

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas - Departamento de Ciência da Computação
Programa de Pós-Graduação em Inovação Tecnológica

Douglas Giovane Nunes Pimenta

**INOVAÇÃO PARA EFICIÊNCIA ENERGÉTICA DE CAMINHÕES FORA DE
ESTRADA EM MINA A CÉU ABERTO**

Belo Horizonte

2024

Douglas Giovane Nunes Pimenta

**INOVAÇÃO PARA EFICIÊNCIA ENERGÉTICA DE CAMINHÕES FORA DE
ESTRADA EM MINA A CÉU ABERTO**

Dissertação apresentada ao Programa de Pós-graduação em Inovação Tecnológica da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Mestre em Inovação Tecnológica, área de concentração: Transformação Digital e Inteligência Artificial.

Orientador: Prof. Dr. Adriano Veloso

Coorientador: Prof. Dr. Emérito Nívio Ziviani

Belo Horizonte

2024

Ficha Catalográfica

P644i Pimenta, Douglas Giovane Nunes.
2024 Inovação para eficiência energética de caminhões fora de estrada em mina a céu
D aberto [manuscrito] / Douglas Giovane Nunes Pimenta. 2024.
1 recurso online (78 f. : il., gráfs., tabs., color.) : pdf.

Orientador: Adriano Alonso Veloso.

Coorientador: Nívio Ziviani.

Dissertação (mestrado) – Universidade Federal de Minas Gerais – Departamento de Química (Programa de Pós-Graduação em Inovação Tecnológica).

Bibliografia: f. 75-78.

1. Inovações tecnológicas – Teses. 2. Inteligência artificial – Teses. 3. Indústria mineral – Teses. 4. Mineração a céu aberto – Teses. 5. Desenvolvimento sustentável – Teses. 6. Aprendizado do computador – Teses. 7. Algoritmos genéticos – Teses. 8. Eficiência energética – Teses. I. Veloso, Adriano Alonso, Orientador. II. Ziviani, Nívio, Coorientador. III. Título.

CDU 043



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM INOVAÇÃO TECNOLÓGICA

ATA DA SESSÃO DE DEFESA DA 24ª DISSERTAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM INOVAÇÃO TECNOLÓGICA DA UNIVERSIDADE FEDERAL DE MINAS GERAIS, DO DISCENTE DOUGLAS GIOVANE NUNES PIMENTA Nº DE REGISTRO 2021729502.

Aos 23 (vinte e três) dias do mês de fevereiro de 2024, às 09 horas e 30 minutos, no Auditório II do Departamento de Química da UFMG, reuniu-se, a Comissão Examinadora composta pelos Professores Doutores: Adriano Alonso Veloso do Programa de Pós-graduação em Inovação Tecnológica da UFMG (Orientador), Nivio Ziviani da Universidade Federal de Minas Gerais – UFMG (Coorientador), Allan Claudius Queiroz Barbosa do Programa de Pós-graduação em Inovação Tecnológica da UFMG e Adriano César Machado Pereira do Programa de Pós-graduação em Inovação Tecnológica da UFMG, para julgamento da Dissertação do Mestrado em Inovação Tecnológica - Área de Concentração: Transformação Digital e Inteligência Artificial, do discente Douglas Giovane Nunes Pimenta, Dissertação intitulada: **“INOVAÇÃO PARA EFICIÊNCIA ENERGÉTICA DE CAMINHÕES FORA DE ESTRADA EM MINA A CÉU ABERTO”** O Presidente da Banca abriu a sessão e apresentou a Comissão Examinadora, bem como esclareceu sobre os procedimentos que regem da defesa pública de dissertação. Após a exposição oral do trabalho pelo discente, seguiu-se com arguição pelos membros da Banca Examinadora, com a respectiva defesa do candidato. Finda a arguição, a Banca Examinadora se reuniu, sem a presença do discente e do público, tendo deliberado unanimemente pela sua **APROVAÇÃO**. Nada mais havendo para constar, lavrou-se e fez a leitura pública da presente Ata que segue assinada por mim e pelos membros da Comissão Examinadora e pela coordenação do PPGIT. Belo Horizonte, 23 de fevereiro de 2024.

Professor Doutor Adriano Alonso Veloso (Orientador)
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Nivio Ziviani (Coorientador)
(Universidade Federal de Minas Gerais – UFMG)

Professor Doutor Allan Claudius Queiroz Barbosa
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Adriano César Machado Pereira
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Allan Claudius Queiroz Barbosa
Coordenador do PPG em Inovação Tecnológica da UFMG



Documento assinado eletronicamente por **Allan Claudius Queiroz Barbosa, Professor(a)**, em 11/03/2024, às 14:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Cesar Machado Pereira, Professor do Magistério Superior**, em 11/03/2024, às 20:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Alonso Veloso, Professor do Magistério Superior**, em 18/03/2024, às 15:45, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Nivio Ziviani, Professor Magistério Superior - Voluntário**, em 20/03/2024, às 08:19, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3056590** e o código CRC **1DB5A8AF**.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM INOVAÇÃO TECNOLÓGICA

"INOVAÇÃO PARA EFICIÊNCIA ENERGÉTICA DE CAMINHÕES FORA DE ESTRADA EM MINA A CÉU ABERTO."

DOUGLAS GIOVANE NUNES PIMENTA Nº DE REGISTRO 2021729502.

Dissertação **Aprovada** pela Banca Examinadora constituída pelos Professores Doutores:

Professor Doutor Adriano Alonso Veloso (Orientador)
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Nivio Ziviani (Coorientador)
(Universidade Federal de Minas Gerais – UFMG)

Professor Doutor Allan Claudius Queiroz Barbosa
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Adriano César Machado Pereira
(PPG em Inovação Tecnológica da UFMG)

Professor Doutor Allan Claudius Queiroz Barbosa
Coordenador do PPG em Inovação Tecnológica da UFMG

Belo Horizonte, 23 de fevereiro de 2024.



Documento assinado eletronicamente por **Allan Claudius Queiroz Barbosa, Professor(a)**, em 11/03/2024, às 14:12, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Cesar Machado Pereira, Professor do Magistério Superior**, em 11/03/2024, às 20:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Adriano Alonso Veloso, Professor do Magistério Superior**, em 18/03/2024, às 15:45, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Nivio Ziviani, Professor Magistério Superior - Voluntário**, em 20/03/2024, às 08:19, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3056603** e o código CRC **8228C2A5**.

AGRADECIMENTOS

Agradeço a Deus por me proporcionar a oportunidade de realizar este sonho e por conceder-me saúde, sabedoria e perseverança para superar os desafios.

À minha esposa, Pâmela, expresso minha profunda gratidão pelo amor, apoio, compreensão e paciência dedicados a mim ao longo desta jornada. Sua presença foi essencial para a concretização deste trabalho.

Ao meu filho Theo, agradeço a constante alegria, carinho e motivação que traz todos os dias. Você é a razão do meu viver.

Aos meus pais, José Pimenta e Dulcenéia, manifesto meu agradecimento pelo exemplo de vida, educação, valores e incentivo que me proporcionaram desde a infância. Vocês são meus heróis e minha base.

Expresso minha gratidão à minha família, em especial aos meus irmãos, Nicolas e Sara, meus sobrinhos, Noah, e aos gêmeos que estão por vir. A jornada de vocês é uma fonte constante de inspiração para mim.

Ao meu orientador, Professor Adriano Veloso, e ao coorientador, Professor Emérito Nívio Ziviani, agradeço sinceramente por suas orientações cuidadosas, valiosos insights e paciência ao me guiar durante esta trajetória acadêmica.

Agradeço imensamente aos amigos e colegas de trabalho, Alexandre Pigatti, Ali Soofastaei, Izabela Silva e Jodelson Sabino, pelo valioso apoio em minha dissertação de mestrado. A colaboração, suporte e mentoria de vocês foram fundamentais para o sucesso deste trabalho.

A todos vocês, dedico este trabalho, que é resultado não apenas do meu esforço, mas também do apoio incondicional que recebi de cada um de vocês. Muito obrigado!

“Se vi mais longe foi por estar de pé sobre ombros de gigantes.”
(Isaac Newton, 1676).

RESUMO

A mineração é uma atividade econômica que contribui para o desenvolvimento, geração de empregos e renda, no Brasil e no mundo. A mineração representa cerca de 4% do Produto Interno Bruto (PIB) e 40% do saldo comercial brasileiro, com US\$ 24,9 bilhões em 2022. Além disso, a indústria extrativa mineral é essencial para a produção de diversos bens e serviços que fazem parte da vida das pessoas, como energia, tecnologia, construção civil, agricultura, indústria química, dentre outras. Pode ser considerada a 'indústria das indústrias'.

No entanto, a mineração também pode trazer impactos ambientais e sociais negativos, especialmente quando não são adotadas práticas sustentáveis em seu processo. A indústria mineral mal gerida pode causar danos severos ao meio ambiente e impactar as comunidades locais, como rompimento de barragens, poluição, desmatamento, erosão, perda de biodiversidade, mudanças climáticas, dentre outras.

Por isso, é fundamental que a mineração seja realizada de forma responsável e sustentável, buscando minimizar os danos ambientais e maximizar os benefícios sociais da atividade. A mineração sustentável envolve a adoção de medidas como a redução do consumo de água e energia, a recuperação de áreas degradadas, a gestão de resíduos e rejeitos e a promoção da participação social e do diálogo com as comunidades envolvidas.

A mineração sustentável também está relacionada à utilização de energias renováveis e à descarbonização da atividade. Nesse contexto, a Inovação Tecnológica, mais especificamente o uso de dados e Inteligência Artificial, podem ser alavancas para a busca da mineração sustentável.

O estudo se propõe a verificar a aplicação de técnicas de Análise Avançada de Dados na redução das emissões de gases de efeito estufa para caminhões fora de estrada em minas de céu aberto, gases esses que contribuem para o aquecimento global e as mudanças climáticas.

Portanto, é sabido que a mineração é uma atividade de grande importância para o Brasil e para o mundo, mas que deve ser realizada de forma sustentável. A inovação tecnológica pode ser o caminho para a transformação da indústria mineral que gere mais valor econômico, social e ambiental para a sociedade.

Palavras-Chave: Mineração sustentável; inovação tecnológica; inteligência artificial; análise avançada de dados; descarbonização.

ABSTRACT

Mining is an economic activity that contributes to development, job creation, and income in Brazil and worldwide. Mining represents approximately 4% of the Gross Domestic Product (GDP) and 40% of the Brazilian trade balance, with US\$ 24.9 billion in 2022. Additionally, the extractive mineral industry is essential for producing various goods and services that are part of people's lives, such as energy, technology, construction, agriculture, chemical industry, among others. It can be considered the 'industry of industries.'

However, mining can also bring negative environmental and social impacts, especially when unsustainable practices are not adopted. Poorly managed mining industries can cause severe damage to the environment and impact local communities, including dam failures, pollution, deforestation, erosion, loss of biodiversity, climate change, among others.

Therefore, it is crucial for mining to be conducted responsibly and sustainably, aiming to minimize environmental damage and maximize social benefits. Sustainable mining involves adopting measures such as reducing water and energy consumption, reclaiming degraded areas, managing waste and tailings, and promoting social participation and dialogue with involved communities.

Sustainable mining is also linked to the use of renewable energies and decarbonization of the activity. In this context, Technological Innovation, specifically the use of data and Artificial Intelligence, can be leveraged in the pursuit of sustainable mining. The study aims to examine the application of Advanced Data Analysis techniques in reducing greenhouse gas emissions for off-road trucks in open-pit mines—gases that contribute to global warming and climate change.

