGA, LUDERMIR e CARVALHO, 2000).

#### 2.5.3.1. Tipos de aprendizagem

A aprendizagem de máquina pode ser classificada em alguns tipos. Segundo Zhang (2020), são quatro categorias de aprendizagem, caracterizadas pela natureza dos dados de treinamento. São elas: supervisionada, não supervisionada, semi-supervisionada e por reforço. Na aprendizagem supervisionada, foco desta pesquisa, os dados de treinamento são rotulados e o modelo aprende uma regra geral para mapear entradas e saídas. O supervisor fornece um conjunto de dados de entrada e saída já rotulados e o modelo precisa identificar a relação entre esses dados e suas soluções de modo a aprender como resolver outros problemas similares.

Para a aprendizagem não supervisionada, ainda conforme Zhang (2020), os dados de treinamento não são rotulados, desse modo, a máquina necessita reconhecer por conta própria os padrões existentes. A aprendizagem semi-supervisionada consiste no híbrido das duas anteriores, sendo caracterizada pela utilização de uma pequena parcela de dados rotulados e uma grande parcela de não rotulados. Por último, a aprendizagem por reforço é baseada em um feedback de dados, os quais tem a função de recompensar ou punir o sistema como forma de melhorar o desempenho do modelo acerca do aprendizado em questão.

#### 2.5.3.2. Modelagem preditiva

A análise preditiva é amplamente utilizada como apoio para a tomada de decisão e gerenciamento de riscos. Segundo Ambika (2020), é um tipo de análise de dados empregada para realizar previsões acerca de eventos futuros ou desconhecidos, através da aplicação de técnicas de mineração de dados, estatística, modelagem e aprendizagem de máquina.

A elaboração de uma análise preditiva pode ocorrer em sequência de etapas, seguindo processos como o KDD (Knowledge Discovery in Databases) e o CRISP-DM (Cross Industry Standard Process for Data Mining). De modo geral, estas metodologias

incluem definição do projeto, pré-processamento de dados, modelagem, validação e implantação do modelo.

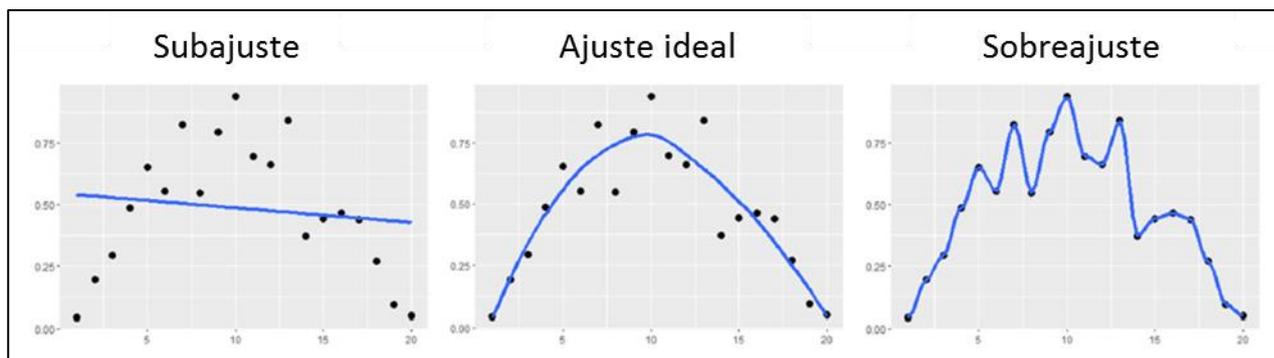
A fase de pré-processamento dos dados apresenta grande importância por possibilitar a modificação dos dados antes da aplicação das técnicas de aprendizagem de máquina, permitindo, muitas vezes, o aumento da eficiência do modelo (BAUMANN; HOCHBAUM; YANG, 2019; HUBER et al, 2019).

Os modelos preditivos de aprendizagem de máquina são desenvolvidos através de algoritmos, dados e métodos estatísticos, diferenciando-se em modelos que solucionam problemas de classificação e problemas de regressão. Zhang (2020) declara que os problemas de classificação se relacionam com a identificação de padrões, de modo que o modelo é um classificador treinado para fazer reconhecimento de amostras desconhecidas, enquanto modelos de regressão, utilizados nesta pesquisa, são treinados para realizar previsões acerca de amostras numéricas. Zhang (2020) declara ainda que os problemas de regressão se relacionam com a identificação de padrões, de modo que o modelo é um classificador treinado para fazer reconhecimento de amostras desconhecidas, enquanto modelos de regressão são treinados para realizar previsões acerca de amostras numéricas.

#### 2.5.3.3. Sobreajuste e subajuste

Conforme dito anteriormente, validade de um modelo é determinada por sua habilidade em generalizar, no entanto, no processo de generalização podem ocorrer dois extremos: o sobreajuste e o subajuste. Ambos são desfavoráveis à confiabilidade dos resultados preditivos. No sobreajuste, o algoritmo aprende extremamente bem e sua alta capacidade de memorização de dados acarreta na definição de predições bastante acertadas para cada amostra da base de treinamento, resultando na criação de novos padrões para acomodar novos dados ao invés de generalizá-lo em uma previsão. O subajuste é o oposto, no qual o modelo por ser muito fraco, é incapaz de fazer distinções entre os dados. Por isso, para que um modelo generalize bem, é necessário encontrar um equilíbrio entre o sobreajuste e o subajuste (COHEN, 2020). A Figura 8 apresenta uma representação para sobreajuste, subajuste e o que pode ser considerado um bom ajuste.

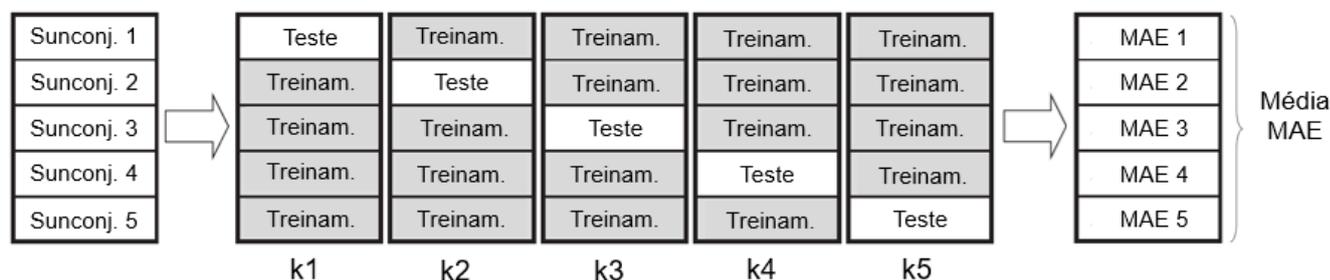
Figura 8 - Subajuste, sobreajuste e ajuste ideal.



Fonte: Adaptado de Mutasa, Sun e Ha (2020).

Para atestar a eficiência de um modelo preditivo, é necessário testá-lo com dados diferentes daqueles utilizados para o ajuste. Essa metodologia de aprendizado é denominada validação cruzada e possui diversas variações, que dentre elas estão as técnicas de treino-teste, *p-leave out* (deixar- $p$  elementos de fora) e *k-fold*. A primeira técnica baseia-se na divisão do conjunto de dados em proporções pré-definidas, como por exemplo, 70% para treinamento e 30% para teste de desempenho (70-30). Esta é uma metodologia simples e facilmente aplicável, mas que possui a desvantagem da perda de uma fração dos dados de treino. A validação cruzada pela técnica Deixar- $p$  Elementos Fora possui funcionamento semelhante. Do conjunto de dados remove-se a quantidade de  $p$  elementos a cada interação, executa o treinamento com os dados restantes e o teste com os removidos, ao final o erro da validação é calculado a partir da média dos erros de todos os casos. Esta possui a vantagem de obter resultados mais robustos, no entanto, seu preço é elevado, em função do tempo e esforço elevado no processamento de dados realizado em cada interação (SCIKIT-LEARN, n.d.c).

A validação cruzada *k-vezes* (*k-fold*) é a técnica mais empregada, pois une benefícios de ambas anteriores. A sua execução ocorre da seguinte forma, o conjunto de dados é dividido em  $k$  subconjuntos de mesmo tamanho, destes, um subconjunto é reservado para o teste e os demais  $k-1$  são aplicados no treinamento do modelo. Posteriormente, repete-se esse procedimento por  $k$  vezes e, por fim, se observa a média do desempenho alcançado em todas as  $k$  rodadas (ZHANG; SUN; WU, 2019). A Figura 9 expõe uma representação da Validação cruzada em *k-vezes*.

Figura 9 - Validação cruzada em  $k$ -folds.

Fonte: Adaptado de Donate *et al* (2013).

#### 2.5.3.4. Algoritmos de Aprendizado de máquina

Os modelos preditivos são desenvolvidos através de algoritmos, dados e métodos estatísticos, diferenciando-se em modelos que solucionam problemas de classificação e regressão, conforme explicado anteriormente.

No geral, há algumas classes que subdividem os algoritmos de regressão de acordo com o princípio existente em sua formulação.

Os algoritmos tipicamente de regressão, por exemplo, modelam a relação entre variáveis ajustando as medidas de erro nas previsões feitas pelo modelo. São exemplos: *Ordinary Least Square Regression (OLSR)*, *Linear Regression*, *Logistic Regression*, *Stepwise Regression*, *Multivariate Adaptive Regression Splines (MARS)*, *Locality Estimated Scatterplot Smoothing*;

Os algoritmos regulatórios penalizam os modelos baseados em sua complexidade. Esses são caracterizados por sua simplicidade e alta capacidade de generalização. São exemplos: *Ridge Regression*, *Least Absolute Shrinkage and Selecion Operator (LASSO)*, *Elastic Net*, *Least-Angle Regression*.

Os algoritmos baseados em Instância, utilizam bancos de dados de exemplo e comparam novos dados com a base já existente, usando medidas de similiaridade para encontrar a melhor correspondência realizar previsões. Ou seja, todo o foco é colocado sobre a representação nas instâncias armazenadas e medidas de similaridade usadas entre as instâncias. São exemplos: *k-Nearest Neighbor*, *Learning Vector Quantization (LVQ)*, *Self-Organizing Map (SOM)* e *Locally Weighted Learning (LWL)*.

Os algoritmos de árvore de decisão constroem modelos com base em valores reais dos atributos dos dados, onde a decisão a ser tomada segue o resultado predito. São

exemplos: *Classification and Regression Tree (CART)*, *Conditional Decision Tree*, *Interactive Dichotomizer (ID3)*, *C4.5 and C5.0 (different versions of a powerful approach)*, *Chi-squared Automatic Interaction Detection (CHAID)*, *Decision Stump* e *M5*.

Os algoritmos baseados em Redes Bayesianas são ferramentas que utilizam o raciocínio probabilista, ou seja, toda sua metodologia é baseada em probabilidades, especialmente a probabilidade condicional. São exemplos: *Naive Bayes*, *Gaussian Naive Bayes*, *Multinomial Naive Bayes*, *Bayesian Network*, *Averaged One-Dependence Estimators (AODE)*, *Bayesian Belief Network (BBN)*.

Os algoritmos de Clusterização (*Clustering*) funcionam com a formação de grupos baseados em similaridades. São exemplos: *K-means*, *K-means ++*, *K-medians*, *Expectation Maximization (EM)*, *Hierarchical Clustering*.

Os algoritmos baseados em Regras de Associação utilizam métodos que extraem regras que melhor explicam relações observadas entre variáveis. São exemplos: *Apriori Algorithm* e *Eclat Algorithm*.

Os algoritmos de redes neurais artificiais possuem estruturas ou funcionamento baseados no funcionamento do cérebro e possuem classes de correspondência de padrões. São exemplos: *Perceptron*, *Multilayer Perceptron*, *Back Propagation*, *Hop Field Network*, *Radial Basis Function Network (RBFN)*.

Os algoritmos de *Deep Learning* funcionam como redes neurais bem maiores que as redes convencionais e com funcionamento mais complexo. São exemplos: *Convolutional Neural Network (CNN)*, *Recurrent Neural Network (RNN)*, *Deep Boltzmann Machine (DBM)*, *Deep Belief Networks (DBN)*, *Stacked Auto-Encoders*, *Generative Adversarial Network*.

Os algoritmos de redução de dimensionalidade atuam para diminuir a grandeza das dimensões de bases de dados. São exemplos: *Principal Component Analysis (PCA)*, *Principal Component Regression*, *Partial Least Squares Regression (PLSR)*, *Multidimensional Scalling (MDS)*, *Linear Discriminant Analysis (LDA)*, *Quadratic Discriminant Analysis (QDA)*, *Flexible Discriminant Analysis (FDA)*.

Os algoritmos Ensemble se baseiam na combinação de diversos modelos de predição mais simples que são treinados para uma mesma tarefa, e são agrupados para a criação de um modelo mais complexo que é a soma de suas partes. São exemplos: *Boosting*, *Bootstrapped Aggregation (Bagging)*, *Adaboost*, *Stacked Generalization (Blending)*, *Gradient Boosting Machines (GBM)*, *Random Forest*, *Gradient Boosted Regression Trees (GBRT)*.

Outros exemplos de algoritmos que não são parte das categorias supracitadas são *Support Vector Machine (SVM)*, *Computer Vision*, *Natural Language Processing*, *Sistema de Recomendação* e *Graphical Models*.

Alguns destes algoritmos serão melhor detalhados a seguir.

- *kNN*:

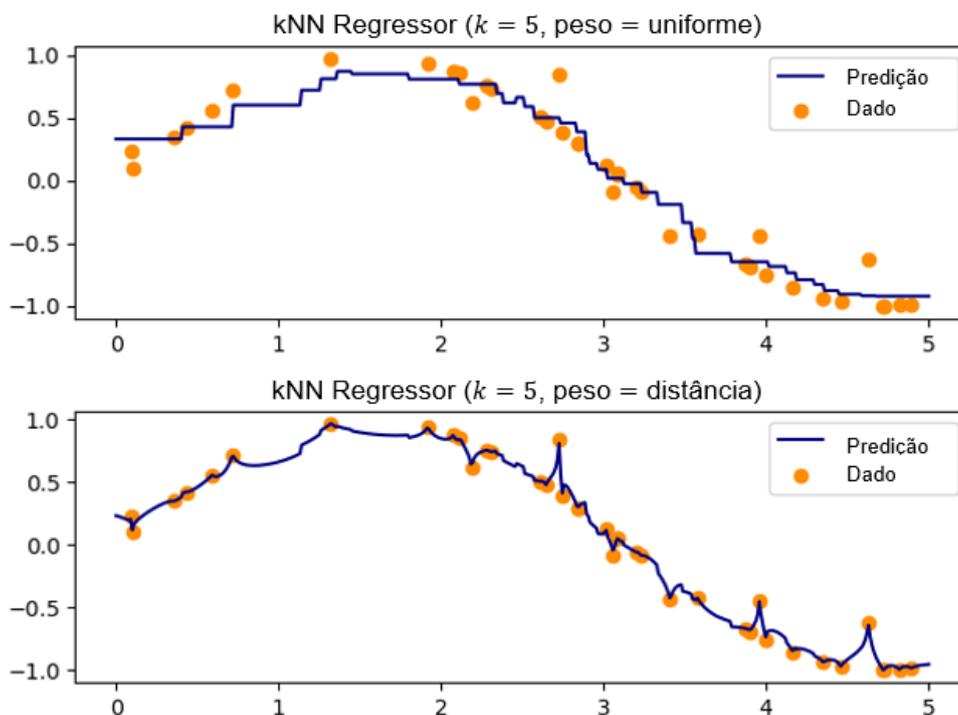
O algoritmo *kNN* (k-Nearest Neighbor), ou *k-Vizinhos Próximos*, é um método de classificação paramétrica baseado em medidas de distância, que surgiu em 1951 a partir de um trabalho técnico não publicado de Evelyn Fix e Ji Hodges. Os dados são representados como vetores  $n$ -dimensionais, sendo  $n$  correspondente ao número de parâmetros. A classificação ou regressão desses dados é obtida através da distância de seus vizinhos, a qual pode ser definida por distância euclidiana, cosseno ou de Chebyshev (ALI et al, 2020; TSALERA; PAPADAKIS; SAMARAKOU, 2020).

Algoritmos da família k-Nearest Neighbor (*kNN*) realizam predições de valores de novos exemplos comparando as características deste mesmo exemplo desconhecido com outros, cujos valores são conhecidos, atribuindo uma predição dos valores preditos.

Ferrero (2009) afirma que a ideia geral do *kNN* consiste em encontrar os  $k$  exemplos que estejam rotulados e mais próximos do exemplo não rotulado e, com base no rótulo prévio desses exemplos mais próximos, se prevê o exemplo não conhecido.

O procedimento de classificação é empregado para dados contínuos. Nos casos em que os dados são discretos deve-se aplicar o algoritmo *kNN Regressor*. O *kNN* para regressão atribui rótulos aos dados por meio da média obtida dos seus  $k$  vizinhos mais próximos, aos quais são atribuídos pesos que podem ser de dois tipos a depender do que é mais vantajoso para a situação. O primeiro tipo de peso é o uniforme, em que o valor de um dado é definido pela maioria de votos dos vizinhos mais próximos, sem diferenciação, enquanto o outro é a distância, ou seja, quanto mais próximo o vizinho, maior a sua contribuição (SCIKIT-LEARN, n.d.a).

Figura 10 - Representação de uma aplicação para o algoritmo *kNN*.



Fonte: Adaptado de Scikit-Learn (n.d.a.).

O desempenho desse algoritmo é dependente da representatividade e extensão dos dados. Além disso, o *kNN* clássico pode apresentar problemas como, sensibilidade a ruídos, soluções instáveis e dependência ao hiperparâmetro  $k$ , isto é, aos  $k$  vizinhos mais próximos, obtendo desempenhos diferentes para  $k$  variados (Hu et al, 2020; LIN; LIN; CAO, 2021).

Algoritmos do tipo *kNN* requerem poucos esforços durante o seu treinamento, entretanto o custo computacional para rotular um novo exemplo pode se tornar relativamente alto, tendo em vista que no pior dos casos, esse exemplo deverá ser comparado com todos os exemplos contidos no conjunto de dados de treinamento.

- Árvores de decisão:

Árvores de decisão são baseadas na criação de um modelo para previsão de uma variável dependente a partir da entrada de um conjunto de dados hierarquizados. O CART, mais conhecido algoritmo de árvore de decisão, é responsável pela criação de árvores de classificação ou regressão, as quais diferenciam-se entre si pelo tipo de

variável de destino. (BAUMANN; HOCHBAUM; YANG, 2019; KOULINAS; PARASCHOS; KOULOURIOTIS, 2020; MIENYE; SUN; WANG, 2019).

Patel *et al* (2012), define árvores de decisão como sendo um fluxograma que segue a estrutura de uma árvore, onde cada nó denota um teste em um determinado atributo, e cada ramo, por sua vez, denota um resultado desses testes. Cada nó da folha é etiquetado com valores dos atributos.

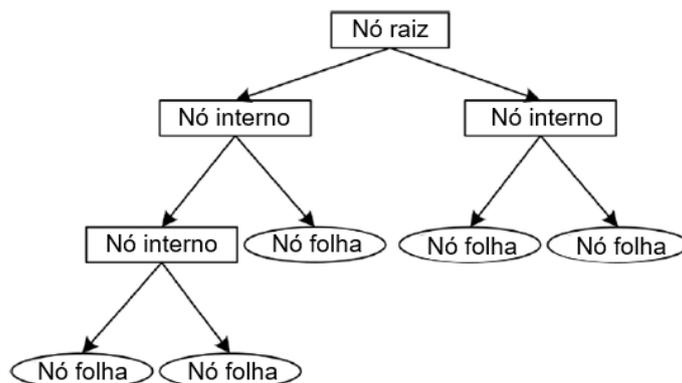
As etapas de aprendizagem e classificação por meio da indução da árvore de decisão são simples e rápidas. A representação desta indução em forma de árvores é de fácil assimilação dos usuários, e em geral, os classificadores e regressores baseados em árvores de decisão têm boa precisão.

De acordo com Zhou *et al* (2021), algoritmos de árvore de decisão apresentam a vantagem de alta precisão de classificação, entretanto, conforme Antoniadis *et al* (2021), são instáveis e pequenas variações nos dados de treinamento são capazes de acarretar grandes mudanças na previsão e estrutura da árvore.

Algoritmos de árvores de decisão constroem uma estrutura constituída de nós e ramos, responsável por refletir informações. Cada nó representa um teste aplicado a um ou mais atributos, os resultados destes testes são expressos nos ramos. Ocorrem também os nós terminais, chamados de folhas, decorrentes da rotulagem dos dados. Uma característica importante dessas estruturas é que os ramos que partem de um nó são mutuamente exclusivos e configuram uma probabilidade (CZAJKOWSKI; KRETOWSKI, 2019).

Meng *et al* (2020) apresenta um esquema genérico da árvore de decisão, em que é possível observar que o início advém de um único nó, denominado de raiz, que se ramifica em nós internos e folhas. A Figura 11 apresenta uma representação dos algoritmos de árvores de decisão.

Figura 11 - Esquema para algoritmo de árvores de decisão.



Fonte: Adaptado de Meng (2020).

Desde o nó raiz, cada árvore vai criando ramificações produzindo uma espécie de regras de divisão das variáveis a partir do chamado grau de pureza. Geralmente, se utiliza como medida de pureza a entropia e o ganho de informação onde o critério de interrupção para que uma árvore cresça ou não é a ocorrência de pureza em um nó, ou quando o número de pontos de uma região está abaixo de um limite predeterminado para a divisão do nó.