Therefore, it is known that mining is an activity of great importance for Brazil and the world, but it must be carried out sustainably. Technological innovation may be the path to transforming the mining industry, generating more economic, social, and environmental value for society.

Keywords: Sustainable mining; technological Innovation; artificial intelligence; advanced data analysis; decarbonization.

LISTA DE FIGURAS

Figura 1 - Objetivos de desenvolvimento sustentável	25
Figura 2 - Itabirito com domínios ricos em quartzo (parte clara) intercalado com níveis ricos em hematita configurando uma dobra	28
Figura 3 - Etapas do Minério, de sua Exploração até o Beneficiamento	30
Figura 4 - <i>Four Types of Analytics Capability</i>	35
Figura 5 - Visão geral de técnicas de interpretabilidade.....	41
Figura 6 - Vale no Mundo	44
Figura 7 - Modelo de Negócio	45
Figura 8 - Trechos da cava da Mina em estudo utilizados para treinamento e execução do modelo	49
Figura 9 - Sessões da Mina.....	50
Figura 10 - Parâmetros de consumo de combustível	51
Figura 11 - Curva de aprendizado (XGboost).....	56
Figura 12 - Previsões vs Observações (XGBoost).....	57
Figura 13 - Importância das variáveis (XGboost)	58
Figura 14 - Importância das variáveis XGBoost (treinamento).....	59
Figura 15 - Importância das variáveis XGBoost (teste)	59
Figura 16 - Shap Decision Plot (XGBoost).....	61
Figura 17 - Shap Dependence Plot (XGBoost).....	62
Figura 18 - Curva de aprendizado (Random Forest).....	64
Figura 19 - Previsões vs Observações (Random Forest).....	64
Figura 20 - Importância das variáveis (Random Forest)	65
Figura 21 - Importância das variáveis Random Forest (treinamento).....	66
Figura 22 - Importância das variáveis Random Forest (teste).....	66
Figura 23 - Shap Decision Plot (Random Forest).....	68
Figura 24 - Shap Dependence Plot (Random Forest)	68

LISTA DE TABELAS

Tabela 1 - Resumo da metodologia (continua).....	46
Tabela 2 - Especificações dos equipamentos Catterpillar 793D e 797F	48
Tabela 3 - Exemplo de recomendação de velocidade por trecho.....	52
Tabela 4 - Benefício potencial com a utilização de técnicas de inteligência artificial para recomendação de velocidade.	53
Tabela 5 - Benefícios realizados com a utilização do modelo (2023).....	54
Tabela 6 - Características dos modelos XBoost e Random Forest	69
Tabela 7 - Performance dos modelos XBoost e Random Forest	70

LISTA DE ABREVIATURAS

CO ₂	Dióxido de Carbono
Fe ₂ O ₃	Óxido férrico (Hematita)
R ²	Coefficiente de determinação
SiO ₂	Dióxido de silício (Quartzo)

LISTA DE SIGLAS

AG	Algoritmo Genético
ANM	Agência Nacional de Mineração
CAGED	Cadastro Geral de Empregados e Desempregados
CAT	Referência à empresa Caterpillar.
CONFEA	Conselho Federal de Engenharia e Agronomia
COP26	Conferência das Nações Unidas sobre Mudanças Climáticas de 2023
DMT	Distância Média de Transporte
DX	Coordenada de GPS (eixo x)
DY	Coordenada de GPS (eixo y)
GEE	Gases de Efeito Estufa
GitHub	Plataforma de desenvolvimento colaborativo
GVW	<i>Gross Vehicle Weight (Peso do veículo vazio mais a carga útil)</i>
IA	Inteligência Artificial
IBRAM	Instituto Brasileiro de Mineração
MAE	Erro Médio Absoluto
MAPE	Erro Médio Absoluto Percentual
OC	Observatório do Clima
ODS	Objetivos de Desenvolvimento Sustentável
OECD	Organização para a Cooperação e Desenvolvimento Econômico
ONU	Organização das Nações Unidas
PIB	Produto Interno Bruto
PSMA	Planos Setoriais de Mitigação e Adaptação
RIO+20	Conferência das Nações Unidas sobre Desenvolvimento Sustentável realizada no Rio de Janeiro em 2012
RPM	Rotação por minuto
TR	Resistência total de rolamento
SHAP	<i>SHapley Additive exPlanations.</i>
SQL	<i>Structured Query Language</i> (Linguagem de Consulta Estruturada)
UO	Unidade Operacional
Vale	Referência à empresa Vale S/A.
XAI	<i>Explainable Artificial Intelligence.</i>
XGBoost	<i>Extreme Gradient Boosting</i>

SUMÁRIO

1	INTRODUÇÃO	18
1.1	Definição do Problema	19
1.2	Pergunta de pesquisa	20
1.3	Objetivo geral	20
1.4	Organização do Trabalho	21
1.5	Justificativas	22
2	REFERENCIAL TEÓRICO	24
2.1	Sustentabilidade	24
2.2	Acordos internacionais e ODS	25
2.3	Industria Extrativa Mineral	26
2.3.1	<i>Importância da Mineração para o Brasil</i>	26
2.3.2	<i>Conceitos básicos da Mineração</i>	27
2.3.3	<i>Etapas do Minério: da sua exploração até o beneficiamento</i>	29
2.3.4	<i>Mineração a céu aberto</i>	29
2.3.5	<i>Uso de caminhões fora de estrada em Minas de céu aberto de Minério de Ferro</i>	31
2.4	Industria e a inovação orientada à sustentabilidade	32
2.5	Inteligência Artificial e o aprendizado de máquina	33
2.6	Análise Avançada de dados	34
2.6.1	<i>Ferramentas e Plataformas de análise de dados</i>	35
2.6.2	<i>Algoritmos e análise de dados</i>	36
2.7	Técnicas de interpretabilidade	39
2.7.1	<i>Análise SHAP (SHapley Additive exPlanations)</i>	40
3	METODOLOGIA	43
4	ESTUDO DE CASO MODELO DE ANÁLISE AVANÇADA PARA AUMENTO DA EFICIÊNCIA ENERGÉTICA DOS CAMINHÕES FORA DE ESTRADA	48
4.1	Detalhamento do estudo de caso	48
4.2	Estrutura do modelo e suas variáveis	50
4.2.1	<i>Consumo de combustível em caminhões fora de estrada</i>	51
4.3	Recomendações do modelo por trecho	52
4.4	Resultados potenciais	53

4.4.1	<i>Pessoas no centro</i>	54
4.5	Interpretabilidade dos modelos e resultados	55
4.5.1	<i>XGBoost</i>	55
4.5.2	<i>Random Forest</i>	62
4.5.3	<i>Quadro Resumo</i>	69
5	CONSIDERAÇÕES FINAIS	71
5.1	Limitações do trabalho	73
5.2	Trabalhos Futuros	73
	REFERÊNCIAS	75

1 INTRODUÇÃO

A mineração é uma atividade econômica de grande importância para o Brasil e para o mundo. Os recursos minerais estão entre os mais relevantes do planeta, capazes de influenciar o desenvolvimento econômico e industrial de um país, fornecendo matérias-primas para os setores primário, secundário e terciário da economia (CONFEA, 2021). Segundo o infográfico anual do IBRAM (Instituto Brasileiro de Mineração) para o ano de 2022, o setor mineral foi decisivo para manter positivo o saldo da balança comercial brasileira, representando 4% do Produto Interno Bruto (PIB) (ANM, 2023) e 40% do saldo comercial brasileiro, totalizando US\$ 24,9 bilhões em 2022 (IBRAM, 2022).

Entretanto, é evidente que garantir demanda, lucro e investimentos não é suficiente para a continuidade e sustentabilidade da indústria, como demonstram os desastres ocorridos com barragens de rejeito em Mariana (2015), Brumadinho (2019) e Maceió (2023). O conceito de sustentabilidade deixou de ser uma ideia utópica de longo prazo e tornou-se um tema essencial para a sobrevivência da indústria. Atualmente, espera-se que as organizações assumam a responsabilidade global, considerando os aspectos e impactos sociais, econômicos e ambientais.

O desafio da indústria mineral é traduzir as expectativas da sociedade e assumir o tripé da sustentabilidade com ações práticas no dia a dia. Nesse contexto, alguns autores veem a Inovação Tecnológica como a forma de alinhar propósitos. Conforme citado por Barbieri et al. (2010), o modelo de organização inovadora sustentável é uma resposta às pressões institucionais por uma organização que seja capaz de inovar com eficiência em termos econômicos, mas com responsabilidade social e ambiental.

As organizações devem estabelecer metas claras e criar um ambiente de inovação que incorpore suas ambições de maneira disruptiva, ágil, questionando constantemente seu "*status quo*". Para o sucesso da mudança de paradigma, é necessário fomentar a inovação tecnológica em busca dos propósitos estabelecidos, nutrindo conexões em todos os lugares com diferentes níveis de parcerias, dentro e fora da empresa, entre universidades, empresas do mesmo setor e grupos de fomento à inovação aberta e/ou privada, etc. As expectativas das comunidades, instituições,

investidores, parceiros e colaboradores mudaram, e o "*modus operandi*" da empresa não pode mais ser o mesmo, devendo acompanhar essas ambições.

Para direcionar os avanços necessários e estabelecer metas e definições de sustentabilidade, a ONU (Organização das Nações Unidas) desenvolveu, em 2015, em conjunto com países membros, um plano de ação global denominado ODS (Objetivos de Desenvolvimento Sustentável), que consiste em 17 objetivos e 169 metas a serem alcançados até 2030, conhecido como "Agenda 2030" (ONU, 2023).

Nesse contexto, a dissertação propõe uma análise de estudo de caso único para verificar a atuação da organização em estudo no desenvolvimento de inovação e parcerias para a aplicação de soluções de análise avançada de dados com o objetivo de reduzir as emissões de gases de efeito estufa, diminuir o consumo de óleo diesel e atingir metas de sustentabilidade voltadas para a eficiência energética de caminhões fora de estrada, conforme citado pela ONU (2023) por meio do ODS número 7 - Energia Limpa e Acessível, item 7.a: "Até 2030, reforçar a cooperação internacional para facilitar o acesso a pesquisas e tecnologias de energia limpa, incluindo energias renováveis, eficiência energética e tecnologias de combustíveis fósseis avançadas e mais limpas, e promover o investimento em infraestrutura de energia e em tecnologias de energia limpa" e ODS 13 - Ação contra a mudança global do clima.

O estudo também se propõe a verificar os modelos de análise avançada de dados para aumento da eficiência energética, incluindo abordagens, resultados, correlações, comportamentos e a importância das variáveis envolvidas, por meio do método *SHAP* (*SHapley Additive exPlanations*) implementado em linguagem de programação *Python*.

1.1 Definição do Problema

O transporte por caminhão é responsável pela maior parte do consumo de combustível fóssil (óleo diesel) em uma operação de mina de céu aberto. A redução do consumo de diesel levaria à diminuição das emissões de gases de efeito estufa (GEE), e alcançando as metas de redução de carbono, além da diminuição dos custos de transporte. A determinação do consumo de combustível é complexa e requer vários

parâmetros, incluindo as condições da mina, da frota, do caminhão, do combustível, do clima e da estrada, como dados de entrada (SOOFASTAEI, 2016).

Nesse sentido, é necessária a aplicação de uma solução de análise avançada para compreender o comportamento desses parâmetros e fornecer recomendações de ajustes para aumentar a eficiência energética dos caminhões.

1.2 Pergunta de pesquisa

Como a análise avançada de dados pode apoiar o aumento da eficiência energética dos caminhões fora de estrada, reduzindo o consumo de óleo diesel e as emissões de gases de efeito estufa?