- Floresta aleatória (*Random Forest*):

O algoritmo Floresta Aleatória foi criado por Breiman (2001) e consiste na combinação de árvores de decisão geradas por um mesmo conjunto de dados, isto é, são árvores de classificação ou regressão construídas com treinamento a partir de subconjuntos de dados selecionados aleatoriamente do todo. Desse modo, as árvores apresentam uma previsão para cada observação, configurando-se como um modelo de maior precisão (BREIMAN, 2001; SPEISER *et al*, 2019).

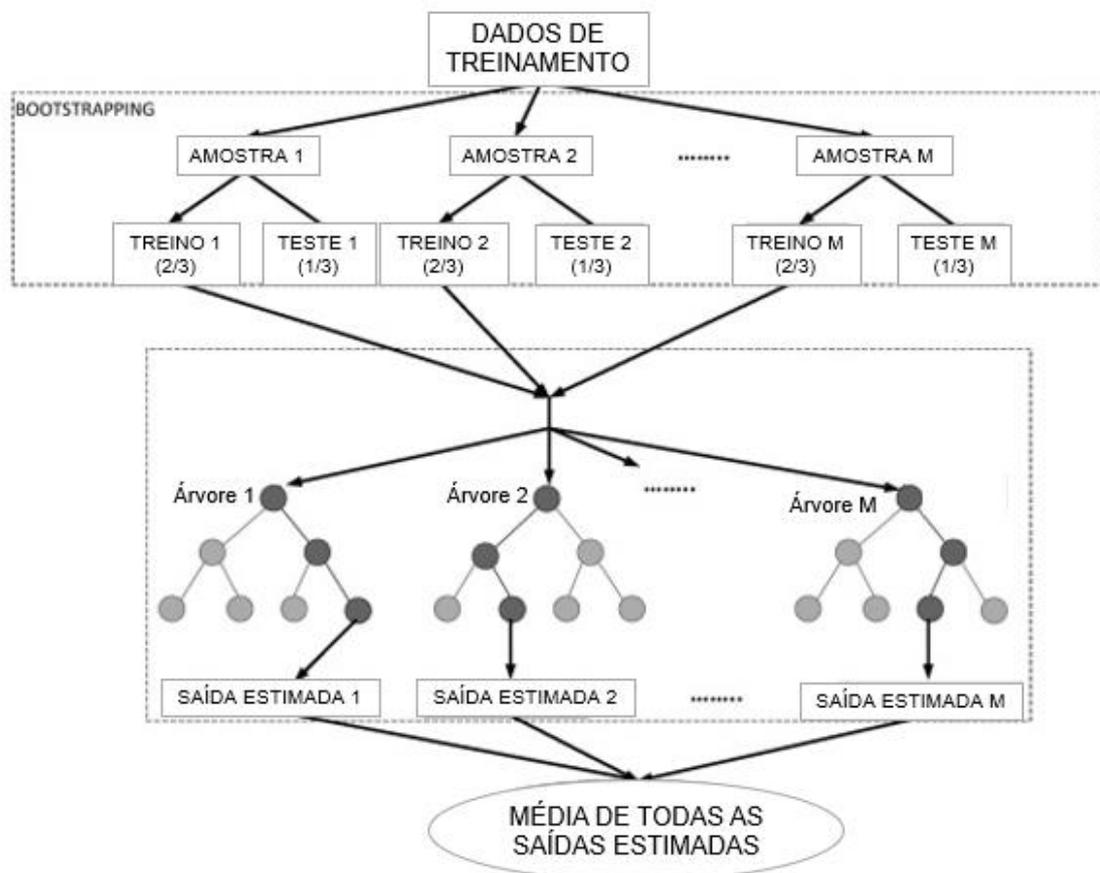
Em conformidade com Abellán *et al* (2018), o Floresta Aleatória possui um classificador base conhecido por Árvore Aleatória (*Random Tree*), de modo que o algoritmo constrói uma floresta a partir desse classificador. Considerando  $M$  o número de parâmetros do conjunto de dados, determina-se  $m$ , um número muito menor que  $M$ , o qual será fixado como constante para a elaboração da floresta e empregado para selecionar de forma aleatória os recursos em cada nó das árvores. A construção de uma Árvore Aleatória ocorre da seguinte forma: seja  $N$  o número de instâncias do conjunto de dados, o Floresta Aleatória escolhe aleatoriamente uma amostra com substituição de  $N$

dos dados originais, constituindo um subconjunto que será utilizado para o treinamento da árvore. Em cada nó dessa árvore serão selecionados  $m$  recursos do  $M$  original e, logo após, é aplicado o critério de divisão sobre  $m$  e o recurso que apresentar melhor valor será utilizado para dividir o nó. Por fim, cada árvore de decisão irá veicular um valor de classificação e o algoritmo irá tomar o valor mais votado da classe.

A seguir um esquema simplificado de regressão no algoritmo Floresta Aleatória. A partir do conjunto de dados de treinamento são obtidas as amostras de *bootstrap* que são subdivididas aleatoriamente e constituem conjuntos de treino e teste com  $2/3$  e  $1/3$  dos dados de cada amostra, respectivamente. Dessa forma, os dados de treinamento geram uma árvore de regressão por vez e sua precisão é testada através dos dados de teste da mesma amostra. Cada árvore criada fornece um resultado estimado e a predição final é calculada através da média da saída de todas as árvores (ISKENDEROĞLU *et al*, 2020).

A Figura 12 expõe uma representação do algoritmo Floresta Aleatória.

Figura 12 - Representação do algoritmo Floresta Aleatória.



- Rede Neural Perceptron de Multicamadas (*Multi-Layer Perceptron – MLP*):

Saracoglu (2008) apresenta as redes neurais artificiais como algoritmos que têm elementos de processamento semelhantes a neurônios biológicos, chamados de nós ou neurônios artificiais e conectados uns aos outros por ponderação. Os pesos para essas ponderações em cada conexão são ajustados dinamicamente até que a saída desejada seja gerada para uma determinada entrada.

Petermann (2006), indica que as funções básicas dos neurônios são 3:

- Avaliar valores de entrada;
- Calcular o total para valores de entrada combinados;
- Comparar o total com um valor limiliar;
- Determinar o que será a saída.

A Rede Neural MLP (Multi-Layer Perceptron) é um algoritmo aplicado ao aprendizado de máquina para classes não linearmente separáveis e caracteriza-se por ser uma generalização da Rede Perceptron, destinada para problemas de discriminações lineares.

Boscarioli *et al* (2008), define a arquitetura de redes neurais *MLPs* consistem tipicamente de uma especificação do número de camadas, tipo de função de ativação de cada unidade e pesos de conexões entre as diferentes unidades que devem ser definidas para a construção desta arquitetura neural. As redes construídas se dão por múltiplas entradas e as saídas, onde as saídas dependem das funções de transferência dos pesos sinápticos que serão ajustados pelo algoritmo de aprendizado, que ocorre com o uso de um conjunto de padrões, conjunto de pares de entrada e saída (X;y) que definem, através de exemplos, o sistema a modelar.

Segundo Feng *et al* (2020), a MLP possui três tipos de camadas. As camadas de entrada, as ocultas e as de saída. A atribuição dessa rede é aprender uma função de ativação linear a partir de um conjunto de entradas e saídas para solucionar problemas de classificação ou regressão.

Ramezani *et al* (2019) descrevem a estrutura de uma Rede Neural MLP da seguinte forma: há uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída, cada uma dessas camadas possui um ou mais neurônios. Na camada de entrada a quantidade de neurônios é igual a de variáveis independentes e cada um

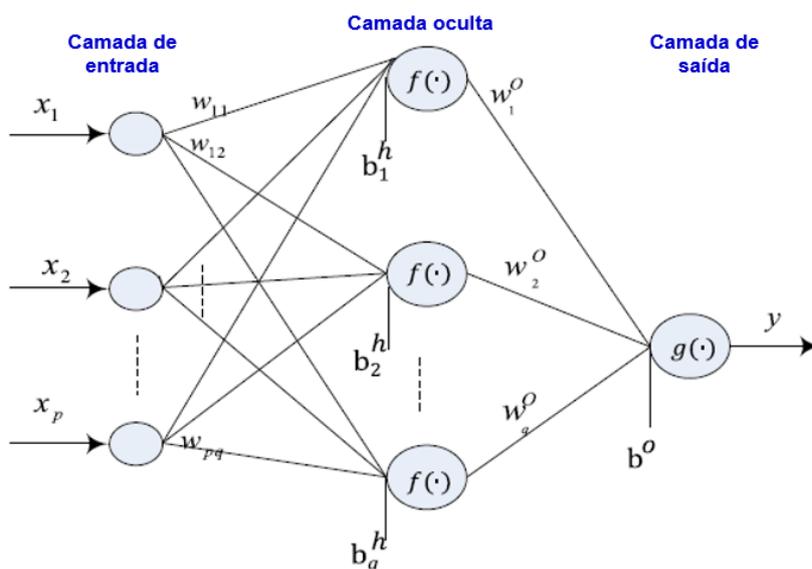
deles conecta-se a um ou vários neurônios na camada oculta e esta, por sua vez, conecta-se a camada de saída.

Os dados de entrada são convertidos em saída através da função de ativação e a quantidade de neurônios na camada de saída será a mesma de variáveis independentes. As camadas ocultas, no entanto, possuem formação dependente dos parâmetros de execução do modelo, como tipo, tamanho e velocidade do programa.

Zhang, Sun e Wu (2019) ilustram uma rede neural MPL de três camadas como a seguir, onde  $W_I = [w_1^I, w_2^I, \dots, w_q^I]$  e  $W_O = [w_1^O, w_2^O, \dots, w_q^O]^T$  são, respectivamente, as matrizes de pesos de entrada e saída, enquanto  $B^h = [b_1^h, b_2^h, \dots, b_q^h]^T$  é o vetor de polarização da camada oculta,  $b^o$  é a polarização da camada de saída, e as funções  $f$  e  $g$  correspondem às funções de ativação das camadas oculta e de saída, nesta ordem.

A Figura 13 expõe uma representação de uma rede neural *Perceptron* de multi-camadas.

Figura 13 - Representação de uma Rede Neural MLP.



Fonte: Adaptado de Zhang, Sun e Wu (2019).

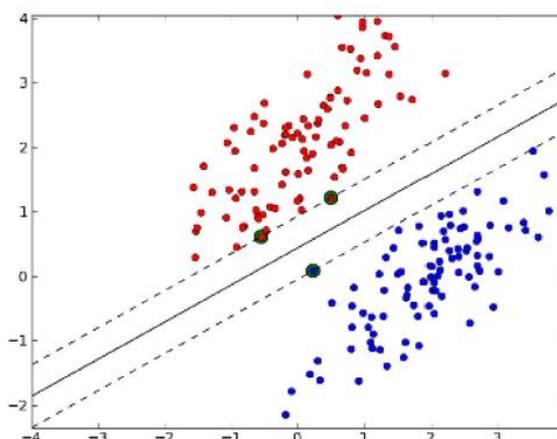
- Máquinas de Vetores de Suporte (SVM):

O algoritmo SVM (Support Vector Machines) foi proposto por Cortes e Vapnik na década de 1990, inicialmente para aplicação em problemas de classificação binária, sendo posteriormente estendido para estimativas de regressão, cuja extensão é o SVR (Support Vector Regression). A base do algoritmo está firmada no princípio de

minimização de riscos estruturais e sua ideia principal é determinar o melhor hiperplano que separa duas classes de padrões. O SVM é conhecido como um dos classificadores mais poderosos, no entanto, pode apresentar uma alta complexidade de treinamento conforme o número de padrões, podendo ser limitado para grandes conjuntos de dados. Enquanto que o SVR possui a desvantagem de apresentar sensibilidade a ruídos e outliers (KIM; KANG; CHO, 2020; SABZEKAR; HASHEMINEJAD, 2021).

Conforme Hussain (2019), a determinação do hiperplano ocorre por meio da maximização da margem entre os pontos mais próximos da fronteira, os quais são conhecidos por vetores de suporte, são eles que possibilitarão encontrar o hiperplano de separação ideal. O algoritmo poderá ser aplicado para dados separáveis linearmente e não linearmente, bem como para problemas de duas ou diversas classes. O autor exemplifica o hiperplano de separação ideal conforme a Figura 14.

Figura 14 - Representação do hiperplano de algoritmo SVR.



Fonte: Hussain (2019)

Para os problemas de regressão, ao invés de ser alcançado um hiperplano de separação de classes, é definida uma curva, superfície ou hiperplano de regressão para a solução do problema proposto.

- Regressão Ridge:

Uma forma de otimização do aprendizado de máquina é a regularização, cujo objetivo é controlar os problemas de sobre-ajuste. Esta técnica foi sugerida primeiramente pelo matemático russo Tikhonov, de forma que um método de regularização muito conhecido é denominado de Regularização de Tikhonov e consiste

na substituição do problema de minimização (I) pelo problema de mínimos quadrados (II), conforme mostrado a seguir:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2, A \in \mathbb{R}^{m \times n}, m \geq n \quad (\text{I})$$

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2 + \|L_\lambda x\|^2 \quad (\text{II})$$

Onde,  $L_\lambda$  é matriz de regularização e  $\lambda$  é parâmetro de regularização (MOHAMMADY; ESLAHCHI, 2020).

O algoritmo Regression Ridge, também chamado de Regularização do Parâmetro  $L_2$  é o caso particular de regulador quadrático, consistindo em adicionar um termo de regularização à função objetivo. Sua aplicação possui o intuito de distinguir as entradas com maior variância, em vista disso, quando a covariância entre um recurso e o alvo é irrelevante se comparada a variância adicionada. O peso daquele será reduzido na etapa de treinamento do modelo (FATHI; SHOJA, 2018).

- Aprendizado profundo (*Deep Learning*)

O algoritmo Aprendizado Profundo é baseado em uma rede neural artificial de alimentação em várias camadas que é treinada utilizando aprendizado de propagação reversa (*back-propagation*).

LeCun *et al* (2015) afirmam que o uso de Aprendizado Profundo permite modelos computacionais compostos de múltiplas camadas de processamento aprenderem representações de dados com múltiplos níveis de abstração.

LeCun *et al* (2015) também declara que Aprendizado Profundo utiliza *backpropagation* para indicar como uma máquina deve alterar seus parâmetros internos que são usados para calcular a representação em cada uma das camadas anteriores.

A rede pode conter um grande número de camadas ocultas que permitem processar milhares de atributos. Cada nó de cálculo treina uma cópia dos parâmetros do modelo global em seus dados locais e contribui periodicamente para o modelo global por meio da média do modelo na rede.

#### 2.5.3.5. Hiperparametrização de algoritmos

No aprendizado de máquina existem os conceitos de parâmetro e hiperparâmetro de algoritmo. O primeiro deles diz respeito aqueles que podem ser ajustados durante o

processo de aprendizagem e execução do modelo, como por exemplo o peso de neurônios em redes neurais. O segundo refere-se aos parâmetros que devem ser estimados anteriormente ao treinamento, isto é, as configurações do algoritmo, como, taxa de aprendizado e função de ativação (YANG; SHAMI, 2020).

O desempenho dos modelos preditivos é influenciado pelos métodos de otimização empregados na aprendizagem. A otimização dos hiperparâmetros, também conhecida por parametrização de algoritmo, será responsável por aperfeiçoar a estrutura e precisão do modelo de previsão (DIAZ et al, 2017; SUN *et al*, 2020).

Há uma grande variedade de métodos de parametrização de algoritmos como a Busca Exaustiva em Grids (*Grid Search*), Busca Aleatória de Parâmetros, Otimização Bayesiana, Redução Sucessiva, *Hyperband*, Algoritmos Metaheurísticos e outros. Entre os métodos supracitados, o grid search e a busca aleatória serão especificados.

- Busca exaustiva em grids (*Grid Search*):

O método de grid search ou busca exaustiva em grids é amplamente conhecido e consiste na determinação do melhor hiperparâmetro por meio da avaliação manual, intervalo a intervalo de grade, testando todas as combinações possíveis no espaço de pesquisa pré-determinado. Essa técnica possui a vantagem da fácil implementação, porém, demanda muito tempo e perde eficiência à medida que a dimensão do espaço de pesquisa aumenta, pois o número de pontos de grade cresce em ritmo exponencial, de modo que torna inviável sua aplicação para casos de alta dimensionalidade (YANG; SHAMI, 2020; YOO, 2019).

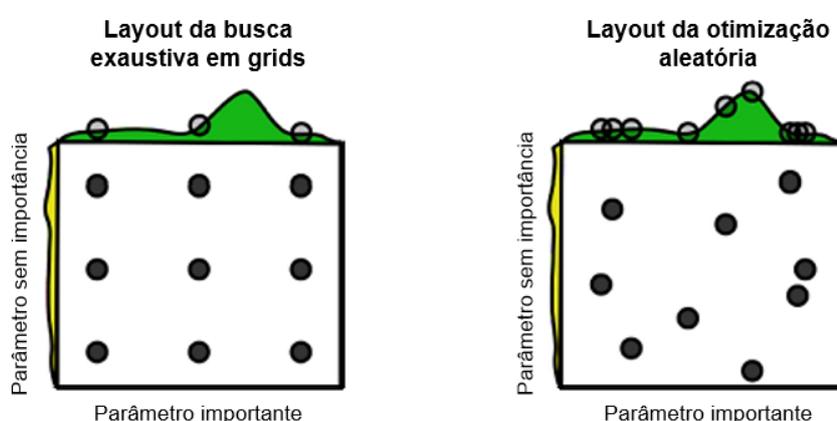
- Otimização aleatória de parâmetros:

Diante às limitações da busca exaustiva em grids foi desenvolvido, por Bergstra e Bengio (2012), o método de otimização aleatória, ou busca aleatória, semelhante ao anterior, mantendo a simplicidade de aplicação, porém demonstrando viabilidade de implantação para problemas de alta dimensionalidade, pois não necessita testar todos os pontos de grade. Consiste em selecionar aleatoriamente um número de amostras e testá-las como candidatos à hiperparâmetros, a distribuição dos testes é especificada conforme os resultados apresentados, de modo a direcioná-los para regiões melhores no espaço

de busca, reduzindo o tempo computacional. Apesar de encontrar bons resultados, a otimização aleatória não assegura o melhor conjunto de hiperparâmetros.

Bergstra e Bengio (2012) apresentam a comparação a seguir entre os métodos de parametrização de busca exaustiva em grids e de otimização aleatória. Ambas as imagens incluem a otimização em nove tentativas de uma função  $f(x, y) = g(x) + h(x) \approx g(x)$  de baixa dimensionalidade. A área verde corresponde a  $g(x)$  e a amarela a  $h(x)$ , ambas as curvas remetem a precisão do modelo, visto que no caso da pesquisa aleatória os nove testes analisam valores em  $g$ , não acontecendo o mesmo para a busca exaustiva.

Figura 15 - Comparação entre Busca Exaustiva em Grids e Otimização Aleatória.



Fonte: Adaptado de Bergstra e Bengio (2012).

### 2.5.3.6. Métricas de desempenho

As métricas de desempenho são utilizadas com o objetivo de avaliar os algoritmos de aprendizado de máquina, para os algoritmos de regressão são comuns as métricas de erro médio absoluto, erro médio quadrático, erro mediano absoluto e coeficiente de determinação, as quais são apresentadas a seguir (AGWU; AKPABIO; DOSUNMU, 2021).

O erro médio absoluto (*MAE*) é expresso pela função:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - X_i| \quad (8)$$

Onde,  $n$  é o número de dados e,  $Y$  e  $X$  são, respectivamente, os valores previstos e reais. Ele representa a razão entre o somatório de todos os erros e o número de pontos, cujos erros correspondem a distância de cada ponto à regressão. Para que um modelo preditivo seja perfeito, o *MAE* deve ser zero.

O erro médio quadrático (*MSE*) tem como base o mesmo princípio do erro médio absoluto, no entanto, leva em consideração o quadrado da distância e quanto menor seu valor, melhor será a previsão realizada pelo modelo.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - X_i)^2 \quad (9)$$

Uma variante do MSE é a chamada Raiz Quadrada do Erro Médio Quadrático (RMSE), que retorna um valor na escala da variável a ser predita.

O erro mediano absoluto corresponde a mediana dos erros obtidos na regressão, isto é, ao valor central da sequência crescente dos erros, e irá demonstrar a sua tendência geral.

Enquanto as três métricas anteriores apontam o quão distantes os dados estão da regressão, o coeficiente de determinação ( $R^2$ ) irá demonstrar o grau de correlação entre as variáveis. Este é calculado através de:

$$R^2 = \left( \frac{\sum_{i=1}^n (A_i - \bar{A}_i)(M_i - \bar{M}_i)}{\sqrt{\sum_{i=1}^n (A_i - \bar{A}_i)^2 \sum_{i=1}^n (M_i - \bar{M}_i)^2}} \right)^2 \quad (10)$$

$M_i$  e  $A_i$  são os valores previsto e real, e  $\bar{M}_i$  e  $\bar{A}_i$  são suas médias, respectivamente. Valores de  $R^2$  no intervalo de 0,7 a 0,9 descrevem uma alta correlação, enquanto entre 0,0 e 0,3 a correlação é irrelevante.