1.3 Objetivo geral

Verificar a aplicação de técnicas de Análise Avançada de Dados e Inteligência Artificial na busca da maior eficiência energética, redução do consumo de combustível fóssil e diminuição das emissões de gases de efeito estufa dos caminhões de grande porte, conhecidos como 'caminhões fora de estrada', em operações de mina de céu aberto no Brasil. Isso inclui um estudo de caso da empresa Vale S/A, bem como a avaliação da contribuição dessa solução para as metas de sustentabilidade dos Objetivos de Desenvolvimento Sustentável (ODS) 7 - Energia Acessível e Limpa e ODS 13 - Ação Contra a Mudança Global do Clima, conforme declarado pela ONU (2023), destacando esses dois objetivos como parte integrante do desenvolvimento sustentável dos países membros.

1.3.1. Objetivos específicos

Considera-se também os seguintes objetivos específicos:

- a) Verificar e comparar as técnicas de análise avançadas de dados e inteligência artificial empregadas;
- b) Identificar as principais variáveis e entender características relevantes para os modelos, através do método *SHAP* (*SHapley Additive exPlanations*) em *Python*.

1.4 Organização do Trabalho

No Capítulo 1, é contextualizado o problema e delineados os objetivos principais do trabalho, incluindo a necessidade de uma mineração cada vez mais responsável e sustentável, a inovação como catalisadora do aumento da eficiência energética e descarbonização. O capítulo encerra com a apresentação da pergunta de pesquisa e sua justificativa.

No Capítulo 2, é apresentado o embasamento teórico para o desenvolvimento do trabalho, abrangendo desde os Objetivos de Desenvolvimento Sustentável (ONU, 2023) até o contexto do setor extrativo mineral no Brasil, a Indústria 4.0 e a Inovação orientada para a sustentabilidade. Nesse capítulo, detalham-se também as soluções de Análise Avançada de Dados, Inteligência Artificial, bem como os modelos estudados, *XGBoost/Random Forest*, e a técnica *SHAP* (*SHapley Additive exPlanations*) e suas ramificações.

No Capítulo 3, são apresentados a metodologia, o estudo de caso único e os resultados esperados, além da descrição da base de dados em estudo e suas características.

O Capítulo 4 verifica o estudo de caso da aplicação do modelo de Análise Avançada de Dados para Eficiência Energética dos caminhões fora de estrada, assim como detalhamento das suas variáveis, fases e estrutura. Também analisa os resultados relacionados à eficiência energética, correlação entre as variáveis e relevância para o modelo. O capítulo encerra com quadro resumo comparativo dos modelos e resultados.

No Capítulo 5, são apresentadas as conclusões e análises dos resultados, destacando suas limitações e propondo estudos de continuidade para o trabalho.

1.5 Justificativas

A indústria é um dos setores importantes que contribuem significativamente para as emissões de gases de efeito estufa no Brasil e no mundo (CLIMA, 2022), especialmente o setor extrativo mineral. Este setor utiliza, entre outras fontes de energia, combustíveis derivados do petróleo para a exploração de recursos naturais como minério de ferro, carvão, cobre, etc. O uso de combustíveis fósseis é responsável por grande parte da geração de energia e do desenvolvimento econômico, mas também intensifica o efeito estufa e o aquecimento global, causando impactos negativos para o meio ambiente e a sociedade.

Segundo CLIMA (2022), o setor de transporte emitiu 203,8 milhões de toneladas de CO₂ em 2021, configurando-se como o maior emissor dentro dos setores de energia e PIUP (Processos industriais e uso dos produtos). Para dar uma ideia de grandeza, as emissões do transporte no Brasil já correspondem a quase o dobro do valor emitido devido ao desmatamento no bioma Cerrado. Também se destaca a variação de emissões na atividade de geração de eletricidade. Com uma variação positiva de 46%, essa foi a categoria com maior aumento percentual em energia e processos industriais.

Por isso, é imprescindível que a indústria invista em ações de redução de emissões de GEE e de descarbonização, ou seja, aumento da eficiência energética e transição para uma matriz energética mais limpa e renovável, que minimize a dependência de carbono. Essas ações podem trazer benefícios não só para o clima, mas também para a competitividade, a inovação, a sustentabilidade e a responsabilidade social das empresas.

Nesse sentido, a tecnologia, mais especificamente a análise e uso dos dados para tomada de decisão, assim como técnicas de inteligência artificial aplicadas, podem ser impulsionadoras para o aumento da eficiência energética dos processos produtivos, reduzindo o consumo, o desperdício e aumentando a eficiência no uso de recursos naturais.

Essas ações estão alinhadas com os compromissos internacionais assumidos pelo Brasil, como o Acordo de Paris, que visa limitar o aumento da temperatura global a 1,5°C até o final do século (CLIMA, 2023), e os ODS da ONU: ODS 7 – Energia limpa e acessível – que, entre outros pontos, visa garantir o acesso a fontes de energia

confiáveis, sustentáveis e modernas para todos. Busca também reforçar a cooperação internacional para facilitar o acesso a pesquisa e tecnologias de energia limpa, incluindo energias renováveis, eficiência energética e tecnologias de combustíveis fósseis avançadas e mais limpas. ODS 13 - Ação contra a mudança global do clima – que tem reunido esforços para elaborar e executar os planos setoriais de mitigação e de adaptação à mudança do Clima (PSMAs) (BRASIL, 2023), estabelecendo metas e medidas para reduzir as emissões e os impactos das mudanças climáticas em setores como agricultura, energia, indústria, transportes, saúde, entre outros (IPEA, 2023). Em resumo, são acordos com o intuito de promover o desenvolvimento humano e ambiental de forma integrada e equilibrada (ONU, 2023).

2 REFERENCIAL TEÓRICO

2.1 Sustentabilidade

Segundo a ONU, desenvolvimento sustentável é definido como "aquele que satisfaz as necessidades presentes, sem comprometer a capacidade das gerações futuras de suprir suas próprias necessidades" (RIO+20, 2012). A demonstração de resultados por meio de pesquisas científicas sobre o meio ambiente e o impacto do ser humano no planeta tem amadurecido a sociedade sobre o tema e mostrado cada vez mais a importância do equilíbrio dos fatores sociais, econômicos e ambientais, se quisermos um futuro melhor para a humanidade.

Nessa linha, o interesse sobre sustentabilidade (ou Desenvolvimento Sustentável) e, mais recentemente, as abordagens referentes a estratégias, produção mais limpa, controle da poluição, ecoeficiência, gestão ambiental, responsabilidade social, ecologia industrial, investimentos éticos, economia verde, ecodesign, reuso, consumo sustentável, resíduos zero, entre inúmeros outros termos, é crescente, conforme Glavić e Lukman (2007), citado por Sartori, Latrônico e Campos (2014).

Para Elkington (1997), criador do termo *Triple Bottom Line*, a sustentabilidade é o equilíbrio entre os três pilares: ambiental, econômico e social. A expectativa de que as empresas devem contribuir progressivamente com a sustentabilidade surge do reconhecimento de que os negócios precisam de mercados estáveis e devem possuir habilidades tecnológicas, financeiras e de gerenciamento necessárias para possibilitar a transição rumo ao desenvolvimento sustentável, conforme Elkington (2001), citado por Sartori (2014).

Para dar foco aos avanços necessários, estabelecer metas e definições de sustentabilidade, a ONU desenvolveu, juntamente com países membros, um plano de ação global chamado ODS, com 17 objetivos e 169 metas - Agenda 2030, conforme mostra a Figura 1.

Segundo a própria organização, são um apelo global à ação para acabar com a pobreza, proteger o meio ambiente e o clima e garantir que as pessoas, em todos os lugares, possam desfrutar de paz e prosperidade (ONU, 2023).

Figura 1 - Objetivos de desenvolvimento sustentável



Fonte: ONU (2023).

2.2 Acordos internacionais e ODS

Conforme citado anteriormente, os ODS são um conjunto de metas globais que visam, entre outros pontos, gerar ações práticas para mitigar o impacto no meio ambiente e no clima. Mais especificamente, os ODS 7 e 13 se referem, respectivamente, à energia limpa e acessível e à ação contra a mudança global do clima (ONU, 2023).

O Brasil é signatário da Agenda 2030 e do Acordo de Paris, que são os principais acordos internacionais relacionados aos ODS 7 e 13. O Acordo de Paris, assinado em 2015, é um tratado que busca limitar o aumento da temperatura média global a 2°C em relação aos níveis pré-industriais, por meio da redução das emissões de gases de efeito estufa (GEE). O Brasil se comprometeu a reduzir suas emissões de GEE em 50% até 2030, em comparação com os níveis de 2005 (COP26, 2021).

O aquecimento global é um tema de extrema relevância para nossa sobrevivência, e o envolvimento da sociedade no combate ao aquecimento global é muito importante, pois as ações individuais e coletivas podem contribuir para reduzir as emissões de gases de efeito estufa e os impactos negativos das mudanças climáticas. Tendo o Brasil se comprometido com as metas de redução, não somente

o Estado, mas toda a sociedade, incluindo as organizações, são corresponsáveis pelo cumprimento do acordo.

Segundo o relatório do Observatório do Clima (CLIMA, 2022), a cadeia minero-siderúrgica emitiu 107,6 milhões de toneladas de CO₂ em 2020, representando 5% do total de emissões do Brasil. Acima do esperado no período, principalmente considerando a pandemia global, o consumo de energia (queima de combustíveis fósseis) está entre os principais fatores, representando 29,9% no total das emissões da cadeia minero-siderúrgica. Por isso, é importante a participação da indústria mineral, principalmente as de grande porte, nas ações que envolvem o aquecimento global e a emissão de GEE. Além de contribuírem para a redução direta das emissões, conseguem envolver e fomentar seu entorno em um ciclo positivo de ações para mitigar o impacto das mudanças climáticas.

2.3 Indústria Extrativa Mineral

2.3.1 Importância da Mineração para o Brasil

A mineração é uma atividade econômica de grande relevância para o Brasil e para o mundo, contribuindo para o desenvolvimento, geração de empregos e renda. Além disso, as exportações de produtos minerais desempenham um papel crucial no equilíbrio da balança comercial brasileira.

A mineração é uma das atividades mais antigas da humanidade, responsável pela construção das primeiras civilizações e pelo progresso tecnológico ao longo da história. Envolve a extração de recursos naturais utilizados em diversos setores industriais, como metalurgia, siderurgia, petroquímica, fertilizantes, construção civil, entre outros. Também é uma fonte significativa de energia, fornecendo carvão, petróleo e gás natural.

No contexto brasileiro, a mineração desempenha um papel crucial na economia, sendo um dos setores que mais contribuem para o PIB do país. Segundo a Agência Nacional de Mineração (ANM), a indústria da mineração representou

aproximadamente 4% do PIB brasileiro em 2020 (ANM, 2023). Esse setor é reconhecido como dinâmico e altamente competitivo na economia nacional. Adicionalmente, a mineração é responsável por uma parcela significativa das exportações do Brasil, alcançando cerca de US\$ 30 bilhões em produtos minerais em 2020, contribuindo com 40% do saldo comercial brasileiro em 2022 (IBRAM, 2022). Entre os principais produtos minerais exportados estão minério de ferro, alumínio, cobre, ouro e nióbio (ANM, 2023).

A mineração também exerce um impacto positivo na geração de empregos e renda, proporcionando cerca de 207 mil empregos formais no país, conforme o Cadastro Geral de Empregados e Desempregados (CAGED) em janeiro de 2021. Além disso, a distribuição de royalties para os municípios onde ocorre a extração mineral beneficia áreas como educação, saúde, infraestrutura e meio ambiente, promovendo o desenvolvimento sustentável das regiões (IBRAM, 2022).