#### 2.5.3.7. Seleção de variáveis por análise de importância (*Feature Selection*)

Para a construção de modelos de aprendizado de máquina é comum a utilização de amostras de alta dimensão, as quais podem incluir dados irrelevantes e/ou redundantes. Neste cenário existem métodos de redução de dimensionalidade, como o Extração de Recursos (*Feature Extraction*) e o Seleção de Recursos (*Feature Selection*). O primeiro método consiste na extração de recursos concisos do conjunto original, enquanto o segundo avalia as combinações de recursos disponíveis e seleciona o melhor subconjunto de dados (AGHAEIPOOR; JAVIDI, 2020).

De acordo com Effrosynidis e Arampatzis (2021), o método Seleção de Recursos pode ser tomado como um pré-processamento dos dados.

O método também pode ser compreendido com parte da fase de modelagem dos dados, já que seleciona as variáveis que comporão o modelo preditivo. Ele não cria novos recursos e proporciona diversos benefícios como aumento no nível de interpretação,

eliminação de recursos desnecessários e ruídos, redundância mínima, redução do tempo de cálculo e da complexidade do treinamento, bem como melhorias no desempenho dos algoritmos e menor possibilidade de sobreajuste.

Conforme os mesmos autores, a seleção de recursos possui três categorias, são elas o *filter method*, *wrapper method* e *embedded method*. O *filter method* atua como um classificador, ordenando os recursos de acordo com as propriedades intrínsecas dos dados (variância, distância, correlação, etc), mantendo os melhores e removendo os piores. O *wrapper method* utiliza um algoritmo para encontrar os melhores subconjuntos a partir de uma combinação de estratégia de busca. Por fim, o *embedded method* utiliza critérios estatísticos como filtro para escolher recursos e, logo após, aplica um algoritmo de aprendizado de máquina para escolher o subconjunto de melhor desempenho.

Há uma grande variedade de métodos de seleção pertencentes às categorias supracitadas. Em especial, para o *filter method*, há os testes estatísticos *f\_regression* e *mutual\_info\_regression*. O *f\_regression* é uma função de pontuação empregada para testar de forma individual o efeito dos regressores. Ele calcula a correlação entre o regressor e o alvo e, posteriormente, a converte em uma pontuação *F*. Enquanto o *f\_regression* estima apenas a dependência linear entre duas variáveis, enquanto o *mutual\_info\_regression* pode estimar quaisquer tipos de dependências estatísticas (SCIKIT-LEARN, n.d.b).

## 2.6. MODELAGEM PREDITIVA PARA PLANEJAMENTO OPERACIONAL

Temkin *et al.* (2020) apontam a importância da análise preditiva na mineração, especificamente, nas operações de carregamento e transporte, em que o desempenho é dependente do monitoramento das condições dos caminhões, distribuição dos equipamentos de carga e pontos de descarga, entre outros parâmetros. Nesse contexto, a eficiência da mineração pode ser melhorada pela operação ideal dos equipamentos e otimização do tráfego de carga, tarefas que podem ser realizadas por sistemas de gerenciamento de frotas.

Baek e Choi (2019) utilizaram análise preditiva para simular operações de transporte com caminhões em uma mina subterrânea de calcário. A metodologia empregada no estudo consistiu, essencialmente, em duas etapas. Em primeiro lugar, foi realizada a medição do tempo de viagem dos caminhões a partir dados armazenados em

universo *big data*, adquiridos por meio do sistema de gerenciamento de segurança da mina. Posteriormente, foi empregado um algoritmo de simulação de eventos discretos, onde por meio deste foi simulado o sistema de transporte por caminhões. Por fim, nas análises estatísticas, foi observado um coeficiente de determinação de 0,99 entre os resultados dos tempos médios de viagem do caminhão obtidos por simulação e os resultados reais, apontado que a simulação desenvolvida poderia prever a carga de minério do caminhão, a produção total de minério e os valores de tempos de espera dos caminhões.

Somua-Gyimah *et al.* (2019) desenvolveram um modelo de visão computacional para escavadeiras *draglines*, objetivando prever colisões e manuseio de rocha em operações de escavação na mineração e construção civil, de modo a fornecer total autonomia aos sistemas automatizados. O modelo foi desenvolvido através de aprendizado de máquina, empregando redes neurais convolucionais para reconhecer objetos diferentes e determinar sua localização. Nos testes aplicados, o modelo apresentou um bom funcionamento, no entanto, necessita de melhorias para atender os requisitos para implantação em operações. Além disso, também apresenta possibilidade de ser modificado para aplicação com outros tipos de escavadeiras automatizadas.

## 2.7. VISÃO GERAL DA MINA DO SOSSEGO

A Mina do Sossego é um empreendimento mineral de cobre que está situado na região sudeste do estado do Pará, a uma distância de 20 km da sede do município de Canaã dos Carajás (PA) e inserido na Província Mineral de Carajás, conforme apresentado na Figura 16. Os dados utilizados para exemplos de aplicação desta pesquisa são oriundos da operação de lavra deste empreendimento mineral.



Lacerda (2019) apresenta um sumário descritivo para a frota de equipamentos de lavra da mina do Sossego, conforme abaixo na Tabela 1.

Tabela 1 – Sumário de frotas da Mina do Sossego.

<b>Equipamento</b>	<b>Tipo</b>	<b>Frota</b>	<b>Unidades</b>
<b>Transporte</b>	Caminhões Fora de Estrada	CAT 793C/D	28
		CAT 785C	12
	Caminhões Articulados	CAT 740A	8
	Caminhões Rodoviários	Mercedes 4844K	13
<b>Escavação</b>	Escavadeiras Elétricas	PH-4100 XPB	1
		Buc 495-HR	2
		PH-2300 XPA	1
	Escavadeiras Hidráulicas	PC5500	1
		PC2000	1
<b>Carregamento</b>	Carregadeiras	WA1200	2

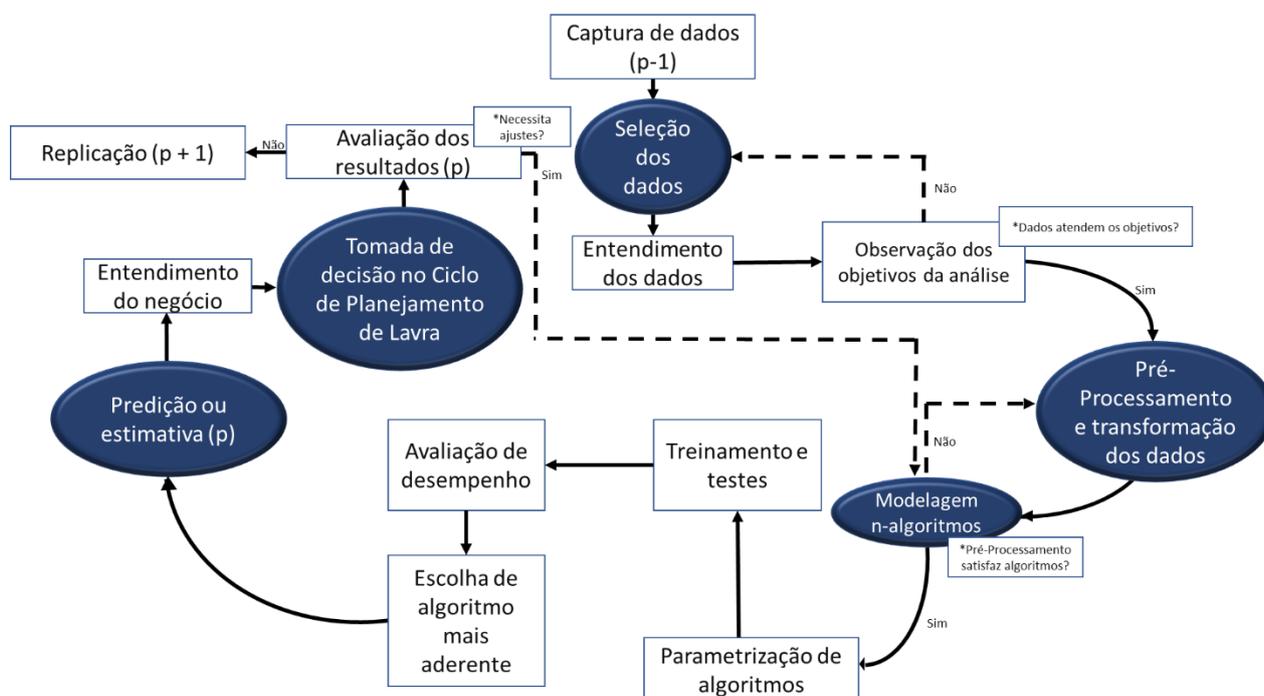
Fonte: Adaptado de Lacerda (2019).

### 3. MATERIAIS E MÉTODOS

O ciclo de planejamento de lavra exige de seus atores tomadas de decisão que sejam ágeis, coerentes, consistentes e dinâmicas. Isto se dá pelo fato de que as decisões precisam ser instantâneas, tendo em vista que idealmente o ciclo produtivo de mina não pode ser interrompido. Assim, baseado na metodologia CRISP-DM, este trabalho apresenta um ciclo ciência de dados compatível com as necessidades da indústria mineral.

Sendo assim, neste capítulo são apresentados os procedimentos metodológicos da pesquisa e as etapas de seu desenvolvimento. O presente estudo se dividiu nas etapas descritas na Figura 17.

Figura 17 - Processo padrão de ciência de dados para Planejamento de Lavra.



Fonte: Elaborado pelo autor.

De forma geral, a metodologia considera que no período que se interessa realizar a predição ( $p = \text{período}$ ), que nesta pesquisa é o turno operacional, o usuário do sistema possa, a partir dos dados históricos registrados até o período anterior ( $p - 1$ ), aplicar uma rotina de análise eficiente composta por etapas de preparação e transformação dos dados, modelagem preditiva, fornecimento e análise de resultados, até a replicação do método para os próximos turnos operacionais ( $p + 1$ ).

A metodologia proposta para esse trabalho, denominada de Processo Padrão de Ciência de Dados para Planejamento de Lavra, visa apoiar a tomada de decisão da equipe de planejadores de mina, satisfazendo a necessidade de conquista de conhecimento acerca das operações por meio de uma rotina prática e eficiente.

### 3.1. SELEÇÃO DOS DADOS

Os dados coletados são oriundos do banco de dados do sistema de despacho eletrônico da Mina do Sossego. Esse sistema registra, monitora e controla toda a atividade operacional dos equipamentos de mina. A captura é feita por um pacote de ativos espalhados na mina, que recolhem informações e permitem a realização de controle, monitoramento e gerenciamento dos equipamentos.

Nesta etapa, se buscou selecionar dados históricos, anteriores ao período  $p$  que se desejou realizar previsões. Ou seja, para as previsões realizadas no horizonte diário de planejamento, os dados utilizados para as análises estarão sempre compreendidos do dia anterior ( $p - 1$ ) até os registros mais antigos.

Assim, a partir desta robusta base de dados, alguns dados específicos foram selecionados para compor o conjunto que foi analisado.

Todos os dados capturados foram estruturados em planilhas, representadas como matrizes do tipo  $X_{n \times m}$ , onde  $n$  é a indicação do número de objetos observados dispostos nas linhas das matrizes, e  $m$  significa o número dos atributos (variáveis) que caracterizam cada objeto, dispostos nas colunas

O número de objetos observados é a representação do número de eventos coletados pelo sistema de despacho. Estes eventos também podem ser chamados de exemplos. O valor de cada variável é denominado como registro.

Os dados escolhidos para análise são referentes a 15 anos de operação de uma frota de 4 equipamentos de carga, doravante chamadas de escavadeiras, incluindo variáveis que descrevem as condições operacionais em que estes equipamentos trabalharam, bem como a política de horas operacionais associadas a cada equipamento, equipes de trabalho que os operou e frotas de equipamentos de transporte que demandaram a sua produção.

Os equipamentos estão denominados pelos códigos ES3001, ES3101, ES3102 e PM3401.

A base de dados resultante contém um total de 35 atributos, 69557 exemplos e 2 434 495 registros. Os atributos contidos na base de dados original são os descritos abaixo:

- ANO: Ano de realização do evento;
- MÊS: Mês de realização do evento (expressa numericamente);
- MÊS\_: Mês de realização do evento (expressa nominalmente);
- DIA: Dia de realização do evento (expressa numericamente);
- TURNO: Turno de realização do evento (expressa numericamente);
- TURNO\_: Turno de realização do evento (expressa nominalmente);
- TURMA: Turma de operadores que executou a operação (expressa numericamente);
- TURMA\_: Turma de operadores que executou a operação (expressa nominalmente);
- EQUIPAMENTO: Escavadeira utilizada no evento registrado;
- TIPO: Modelo de cada escavadeira utilizada no evento registrado;
- MASSA: Quantidade em toneladas, produzida por cada escavadeira em toneladas, por turno;
- LOCAL DE CARREGAMENTO: Cota da bancada em que cada escavadeira foi alocada no turno (metros);
- QUALIDADE: Tipo de litologia do material que foi manuseado pela escavadeira no turno;
- HORA TOTAL: Total de horas programadas para o turno, dado em segundos;
- HORA TRABALHANDO: Horas de utilização física no turno, desconsiderando atrasos operacionais, dado em segundos;
- TEMPO DE CARREGAMENTO: Tempo de utilização física no turno destinada apenas para carregamento de material, dado em segundos;
- TEMPO DE CARREGAMENTO (HORAS): Tempo de utilização física no turno destinada apenas para carregamento de material, dado em horas;
- ATRASO OPERACIONAL: Quantidade de horas em que as escavadeiras estiveram momentaneamente paradas por causas naturais ao ciclo

(aguardando caminhões, aguardando forro de piso, etc), excluindo-se as paradas de natureza externa, dado em segundos;

- HORA TRABALHADA: Horas de utilização física no turno, considerando também horas de atrasos operacionais, dado em horas;
- PARADA EXTERNA: Paradas provocadas por razões externas, dado em segundos;
- PARADA INTERNA: Paradas provocadas por razões internas, dado em segundos;
- EQUIPAMENTO LIBERADO: Horas pós manutenções em que o equipamento foi disponibilizado para uso mas não foi utilizado (dado em segundos);
- MANUTENÇÃO PREVENTIVA: Horas no turno gastas com manutenção preventiva, dado em segundos;
- MANUTENÇÃO CORRETIVA: Horas no turno gastas com manutenção corretiva, dado em segundos;
- PARADA DETERMINADA: Quantidade de horas programadas para paradas por razões diversas para ajustes de processos, dado em segundos;
- PARADA TROCA DE TURNO: Quantidade de horas destinadas para trocas de turno e turmas, dado em segundos;
- MANUTENÇÃO CORRETIVA PROG: Quantidade de horas programadas para manutenção corretiva no turno, dado em segundos;
- MANUTENÇÃO PREVENTIVA JO: Quantidade de horas programadas para manutenção preventiva no turno, dado em segundos;
- OCIOSIDADE: Quantidade de horas aplicadas para quando as escavadeiras estavam disponíveis, mas não foram utilizadas, dado em segundos;
- OCIOSIDADE (HORAS): Quantidade de horas aplicadas para quando as escavadeiras estavam disponíveis, mas não foram utilizadas, dado em horas;
- FROTA DE TRANSPORTE: Frota de transporte que atendeu a escavadeira no turno (expressa numericamente);
- FROTA DE TRANSPORTE\_: Frota de transporte que atendeu a escavadeira no turno (expressa nominalmente);

- PRODUTIVIDADE: Quantidade mássica produzida por hora em que as escavadeiras estiveram trabalhando, dado em toneladas por hora (t/h);
- TAXA DE PRODUÇÃO: Produtividade desconsiderando as horas de atraso operacional, dado em toneladas por hora (t/h);
- DIFERENÇA: Diferença entre a Hora Trabalhada e o Tempo de Carregamento.

Assim, considerando a grande quantidade de dados contidos na base, após a etapa de captura e seleção dos dados, foi realizado um trabalho amostral onde se buscou reduzir, de forma considerável, o tamanho total da base de dados a ser analisada. Assim, a amostragem estatística aplicada preservou a representatividade da amostra em relação ao banco de dados original, reduzindo o tempo de processamento de informações na etapa de modelagem.

O trabalho de amostragem foi realizado em duas etapas. A primeira consistiu na discretização da variável numérica MASSA (variável contínua e alvo da análise). Dentre as diversas técnicas de discretização existentes, foi-se utilizada a técnica *Equal-with Binning* (binarização), que consiste na divisão de um intervalo original de valores por faixas de mesma largura (intervalos iguais).

Desse modo, foram escolhidos 5 intervalos (*bins*) que separaram os valores originais em 5 classes de valores para a nova variável chamada MASSA\_Disc. As faixas de valores foram posteriormente nomeadas como:

- Baixíssima produção: de 0 a 11377,04 toneladas por turno, 41534 exemplos;
- Baixa produção: de 11377,05 a 22754,08 toneladas por turno, 11794 exemplos;
- Média produção: de 22754,09 a 34131,12 toneladas por turno, 4739 exemplos;
- Alta produção: de 34131,13 a 45508,16 toneladas por turno, 962 exemplos;
- Altíssima produção: de 45508,17 à 56885,20 toneladas por turno, 93 exemplos.

A segunda etapa consistiu na amostragem propriamente dita. A partir das faixas de valores determinadas pela discretização, foi observado o balanceamento dos registros para cada classe. A partir desta observação, foi aplicada uma amostragem estratificada,

onde foram selecionados da base de dados original, uma quantidade balanceada de registros para a variável MASSA, de acordo com o balanço observado na variável MASSA\_Disc. Desse modo se garantiu que a amostragem teria a mesma representatividade da base de dados original.

A base de dados amostrada passou a conter 5% dos exemplos contidos no conjunto original, resultando em 2957 eventos, conforme Tabela 2.

Tabela 2 - Discretização do atributo MASSA.

CLASSE DE VALORES	QUANTIDADE DE EXEMPLOS
Baixíssima produção	2077
Baixa produção	590
Média produção	237
Alta produção	48
Altíssima produção	5

Fonte: Elaborado pelo autor.

Após isso, foi realizada a fase de entendimento dos dados a serem analisados, de modo que fossem observadas todas as características dos atributos, seus significados, como são representados, bem como identificando os dados relevantes, etc. Além da compreensão dos dados, foi verificado se os bancos de dados resultantes seriam coerentes para os objetivos das análises, estimativas e previsões posteriores.

Conforme a metodologia proposta, o entendimento dos dados deve ser uma tarefa realizada já na etapa de Seleção dos dados. Por isso, nesta fase já foi realizada uma análise descritiva da base de dados, para explicar e apoiar a compreensão dos atributos associados ao problema. A etapa de análise descritiva consistiu na aplicação de métodos estatísticos para tornar possível reconhecer os primeiros padrões operacionais

### 3.2. PRÉ-PROCESSAMENTO E TRANSFORMAÇÃO DOS DADOS

A etapa de pré-processamento foi realizada a fim de preparar os dados para as etapas de análises posteriores, de modo a obter qualidade suficiente para os dados, manter sua representatividade para o mundo real e tornar possível a extração de conhecimento por meio de diversos algoritmos.

Inicialmente, alguns atributos desnecessários foram excluídos, como por exemplo a variável MÊS\_, que representa de forma categórica as mesmas informações contidas na variável MÊS, evitando assim redundância de informações. TURNO, TURMA, EQUIPAMENTO e TIPO são variáveis que foram excluídas pelo fato de representarem de forma numérica as mesmas informações contidas nos atributos categóricos TURNO\_, TURMA\_, EQUIPAMENTO\_ e TIPO\_.

Outro tipo de preparação realizada foi a observação de dados faltantes na base original para o atributo LOCAL DE CARREGAMENTO. Esta observação resultou na eliminação de exemplos que não possuíam a informação da cota de carregamento. Neste mesmo atributo, também foram corrigidos erros de digitação, bem como foram excluídos tanto registros que não se referiam a cota de carregamento, mas a locais espalhados pela mina, quanto registros que não seguiam o padrão de apontamento da variável.

Ainda para o atributo LOCAL DE CARREGAMENTO, foram excluídos registros de cota com valores acima de 258 m e abaixo de -264 m. Esta ação foi realizada pelo fato de que no período analisado, não houve nenhuma operação realizada em cotas fora do intervalo mencionado.

Quanto ao atributo QUALIDADE, foram excluídos mais de 700 registros que não possuíam identificação do tipo de material carregado. Além disso, visando realizar uma redução na quantidade de categorias deste atributo, foi observado o balanceamento das classes registradas, conforme apresentado na Tabela 3.

Tabela 3 - Classes do atributo QUALIDADE.

CLASSE	Nº DE REGISTROS
Granito	37355
Sulf. AT	10960
Sulf. BT	5864
Saprolito	3892
Sulf. SAT	763
Misto Sulfetado	163
Oxid. Verde	61
Oxid. Verm. Seq	52
Misto Oxidado	5
Remanejo	3
Gabro	2
Oxid. Verm. Sos	1

Fonte: Elaborado pelo autor.

Tendo em vista que algumas classes apareceram pouquíssimas vezes ao longo dos 15 anos avaliados, foi realizado um agrupamento dos registros que continham as classes 'Misto Sulfetado', 'Oxid. Verde', 'Oxid. Verm. Seq', 'Misto Oxidado', 'Remanejo', 'Gabro', 'Oxid. Verm. Sos' e 'Vulc. Xisto'. Desse modo, para esse agrupamento, foi criada uma nova classe denominada como 'Outros', que passou a conter 288 registros.

Em diversos casos no *Data Mining* e *Machine Learning*, muitos algoritmos apresentam restrições para os tipos de atributos utilizados. Por exemplo, para casos em que os algoritmos utilizam apenas atributos com valores numéricos, é necessário transformar os atributos com valores qualitativos (nominais, ordinais, categóricos).

As variáveis categóricas TURNO\_, TURMA\_, EQUIPAMENTO\_, TIPO\_, QUALIDADE\_ e FROTA DE TRANSPORTE\_ passaram por um processo que as transformou para variáveis numéricas. A abordagem de transformação utilizada busca eliminar qualquer ordem para as classes existentes. Ou seja, se buscou não simplesmente numerar as classes de forma ordinal para evitar que os algoritmos de mineração atribuíssem peso para as classes em função da ordem que lhes seria atribuída. Assim, a abordagem utilizada para isso foi a Codificação 1-de-c, que considera primeiramente que um atributo possui c valores possíveis. Em seguida, são criados c novos atributos binários. Cada novo atributo indica um possível valor do atributo nominal original. Assim, apenas uma posição da nova sequência de atributos binários de cada objeto poderá ter valor igual a 1. A Figura 18 é apresentada para fins de explicação, ilustrando a ação, por exemplo, para a variável FROTA DE TRANSPORTE\_.

Figura 18 - Exemplo de transformação pela técnica 1-de-c.

Antes da transformação		Depois da Transformação		
EXEMPLOS	FROTA DE TRANSPORTE_	EXEMPLOS	FROTA DE TRANSPORTE_ Cat 793C	FROTA DE TRANSPORTE_ Cat 785C
1	Cat 793C	1	1	0
2	Cat 785C	2	0	1

Fonte: Elaborado pelo autor.

Ou seja, considerando que o atributo FROTA DE TRANSPORTE\_ possui duas classes possíveis (valor c), foram criados dois atributos, cada um para representar o evento em que foi utilizada uma classe ou outra, por meio de caracteres binários, onde o

valor “1” significa a ocorrência de cada classe, especificamente. Estas alterações pelo método 1-de-c resultaram em uma base de dados com 52 atributos.

Pelo fato de alguns algoritmos de aprendizado de máquina necessitarem de dados normalizados para alcançarem um uso adequado, também se buscou a normalização dos dados. Em função do método proposto nesta pesquisa indicar o teste de vários modelos para que seja possível buscar aquele com melhor desempenho, escolheu-se criar dois modelos de dados para testes: o primeiro consistiu numa base de dados normalizada por reescala, o segundo se trata de uma base de dados normalizada por padronização, conforme abordado no capítulo 2.5, quando trata sobre técnicas de normalização.

Após a etapa de pré-processamento foi verificada a ocorrência de ruídos. A geração de ruído pode ser em virtude de erros na compilação de dados, mas também erros na captura, ou da existência de dados reais, mas que não são frequentes. Assim, os ruídos foram identificados por meio de gráficos de box-plots, bem como análises de intervalos interquartis, a fim de observar valores máximos e mínimos para serem inclusos na análise. Foram identificados dados extremos (*outliers*), que foram excluídos, a fim de garantir a melhor modelagem para o problema em questão.

### 3.3. MODELAGEM E AVALIAÇÃO DE DESEMPENHO DOS ALGORITMOS

A etapa de modelagem foi dividida em duas fases. A primeira consistiu na chamada Engenharia de Recursos (*Features*), que observou o comportamento dos atributos e adequou os bancos de dados para a fase de modelagem. A segunda consistiu na definição dos hiperparâmetros utilizados pelos algoritmos escolhidos para serem avaliados.

Durante a Engenharia de Recursos, se optou, a partir das duas bases normalizadas, definir um banco de dados reduzido. Para isso, foi escolhido o método *SelectKBest* onde  $k = 30$  selecionou os 30 atributos que mais impactam no resultado do modelo. As 30 variáveis mais importantes foram selecionadas por meio do teste estatístico *f\_regression*.

Após a seleção das melhores features e observação das suas correlações com a variável alvo MASSA, foram escolhidos diversos algoritmos de aprendizado de máquina, cada um com a capacidade de realizar determinadas tarefas de Ciência de Dados. Os algoritmos escolhidos são k-Vizinhos Próximos (*kNN*), Árvores de Decisão, Floresta

Aleatória, Rede Neural Perceptron de Multi-camadas (*MLP*), Máquina de Vetor de Suporte - Regressão (*SVR*), Regressão de Ridge.

A técnica utilizada para a hiperparametrização dos algoritmos foi a Busca Exaustiva em Grids, a fim de submeter diversos modelos para a fase de treinamento, de modo que existisse também uma possível diversificação nos hiperparâmetros de cada modelo.

Os modelos criados foram conduzidos para a etapa de treinamento, onde dados históricos são primordiais para capacitar os algoritmos a entender, aprender e realizar ações acerca das situações expressas pelos dados. A metodologia utilizada para o aprendizado foi a Validação Cruzada com *10-folds*, a fim de preparar melhor os algoritmos para atuarem em situações reais, bem como reduzir o tempo de processamento no treino, evitar o fenômeno de sobreajuste e garantir a representatividade do modelo.

Após a etapa de treinamento, os sistemas modelados passaram pelos testes. Nesta fase, os dados já conhecidos foram submetidos a uma comparação diante dos resultados preditos pelos modelos. Ou seja, os resultados obtidos por cada algoritmo na fase de testes foram comparados aos resultados reais, de modo que fosse possível analisar o desempenho dos algoritmos ao realizar predições durante a fase de aprendizagem.

A métrica de qualidade utilizada para avaliar o desempenho de todos os modelos criados foi o Erro Médio Absoluto (*MAE*). Ao fim, dentre todos os modelos avaliados, se escolheu o que atingiu maior aderência aos resultados reais a fim de submetê-lo à etapa seguinte, de predição de situações reais e desconhecidas.

### 3.4. ANÁLISE PREDITIVA

Esta fase consistiu na aplicação do algoritmo de aprendizado de máquina que obteve melhor desempenho na etapa de testes. Este foi submetido ao ambiente real de mina para gerar predições acerca do dia-a-dia da operação. Nesta etapa, esperou-se que o modelo fosse capaz de prever a produção mássica de cada escavadeira que seria alocada nos turnos de trabalho, levando em consideração as mais variadas condições operacionais representadas pelos atributos contidos no banco de dados.

Foram escolhidos 75 turnos reais de operação para avaliar a atuação do sistema preditivo criado, comparando os resultados das predições com os resultados que

ocorreram na prática, considerando as particularidades da configuração operacional de cada turno analisado.

Desse modo, também foi observado o Erro Médio Absoluto (MAE) para as situações reais que o algoritmo foi submetido. Esta observação foi relevante para avaliar se o algoritmo escolhido na etapa de testes também obteve um bom desempenho na rotina operacional de planejamento.

### 3.5. TOMADA DE DECISÃO NO CICLO DE PLANEJAMENTO DE LAVRA

A última etapa desta metodologia foi dividida em duas partes: a tomada de decisão apoiada pelos resultados da análise preditiva e, a replicação do ciclo de análise para os períodos seguintes ( $p + 1$ ).

A metodologia proposta para a etapa de tomada de decisão consistiu em 3 passos que têm como principal objetivo, aumentar o valor de massa produzida por cada equipamento no turno, para os casos em que a predição inicial fornecida tiver valores abaixo dos definidos como meta. Seguem passos sugeridos:

- 1) Observar se a predição fornecida pelo sistema atende as metas definidas.  
Em caso negativo, passar para o passo 2;
- 2) Observar os recursos de maior importância para o modelo, bem como observar dentre as principais recursos, quais possuem maior correlação com a variável alvo;
- 3) Alterar o valor destas variáveis de entrada, visando o aumento de produção, de acordo com o maior valor de importância, nas maiores correlações e na aplicabilidade da decisão tomada na rotina operacional.

Ou seja, considerando os dados fornecidos pelo sistema preditivo, é possível tomar decisões para a realização de possíveis ações que resultarão no aumento de produção de cada escavadeira por turno trabalhado.

## 4. RESULTADOS

Os resultados apresentados neste capítulo foram gerados a partir da aplicação do Processo padrão de ciência de dados para Planejamento de Lavra, apresentado no capítulo 3, “Materiais e Métodos”.

### 4.1. RESULTADOS PARA SELEÇÃO DOS DADOS E ANÁLISE DESCRITIVA

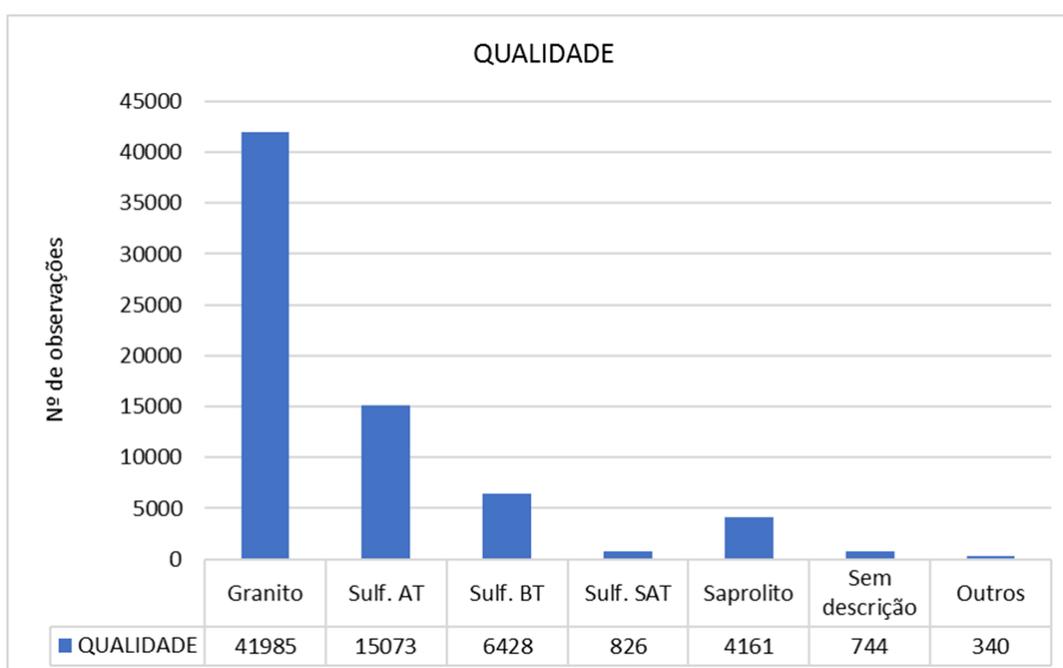
Assim como propõe a metodologia, além da seleção dos dados explicada no capítulo de Materiais e Métodos, nesta etapa também foi realizado o trabalho de entendimento aprofundado dos dados, a fim de que estes sejam adequados para o objetivo da modelagem.

Desse modo, algumas técnicas estatísticas de visualização de dados foram adotadas para alguns atributos contidos na base de dados, conforme os tópicos a seguir.

#### 4.1.1. Litologia da jazida

A variedade litológica observada no complexo minerário é bastante alta. É possível identificar que os equipamentos atuaram em mais de 10 qualidades de rocha diferentes ao longo dos 16 anos de operação registrados.

Figura 19 - Registros de alocação da ES3001 conforme litologias da mina (2016).



Fonte: Elaborado pelo autor.

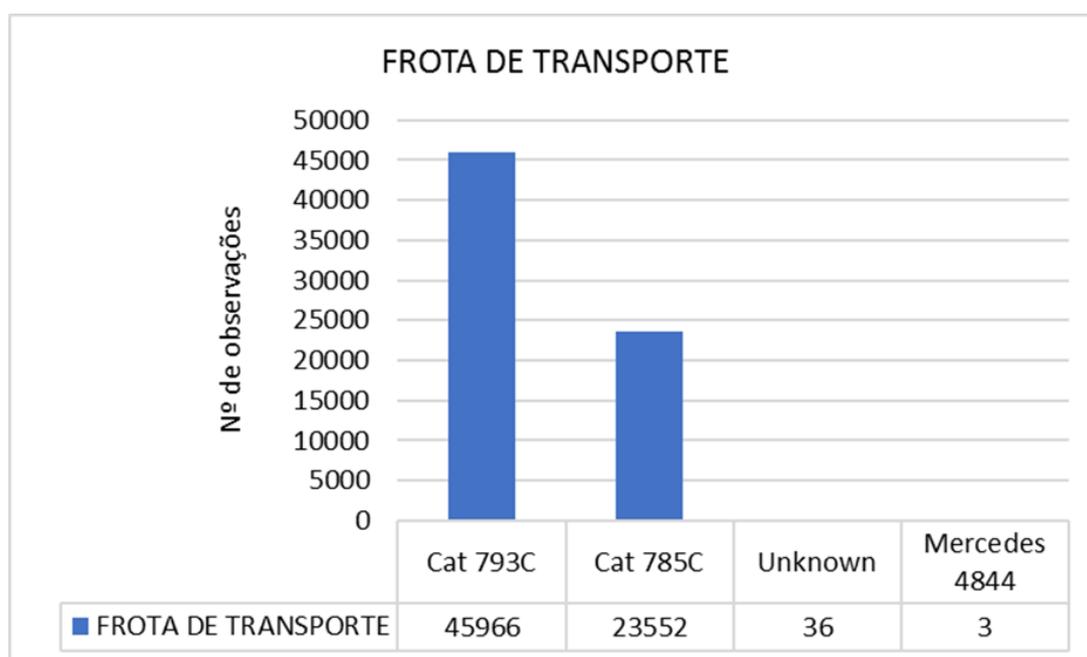
Os registros das litologias Misto Sulfetado, Oxid. Verde, Oxid. Verm. Seq, V Sulf. AT, Misto Oxidado, Remanejo, Gabro, Vulc. Xisto, Oxid. Verm. Sos e 4 Sulf. AT foram agrupadas na categoria “Outros”, em função da baixa quantidade de registros (abaixo de 400).

Conforme a Figura 19, as litologias mais frequentes na base de dados são Granito, Sulfetado de Alto Teor, Sulfetado de Baixo Teor, Sulfetado de Super Alto Teor e Saprolito.

#### 4.1.2. Frota de transporte

O escoamento de material lavrado na Mina do Sossego se dá por meio de sistema escavadeira-caminhão. Nos anos analisados, as escavadeiras foram atendidas, principalmente, por duas frotas de transporte, como mostrado na Figura 20.

Figura 20 - Frota de transporte utilizada nos períodos analisados.



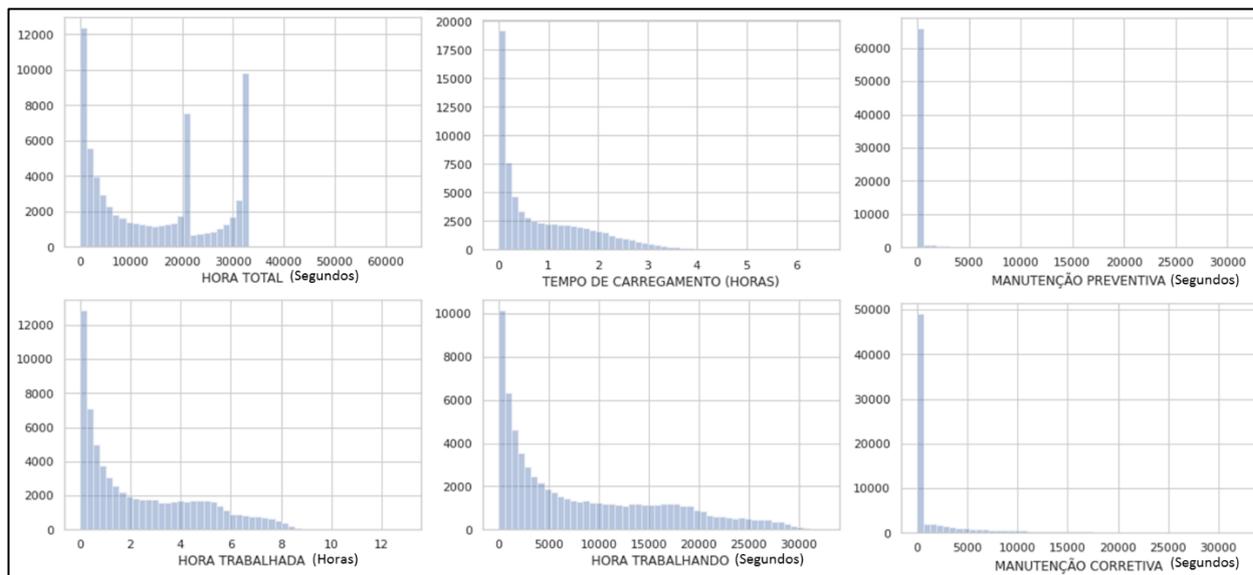
Fonte: Elaborado pelo autor.

A Figura 20 mostra que as frotas de caminhões Cat 793C e Cat 785C foram as mais atuantes para atender o trabalho das escavadeiras ES3001, ES3101, ES3102 e PM3401.

### 4.1.3. Registros de horas

O banco de dados apresenta registros de horas da atuação dos equipamentos ao longo dos anos.

Figura 21 - Registros de horas das escavadeiras.



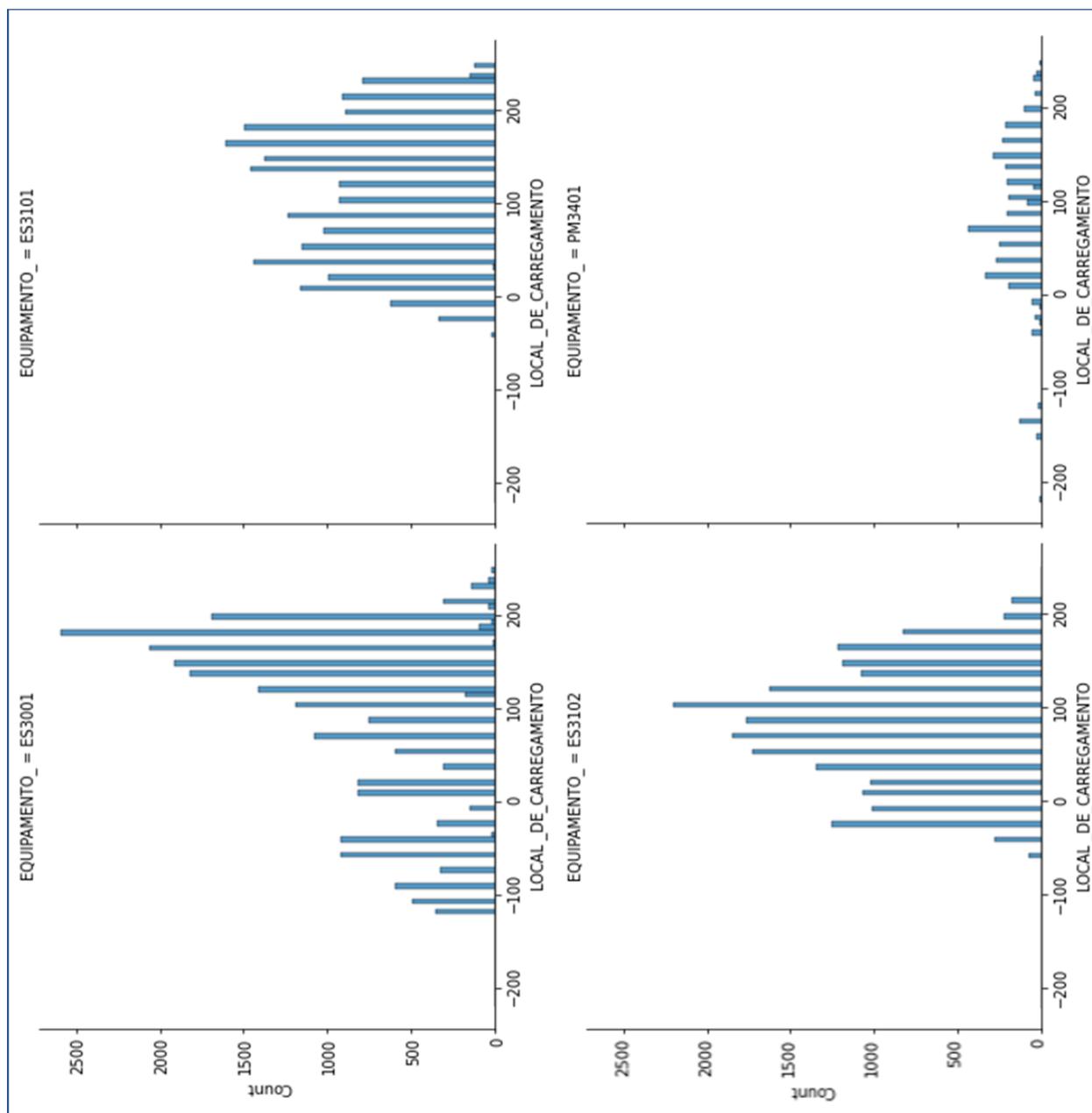
Fonte: Elaborado pelo autor.

Como observado na Figura 21, a realização de manutenções preventivas não foram frequentes ao longo dos anos de operações.

### 4.1.4. Local de carregamento

O atributo Local de carregamento identifica a cota dos bancos em que as escavadeiras atuaram. Esta informação é bastante relevante pois, em geral, as cotas superiores no interior da mina possuem mais espaço para que os equipamentos possam manobrar e trabalhar adequadamente, em relação às cotas inferiores. Portanto, construir um modelo que preveja o desempenho dos equipamentos considerando também a informações das cotas das bancadas, potencializa a eficiência da decisão de alocação por meio de um sistema preditivo.

Figura 22 - Registros das cotas em que as escavadeiras atuaram.