Entretanto, a mineração pode acarretar impactos ambientais e sociais negativos quando não são adotadas práticas sustentáveis de extração mineral. Por esse motivo, é crucial que o setor seja regulamentado e monitorado por órgãos competentes, assegurando a proteção do meio ambiente e a segurança dos trabalhadores. Os principais desafios incluem a redução do consumo de água e energia, a recuperação de áreas degradadas, a gestão de resíduos e rejeitos, a prevenção de acidentes e desastres, e a promoção da participação social e do diálogo com as comunidades afetadas.

Em síntese, a mineração é uma atividade econômica de suma importância para o Brasil e o mundo, contribuindo para o desenvolvimento, a geração de empregos e renda, além das exportações minerais. No entanto, é imperativo que a mineração seja conduzida de maneira responsável e sustentável, garantindo a preservação do meio ambiente e a segurança dos trabalhadores.

2.3.2 Conceitos básicos da Mineração

Mineração é um conjunto de processos industriais cujo objetivo principal é extrair substâncias minerais de depósitos ou massas minerais. Pode-se citar como produtos derivados da mineração desde os metais até as cerâmicas e os

combustíveis, passando por componentes elétricos, computadores, carros, estradas, entre outros tantos produtos. Conforme destacado por Menezes (2008), a mineração está presente em todos os aspectos do nosso dia a dia, e sua evolução trouxe desenvolvimento para a sociedade em que vivemos.

Os minérios são rochas ou materiais nos quais os elementos ou substâncias de valor econômico foram concentrados por processos naturais, podendo ser extraídos e beneficiados, como explica Menezes (2008). A mineração visa à extração e beneficiamento desses minérios. Exemplos de minérios extraídos da natureza incluem cobre, carvão, níquel, alumínio, manganês, ferro, entre outros.

O estudo de caso verifica a aplicação de análise de dados avançada na exploração do minério de ferro, mais especificamente no aumento da eficiência energética (óleo diesel) na frota de transporte do minério. O minério de ferro é muito comum no Brasil, principalmente nos estados de Minas Gerais e Pará. Conforme mencionado por Menezes (2008), o minério de ferro é composto por dois minerais comuns na natureza, a hematita e o quartzo, que constituem rochas como o itabirito (Quadrilátero Ferrífero, Minas Gerais) ou o jaspilito (Carajás, Pará). Depois de retirados do subsolo, são transportados, geralmente por caminhões de grande porte, para operações industriais (beneficiamento) onde se separa a parte útil (hematita – Fe_2O_3) da inútil (quartzo – SiO_2). A Figura 2 mostra um exemplo de rocha Itabirito.

Figura 2 - Itabirito com domínios ricos em quartzo (parte clara) intercalado com níveis ricos em hematita configurando uma dobra



Fonte: Menezes (2008)

O processo de exploração e beneficiamento do minério são etapas que constituem a indústria de exploração mineral. No entanto, antes da exploração da área propriamente dita, ocorrem pesquisa e estudos geológicos para determinar a viabilidade da exploração naquela área, conforme apontado por Menezes (2008).

2.3.3 Etapas do Minério: da sua exploração até o beneficiamento

Conforme mencionado por Menezes (2008), quase todos os elementos metálicos úteis ao homem são raros. Para viabilizar seu aproveitamento, procura-se descobrir, por meio de pesquisas geológicas, ocorrências naturais com concentrações bem acima da média.

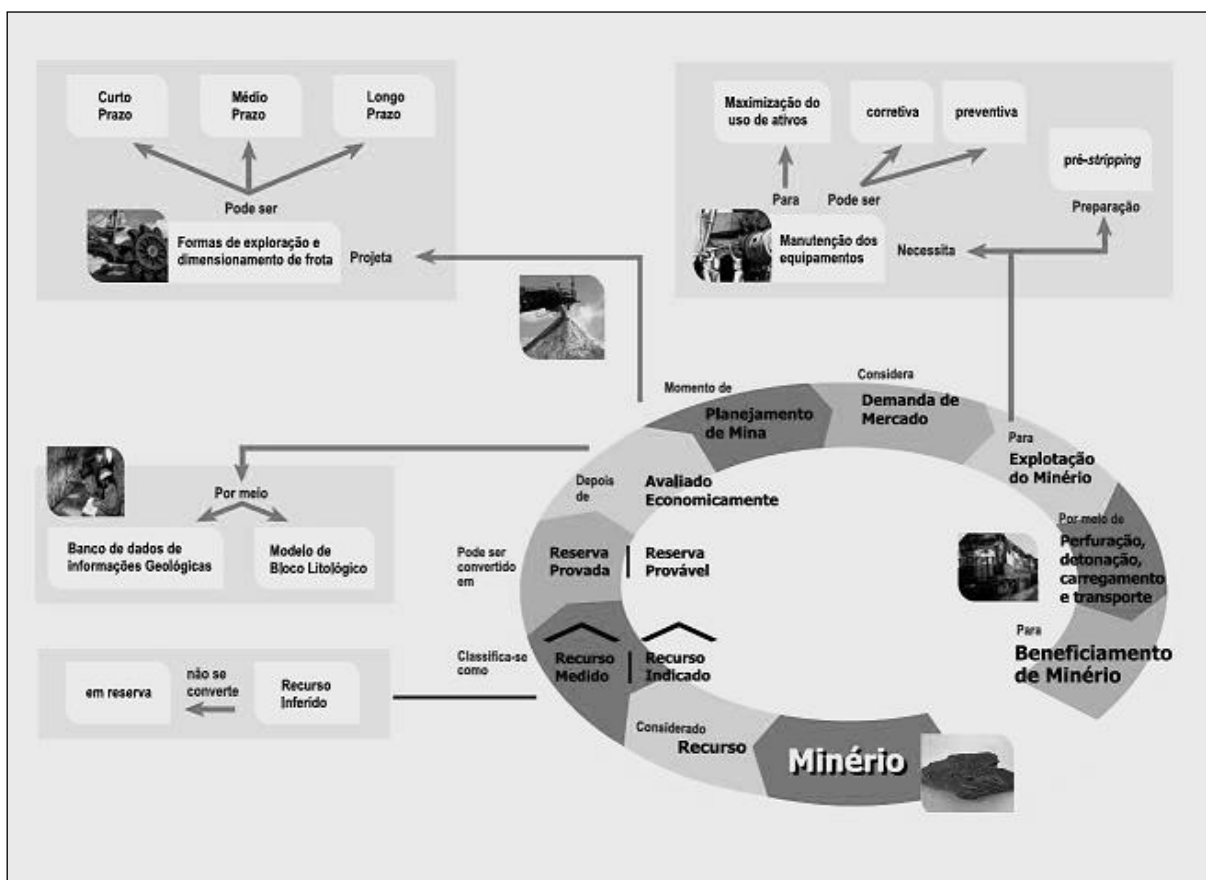
Dessa forma, a exploração do minério possui diversas etapas, que vão desde a definição do minério, planejamento da mina, avaliação econômica, perfuração, desmonte, transporte e beneficiamento. Uma vez caracterizado o potencial mineral, são iniciados os trabalhos de mapeamento detalhado e a execução de sondagens geológicas, objetivando conhecer a forma e o modo de ocorrência do corpo de minério, além de caracterizá-lo mineralogicamente, texturalmente, fisicamente, quimicamente e metalurgicamente, como destacado por Menezes (2008).

Nos trabalhos de pesquisa e avaliação, busca-se conhecer o tamanho do depósito e a qualidade e teor do minério (teor x volume), pois é preciso comprovar que o depósito mineral atende às necessidades das indústrias siderúrgicas, químicas, entre outras, e se é economicamente viável, como observado por Menezes (2008). A Figura 3 ilustra as etapas do processo de mineração.

2.3.4 Mineração a céu aberto

Conforme citado em Wikipedia ("Mineração a Céu Aberto", 2023), a mineração a céu aberto refere-se ao método de extração de rochas ou minerais da terra pela remoção de um poço aberto ou de uma escavação em empréstimo.

Figura 3 - Etapas do Minério, de sua Exploração até o Beneficiamento



Fonte: Sousa (2008)

O termo é utilizado para diferenciar essa forma de mineração dos métodos extrativos que requerem perfuração de túneis na terra - mineração subterrânea. A mineração a céu aberto é empregada quando depósitos de minerais ou rochas comercialmente úteis são encontrados perto da superfície; ou seja, onde a espessura do terreno de cobertura (localizado acima do material de interesse, que deve ser removido para alcançá-lo) é relativamente pequena, ou o material de interesse é estruturalmente impróprio para a abertura de túneis (como é o caso de areias, cinzas vulcânicas e cascalhos). Onde os minerais ocorrem muito abaixo da superfície, e a espessura dos terrenos de cobertura é grande, ou o mineral ocorre em veios na rocha, o material de interesse é extraído usando métodos de mineração subterrânea. As minas a céu aberto são geralmente expandidas até que o recurso mineral (ou a área de terra de propriedade da empresa de mineração) seja esgotado.

São exemplos de materiais extraídos na mineração a céu aberto: argila, carvão, coquina, granito, gravilha, gesso, calcário, mármore, metais (cobre, ferro, por exemplo), areia, cascalho e arenito.

2.3.5 *Uso de caminhões fora de estrada em Minas de céu aberto de Minério de Ferro*

Os caminhões fora de estrada são veículos projetados para transportar grandes quantidades de material em terrenos irregulares e de difícil acesso, como as minas de céu aberto de minério de ferro. Esses veículos desempenham um papel crucial na operação de mineração, influenciando a produtividade, o custo e a segurança da atividade.

Conforme mencionado anteriormente, a mineração de céu aberto é uma técnica de extração de minerais que envolve a remoção do solo e das rochas que cobrem o depósito mineral, formando uma cava ou uma bancada. Esse método é utilizado quando o minério está próximo à superfície e tem uma forma regular. A mineração de céu aberto requer o uso de equipamentos de grande porte, como escavadeiras, carregadeiras e caminhões fora de estrada, para realizar as operações de desmonte, carregamento e transporte do material (CESAR, 2019).

Os caminhões fora de estrada são veículos especiais com uma capacidade de carga elevada, variando de 40 a 400 toneladas, capazes de trafegar em estradas não pavimentadas, com aclives, declives e curvas acentuadas. Esses veículos são essenciais para o transporte do minério da frente de lavra até a planta de beneficiamento ou do estéril até as pilhas de disposição. Representam uma parcela significativa do custo operacional de uma mina de céu aberto, devido ao consumo elevado de combustível (aproximadamente 170L por hora trabalhada), pneus e peças de reposição. Além disso, estão sujeitos a falhas mecânicas, acidentes e paradas não programadas, comprometendo eficiência e segurança da operação (METSO, 2019).

O artigo de Bullivant (1987 apud METSO, 2019), "*Current Surface Mining Techniques*", destaca que, em termos de eficiência energética, os caminhões usam apenas 40% da energia para mover a carga, enquanto os 60% restantes são utilizados para mover a massa de tara do próprio caminhão.

Portanto, a necessidade de utilizar caminhões fora de estrada em minas de céu aberto está relacionada à demanda de transporte de grandes volumes de material em condições adversas de terreno. Contudo, o uso desses veículos implica em desafios técnicos, econômicos e ambientais, que devem ser considerados no planejamento e na gestão da mineração. Algumas alternativas para otimizar o uso de caminhões fora

de estrada incluem o dimensionamento adequado da frota, a manutenção preventiva, o treinamento dos operadores, a melhoria das condições das vias e a adoção de tecnologias para aumentar a eficiência e o controle, conforme verificado no estudo.