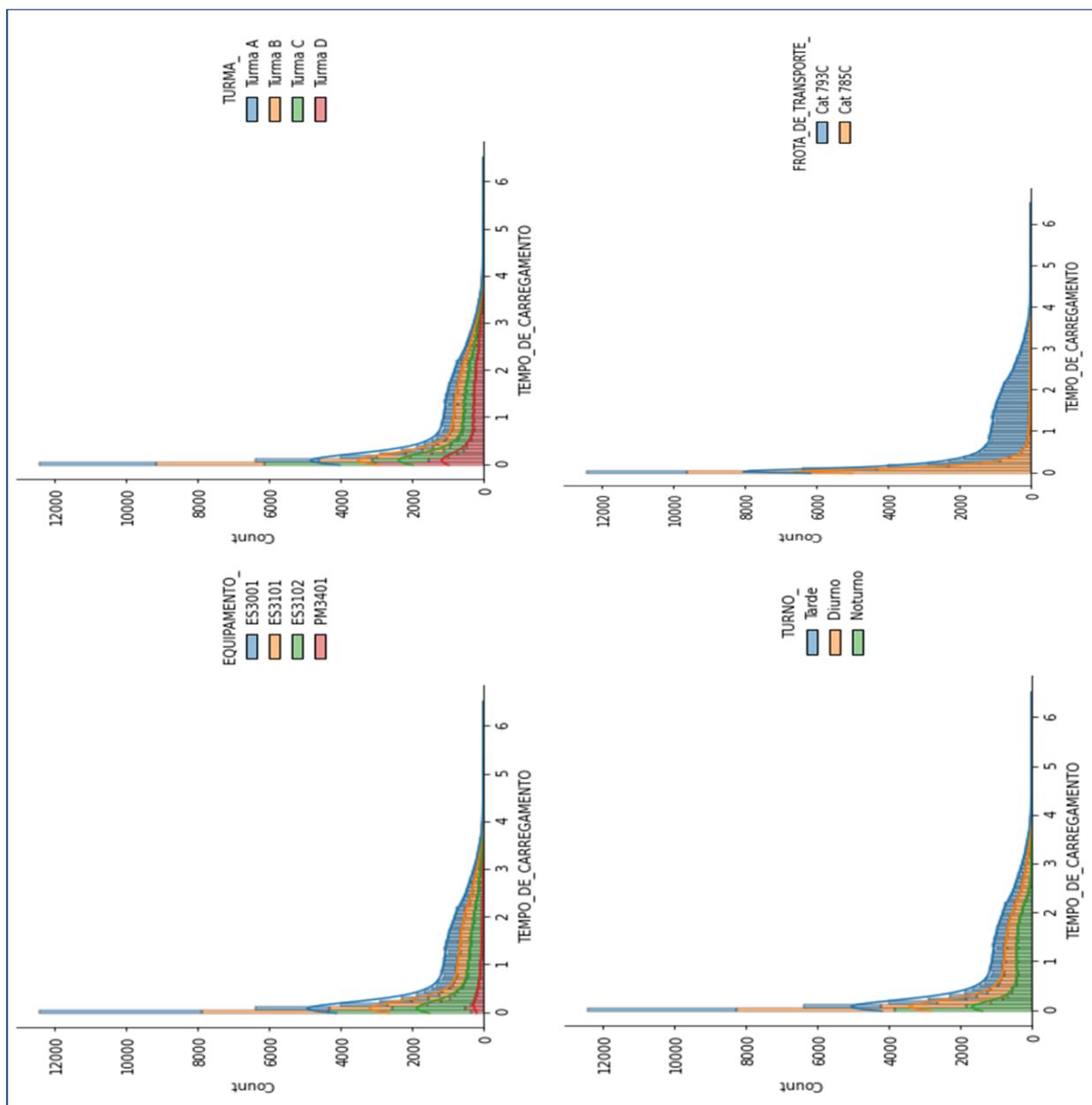
Fonte: Elaborado pelo autor.

A Figura 22 mostra que as escavadeiras realizaram manuseio de material em bancos localizados nas mais diversas cotas. As escavadeiras ES3001, ES3101 e ES3102 atuaram com mais frequência nas cotas mais superiores. A possível razão para isto é o fato de serem equipamentos de maior porte, quando comparadas à escavadeira PM3401, e por isso necessitam atuar em praças mais amplas para a melhor realização das manobras.

#### 4.1.5. Tempo de carregamento (h)

Foram avaliadas as horas destinadas ao tempo total de carregamento gasto pelos equipamentos de carga a cada turno. Este conhecimento é imprescindível para a tomada de decisão a respeito da alocação dos equipamentos. A Figura 23 apresenta a distribuição das horas destinadas para tempo de carregamento, de acordo com cada escavadeira, turno, turma e frota de transporte.

Figura 23 - Registros das horas destinadas para tempo de carregamento (h).



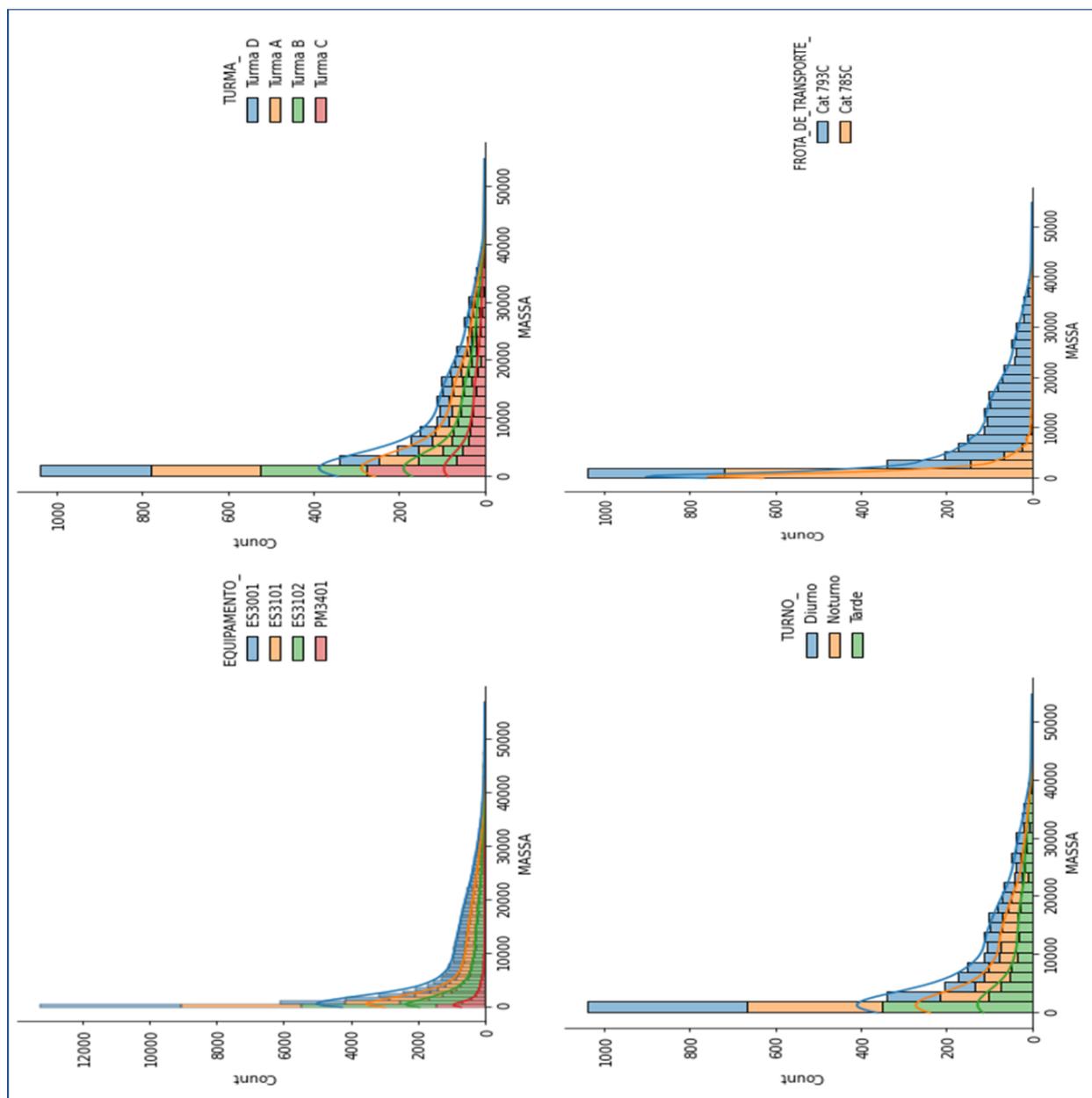
Fonte: Elaborado pelo autor.

É possível observar que o tempo de carregamento na PM3401, em geral, gastaram menos horas que os demais equipamentos. As Turmas e Turnos de trabalho, no geral, possuem o mesmo padrão de dispersão para horas gastas no tempo de carregamento. Os registros de tempo total de carregamento foram maiores quando as escavadeiras atuaram com a Frota de Transporte Cat 793C. Isso pode ser explicado pelo fato de que a frota de caminhões Cat 785C possuem capacidade menor que a frota Cat 793C.

#### **4.1.6. Massa (t)**

Na Figura 24, foi observada a distribuição da produção considerando cada equipamento, e também considerando os eventos vinculados à atuação das escavadeiras associadas a cada turno, turmas de trabalho e frotas de transporte que atenderam os equipamentos.

Figura 24 - Massa produzida (t).



Fonte: Elaborado pelo autor.

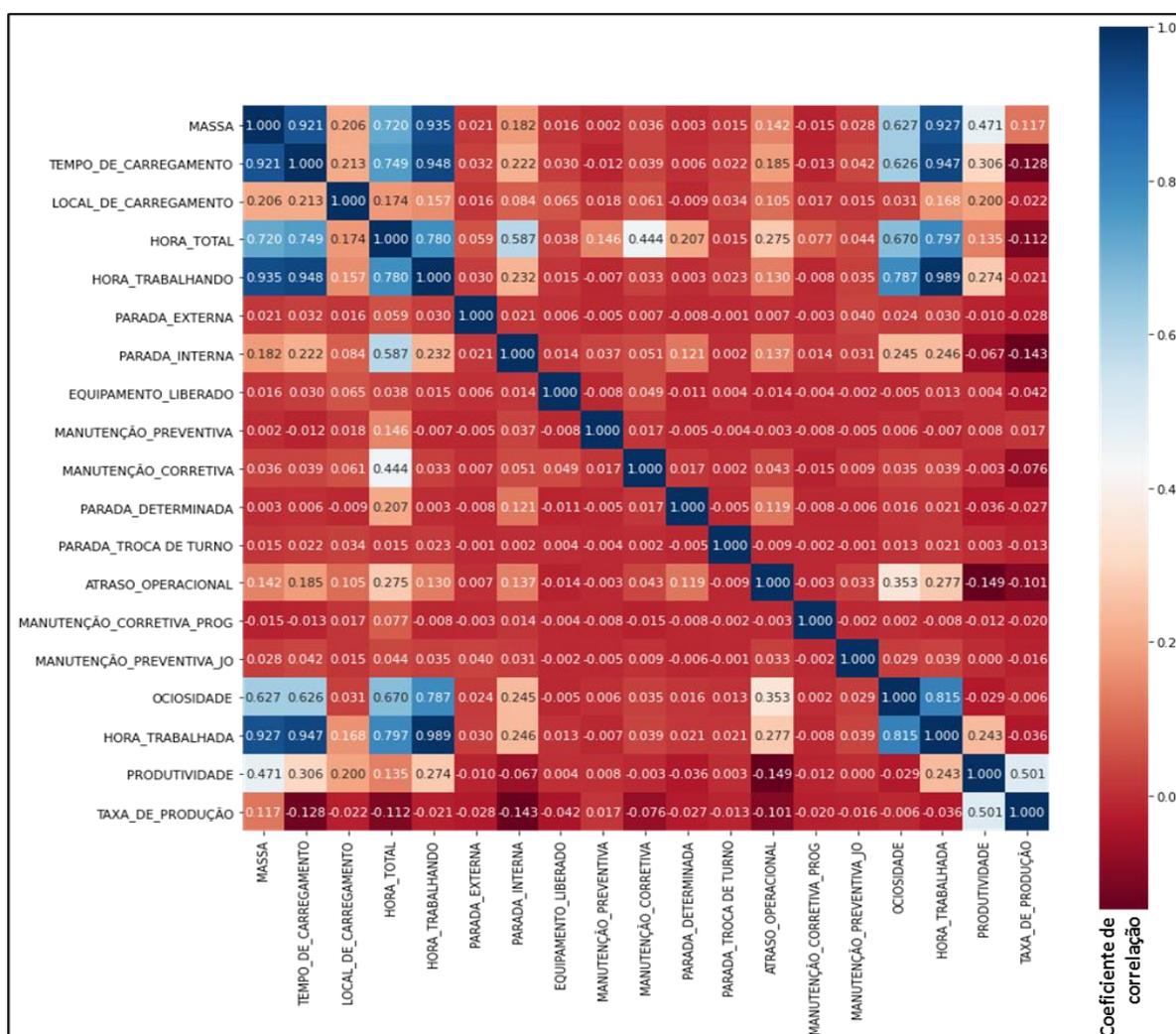
Por ser um equipamento de menor porte, a PM3401 tem um total de massa produzida ao fim do turno menor que as demais escavadeiras, que por sua vez, possuem padrões semelhantes. O padrão das Turmas de trabalho também não possui padrões distintos entre si. De maneira diferente, os turnos operacionais Tarde e Diurno possuem padrão de maior quantidade mássica produzida em relação ao turno da noite. Uma outra observação importante é que as escavadeiras alcançam maiores valores de produção quando atuam com a frota de transporte 795C. Uma possível explicação para isso, é o fato de os caminhões desta frota serem de maior porte, e portanto, mais adequados para

atuarem com as escavadeiras maiores (ES3001, ES3101 e ES3102) e consequentemente, atingem maior massa produzida quando atuam em conjunto.

#### 4.1.7. Correlação das variáveis

Foram observadas as correlações lineares existentes entre cada variável da base de dados, incluindo a correlação de cada uma delas com a variável MASSA. A Figura 25 mostra um Mapa de Calor que representa a matriz de correlação gerada pelos valores do  $R^2$  para cada par de variável analisado.

Figura 25 - Correlação entre as variáveis da base de dados.



Fonte: Elaborado pelo autor.

Por se tratar de um mapa de calor, as cores com tons mais próximos do azul escuro representam uma maior correlação. Por outro lado, as cores mais próximas de vermelho escuro, representam correlações fracas.

Sendo assim, as correlações foram avaliadas conforme a classificação abaixo.

- Correlação fraca: de -0,3 até 0,3;
- Correlação média: de -0,7 até 0,3 e de 0,3 até 0,7;
- Correlação forte: menor que -0,7 e maior 0,7.

É possível observar que algumas variáveis apresentam forte correlação entre si, mas que não ajudam a levantar muitas hipóteses sobre o padrão operacional dos dados. Esta afirmativa pode ser explicada pela alta colinearidade encontrada entre alguns pares de variáveis que se explicam por si só como: massa produzida, tempo de carregamento, hora total e hora trabalhando, que se correlacionam fortemente entre si; massa produzida com forte correlação negativa com a variável ociosidade; produtividade tendo média correlação com a taxa de produção, etc.

Desse modo, tornou-se imprescindível realizar uma análise aprofundada de avaliação da importância de cada variável para a explicação da variável alvo, a fim de escolher o melhor conjunto de atributos para compor o modelo preditivo.

#### 4.2. RESULTADOS PARA PRÉ-PROCESSAMENTO E TRANSFORMAÇÃO

O pré-processamento foi realizado na base de dados, conforme descrito no Capítulo 3. Esta base de dados passou pelo processo de amostragem, exclusão de dados redundantes, transformação de variáveis categóricas em variáveis numéricas e agrupamento de classes para maior generalização de variáveis. A Tabela 4 apresenta as variáveis da base de dados resultante, bem como a checagem de valores nulos nos registros, além do tipo de cada variável.

Tabela 4 - Resumo da base de dados após pré-processamento.

VARIÁVEL	Nº DE EXEMPLOS	Nº VALORES NULOS	TIPO DA VARIÁVEL
ANO	2957	0	Inteira
MÊS	2957	0	Inteira
DIA	2957	0	Inteira
MASSA (t)	2957	0	Real
TEMPO_DE_CARREGAMENTO (seg)	2957	0	Real
LOCAL_DE_CARREGAMENTO (m)	2957	0	Inteira
HORA_TOTAL (seg)	2957	0	Real
HORA_TRABALHANDO (seg)	2957	0	Real
PARADA_EXTERNA (seg)	2957	0	Real
PARADA_INTERNA (seg)	2957	0	Real
EQUIPAMENTO_LIBERADO (seg)	2957	0	Real

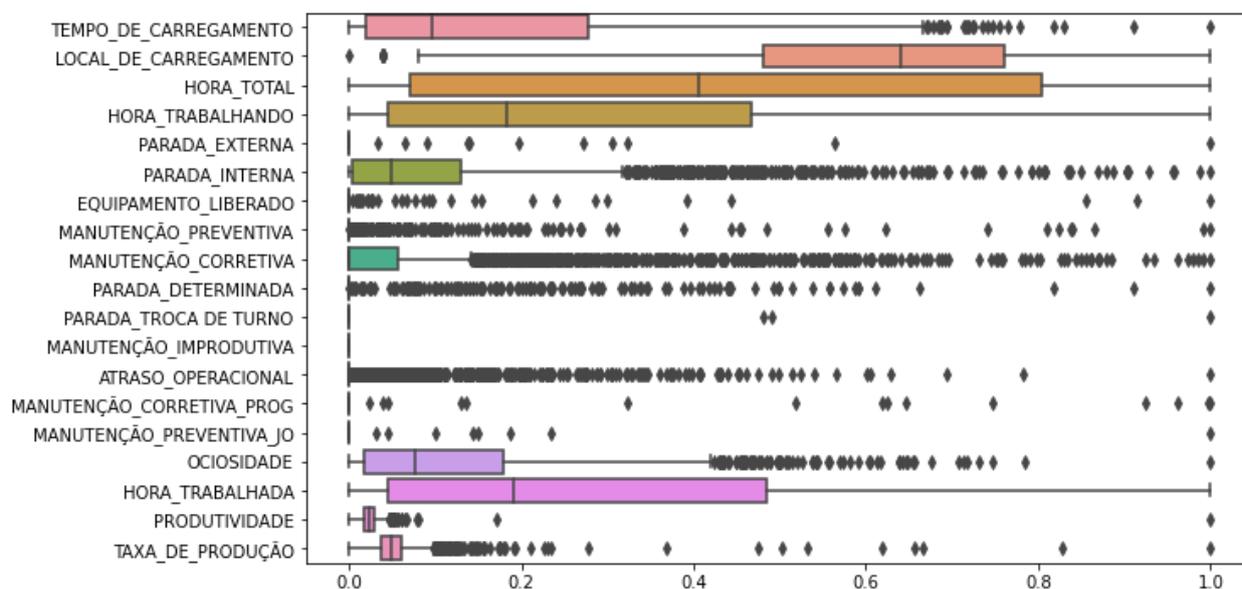
MANUTENÇÃO_PREVENTIVA (seg)	2957	0	Real
MANUTENÇÃO_CORRETIVA (seg)	2957	0	Real
PARADA_DETERMINADA (seg)	2957	0	Real
PARADA_TROCA DE TURNO (seg)	2957	0	Real
MANUTENÇÃO_IMPRODUTIVA (seg)	2957	0	int64
ATRASO_OPERACIONAL (seg)	2957	0	Real
MANUTENÇÃO_CORRETIVA_PROG (seg)	2957	0	Real
MANUTENÇÃO_PREVENTIVA_JO (seg)	2957	0	Real
OCIOSIDADE (seg)	2957	0	Real
HORA_TRABALHADA (horas)	2957	0	Real
PRODUTIVIDADE (t/h)	2957	0	Real
TAXA_DE_PRODUÇÃO (t/h)	2957	0	Real
TURNO_Diurno	2957	0	Binária
TURNO_Noturno	2957	0	Binária
TURNO_Tarde	2957	0	Binária
TURMA_Turma A	2957	0	Binária
TURMA_Turma B	2957	0	Binária
TURMA_Turma C	2957	0	Binária
TURMA_Turma D	2957	0	Binária
EQUIPAMENTO_ES3001	2957	0	Binária
EQUIPAMENTO_ES3101	2957	0	Binária
EQUIPAMENTO_ES3102	2957	0	Binária
EQUIPAMENTO_PM3401	2957	0	Binária
TIPO_Buc 495-HR	2957	0	Binária
TIPO_PH-4100 XPB	2957	0	Binária
TIPO_WA 1200	2957	0	Binária
QUALIDADE_Granito	2957	0	Binária
QUALIDADE_Outros	2957	0	Binária
QUALIDADE_Saprolito	2957	0	Binária
QUALIDADE_Sulf. AT	2957	0	Binária
QUALIDADE_Sulf. BT	2957	0	Binária
QUALIDADE_Sulf. SAT	2957	0	Binária
FROTA_DE_TRANSPORTE_Cat 785C	2957	0	Binária
FROTA_DE_TRANSPORTE_Cat 793C	2957	0	Binária

Fonte: Elaborado pelo autor.

Em seguida foi realizada uma transformação de dados para análise de ruídos, bem como suas limpezas. Para esta análise, optou-se por primeiramente, normalizar a base de dados e avaliar possíveis valores discrepantes para as variáveis numéricas por meio de gráficos do tipo *boxplot*.