2.4 Indústria e a inovação orientada à sustentabilidade

A inovação é a capacidade de questionar o "*status quo*". De acordo com o Manual de Oslo, inovação é a implementação de um produto (bem ou serviço) novo ou significativamente melhorado, um processo, um novo método de marketing ou um novo método organizacional nas práticas de negócios, nas organizações do local de trabalho ou nas relações externas, conforme citado por Barbieri et al. (2010 apud OECD, 1997, p.55).

O percentual de mudança tecnológica e inovação é determinado pela interação entre um grupo de organizações dos setores público e privado, incluindo empresas, universidades, organizações de pesquisa, governo, entidades de ensino e organizações de fomento, que se combinam para criar, desenvolver e difundir novas tecnologias e inovações, cabendo um papel chave aos governos regional e nacional. As fronteiras da inovação são a abertura e a conectividade (CARMIGNANO, 2021).

Alguns autores defendem que, sem inovar, a empresa reduz significativamente sua capacidade competitiva (OLIVEIRA, 2020 apud RADOS; DIAS, 2015).

Em outras palavras, para inovar é necessário um ambiente aberto, colaborativo, que incentive a criatividade e, ao mesmo tempo, incentive a competição de mercado e/ou um propósito maior para que paradigmas sejam quebrados.

Segundo Barbieri et al. (2010), o modelo de organização inovadora sustentável é uma resposta às pressões institucionais por uma organização capaz de inovar com eficiência em termos econômicos, mas com responsabilidade social e ambiental. Esse tipo de organização busca vantagem competitiva desenvolvendo produtos, serviços, processos e negócios, novos ou modificados, com base nas dimensões social, ambiental e econômica. Ela reúne duas características essenciais: é inovadora e orientada para a sustentabilidade. Institucionaliza-se uma nova lógica de produção na qual a sustentabilidade e a inovação caminham juntas.

Nesse contexto, conforme Vasconcelos (2022), hoje é praticamente impossível dissociar a inovação do uso de dados para tomada de decisão. Para ser inovadora, uma empresa necessariamente precisa promover a transformação digital em suas atividades, e as mudanças necessárias incluem o entendimento das preferências dos clientes, consumidores e sociedade, aprendizagem a partir de erros cometidos para melhor qualificar serviços e produtos, entregas mais ágeis para atender às mudanças rápidas de contexto, e ter resiliência e capacidade de flexibilizar processos. Nesse contexto, o uso da Inteligência Artificial (IA), análise avançada dos dados e técnicas de aprendizagem de máquina tornam-se mecanismos cada vez mais importantes para a indústria que busca ser inovadora e sustentável.

2.5 Inteligência Artificial e o aprendizado de máquina

A Inteligência Artificial, conforme citado por Tableau (2023), é a capacidade que soluções tecnológicas têm de simular a inteligência humana, realizando determinadas atividades de maneira autônoma e aprendendo por si mesmas, graças ao processamento de um grande volume de dados que recebem de seus usuários ou de aplicações.

Importante ressaltar, conforme Vasconcelos (2022), que dentre as novas tecnologias, o aprendizado de máquina, uma subárea da Inteligência Artificial, está dentro do que é chamado tecnologia de propósito geral: aquela que promove o crescimento econômico de forma pervasiva (passa a fazer parte do nosso cotidiano), é capaz de ser melhorada ao longo do tempo e gera inovações complementares. Dois exemplos de tecnologia de propósito geral são a eletricidade e os motores de combustão interna, sem os quais seria muito difícil imaginar a vida humana na atualidade.

Ainda citando Vasconcelos (2022), o aprendizado de máquina é a tecnologia de propósito geral do momento, pelas seguintes razões: i) ela está no nosso dia a dia, pois tem as capacidades de classificar, rotular e predizer coisas centrais para ocupações e indústrias; ii) mostra-se capaz de ser melhorada continuamente; iii) gera inovações complementares, como percepção, visão computacional, reconhecimento de emoções pela voz e cognição. A principal vantagem do aprendizado de máquina,

do ponto de vista técnico, é oferecer a possibilidade de os computadores aprenderem sem serem explicitamente programados; eles aprendem com os dados.

2.6 Análise Avançada de dados

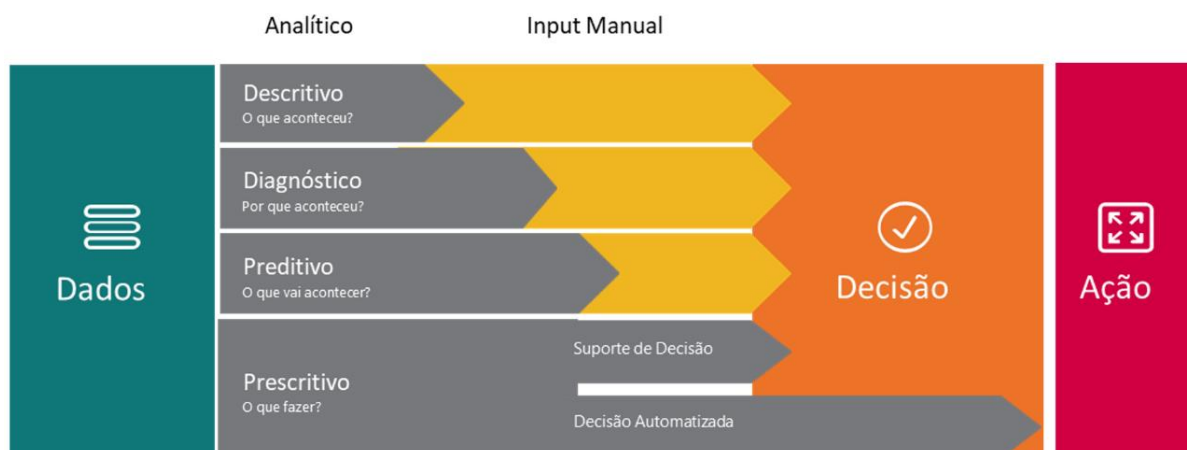
A análise avançada dos dados é um processo de examinar, transformar e modelar dados complexos e volumosos, usando técnicas estatísticas, computacionais e de aprendizado de máquina, com o objetivo de extrair informações, padrões e *insights* relevantes para a tomada de decisão, a solução de problemas ou a geração de conhecimento.

Existem quatro principais tipos de análise avançada dos dados, segundo Gartner (2014) (ver Figura 4):

- a) análise descritiva: resume o que aconteceu com os dados, usando medidas como média, frequência, desvio padrão, etc. É útil para entender o passado e o presente;
- b) análise diagnóstica: investiga o porquê de algo ter acontecido com os dados, usando técnicas como análise de correlação, teste de hipóteses, análise de causa raiz, etc. É útil para identificar problemas e oportunidades;
- c) análise preditiva: usa técnicas como regressão, séries temporais, aprendizado de máquina, etc., para estimar o que pode acontecer no futuro com base nos dados históricos. É útil para antecipar tendências e cenários;
- d) análise prescritiva: usa técnicas como otimização, simulação, análise de decisão, *deep learning*, etc., para sugerir o que deve ser feito para alcançar um objetivo ou resolver um problema com base nos dados. É útil para planejar ações e estratégias.

À medida que a organização intensifica seus investimentos em fundação, integração, qualidade e categorização de dados, habilita novas etapas na jornada analítica. Essa progressão constante a aproxima cada vez mais da tomada de decisão, promovendo ação ágil e precisa, orientada por dados refinados.

Figura 4 - Four Types of Analytics Capability



Fonte: Gartner (2014)

O compromisso contínuo com aprimoramentos coloca a organização em posição estratégica para inovação maior eficiência em seus processos.

2.6.1 Ferramentas e Plataformas de análise de dados

Existem diversas ferramentas e linguagens de programação disponíveis para análise de dados, variando conforme o objetivo, tipo de dados e complexidade. Algumas das mais utilizadas incluem:

- a) *Python*: Linguagem de alto nível, interpretada e de propósito geral, conhecida pela simplicidade, legibilidade e extensibilidade. Bibliotecas como *NumPy*, *pandas*, *matplotlib* e *scikit-learn* a tornam ideal para manipulação, processamento e visualização de dados (PYTHON, 2024);
- b) *R*: Linguagem e ambiente de software para computação estatística. Amplamente reconhecida como uma poderosa ferramenta *open source* para análise estatística, com pacotes como *dplyr*, *ggplot2* e *tidyr*. (R-PROJECT, 2024);
- c) *SQL*: Linguagem de consulta estruturada para manipular dados em bancos de dados relacionais. Padrão para interação com bancos de

dados como *MySQL* e *PostgreSQL*, permitindo diversas operações e análises complexas. (SQL, 2024);

- d) *Java*: Linguagem compilada, orientada a objetos e versátil, sendo crucial para análise de dados em Big Data e sistemas distribuídos. Ecossistema robusto com ferramentas como *Hadoop*, *Spark* e *Kafka*. (JAVA, 2024).

Essas são apenas algumas das opções disponíveis, e há muitas outras ferramentas e linguagens úteis para diversos propósitos e contextos na análise de dados.

2.6.2 Algoritmos e análise de dados

Conforme Ziviani (2010), os algoritmos fazem parte do dia a dia das pessoas. Exemplos de algoritmos:

- a) instruções para o uso de medicamentos;
- b) indicações de como montar um aparelho;
- c) uma receita de culinária.

Ainda citando Ziviani (2010), algoritmos são sequência de ações executáveis para a obtenção de uma solução para um determinado tipo de problema.

Ziviani (2010 apud DIJKSTRA, 1971), reforça que um algoritmo corresponde a uma descrição de um padrão de comportamento, expresso em termos de um conjunto finito de ações.

Desde a década de 1940, quando os primeiros computadores surgiram, os engenheiros de software programam utilizando lógica matemática. Diante de um conjunto de dados, eles elaboram um programa com um passo a passo para a máquina executar. Nesse modelo clássico de programação, o resultado é previsível e conhecido (VASCONCELOS, 2022).

Contudo, a análise de dados tornou-se progressivamente mais complexa, exigindo novos métodos para obter resultados aprimorados. Conforme destacado por Vasconcelos (2022), com o aprendizado de máquina, o paradigma mudou.

Transcendemos de uma ciência matemática para uma ciência natural, realizando observações em um mundo incerto, conduzindo experimentos com dados coletados e aplicando estatística e probabilidade para classificação, otimização e previsão.

2.6.2.1 *Random Forest*

O modelo *Random Forest* é um algoritmo de aprendizado de máquina supervisionado que utiliza múltiplos modelos de árvores de decisão para realizar previsões ou classificações. Ele combina os resultados das árvores de decisão individuais para produzir uma resposta final mais precisa e robusta. O *Random Forest* é considerado um método *ensemble*, pois combina vários modelos simples para formar um modelo mais complexo (TECH, 2024).

2.6.2.2 *XGBoost*

O algoritmo *XGBoost* foi desenvolvido como um projeto de pesquisa na Universidade de Washington. Seus criadores, Tianqi Chen e Carlos Guestrin, apresentaram o algoritmo por meio de um artigo na Conferência SIGKDD (Comunidade para *data mining*, *data science* e *analytics*) ((CHEN; GUESTRIN, 2016).

Desde a sua introdução, esse algoritmo tem se destacado no *Kaggle*, uma plataforma de competições onde cientistas de dados resolvem desafios por meio de *machine learning*. Além disso, também tem sido utilizado em várias aplicações de ponta na indústria pois sai capacidade de escala (CHEN; GUESTRIN, 2016).

Como resultado, existe uma forte comunidade de cientistas de dados contribuindo para os projetos de código aberto *XGBoost*, com mais de 350 colaboradores e mais de 3.500 *commits* no *GitHub*.

O algoritmo se diferencia de outros de várias maneiras (CHEN; GUESTRIN, 2016):

- a) abrangência em várias aplicações: pode ser usado para resolver problemas de regressão, classificação, ranqueamento, entre muitos outros;
- b) Portabilidade: Funciona em *Windows*, *Linux* e *OS X*;
- c) Suporte a múltiplas linguagens de programação: *C++*, *Python*, *R*, *Java*, *Scala* e *Julia*, etc;
- d) Integração em nuvem: Oferece suporte a *AWS*, *Azure*, *Clusters Yarn* e integra-se bem com *Flink*, *Spark*, entre outros ecossistemas.

2.6.2.3 Validação cruzada do coeficiente de determinação (R^2)

O coeficiente de determinação (R^2) mensura a proporção da variabilidade na variável de resposta que pode ser explicada pelo modelo, variando de 0 a 1. Um valor de 1 indica que o modelo explica toda a variabilidade, enquanto 0 sugere que não há explicação (DETERMINAÇÃO, 2024).

A validação cruzada é uma técnica para avaliar a capacidade de generalização de um modelo, dividindo o conjunto de dados em partes, treinando o modelo em algumas dessas partes e avaliando-o nas restantes.

Para validar a qualidade da previsão e a métrica de desempenho dos modelos *Random Forest* e *XGBoost*, utilizou-se a técnica de validação cruzada do coeficiente de determinação R^2 através da função *cross_val_score* da biblioteca *scikit-learn* do *Python* (SKLEARN, 2024). Essa função permite avaliar o desempenho de um modelo utilizando a técnica de validação cruzada, para explicar a variabilidade e o comportamento das previsões conforme se aproxima de 1, com base nos regressores.

2.6.2.4 Algoritmo genético (GA)

Um algoritmo genético é uma técnica de busca e otimização inspirada na biologia evolutiva, simulando o processo de seleção natural e evolução.

Ele opera com uma população de soluções candidatas, chamadas indivíduos, codificadas por representações abstratas, como sequências de *bits*. A cada geração, os indivíduos são avaliados por uma função de aptidão, medida da qualidade da solução que representam. Os mais aptos são selecionados para gerar novos indivíduos por operadores genéticos, como cruzamento e mutação, que modificam o código genético dos pais. Esse processo continua até que uma condição de parada seja alcançada, como um número máximo de gerações, descoberta de uma solução ótima, ou convergência da população. O algoritmo genético é uma técnica meta-heurística, não garantindo a solução ótima global, mas oferecendo soluções satisfatórias para problemas complexos e desafiadores, difíceis de resolver por métodos convencionais (GENÉTICO, 2024).

2.7 Técnicas de interpretabilidade

A maldição da dimensionalidade afirma que a quantidade de dados necessários para atingir o conhecimento desejado impacta exponencialmente o número de variáveis necessárias (VASCONCELOS, 2022 apud BELLMAN, 1961). Nós, humanos, lidamos bem com duas ou três variáveis, mas nos deparamos com dificuldades diante do grande número de variáveis presentes nos problemas da vida real. Esse ponto se aplica tanto na construção de soluções de tecnologia, na mudança das técnicas de desenvolvimento e engenharia de *software* tradicionais para o uso de *machine learning* e inteligência artificial, quanto na compreensão e diagnóstico dos resultados, através das técnicas de "interpretabilidade" e "explicabilidade".

Nesse mesmo sentido, o relatório *Explainable Artificial Intelligence* (XAI) (SOLUTIONS, 2023) cita diversos autores, incluindo Gall (2018), Editor da *Thoughtworks* e da *The New Stack*, e Broniatowsky (2021), Professor Associado do Departamento de Gestão de Engenharia e Engenharia de Sistemas da Universidade George Washington, que demonstram os desafios na interpretabilidade dos modelos.

As definições de "interpretabilidade" e "explicabilidade" de um modelo tendem a fazer uma certa distinção entre elas, embora, na prática, esses conceitos sejam frequentemente usados de forma intercambiável. Em termos gerais, a interpretabilidade está ligada à capacidade de explicar a um ser humano os resultados

de um modelo (sua relação de causa e efeito), enquanto a explicabilidade está associada à compreensão da lógica interna do algoritmo, como ele é projetado e treinado, e as etapas envolvidas na tomada de decisões para chegar a um determinado resultado.

Ainda conforme Solutions (2023), seguem algumas definições acadêmicas a esse respeito:

- a) interpretabilidade é a capacidade de explicar ou apresentar em termos compreensíveis para um ser humano;
- b) interpretabilidade é o grau em que um ser humano pode entender a causa de uma decisão;
- c) a explicabilidade do resultado de um modelo é a descrição de como o resultado do modelo foi produzido;
- d) explicabilidade é o grau em que a mecânica interna de um sistema de *machine learning* pode ser explicada em termos humanos.

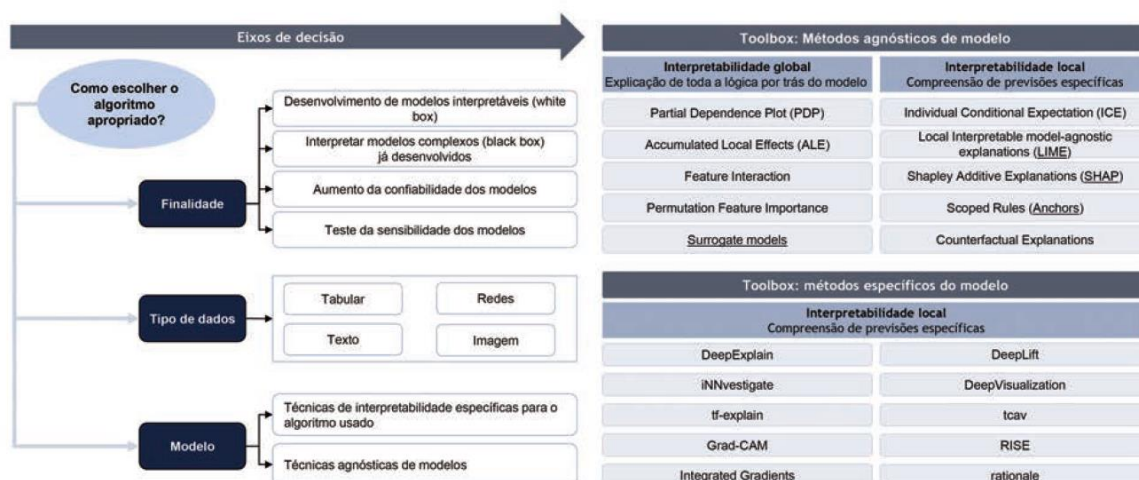
A necessidade de explicabilidade e interpretabilidade dos modelos favoreceu o surgimento de técnicas cada vez mais sofisticadas para a interpretabilidade local e global dos resultados dos modelos, e a situação atual é de certa padronização e convergência no uso de determinadas técnicas (por exemplo, *PDP*, *LIME* ou *SHAP*).

A Figura 5 traz uma visão geral de técnicas de interpretabilidade. Como o estudo em questão tem por finalidade "interpretar modelos complexos já desenvolvidos", compreendendo previsões locais específicas e a importância de suas variáveis, o estudo considera a análise *SHAP* (*SHapley Additive exPlanations*) como técnica mais adequada para a interpretabilidade do modelo em questão.

2.7.1 Análise SHAP (*SHapley Additive exPlanations*)

Conforme citado por (SOLUTIONS, 2023), *SHAP* (*SHapley Additive exPlanations*) é um método de explicação de modelo baseado no Teorema do Valor de Shapley, proposto em 1952 para distribuir o valor de um jogo entre os jogadores. O SHAP é utilizado para explicar a importância de cada variável (medida como a alteração média na previsão do modelo quando o valor da variável varia).

Figura 5 - Visão geral de técnicas de interpretabilidade



Fonte: Solutions (2023),

Especificamente, o *SHAP* usa uma combinação de linhas de base, funções de importância local e o Teorema do Valor de *Shapley* para calcular a importância de cada variável em uma previsão individual.

Nesse método:

- os valores de *Shapley* são calculados, onde as variáveis independentes são interpretadas como jogadores que cooperam para receber o pagamento;
- os valores de *Shapley* correspondem à contribuição de cada variável para a previsão do modelo;
- o pagamento é a previsão real feita pelo modelo menos o valor médio de todas as previsões;
- os jogadores "dividem" esse pagamento de acordo com sua contribuição, e essa divisão é calculada pelos valores de *Shapley* e reflete a importância de cada variável;

Esse método também permite interpretações globais, calculando a média das contribuições de cada variável para cada previsão de modelo.

Formalmente, os valores de *Shapley* podem ser definidos como a contribuição de cada variável para o resultado do modelo, ponderada em relação a todas as combinações possíveis de variáveis usadas, em que val corresponde à previsão do modelo para variáveis incluídas no conjunto s com relação à previsão para variáveis não incluídas em s :

$$val = \int f(x_1 \dots x_p) dP_{x \notin s} E_x(F(X))$$

onde, x é o vetor de variáveis usadas no modelo, s é um subconjunto de x , p é o número de variáveis usadas no modelo, $dP_{(x \notin s)}$ representa o conjunto de variáveis não incluídas em s para as quais a integração é realizada, E é o valor esperado da previsão de X com o modelo f .

Usando esses valores, o SHAP pode ser usado para obter uma explicação local para o modelo como:

$$Expl(x) = Ex(f(X)) + \sum \phi_j x_j$$

Por fim, o *SHAP* também é capaz de calcular explicações locais por meio da agregação de valores de *Shapley* em um conjunto de dados.

3 METODOLOGIA

O projeto de pesquisa, conforme citado por Yin (2021), no sentido mais elementar, é a sequência lógica que conecta os dados empíricos às questões de pesquisa iniciais do estudo e, em última análise, às suas conclusões. Coloquialmente, um projeto de pesquisa é um plano de ação para sair "daqui" e chegar "lá", onde "aqui" pode ser definido como o conjunto inicial de questões a serem respondidas, e "lá" é um conjunto de conclusões (respostas) sobre essas questões.

Conectado ao objetivo do projeto, o estudo da dissertação revela a necessidade de aprofundar em perguntas de 'Como' e 'Por que', conforme Yin (2021). Perguntas essas que estão em desenvolvimento, mas trazem como objeto entender, dentro do contexto das mineradoras, a necessidade de incorporação da inovação, mais especificamente a adoção da inteligência artificial, para a Sustentabilidade e/ou Desenvolvimento Sustentável. Por exemplo: “Como as mineradoras podem alavancar o desenvolvimento sustentável a partir da inovação tecnológica, mais especificamente na adoção da inteligência artificial como técnica de redução de emissão de GEE?” ou “Como a aplicação da inteligência artificial pode ser orientada ao desenvolvimento sustentável e eficiência na utilização de recursos naturais, considerando o contexto global e compromissos globais com a sustentabilidade?”

O objetivo do estudo também revela que o caráter contemporâneo também é importante para o estudo (YIN, 2001). Isto posto, a estratégia que melhor se adequa ao desenvolvimento do trabalho é utilizar a abordagem de estudo de caso como metodologia para aprofundar melhor no tema proposto. Conforme citado por Yin (2021), o estudo de caso permite uma investigação para preservar as características holísticas e significativas dos eventos da vida real, como ciclos de vida individuais, processos organizacionais e administrativos, mudanças ocorridas em regiões urbanas, relações internacionais e a maturação de alguns setores.