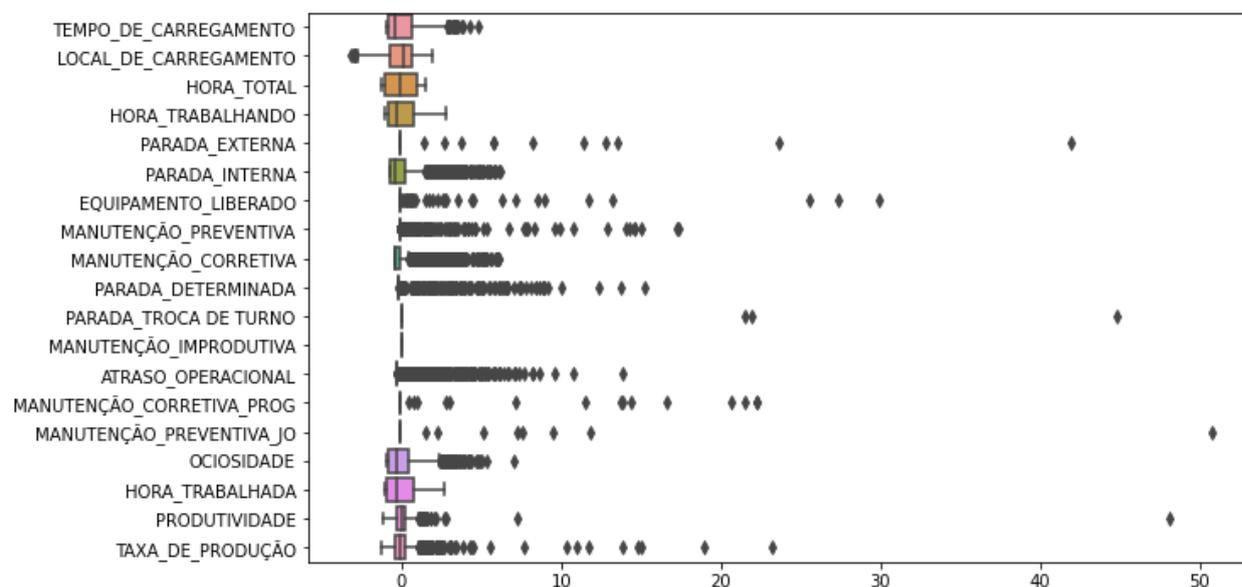
Assim, foram escolhidos dois métodos de normalização para a modelagem: por reescala e por padronização. As Figuras 26 e 27 apresentam os dados transformados pelos dois métodos para a base de dados amostrada.

Figura 26 - Boxplot das variáveis numéricas normalizadas por reescala.



Fonte: Elaborado pelo autor.

Figura 27 - Boxplot das variáveis numéricas normalizadas por padronização.



Fonte: Elaborado pelo autor.

Como observado nas Figuras 26 e 27, diversos valores podem ser considerados potencialmente discrepantes. Entretanto, foi adotada a técnica estatística de análise de intervalos interquartis para definir limites máximos e mínimos de confiabilidade que serão considerados para cada base.

As equações 11a, 11b e 11c apresentam matematicamente como se deu a definição dos intervalos máximos e mínimos das variáveis.

$$IQR = Q3 - Q1 \quad (11a)$$

$$\text{Limite superior} = Q3 + 1,5 * IQR \quad (11b)$$

$$\text{Limite inferior} = Q3 - 1,5 * IQR \quad (11c)$$

Onde,

*IQR: Intervalo interquartil;*

*Q1: 1º quartil que contém 25% dos dados;*

*Q3: 3º quartil que contém 75% dos dados.*

Estes limites foram encontrados para todas as variáveis contidas tanto na base normalizada por reescala, quanto para a base normalizada por padronização. A faixa de valores que classifica valores comuns de valores discrepantes é apresentada na Tabela 5.

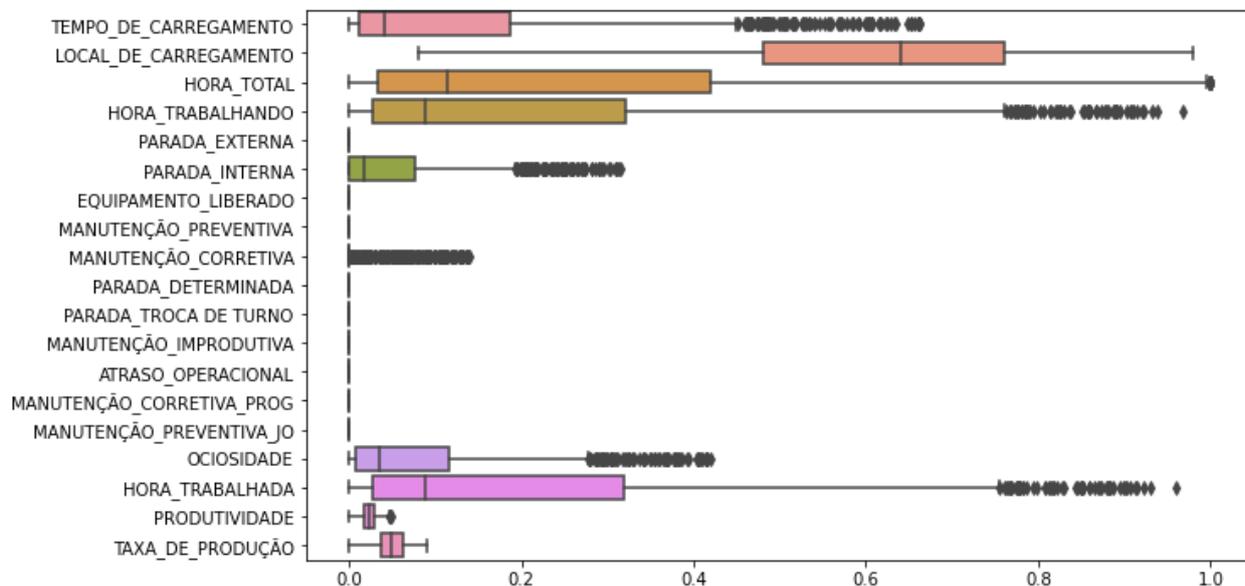
Tabela 5 - Limites superiores e inferiores para variáveis das bases normalizadas.

VARIÁVEL	REESCALA		PADRONIZAÇÃO	
	Lim. Inferior	Lim. Superior	Lim. Inferior	Lim. Superior
TEMPO_DE_CARREGAMENTO (seg)	-0.37	0.67	-5.32	6.38
LOCAL_DE_CARREGAMENTO (m)	0.06	1.18	-6.82	6.18
HORA_TOTAL (seg)	-1.03	1.9	-5.17	5.45
HORA_TRABALHANDO (seg)	-0.59	1.1	-5.5	6.44
PARADA_EXTERNA (seg)	0	0	-0.04	-0.04
PARADA_INTERNA (seg)	-0.18	0.32	-5.01	6.49
EQUIPAMENTO_LIBERADO (seg)	0	0	-0.06	-0.06
MANUTENÇÃO_PREVENTIVA (seg)	0	0	-0.48	0.38
MANUTENÇÃO_CORRETIVA (seg)	-0.08	0.14	-4.62	6.51
PARADA_DETERMINADA (seg)	0	0	-0.18	-0.16
PARADA_TROCA DE TURNO (seg)	0	0	-0.03	-0.03
MANUTENÇÃO_IMPRODUTIVA (seg)	0	0	0	0
ATRASSO_OPERACIONAL (seg)	0	0	-3.33	4.74
MANUTENÇÃO_CORRETIVA_PROG (seg)	0	0	-0.06	-0.06
MANUTENÇÃO_PREVENTIVA_JO (seg)	0	0	-0.03	-0.03
OCIOSIDADE (seg)	-0.22	0.42	-5.35	6.42
HORA_TRABALHADA (horas)	-0.61	1.14	-5.55	6.5
PRODUTIVIDADE (t/h)	0	0.05	-2.92	2.94
TAXA_DE_PRODUÇÃO (t/h)	0	0.09	-3.58	3.64

Fonte: Elaborado pelo autor.

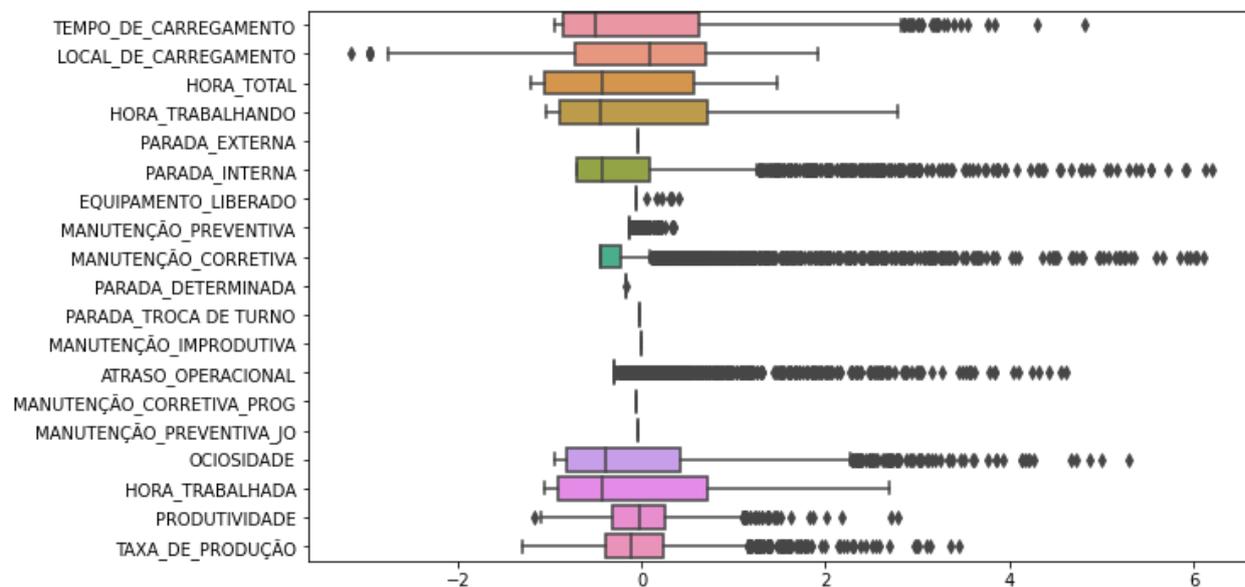
Após isso, foram excluídos todos os dados contidos fora dos limites superiores e inferiores de cada variável e os resultados desta limpeza de ruídos estão representados nas Figuras 28 e 29.

Figura 28 - Boxplot das variáveis numéricas normalizadas por reescala, após limpeza.



Fonte: Elaborado pelo autor.

Figura 29 - Boxplot das variáveis numéricas normalizadas por padronização, após limpeza.



Fonte: Elaborado pelo autor.

A base de dados normalizada por reescala resultante da limpeza de ruídos contém 1544 exemplos, enquanto que a quantidade de exemplos para a base normalizada por padronização é de 355 exemplos. Após esta primeira rodada de limpeza, os dois conjuntos de dados resultantes foram submetidos para a etapa de modelagem. Conforme a metodologia proposta, se os modelos não obtiverem bons resultados preditivos, serão submetidos a outras rodadas de pré-processamento e transformação de dados.

### 4.3. RESULTADOS PARA MODELAGEM E AVALIAÇÃO

#### 4.3.1. Análise de importâncias (Seleção de Recursos)

Conforme informado no Capítulo 3, a fase de modelagem usou duas bases de dados normalizadas: uma por reescala e outra por padronização. Estas foram submetidas a um processo de Engenharia de Recursos, que consistiu em avaliar as variáveis que mais impactam no resultado final da variável alvo MASSA, a fim de evitar perda de explicação do modelo, reduzir tempo de processamento e reduzir sobreajuste.

A técnica *Select k Best* verificou a importância de cada variável para o modelo, retornando pontuações para cada atributo que representam o quão importante é cada variável para o processo de explicação da variável MASSA. O teste estatístico aplicado foi o *f\_regression*, que é utilizado em problemas de regressão e que define uma pontuação entre cada atributo e a variável alvo. As Tabelas 6 e 7 apresentam as pontuações de importância de cada uma das 30 principais variáveis, escolhidas para compor a base de dados final.

Tabela 6 - Seleção de recursos para a base normalizada por reescala.

RECURSOS (VARIÁVEIS)	IMPORTÂNCIA
HORA_TRABALHANDO	0.6493
HORA_TRABALHADA	0.2053
PRODUTIVIDADE	0.1294
TEMPO_DE_CARREGAMENTO	0.0129
TAXA_DE_PRODUÇÃO	0.0013
OCIOSIDADE	0.0004
PARADA_INTERNA	0.0003
LOCAL_DE_CARREGAMENTO	0.0003
MANUTENÇÃO_CORRETIVA	0.0002
HORA_TOTAL	0.0002
ANO	0.0001
EQUIPAMENTO_ES3001	8.15454E-05
TURNO_Tarde	5.55975E-05
TURMA_Turma A	4.81272E-05
TURMA_Turma D	4.45142E-05
QUALIDADE_Granito	3.05467E-05
TURMA_Turma B	3.00608E-05
TURNO_Diurno	2.92852E-05
EQUIPAMENTO_ES3101	2.49362E-05

EQUIPAMENTO_ES3102	2.3993E-05
QUALIDADE_Sulf. AT	1.78497E-05
TURNO_Noturno	1.55426E-05
QUALIDADE_Saprolito	9.47922E-06
QUALIDADE_Sulf. BT	5.94197E-06
FROTA_DE_TRANSPORTE_Cat 785C	2.89955E-06
EQUIPAMENTO_PM3401	2.23858E-06
QUALIDADE_Outros	1.60906E-06
FROTA_DE_TRANSPORTE_Cat 793C	1.04664E-06
QUALIDADE_Sulf. SAT	1.24998E-07
PARADA_TROCA DE TURNO	4.6563E-11

Fonte: Elaborado pelo autor.

Tabela 7 - Seleção de *features* para a base normalizada por padronização.

RECURSOS (VARIÁVEIS)	IMPORTÂNCIA
HORA_TRABALHANDO	0.8146
PRODUTIVIDADE	0.1226
HORA_TRABALHADA	0.0572
TEMPO_DE_CARREGAMENTO	0.0023
TAXA_DE_PRODUÇÃO	0.0013
OCIOSIDADE	0.0005
PARADA_INTERNA	0.0003
LOCAL_DE_CARREGAMENTO	0.0003
MANUTENÇÃO_CORRETIVA	0.0002
ANO	0.0002
HORA_TOTAL	0.0001
EQUIPAMENTO_ES3001	8.47188E-05
EQUIPAMENTO_ES3101	5.9622E-05
TURMA_Turma A	4.86155E-05
TURNO_Tarde	4.68238E-05
TURNO_Diurno	3.55039E-05
EQUIPAMENTO_ES3102	3.04601E-05
TURMA_Turma B	2.43168E-05
TURMA_Turma D	1.89205E-05
QUALIDADE_Sulf. AT	1.83708E-05
QUALIDADE_Granito	1.75959E-05
EQUIPAMENTO_PM3401	1.53833E-05
TURNO_Noturno	1.25463E-05
QUALIDADE_Sulf. BT	8.65709E-06
FROTA_DE_TRANSPORTE_Cat 785C	4.14431E-06

FROTA_DE_TRANSPORTE_Cat 793C	4.02735E-06
QUALIDADE_Saprolito	3.88124E-06
QUALIDADE_Sulf. SAT	4.6865E-07
QUALIDADE_Outros	3.76114E-09
PARADA_TROCA DE TURNO	0

Fonte: Elaborado pelo autor.

A análise de importância selecionou as variáveis para compor a base de dados que foi submetida aos algoritmos de aprendizado de máquina. Apesar de terem sido selecionados em ordens de pontuações diferentes, a Seleção de Recursos identificou as mesmas variáveis como *30-best*, tanto para a base normalizada por reescala, quanto para a base normalizada por padronização.

Os modelos foram criados utilizando a Busca Exaustiva em Grids como técnica de hiperparametrização e avaliados pela métrica de qualidade Erro Médio Absoluto. Além disso, na fase de treinamento e teste, foi utilizada a técnica de validação cruzada em 10-folds.

Desse modo, são apresentados a seguir os valores utilizados para os hiperparâmetros de cada modelo, bem como os resultados obtidos para a métrica do Erro Médio Absoluto (MAE).

#### 4.3.2. k-Vizinhos Próximos (*kNN*)

A Tabela 8 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem para o algoritmo *kNN*.

Tabela 8 - Hiperparâmetros avaliados para algoritmo *kNN*.

ALGORITMO	<i>kNN</i>
Hiperparâmetro	Valor
k (Número de vizinhos)	1, 3, 5, 7, 9, 11, 13, 15, 17, 19
p (Tipo de distância)	Manhattan ou Euclidiana

Fonte: Elaborado pelo autor.

De todas as possibilidades possíveis, a melhor combinação de hiperparametrização para a base de dados normalizada por reescala foi  $k = 17$  e  $p = 1$  (Distância Manhattan) com  $MAE = 3113,03$  t. Já para a base de dados normalizada por padronização, O MAE associado a esta combinação foi de 1668,41 t, com  $k = 7$  e  $p = 1$ .

### 4.3.3. Árvores de Decisão

A Tabela 9 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem utilizando o algoritmo Árvores de Decisão.

Tabela 9 - Hiperparâmetros avaliados para algoritmo Árvore de Decisão.

ALGORITMO	ÁRVORE DE DECISÃO
Hiperparâmetro	Valor
<i>min samples leaf</i>	20, 40, 100
<i>min samples split</i>	10, 20, 40
Profundidade máxima	2, 6, 8
<i>min leaf nodes</i>	5, 20, 100

Fonte: Elaborado pelo autor.

De todas as possibilidades testadas pela busca em grids, a melhor hiperparametrização para a base de dados normalizada por reescala foi utilizar o máximo de profundidade como 8, o número máximo de nós = 100, o nº mínimo de amostras em cada nó = 20 e o número mínimo de amostras para dividir um nó interno como 10. Estes hiperparâmetros fornecem um MAE = 777,43 t. Para a base normalizada por padronização, o melhor MAE atingiu o valor de 777,08 t, utilizando a profundidade máxima = 8, número máximo de nós = 100, *min\_samples\_leaf* = 20 e *min\_samples\_split* = 40.

### 4.3.4. Regressão de Vetor de Suporte (SVR)

A Tabela 10 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem utilizando o algoritmo *Support Vector Regression*.

Tabela 10 - Hiperparâmetros avaliados para algoritmo SVR.

ALGORITMO	SVR
Hiperparâmetro	Valor
C	0.01, 1, 100
<i>Gamma</i>	1, 0.01, 0.001

Fonte: Elaborado pelo autor.

De mesmo modo, por meio da busca exaustiva em grids, foi observada a combinação de hiperparâmetros que resultasse no menor MAE. Para a base de dados normalizada por reescala, o menor MAE foi alcançado com o hiperparâmetro  $C = 100$  e  $\text{Gamma} = 0,01$ , alcançando  $\text{MAE} = 5934,22$  t. Para a base de dados normalizada por reescala, o menor MAE foi alcançado com o hiperparâmetro  $C = 100$  e  $\text{Gamma} = 0,01$ , alcançando  $\text{MAE} = 2686,13$  t.

#### 4.3.5. Florestas aleatórias

A Tabela 11 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem utilizando o algoritmo Florestas aleatórias.

Tabela 11 - Hiperparâmetros avaliados para algoritmo SVR.

ALGORITMO	FLORESTAS ALEATÓRIAS
Hiperparâmetro	Valor
<i>min samples leaf</i>	1, 5, 10
<i>min samples split</i>	2, 6, 10
n estimadores	50, 100, 500
Profundidade máxima	10, 15

Fonte: Elaborado pelo autor.

Ao buscar o melhor grid de hiperparâmetros com o algoritmo *Random Forest*, foi verificado que a melhor combinação tanto para a base de dados normalizada por reescala, quanto para a base normalizada por padronização, foi utilizar o máximo de profundidade das árvores como 15, o nº mínimo de amostras necessárias em cada nó = 1, o número mínimo de amostras necessárias para dividir um nó interno como 2 e o número de árvores (estimadores) utilizadas = 500. Estes hiperparâmetros fornecem um  $\text{MAE} = 247,94$  t para a base normalizada por reescala e,  $\text{MAE} = 246,89$  t para a base normalizada por padronização.

#### 4.3.6. Rede Neural - *Perceptron* de multi camadas (MLP)

A Tabela 12 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem utilizando o algoritmo de Rede Neural do tipo *Perceptron* de multi camadas.

Tabela 12 - Hiperparâmetros avaliados para algoritmo *MLP*.

ALGORITMO	MLP
Hiperparâmetro	Valor
hidden_layer_sizes	35, 45
max_iter	4500, 5500

Fonte: Elaborado pelo autor.

Para a base normalizada por reescala, o grid de hiperparâmetros com melhor desempenho para a rede neural *MLP* ( $MAE = 1546,99$  t) teve o tamanho das camadas ocultas = 35 e  $max\_iter = 5500$ . Para a base normalizada por padronização, o grid de hiperparâmetros com melhor desempenho para a rede neural *MLP* ( $MAE = 789,17$  t) teve o tamanho das camadas ocultas = 45 e  $max\_iter = 5500$ .

#### 4.3.7. Regressão de *Ridge*

A Tabela 13 apresenta os hiperparâmetros escolhidos para busca da melhor modelagem utilizando o algoritmo de Regressão de *Ridge*.

Tabela 13 - Hiperparâmetros avaliados para algoritmo Regressão de *Ridge*.