Ainda segundo Yin (2021), uma distinção básica que deve ser feita ao projetar estudos de caso é entre projetos de caso único e de casos múltiplos. Significa a necessidade de decidir, antes da coleta de dados, se será utilizado um estudo de caso único ou de casos múltiplos ao formular as questões da pesquisa. Nesse sentido, a proposta desse trabalho é seguir o estudo de caso único, considerando a Vale S/A

como seu estudo de caso, devido à proximidade e experiência do autor, abrangência global da organização e o impacto que ela representa globalmente e dentro do setor.

Nesse sentido, conforme citado no Relato Integrado da VALE (2022), a empresa Vale S/A é uma mineradora fundada em 1942, constituindo-se como uma Sociedade Anônima de capital aberto, com sede na cidade do Rio de Janeiro, Brasil. Suas ações são negociadas no Novo Mercado da B3, a bolsa de valores de São Paulo, e também estão listadas nos mercados financeiros de Nova Iorque (NYSE) e Madri (Bolsa de Madrid-Latibex). Ao longo de sua trajetória desde 1942, consolidou-se como uma das maiores mineradoras globalmente, atuando em mais de 30 países distribuídos pelos cinco continentes (ver Figura 6).

Figura 6 - Vale no Mundo

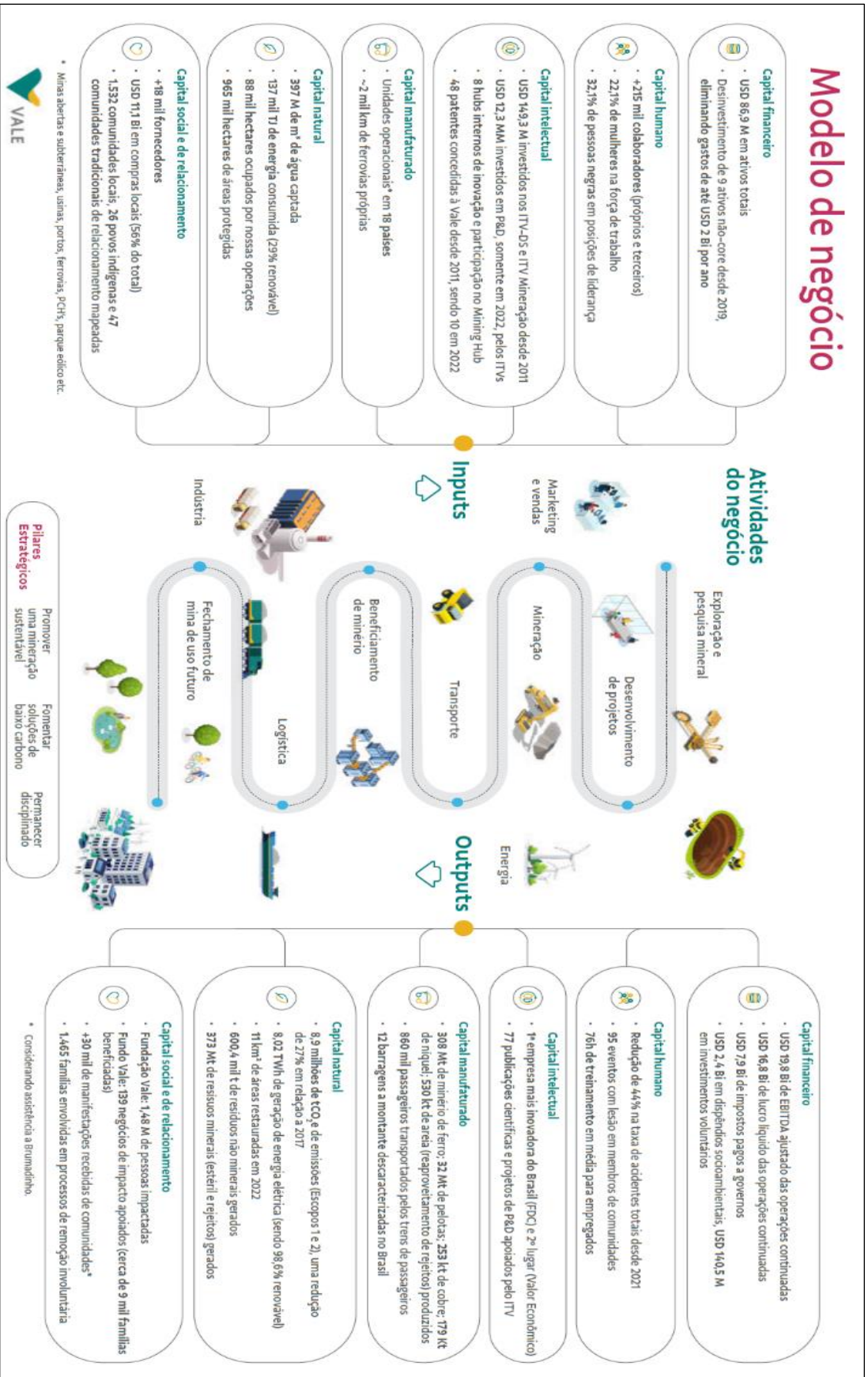


Fonte: Relato integrado (VALE, 2023)

A Vale destaca-se como líder nos mercados mundiais de Minério de Ferro, Pelotas de Minério de Ferro e Níquel. Suas unidades operacionais são integradas a sistemas logísticos abrangentes, incluindo ferrovias, terminais marítimos e portos. Além disso, estende sua presença por meio de coligadas, *joint ventures* e participações diretas em segmentos diversos, como energia, siderurgia e bauxita. Seu modelo de negócio é detalhado na Figura 7.

A Vale S.A. é considerada um estudo de caso único de acordo com Yin (2001), pois representa o caso crítico para testar uma teoria bem formulada.

Figura 7 - Modelo de Negócio



Fonte: Relato integrado (VALE, 2023)

Ainda dentro da Vale S.A., certamente poderemos encontrar diversas iniciativas de estudo voltadas à inovação. A teoria especificou um conjunto claro de proposições, assim como as circunstâncias nas quais se acredita que as proposições sejam verdadeiras. Para confirmar, contestar ou estender a teoria, deve existir um caso único que satisfaça todas as condições para testar a teoria.

O caso único pode, então, ser utilizado para determinar se as proposições de uma teoria são corretas ou se algum outro conjunto alternativo de explicações possa ser mais relevante.

Nesse sentido, para dar foco e atender ao objeto de estudo, a proposta do trabalho é avaliar a aplicação da inteligência artificial para o aumento da eficiência operacional e o atingimento de metas de sustentabilidade, devido à quantidade de ativos e ao impacto ambiental, econômico e social, não só na organização como em todo o ecossistema envolvido (sociedade, empresas externas, parceiros etc.), atendendo aos objetivos gerais e específicos listados com as seguintes etapas (ver Tabela 1).

Tabela 1 - Resumo da metodologia (continua)

Item	Descrição Geral
Objeto de pesquisa	Verificar a aplicação, interpretar as características das variáveis e os possíveis benefícios potenciais, em uma solução de inteligência artificial para aumentar a eficiência energética e reduzir a emissão de Gases de Efeito Estufa (GEE) em caminhões fora de estrada da empresa em estudo.
Natureza da Pesquisa	Pesquisa quantitativa, natureza aplicada, objetivos descritivos, voltada ao estudo de caso único
Método de pesquisa	Obter dados operacionais das minas a céu aberto, observando aspectos operacionais, como, por exemplo, produção, número de equipamentos, dados de eficiência operacional, consumo de combustível, entre outros. Levantamento de dados e tipo de solução de IA implementada.
Unidade de análise; sujeitos de pesquisa	Operadores, analistas, engenheiros, supervisores, gestores das seguintes áreas: Operação de mina, Manutenção e Tecnologia.
Universo e Amostra	Dados operacionais de mina de céu aberto da Vale S/A, com modelos de análise avançada implementados, em períodos de análise que serão detalhados no estudo de caso.
Coleta de dados	Observação sistêmica das áreas operacionais através de sistemas de controle.
Tratamento e análise dos dados	Utilizar a linguagem de programação <i>Python</i> , juntamente com técnicas para análise avançada de dados, a fim de entender e interpretar as correlações das variáveis e definir variáveis relevantes para o comportamento dos modelos e atingimento das metas de sustentabilidade e eficiência energética.

Limitações da pesquisa	Por se tratar de um ambiente operacional externo, não totalmente controlado, não é objetivo do trabalho ser exaustivo no levantamento de fatores que podem ou não contribuir para o atingimento das metas, mas sim verificar técnicas, considerando suas limitações, voltadas à busca de maior eficiência, sustentabilidade operacional e seus resultados.
Resultados	Verificar a aplicação de soluções de <i>Machine Learning</i> em caminhões de grande porte, "fora de estrada", em operações de mina a céu aberto do Brasil. Verificar e comparar, através de métodos de análise de dados, as técnicas de inteligência artificial empregadas, entender características e variáveis relevantes para a busca do aumento da eficiência energética e redução de emissão de gases de efeito estufa (GEE).

Fonte: Próprio Autor

4 ESTUDO DE CASO MODELO DE ANÁLISE AVANÇADA PARA AUMENTO DA EFICIÊNCIA ENERGÉTICA DOS CAMINHÕES FORA DE ESTRADA

4.1 Detalhamento do estudo de caso

Conforme explicitado na metodologia, o estudo de caso considera uma operação em uma mina a céu aberto de minério de ferro da empresa Vale S/A.

A Unidade Operacional está situada na região Norte do país e conta com uma frota de transporte de aproximadamente 120 caminhões fora de estrada em operação.

Para o estudo, foram considerados a operação em uma cava específica da mina com 59 caminhões fora de estrada, sendo 23 caminhões do modelo Caterpillar 793D e 36 do modelo Caterpillar 797F (ver Tabela 2, com as especificações dos equipamentos).

Tabela 2 - Especificações dos equipamentos Catterpillar 793D e 797F

Especificação	Caminhão 793D	Caminhão 797F
Carga útil nominal	231 t	364 t
Peso nominal bruto da máquina	386007 kg	623690 kg
Potência bruta - SAE J1995	1976 kW	2983 kW
Modelo do motor	Cat C175-16	Cat C175-20
Potência líquida - SAE J1349	1641 kW	2828 kW
Velocidade nominal (RPM)	1750 r/min	1750 r/min
Capacidade de combustível	3785 L	3785 L
Comprimento	15,09 m	14,8 m
Altura	7,44 m	6,52 m

Fonte: Adaptado de CAT (2024).

O período de estudo foi de 01/Ago/2023 a 09/Set/2023, e a base conta com 286.357 registros, correspondendo a cada momento em que os caminhões passam carregados pelos trechos determinados.

Foram analisados 296 trechos distintos (ver Figura 8 com o mapa ilustrativo da cava da mina e dos trechos mapeados).

Figura 8 - Trechos da cava da Mina em estudo utilizados para treinamento e execução do modelo



Fonte: Próprio Autor

A base de dados conta ainda com 9 colunas (Data/hora; Trecho; Velocidade do trecho; *GVW* (*Gross Vehicle Weight*); TR (Resistência Total de Rolamento); DMT (Distância Média de Transporte); Dados de posição georreferenciada (DX; DY) e Diesel consumido no trecho (estimado com base na quantidade de litros abastecidos pelo caminhão e trechos rodados no período).

Durante o período do estudo, foram consumidos 426.813,75 litros de Diesel, emitindo aproximadamente 1.140,533 toneladas de CO₂, conforme a Calculadora de Emissões de Gases de Efeito Estufa para Combustíveis (VALE, 2022).