ALGORITMO	REGRESSÃO DE RIDGE
Hiperparâmetro	Valor
alpha	0.001, 0.01, 0.1, 1, 10, 100, 1000

Fonte: Elaborado pelo autor.

Por fim, a busca exaustiva em grids também foi aplicada no algoritmo de Regressão de *Ridge*. A melhor combinação de hiperparâmetros para a base normalizada por reescala alcançou um  $MAE = 1518,99$  com  $alpha = 0,1$ . A melhor combinação de hiperparâmetros para a base normalizada por padronização alcançou um  $MAE = 1527,62$  com  $alpha = 1$ .

#### 4.4. RESULTADOS PARA ANÁLISE PREDITIVA

A etapa que antecedeu a análise preditiva consistiu em observar os modelos os 175 modelos criados por meio da combinação das grades de hiperparâmetros, avaliando

os que alcançaram o melhor desempenho preditivo, medido pelo *MAE*. Para isto, o grid com melhor resultados para cada algoritmo e base de dados foi posicionado em um *ranking*, conforme a Tabela 14.

Tabela 14 - Ranking por MAE dos melhores parâmetros para cada algoritmo.

RESUMO DA MODELAGEM				
RANKING	ALGORITMO	BASE NORMALIZADA	MELHORES HIPERPARÂMETROS	MAE (T)
1	Florestas Aleatórias	Por padronização	{'max_depth': 15, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 500}	246,89
2	Florestas Aleatórias	Por reescala	{'max_depth': 15, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 500}	247,94
3	Árvores de Decisão	Por padronização	{'max_depth': 8, 'max_leaf_nodes': 100, 'min_samples_leaf': 20, 'min_samples_split': 40}	777,08
4	Árvores de Decisão	Por reescala	{'max_depth': 8, 'max_leaf_nodes': 100, 'min_samples_leaf': 20, 'min_samples_split': 10}	777,43
5	<i>MLP</i>	Por padronização	{'hidden_layer_sizes': 45, 'max_iter': 5500}	789,17
6	Regressão de Ridge	Por reescala	{'alpha': 0.1}	1518,99
7	Regressão de Ridge	Por padronização	{'alpha': 1}	1527,62
8	<i>MLP</i>	Por reescala	{'hidden_layer_sizes': 35, 'max_iter': 5500}	1546,99
9	<i>kNN</i>	Por padronização	{'n_neighbors': 7, 'p': 1}	1668,41
10	<i>SVR</i>	Por padronização	{'C': 100, 'gamma': 0.01}	2686,13

11	<i>kNN</i>	Por reescala	{'n_neighbors': 17, 'p': 1}	3113,03
12	<i>SVR</i>	Por reescala	{'C': 100, 'gamma': 0.01}	5934,22

Fonte: Elaborado pelo autor.

De acordo com a Tabela 14, o modelo que obteve o melhor desempenho para o padrão dos dados avaliados, foi gerado com o algoritmo Florestas Aleatórias, com *MAE* = 246,89 t. Entretanto, esse resultado de desempenho foi gerado na etapa de treinamento e teste por meio da técnica de validação cruzada com 10-*fold*s, sendo então necessário avaliar o desempenho do modelo quando submetido a novas situações reais.

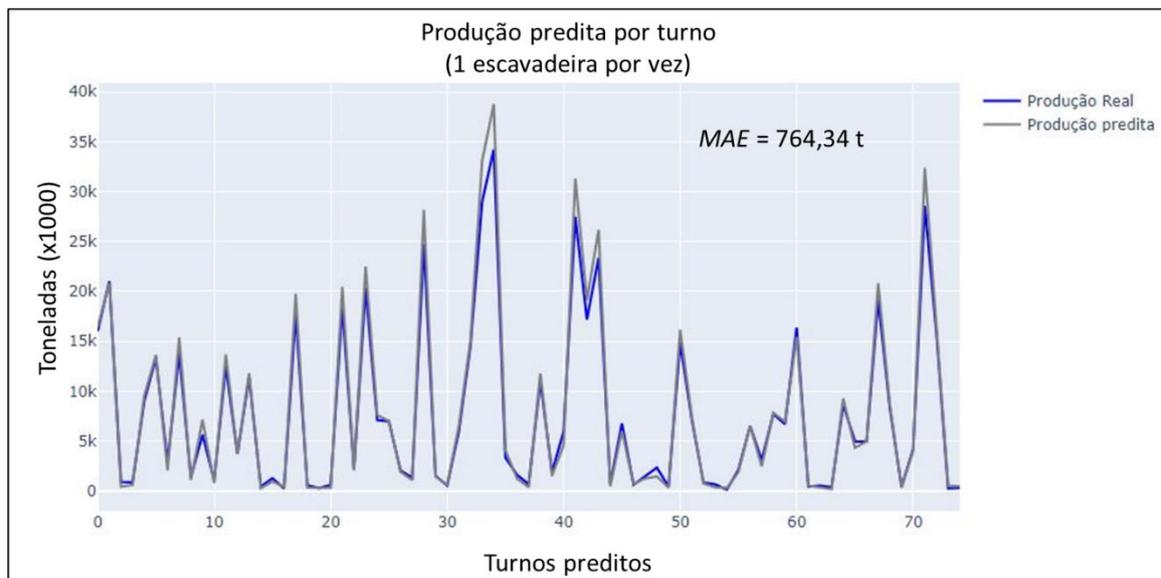
Assim, para que fosse possível avaliar o desempenho do modelo em ambiente de produção, foram selecionados aleatoriamente 75 turnos ao longo do ano de 2018 para quantificar a capacidade do modelo de prever a produção de cada escavadeira no curtíssimo prazo. Ou seja, o modelo precisaria ser capaz de prever, com um baixo erro associado, qual seria a massa produzida no turno por qualquer um dos equipamentos da frota de escavadeiras (ES3001, ES3101, ES3102 e PM3401), considerando as condições operacionais de cada turno como o local de carregamento (cota da bancada que o equipamento será alocado), a qualidade do material a ser manuseado (Granito, Saprolito, Sulfetado de baixo, alto e super-alto teor e Outros), as horas programadas para o turno, as horas previstas de manutenção, a previsão de horas trabalhadas, a turma de operadores escalada para o turno (Turmas A, B, C e D), o respectivo turno a ser previsto (Diurno, Tarde, Noturno), as paradas programadas, etc.

O modelo preditivo recebeu as informações de entrada supracitadas para cada um dos 75 turnos e, a partir de sua capacidade de generalização, retornou uma previsão de produção pra cada cenário de turno planejado para as escavadeiras.

Esta etapa de validação do modelo que foi anteriormente treinado e testado foi chamada de *Back-test*, que consistiu em realizar as previsões dos dados futuros e desconhecidos e, posteriormente, a depender dos resultados obtidos, verificar o erro médio absoluto das previsões.

A Figura 30 apresenta a curva de previsão gerada pelo sistema criado, bem como a curva que apresenta os resultados reais para o turno em que o modelo foi aplicado.

Figura 30 - Resultados preditivos para 75 turnos de operação.



Fonte: Elaborado pelo autor.

Conforme apresentado na Figura 30, o modelo obteve um bom comportamento preditivo, conseguindo gerar uma curva de previsões bastante aderente à curva gerada pelos resultados reais de produção das escavadeiras por turno. As 75 previsões realizadas apresentaram um erro médio absoluto de 764 toneladas. Este resultado de *MAE* comprova que o modelo, quando submetido a um ambiente real de operação, fornece um bom e confiável desempenho.

#### 4.5. RESULTADOS PARA TOMADA DE DECISÃO NO CICLO DE PLANEJAMENTO DE LAVRA

A metodologia proposta foi finalizada com uma rotina de avaliação dos resultados preditivos, seguidos de uma tomada de decisão pelos usuários do sistema preditivo. Tais decisões são associadas ao planejamento do turno, a partir da observação das previsões realizadas pelo modelo. Ou seja, ao verificar o valor predito de produção para uma escavadeira, o planejador pode observar quais os turnos que necessitam de intervenções operacionais, em função do valor predito estar abaixo das metas estabelecidas.

Para testar a eficácia do método de tomada de decisão, foram escolhidos 2 turnos que obtiveram previsão produtiva baixa. O valor planejado inicialmente para as variáveis em cada um dos turnos está descrito nas Tabelas 15 e 16.

Tabela 15 - Cenário para o 1º turno previsto com baixa produção.

VARIÁVEL	CENÁRIO
	1
TURNO PLANEJADO	Diurno
TURMA ESCALADA	Turma C
EQUIPAMENTO A SER ALOCADO	ES3102
LOCAL DE CARREGAMENTO - COTA (m)	24
QUALIDADE – LITOLOGIA	Granito
HORA TOTAL – PROGRAMADA (seg)	6.00
MANUTENÇÃO PREVENTIVA (seg)	0
MANUTENÇÃO CORRETIVA (seg)	0
PARADA TROCA DE TURNO (seg)	0
FROTA DE TRANSPORTE	Cat 785C
HORA TRABALHADA (h)	3.00

Fonte: Elaborado pelo autor.

Tabela 16 - Cenário para o 2º turno previsto com baixa produção.

VARIÁVEL	CENÁRIO
	1
TURNO PLANEJADO	Tarde
TURMA ESCALADA	Turma B
EQUIPAMENTO A SER ALOCADO	ES3101
LOCAL DE CARREGAMENTO – COTA (m)	40
QUALIDADE - LITOLOGIA	Granito
HORA TOTAL – PROGRAMADA (seg)	9.00
MANUTENÇÃO PREVENTIVA (seg)	1
MANUTENÇÃO CORRETIVA (seg)	0
PARADA TROCA DE TURNO (seg)	1
FROTA DE TRANSPORTE	Cat 785C
HORA TRABALHADA (h)	6.00

Fonte: Elaborado pelo autor.

O resultado preditivo para o 1º turno foi de 1245,02 t produzidos pela escavadeira ES3102, enquanto que para o 2º turno, o resultado preditivo foi de 2512,50 t para a

ES3101. Assim, foram tomadas decisões de intervenções a fim de mudar o primeiro cenário planejado para um segundo cenário planejado que retorne resultados preditivos mais satisfatórios. Nas Tabelas 17 e 18 são apresentados os cenários modificados a partir das intervenções previstas (alterações com destaque em verde).

Tabela 17 - Cenário modificado para o 1º turno previsto com baixa produção.

VARIÁVEL	CENÁRIO 1 (MODIFICADO)
TURNO PLANEJADO	Diurno
TURMA ESCALADA	Turma C
EQUIPAMENTO A SER ALOCADO	ES3001
LOCAL DE CARREGAMENTO – COTA (m)	24
QUALIDADE – LITOLOGIA	Granito
HORA TOTAL – PROGRAMADA (seg)	9.00
MANUTENÇÃO PREVENTIVA (seg)	0.00
MANUTENÇÃO CORRETIVA (seg)	0.00
PARADA TROCA DE TURNO (seg)	0.00
FROTA DE TRANSPORTE	Cat 793C
HORA TRABALHADA (horas)	7.00

Fonte: Elaborado pelo autor.

Tabela 18 - Cenário modificado para o 2º turno previsto com baixa produção.

VARIÁVEL	CENÁRIO 2 (MODIFICADO)
TURNO PLANEJADO	Tarde
TURMA ESCALADA	Turma A
EQUIPAMENTO A SER ALOCADO	ES3101
LOCAL DE CARREGAMENTO – COTA (m)	40
QUALIDADE - LITOLOGIA	Granito
HORA TOTAL – PROGRAMADA (seg)	9.00
MANUTENÇÃO PREVENTIVA (seg)	1
MANUTENÇÃO CORRETIVA (seg)	0
PARADA TROCA DE TURNO (seg)	0
FROTA DE TRANSPORTE	Cat 793C
HORA TRABALHADA (horas)	7.00

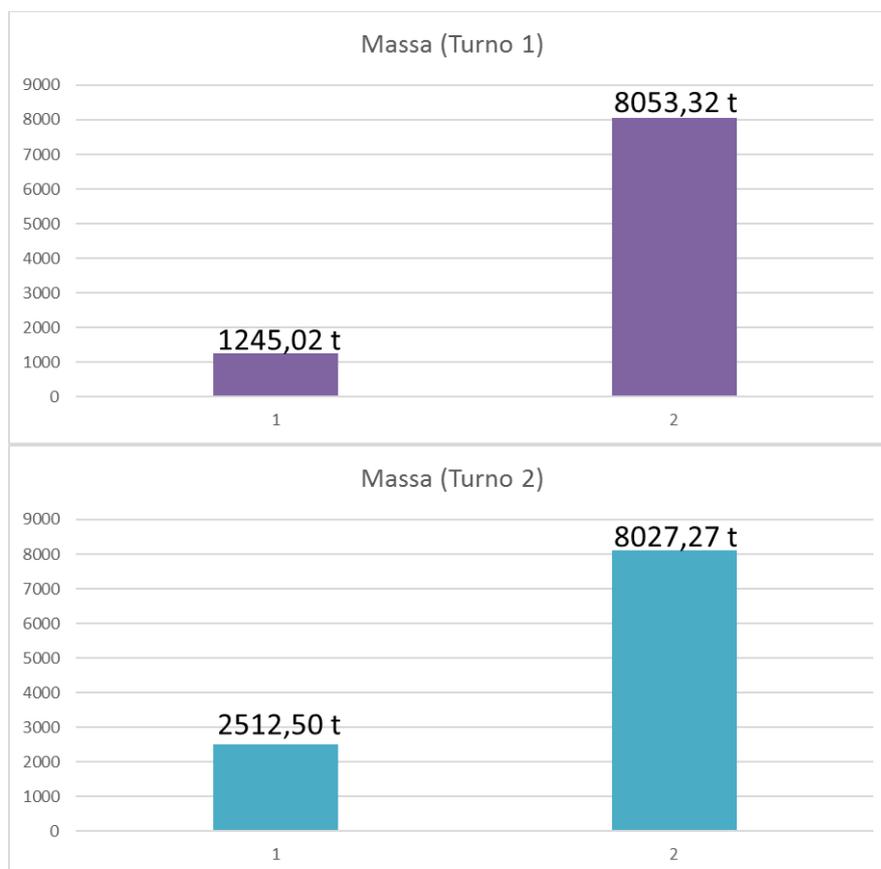
Fonte: Elaborado pelo autor.

As mudanças realizadas em cada variável foram tomadas por base nas informações obtidas na análise descritiva. Ou seja, os valores foram modificados para que fosse possível alcançar maiores níveis de produção, tomando por base os padrões que as observações estatísticas revelaram.

Inicialmente, para o primeiro turno a ser exemplificado, se optou por mudar o equipamento da ES3102 para a ES3001, além de mudar a programação da limpeza da frente do turno noturno (6h programadas) para o turno diurno ou tarde (9h programadas). Outra relevante mudança foi alocar na escavadeira de maior porte a Frota de Transporte Cat 793C, que possui melhor ajuste de passes para a ES3001 (ambos com maiores capacidades). Para o segundo turno exemplificado, se optou por mudar a turma de trabalho, da Turma B para a Turma A, tendo em vista que, conforme a análise descritiva, a segunda possui melhor desempenho no alcance de massa. Por se tratar da programação de operação em frente de trabalho de material estéril (Granito), que naturalmente estão localizadas em praças mais amplas, optou-se pela escolha da Frota de Transporte Cat 793C, com maior capacidade e melhor acerto entre de carga entre os caminhões e a concha da escavadeira. Também se optou por cancelar uma programação de manutenção preventiva, convertendo esta hora em tempo de trabalho efetivo.

Estas alterações causaram melhorias na predição da massa produzida, conforme ilustrado na Figura 31.

Figura 31 - Resultados para cenários modificados para os turnos planejados.



Fonte: Elaborado pelo autor.

É possível observar que o resultado preditivo para o 1º turno foi de 1245 t produzidos pela escavadeira ES3102, enquanto que para o 2º turno, o resultado preditivo foi de 2512 t para a ES3101. Assim, foram tomadas decisões de intervenções a fim de mudar o primeiro cenário planejado para um cenário planejado que retorne resultados mais satisfatórios. Conforme apresentado na Figura 31, após a modificação dos cenários de alocação de equipamentos, foram alcançados os valores preditos de 8053 t e 8027 t para os cenários 1 e 2, respectivamente.

## 5. DISCUSSÃO

A discussão dos resultados desta pesquisa é conduzida pelos 2 questionamentos de pesquisa, apresentados no Capítulo 1, a fim de conectar os resultados obtidos com as questões que fomentaram o desenvolvimento da metodologia proposta.

Questão de pesquisa 1:

- Qual o impacto da análise preditiva no planejamento de lavra de curto prazo?

Questão de pesquisa 2:

- Quais os desafios e benefícios para a implantação de um sistema preditivo no planejamento de curto prazo?

Para que o planejamento de lavra seja executado de forma consistente, é imprescindível que os planos elaborados sejam gerados em tempo hábil, com confiabilidade e de forma realística, para que a sua execução seja garantida.

Entretanto, estes requisitos se tornam desafios ainda maiores em função dos avanços tecnológicos sofridos pela indústria mineral nos últimos anos. O uso de diversos sistemas computacionais, associados ao aumento significativo da geração de dados operacionais e o aumento da capacidade de processamento de informações contribuem diretamente com a elaboração de planos cada vez mais acertados.

Esta afirmação foi sustentada por meio da metodologia apresentada nesta pesquisa e se confirmou com os resultados obtidos. Por meio desta metodologia, milhares de dados operacionais de equipamentos de escavação foram coletados e selecionados, e por meio da capacidade de processamento dos computadores atuais, este gigantesco universo de dados foi tratado e manipulado, de modo a fornecer informações preditivas completamente próximas da realidade.

Neste contexto, foram treinados e testados 175 modelos, a fim de selecionar aquele que melhor se ajustou ao padrão de dados, ao ponto de obter um baixo erro de predição. O modelo com melhor desempenho obteve um Erro Médio Absoluto de 246,89 t, em um cenário que contém equipamentos de grande porte e alta capacidade de produção. Este modelo passou por um processo de validação, a fim de confirmar a sua capacidade de atuar em ambiente real de mina.

O alcance de um  $MAE = 764,34$  t ao ser utilizado em 75 situações operacionais reais comprova que o modelo preditivo criado tem capacidade suficiente de impactar a rotina de análise de equipes de planejamento de curto prazo pelo fato do modelo gerar

informações valiosas de forma rápida, confiável e executável para equipes de planejamento de curto e curtíssimo prazo, que por sua vez são horizontes de planejamento que possuem um alta quantidade de variáveis associadas, bem como a necessidade de tomadas de decisão seguras e ágeis.

Entretanto, existem inúmeros desafios para que sistemas preditivos sejam implementados na rotina de planejamento de curto prazo. Dentre eles está o alcance de um histórico de dados operacionais, antes de tudo, confiável. Para isso, é primordial que os sistemas e rotinas de controle, monitoramento e gestão de frotas sejam suficientemente consistentes e assegurem a geração de informações com alto índice de confiabilidade.

Um outro desafio para esta implementação, é a de que os atores de planejamento precisam estar adaptados e capacitados a trabalhar inseridos na cultura *Data-driven* (cultura dirigida por dados). Isto é, os dados precisam ser encarados, não mais, como fonte de controle operacional, mas principalmente, fonte de aumento da competitividade do negócio.

Desse modo, os gestores precisam entender que cada informação gerada por um modelo preditivo confiável, guiará a construção de planos de lavra cada vez mais realísticos, aumentando constantemente a aderência da operação de lavra a estes planos.

Um outro desafio, é criar um ambiente com infraestrutura adequada para esta realidade. Cada vez mais serão exigidas maiores capacidades de armazenamento e processamento de dados, bem como ambientes colaborativos que permitam que toda a equipe faça parte de um ambiente interativo, dinâmico e eficiente.

Portanto, confirma-se que o Processo Padrão de Ciência de Dados para Planejamento de Lavra (metodologia proposta) impacta positivamente o planejamento de curto e curtíssimo prazo, tendo em vista que por meio de sistemas preditivos, como o criado nesta pesquisa, facilitam a tomada de decisão no que diz respeito à alocação de equipamentos, considerando suas devidas condições operacionais, a definição e alcance das metas estabelecidas, bem como a utilização de método que considere ao mesmo tempo a grande quantidade de variáveis associadas à operação dos equipamentos (informações geológicas, geográficas, política de horas, turmas de trabalho atuando, horários de produção, tipos específicos de equipamentos, etc).

Além disso, os benefícios gerados vão além do desempenho operacional de mina. Compreende-se que os ganhos em eficiência aumentam a competitividade do negócio,

reduzindo despesas e, conseqüentemente, aumento nos lucros associados ao empreendimento mineral.

## 6. CONCLUSÕES

O Processo Padrão de Ciência de Dados para Planejamento de Lavra, metodologia criada nesta pesquisa, alcançou todos os objetivos pré-estabelecidos e apresentou facilidades para ser implementado na rotina operacional de lavra, conforme as conclusões que são apresentadas a seguir.

Comprovou-se que por meio do sistema preditivo desenvolvido, bem como a metodologia proposta, resulta no aumento da previsibilidade associada ao desempenho operacional de equipamentos de mina.

- A execução operacional de mina ocorreu com grande aderência ao que foi definido anteriormente na predição dos turnos. O modelo preditivo foi treinado e testado, realizando predições realísticas sobre a produção de 4 equipamentos de carga, considerando as mais diversas condições de operação;
- Foi possível prever o desempenho de produção de equipamentos no horizonte de curto prazo, mesmo com tantas variáveis envolvidas no processo de produção. Os 75 turnos escolhidos para validação foram preditos com o baixo erro médio absoluto de aproximadamente 765 t;
- Por meio da rotina de análise criada, foram avaliadas e selecionadas 30 variáveis que mais impactam e explicam o padrão operacional produtivo das escavadeiras estudadas;
- As incertezas operacionais podem ser diretamente reduzidas, bem como o fornecimento de informações úteis foram alcançados de forma rápida;
- A integração de dados de operação de lavra, planejamento, geologia e manutenção possibilita a tomada de decisão por meio de visão holística, melhorando a colaboração entre as áreas de atuação.

## REFERÊNCIAS

ABELLÁN, J., MANTAS, C., CASTELLANO, J., MORAL-GARCÍA, S. Increasing diversity in random forest learning algorithm via imprecise probabilities. **Expert Systems With Applications**, v. 97, p. 228-243, 2018.

AGHAEIPOOR, F., JAVIDI, M.M. A hybrid fuzzy feature selection algorithm for high-dimensional regression problems: An mRMR-based framework. **Expert Systems With Applications**, v. 162, 2020.

AGWU, O.E., AKPABIO, J.U., DOSUNMU, A. Modeling the downhole density of drilling muds using multigene genetic programming. **Upstream Oil and Gas Technology**, v.6. 2021.

ALI, M., JUNG, L. T., ABDEL-ATY, A., ABUBAKAR, M. Y., ELHOSENY M., ALI, I. Semantic-k-NN algorithm: An enhanced version of traditional k-NN algorithm. **Expert Systems With Applications**, v. 151, 2020.

AMANKWAH, H. **Mathematical Optimization Models and Methods for Open-Pit Mining**. 2011. 50 p. Dissertação (Mestrado) - Department of Mathematics, Linköping University, Linköping, Sweden, 2011.

AMBIKA, P. Machine learning and deep learning algorithms on the Industrial Internet of Things (IIoT). **Advances in Computers**, v. 117, p. 321-338, 2020.

AMO, S. "**Técnicas de mineração de dados.**" Jornada de Atualização em Informática. Uberlândia, p. 1 – 43, 2004.

ANTONIADIS, A., LAMBERT-LACROIX, S., POGGI, Jean-Michel. Random forests for global sensitivity analysis: A selective review. **Reliability Engineering and System Safe**, v. 206, 2021.

ARAÚJO, Francisco César Rodrigues de. **Planejamento operacional de lavra com alocação dinâmica de caminhões: abordagens exata e heurística**. 2008. 135 p. Dissertação (Mestrado) – Departamento de Engenharia de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2008.

BAEK, J.; CHOI, Y. **Simulation of truck haulage operations in an underground mine using big data from an ict-based mine safety management system**. Applied sciences, v. 9, 2019.

BAUMANN, P. HOCHBAUM, D.S. YANG, Y.T. A comparative study of the leading machine learning techniques and two new optimization algorithms. **European Journal of Operational Research**, v. 272, p. 1041-1057, 2019.

BAZANTE, A. J. **A interação entre o planejamento e controle de qualidade na maximização da função benefício de um empreendimento mineral**. 2004. 123 p. Tese (Doutorado) - Escola Politécnica, Universidade de São Paulo, 2004.

BERGERMAN, M. G. **Modelagem e simulação do circuito de moagem do Sossego**. 2009. 207 p. Dissertação (Mestrados) - Escola Politécnica, Universidade de São Paulo, 2009.

BERGSTRA, J., BENGIO, Y. Random Search for Hyper-Parameter Optimization. **Journal of Machine Learning Research**, v. 13, p. 281-305, 2012.

BLOM, M., PEARCE, A.R., STUCKEY, P.J. Short-term planning for open pit mines: a review. **International Journal of Mining, Reclamation and Environment**, v. 33, p. 318-339. 2018.

BOSCARIOLI, C.; SILVA, L. A.; HERNANDEZ, E. D. M. Uso de Redes Neurais Artificiais na Descoberta de Conhecimento a partir de Dados e Imagens de Sensores. **Revista Iberoamericana de Sensores**, v.6, n. 3, p. 629-635, 2008.

BRAGA, A.P.; LUDERMIR, T.B.; CARVALHO, A.C.P.L.F. Redes Neurais Artificiais: Teoria e Aplicações. Livros Técnicos e Científicos S.A., 2000.

BRANDÃO, R.; DE TOMI, G. Metodologia para estimativa e gestão da produtividade de lavra. **Rev. Esc. Minas**, v.64, n.1, p.77-83, 2011.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5-32, 2001.

CAMILO, C. O.; SILVA, J. C. **Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas** (Relatório Técnico). Goiânia, 2009.

CAMPELO, A. C. M. M.; MARIN, T.; OLIVEIRA, M. L.; SANTOS, C. W. F.; MENIN, R. Q.; SILVA, N. H. C.; LACERDA, A. L. S. Utilização de dados do sistema de despacho para estimativa de produtividade de transporte no plano de lavra de curto prazo. **Tecnol. Metal. Mater. Miner.**, São Paulo, v. 15, n. 2, p. 86-90, 2017.

COHEN, S. The basics of machine learning: strategies and techniques. In.: COHEN, S. **Artificial intelligence and deep learning in pathology**. Elsevier, 2020. p. 13-40.

CÔRTEZ, S. C.; PORCARO, R. M.; LIFSCHITZ, S. **Mineração de Dados – Funcionalidades, Técnicas e Abordagens**. PUC-RJ. Artigo, 2002 Disponível em: [ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02\\_10\\_cortes.pdf](ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_10_cortes.pdf). Acesso em: Ago. de 2018.

COSTA, E. P. **Investigação de técnicas de classificação hierárquica para problemas de bioinformática**. 2008. 160 p. Dissertação (Mestrado) – Instituto de Ciências Matemáticas e de Computação – ICM – USP.

COSTA, F. V. **Análise dos principais indicadores de desempenho usados no planejamento de lavra**. 2015. 98 p. Dissertação (Mestrado) – Departamento de Engenharia de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2015.

CREMONESE, D. T. **Desenvolvimento e aplicação de metodologia para estudos de simulação dinâmica na cadeia do minério de ferro**. 2014. 152 p. Dissertação (Mestrado) - Escola Politécnica, Universidade de São Paulo, 2014.

CZAJKOWSKI, M., KRETOWSKI, M. Decision Tree Underfitting in Mining of Gene Expression Data. Na Evolutionary Multi-Test Tree Approach. **Expert Systems With Applications**, v. 137, p. 392-404, 2019.

DAS, S.; DEY, A.; PAL, A.; ROY, N. Applications of Artificial Intelligence in Machine Learning: Review and Prospect. **International Journal of Computer Applications**, v. 115, N. 9, p. 31-41, 2015

DAVIS, G.; NEWMAN, A. Modern Strategic Mine Planning. **CRC Mining Conference**, p. 13, 2008

DIAS, T. F. **Avaliação de indicadores operacionais: estudo de caso de uma empresa do setor ferroviário**. 2008. 43 p. Monografia - Universidade Federal de Juiz de Fora, Juiz de Fora, 2008.

DIAZ, G.I., FOKOUE-NKOUTCHE, A., NANNICINI, G., SAMULOWITZ, H. An effective algorithm for hyperparameter optimization of neural networks. **IBM Journal of Research and Development**, v. 61, 2017.

DONATE, J. P., CORTEZ, P., SÁNCHEZ, G. G., MIGUEL, A. S. Time series forecasting using a weighted cross-validation evolutionary artificial neural network ensemble. **Neurocomputing**, v. 109, p. 27-32, 2013.

EFFROSYNIDIS, D., ARAMPATZIS, A. An evaluation of feature selection methods for environmental data. **Ecological Informatics**, v. 61, 2021.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F. **Inteligência artificial: Uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2:192. (2011).

FATHI, E., SHOJA, B.M. Chapter 9 - Deep Neural Networks for Natural Language Processing. In.: GUDIVADA, V.N., RAO, C.R. **Handbook of Statistics 38: Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications**, p. 229-316, 2018.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, v. 17, n. 3, p. 37-54, 1996.

FENG, X., MA, G., SU, S., HUANG, C., BOSWELL, M.K., XUE, P. A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan. **Ocean Engineering**, v. 211, 2020.

FERRERO, C. A. **Algoritmo kNN para previsão de dados temporais: funções de previsão e critérios de seleção de vizinhos próximos aplicados a variáveis**

**ambientais em limnologia.** 2009. 103 p. Instituto de Ciências Matemáticas e de Computação - ICMC-USP, São Carlos, 2009.

FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, p. 185 – 201, 2002.

HUBER, S., WIEMER, H., SCHNEIDER, D., IHLENFELDT, S. DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model. **Procedia CIRP**, v. 79, p. 403-408, 2019.

HU, J., PENG. H., WANG, J., YU, W. kNN-P: A kNN classifier optimized by P systems. **Theoretical Computer Science**, v. 817, p. 55-65, 2020.

HUSSAIN, S.F. A novel robust kernel for classifying high-dimensional data using Support Vector Machines. **Expert Systems With Applications**, v. 131, p. 116-131, 2019.

ISKENDEROĞLU, F. C., BALTACIOĞLU, M. K., DEMIR, M. H., BALDINELLI, A., BARELLI, L., BIDINI, G. Comparison of support vector regression and random forest algorithms for estimating the SOFC output voltage by considering hydrogen flow rates. **International Journal of Hydrogen Energy**, v. 45, p. 35023-35028, 2020.

JOHANSSON, B.; JOHANSSON, J.; ABRAHAMSSON, L. Attractive workplaces in the mine of the future: 26 statements. **International Journal of Mining and Mineral Engineering**, v. 2, n. 3, p. 239–252, 2010.

KART, L. Market Trends: Big Data Opportunities in Vertical Industries, **Industry Market Strategies Worldwide**, Gartner, 2012.

KELEMEN, J. From artificial neural networks to emotion machines with marvin Minsky. **Acta Polytechnica Hungarica**, v. 4, n. 4, p. 5-16, 2007.

KIM, D., KANG, S., CHO, S. Expecte d margin–base d pattern selection for support vector machines. **Expert Systems With Applications**, v. 139, 2020.

KOULINAS, G., PARASCHOS, P., KOULOURIOTIS, D. A decision trees-based knowledge mining approach for controlling a comples production system. **Procedia Manufacturing**, v. 51, p. 1439-1445, 2020.

KWIRI, J.; GENÇ. B. Mine planning and optimization techniques used in surface mining – Position Paper. **25th International Mining Congress and Exhibition of Turkey**, At Antalya, Turkey, p. 210 – 221, 2017.

LACERDA, A. L. S. Simulação aplicada a sistema de despacho como forma de avaliação da influência da otimização e integridade de dados na Mina do Sossego. 2019. 75 p. Monografia (Especialização latu sensu) – Instituto Tecnológico Vale, 2019.

LAROSE, D. T. **Discovering Knowlege in Data – An Introduction to Data Mining**. New Jersey: A john wiley & sons, inc., 2005.

LECUN, Y. BENGIO, Y. HINRON. Deep Learning. *NATURE*, v 521, p. 436 – 444, 2015.

LIN, G., LIN, A., CAO, J. Multidimensional KNN algorithm based on EEMD and complexity measures in financial time series forecasting. **Expert Systems With Applications**, v. 168, 2021.

LIPOVETZKY, N.; BURT, C.; PEARCE, A.; STUCKEY, P. Planning for Mining Operations with Time and Resource Constraints. **24th Int. Conf. on Aut. Planning and Scheduling, ICAPS**, p. 9, 2014.

LIU, S., REVIRIEGO, P., MONTUSCHI, P., LOMBARDI, F. Less-is-Better Protection (LBP) for memory errors in kNNs classifiers. **Future Generation Computer Systems**, v. 117, p. 401-411. 2021.

MALHOTRA, Y. P.; LIST, S. J. Short term mine planning at Syncrude. **International Journal of Surface Mining, Reclamation and Environment**, v. 3, p. 11 – 20, 1989.

MCCARTHY, J. What is artificial intelligence? Computer Science Department, Stanford University, p 1-15, 2007.

MCCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mothemnticnl Biology**, v. 52, n. l/2. p. 99-115, 1990.

MEIER H, LAGEMANN H, MORLOCK F, RATHMANN C. Key performance indicators for assessing the planning and delivery of industrial services. **Procedia CIRP**, v. 11, p. 99 – 104, 2013.

MENG, X., ZHANG, P., XU, Y., XIE, H. Construction of decision tree based on C4.5 algorithm for online voltage stability assessment. **Electrical Power and Energy Systems**, v. 118, 2020.

MIENYE, I.D., SUN, Y., WANG, Z. Prediction performance of improved decision tree-based algorithms: a review. **Procedia Manufacturing**, v. 35, p. 698-703, 2019.

MOHAMMADY, S., ESLAHCHI, M.R. Extension of Tikhonov regularization method using linear fractional programming. **Journal of Computational and Applied Mathematics**, v. 371, 2020.

MUTASA, S., SUN, S., HA, R. Understanding artificial intelligence based radiology studies: What is overfitting? **Clinical Imaging**, v. 65, p. 96-99, 2020.

NADER, B.; DE TOMI, G.; PASSOS, A. O. Indicadores-chave de desempenho e a gestão integrada da mineração. **REM - Revista Escola de Minas**, v. 65, n. 4, p.537-542, 2012.

OSANLOO M.; RAHMANPOUR, M. Optimizing short-term production plan using a portfolio optimization model. **REM – International Engineering Journal**, v. 70, n. 1, p. 109-116, 2017

PATEL, B. N.; PRAJAPATI, S. G.; LAKHTARIA, K. I. Efficient Classification of Data Using Decision Tree. **Bonfring International Journal of Data Mining**, V. 2, N. 1, P. 6-12, 2012.

PERAL, J.; MATÉ.; A.; MARCO, M. Application of Data Mining techniques to identify relevant Key Performance Indicator. **Computer standards and Interfaces**, v. 50, p. 55-64, 2017.

PETERMANN, F. J. **Modelo de Mineração de Dados para Classificação de Clientes em Telecomunicação**. 2006. 164 p. Dissertação (Mestrado) – Departamento de Engenharia Elétrica, Pontifca Universidade Católica, Porto Alegre, 2006.

PRINA, B. Z; TRENTIN, R. GMC: Geração de Matriz de Confusão a partir de uma classificação digital de imagem do ArcGIS®. **XVII Simpósio Brasileiro de Sensoriamento Remoto – SBSR**, João Pessoa, Brasil, p. 131 – 139, 2015.

PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. **Big Data**, v. 1, n. 1, p. 51-59, 2013.

QUEVEDO, J. M. G. **Modelo de Simulação para o sistema de carregamento e transporte em mina a céu aberto**. 2009. 133 p. Dissertação (Mestrado) – Departamento de Engenharia Industrial, Pontifca Universidade Católica, Rio de Janeiro, 2009.

RAMEZARIAN, R., PEYMANFAR, A., EBRAHIMI, S.B. An integrated framework of genetic network programming and multi-layer perceptron neural network for prediction of daily stock return: An application in Tehran stock exchange market. **Applied Soft Computing Journal**, v. 82, 2019

ROSENBLATT, F. **The Perceptron — A Perceiving and Recognizing Automaton**. 1957. 33 p. Tese (Doutorado) - Cornell Aeronautical Laboratory, New York, 1957.

ROUMPOS, C.; PAPAOKOSTA, E. Strategic mine planning of surface mining projects incorporating sustainability concepts. **6th International Conference on Sustainable Development in the Minerals Industry (SDIMI 2013)**, At Milos Island, Greece, p. 645 – 651, 2013.

ROUMPOS, C.; PAVOUDAKIS, F.; KOURIDOU, O. Managing Surface Lignite Mining Projects in a Competitive Environment. **IX International Brown Coal Mining Congress**, At Bełchatów, Poland, P. 435 – 444, 2016.

SABZEKAR, M., HASHEMINEJAD, S.M.H. Robust regression using support vector regressions. **Chaos, Solitons & Fractals**, v. 144, 2021.

SACHS, P. F. T.; NADER, B. **Sistemas de Gestão da Produção e a Cadeia de Valor Mineral**. (Trabalho Técnico). São Paulo, 2005, 16 p.

SAMUEL, A. L. Some Studies in Machine Learning Using the Game of Checkers. **IBM journal**, p. 211-229, 1959.

SARACOGLU, O. G. An Artificial Neural Network Approach for the Prediction of Absorption Measurements of an Evanescent Field Fiber Sensor. **Sensors**, v. 8, p. 1585 – 1594, 2008.

SCIKIT-LEARN. Scikit-learn 1.6.3. Nearest Neighbors Regression. **Machine learning in Python**, n.d.a. Disponível em: <https://scikit-learn.org/stable/modules/neighbors.html#nearest-neighbors-regression>. Acesso em: 04 Mar. 2021.

SCIKIT-LEARN. Scikit-learn 1.13.2. Univariate feature selection. **Machine learning in Python**, n.d.b. Disponível: [https://scikit-learn.org/stable/modules/feature\\_selection.html](https://scikit-learn.org/stable/modules/feature_selection.html). Acesso em: 28 Fev. 2021.

SCIKIT-LEARN. Scikit-learn 3.1. Cross-validation: evaluating estimator performance. **Machine learning in Python**, n.d.c. Disponível: [https://scikit-learn.org/stable/modules/cross\\_validation.html#cross-validation-and-model-selection](https://scikit-learn.org/stable/modules/cross_validation.html#cross-validation-and-model-selection). Acesso em: 05 Mar. 2021.

SOMUA-GYIMAH, G.; FRIMPONG, S.; NYAABA, W.; GBADAM, E. **A computer vision system for terrain recognition and object detection tasks in mining and construction environments**. Proceedings of the annual conference of the society for mining, metallurgy & exploration (SME), 2019.

SILVA, N. C. S. **Otimização de cava em minas de calcário para cimento**. 2001.154 p. Dissertação (Mestrado) - Escola Politécnica, Universidade de São Paulo, 2001.

SILVA, N. C. S. **Metodologia de planejamento estratégico de lavra incorporando riscos e incertezas para a obtenção de resultados operacionais**. 2008. 128 p. Tese (Doutorado) - Escola Politécnica, Universidade de São Paulo, 2008.

SILVA, R. S. **Aplicação de tecnologia wireless para controle de qualidade do planejamento de lavra incorporando a incerteza geológica**. 2006. 97 p. Dissertação (Mestrado) - Escola Politécnica, Universidade de São Paulo, 2006.

SILVA, R. S.; DE TOMI, G.; BEZERRA, E. B.; CHAVES, A. P. A mineração de dados como suporte ao planejamento de lavra. **9º Congresso Brasileiro de Mina a Céu Aberto e Mina Subterrânea**. Minas Gerais, Belo Horizonte, p 15, 2018.

SMITH, M. L.; WICKS, S. J. Medium-Term Production Scheduling of the Lumwana Mining Complex. **Interfaces**, v. 44. N. 2, p. 176 – 194, 2014.

SPEISER, J. L., MILLER, M. E., TOOZE, J., IP, E. A comparison of random forest variable selection methods for classification prediction modeling. **Expert Systems With Applications**, v. 134, p. 93-101, 2019.

SUN, S., CAO, Z., ZHU, H., ZHAO, J. A Survey of Optimization Methods From a Machine Learning Perspective. **IEEE Transactions on Cybernetics**, v. 50, p. 3668-3681, 2020.

TEMKIN, I.; KLEBANOV, D.; DERYABIN, S.; KONOV, I. **Predictive analytics in mining. Dispatch system is the core element of creating intelligent digital mine.** In.: Sukhomlin V.; Zubareva E. Modern information technology and it education. Springer, 2018, p. 365-374.

TSALERA, E., PAPADAKIS, A., SAMARAKOU, M. Monitoring, profiling and classification of urban environmental noise using sound characteristics and the KNN algorithm. **Energy Reports**, v. 6, p. 223-230, 2020.

TURING, A. M. **Computing Machinery and Intelligence.** Mind 49: 433-460. COMPUTING MACHINERY AND INTELLIGENCE.

UPADHYAY, S.P., ASKARI-NASAB, H. Simulation and optimization approach for uncertainty-based short-term planning in open pit mines. **International Journal of Mining Science and Technology**, v. 28, p. 153-166, 2018.

UPADHYAY, S.P., ASKARI-NASAB, H. Truck-shovel allocation optimisation: a goal programming approach. **Mining Technology**, v. 125, p. 82-92, 2016.

VASCONCELOS, J. H. **Investigações sobre métodos de classificação para uso em termografia de mama.** 2017. 93 p. Recife, 2017.

VIEIRA, S., LOPEZ, W. H., MECHELLI, P. A. Chapter 1: Introduction to Machine Learning. In.: MECHELLI, P. A., VIEIRA, S. **Machine learning: methods and applications to brain disorders**, 2019. p. 1-20.

YANG, L., SHAMI, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. **Neurocomputing**, v. 415, p. 295-316, 2020.

YOO, Y. Hyperparameter optimization of deep neural network using univariate dynamic encoding algorithm for searches. **Knowledge-Based Systems**, v. 178, p. 74-83, 2019.

ZHANG, F., SUN, K., WU, X. A novel variable selection algorithm for multi-layer perceptron with elastic net. **Neurocomputing**, v. 361, p. 110-118, 2019.

ZHANG, X. **A Matrix Algebra Approach to Artificial Intelligence.** Cingapura: Springer, 2020. p. 223-440.

ZHOU, H. F., ZHANG, J. W., ZHOU, Y. Q., GUO, X. J., MA, Y. M. A feature selection algorithm of decision tree based on feature weight. **Expert Systems With Applications**, v. 164, 2021.