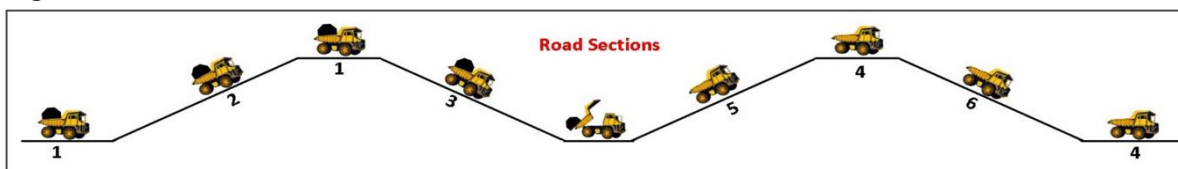
A Vale S/A possui diversas ações e metas de descarbonização e mudança de matriz energética. Dentre elas, destacam-se a redução da emissão de GEE em 33%

e a melhoria em 5% do indicador de eficiência global até 2030 (VALE, 2022). O modelo em questão se propõe a ser uma das iniciativas para atingir tais objetivos.

4.2 Estrutura do modelo e suas variáveis

A operação de mina envolve vários fatores que podem influenciar na eficiência energética e no consumo de combustível dos caminhões, tais como peso do veículo, carga, resistência ao rolamento, inclinação da pista, tração do caminhão, velocidade, tipo de combustível, temperatura ambiente, dentre outros. No entanto, a proposta do modelo é, considerando o estado atual da mina, identificar o ponto ótimo de velocidade por trecho (ver Figura 9) que proporcionará a melhor eficiência energética por caminhão, levando em consideração os parâmetros de Peso Total do Veículo (Peso do caminhão vazio mais a carga - GVW), TR (soma das várias resistências que um veículo ou objeto encontra ao se mover sobre uma superfície) e DMT.

Figura 9 - Sessões da Mina



Fonte: Soofastaei (2016)

Conforme citado por Soofastaei (2016), os modelos de análise de dados avançada para aumento de eficiência energética possuem duas fases:

- a) primeira fase: treinamento, aprendizagem e predição de consumo de combustível. Onde para efeito comparativo, foram utilizados os modelos de *Random Forest* e *XGBoost*, empregando dados de GVW, DMT e TR para treinar e prever o consumo de combustível em litros. Em seguida, aplicou-se a técnica de validação cruzada de R^2 como métrica de desempenho para os modelos de predição;
- b) segunda fase: aplicação do algoritmo genético para busca global de otimização da eficiência energética (consumo de combustível) através do parâmetro de velocidade por trecho do caminhão.

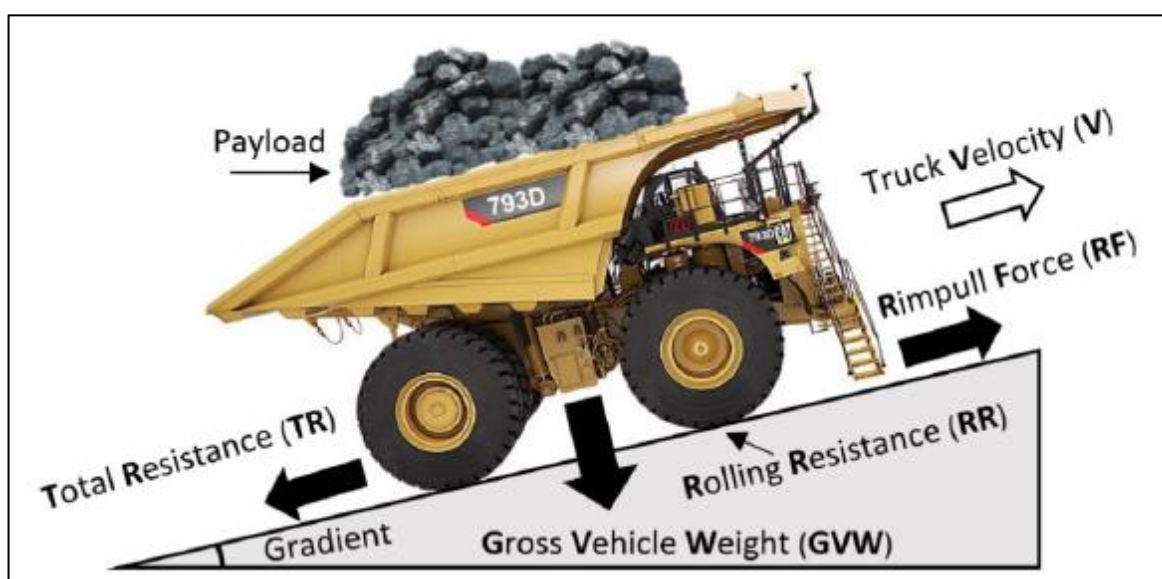
O modelo gera recomendações de velocidade, considerando a carga do caminhão estipulada pela própria área de negócio, para 296 trechos da mina, levando em conta o ciclo em que o caminhão está carregado. Em outras palavras, são fornecidas velocidades específicas por trecho, considerando características da via e GVW.

4.2.1 Consumo de combustível em caminhões fora de estrada

O caminhão fora de estrada CAT 793D e 797f consomem, em média, 170 litros por hora trabalhada (CAT, 2023). O consumo de combustível para caminhões de transporte é determinado com base nos seguintes parâmetros, conforme citado por Soofastaei (2016) (ver Figura 10):

- a) o Peso Total do Veículo (GVW): soma do peso de um caminhão vazio e da carga útil;
- b) a velocidade do caminhão de transporte (V);
- c) a Resistência Total (TR): quando o caminhão se move contra o nivelamento da estrada de transporte;
- d) a Força *Rimpull* (tração) (RF): força disponível entre o pneu e o solo para impulsionar o caminhão.

Figura 10 - Parâmetros de consumo de combustível



Fonte: Soofastaei (2016)

4.3 Recomendações do modelo por trecho

O modelo de análise avançada de dados gera recomendações (ver Tabela 3 para um exemplo de recomendações) de velocidade por trecho da mina. A partir das velocidades recomendadas, é possível ter diversas visões de acompanhamento da aderência às recomendações (por operador, por trecho, por equipamento, por rota, etc.). Essas informações são usadas como *feedback* para o modelo, além de possibilitar ações de qualificação e gamificação com os operadores, e reparação de trechos irregulares da mina, onde era prevista uma determinada velocidade, mas devido a problemas ambientais ou de infraestrutura, está sendo realizada outra. Todas as ações visam a maior aderência às velocidades recomendadas e, por consequência, maior eficiência energética, menor emissão de GEE e consumo de combustível dos caminhões fora de estrada.

Tabela 3 - Exemplo de recomendação de velocidade por trecho

Trecho da Mina	Vel.Média Executada (KM/h)	Vel. Recomendada (KM/h)	Aderência (%)
NT_1_B_1_L_76_para_NT_1_B_0_L_76	19,53442	20	86,34772
NT_12_B_12_L_69_para_NT_12_B_11_L_69	14,452	15	86,27063
NT_1_B_1_L_92_para_NT_1_B_0_L_92	9,324308	10	86,09231
NT_3_B_1_L_32_para_NT_3_B_2_L_32	34,11842	36	85,96491
NT_3_B_2_L_32_para_NT_3_B_1_L_32	34,17121	36	84,66109
NT_1_B_1_L_95_para_NT_1_B_0_L_95	15,2094	16	80,40268
NT_1_B_0_L_92_para_NT_1_B_1_L_92	9,003584	10	79,39068
NT_12_B_2_L_71_para_NT_12_B_1_L_71	12,34468	14	79,38253
NT_1_B_0_L_95_para_NT_1_B_1_L_95	14,96634	16	77,91461
NT_12_B_0_L_69_para_NT_12_B_1_L_69	33,59701	38	75,62189
NT_13_B_1_L_35_para_NT_13_B_0_L_35	11,50951	12	75,28517
NT_3_B_2_L_31_para_NT_3_B_1_L_31	31,35374	36	74,58929
NT_5_B_3_L_74_para_NT_5_B_2_L_74	12,84096	14	74,45783
NT_5_B_4_L_74_para_NT_5_B_3_L_74	13,71931	15	73,3691
NT_3_B_1_L_87_para_NT_3_B_2_L_87	13,86686	15	72,50972
NT_10_B_7_L_79_para_NT_10_B_6_L_79	14,25245	16	70,55495
NT_6_B_5_L_73_para_NT_6_B_4_L_73	10,55172	11	70,11494
NT_12_B_1_L_71_para_NT_12_B_0_L_71	12,6177	14	67,9485
NT_3_B_2_L_87_para_NT_3_B_3_L_87	16,64926	18	67,50335

Fonte: Próprio Autor

4.4 Resultados potenciais

Também faz parte da aplicação do modelo o estudo de estimativa e viabilidade de implantação da solução. Durante o trabalho, busca-se demonstrar ganhos potenciais com base em implantações anteriores e na quantidade de equipamentos na localidade. Neste estudo de caso, considera-se como benefício potencial a previsão da quantidade de litros de diesel consumidos (estimativa), no mesmo período de estudo, levando em conta as mesmas horas trabalhadas e o mesmo volume transportado. Foram utilizados os próprios modelos de previsão de litros (*Random Forest* e *XGBoost*) para estimar o consumo de combustível, substituindo as velocidades reais pelas recomendações de velocidade geradas. Consulte a Tabela 4 para um resumo.

Tabela 4 - Benefício potencial com a utilização de técnicas de inteligência artificial para recomendação de velocidade.

Condição de operação	Consumo em Litros	Emissão de t CO ₂	Comparativo com consumo real
Normal (Sem interferência de modelos de Inteligência Artificial)	426.813,75	1.140,533	-
Utilizando as velocidades geradas pelo modelo <i>Random Forest</i>	321.629,85	859,460	-24,64%
Utilizando as velocidades geradas pelo modelo <i>XGBoost</i>	330.563,3	883,332	-22,55%

Fonte: Adaptado de Calculadora de Emissões de Gases de Efeito Estufa para Combustíveis (VALE, 2022)

Certamente, o benefício potencial está sujeito a incertezas operacionais e tecnológicas para ser atingido, mas a indicação de uma redução potencial acima de 20% no consumo de diesel e nas emissões de GEE pode ser considerada um sinal positivo para os objetivos de maior eficiência energética, considerando as recomendações de velocidades geradas pelos modelos de previsão e otimizadas pelo algoritmo genético.

O estudo teve início em 2017 e, atualmente, conta com seis unidades operacionais implantadas, mais de 900 trechos mapeados para indicação de velocidade e aproximadamente 500 operadores capacitados para seguir as recomendações. O ganho realizado varia conforme a quantidade de equipamentos na

unidade operacional (UO), o tipo de equipamento, assim como as condições da via e a aderência às recomendações pelos operadores. A seguir, a Tabela 5 apresenta os benefícios captados no último ano (2023) com a utilização do modelo.

Tabela 5 - Benefícios realizados com a utilização do modelo (2023)

Unidades operacionais	Redução de consumo de Diesel (Litros)	Redução de emissão de CO₂ (t)
UO-1	26.072,5	69,671
UO-2	22.111,3	59,086
UO-3	17.068,4	45,610
UO-4	106.225,6	283,856
UO-5	571.273,0	1.526,557
UO-6	88.190,8	235,664
Total	830.941,6	2.220,444

Fonte: Próprio Autor

4.4.1 Pessoas no centro

Importante ressaltar a relevância da participação dos operadores de caminhões fora de estrada nos resultados de aumento da eficiência energética. Apesar das recomendações de velocidades mais eficientes a partir do modelo de análise avançada, para atingir o resultado esperado é necessário o total engajamento e empoderamento das áreas operacionais envolvidas.

Durante a implantação do modelo nas unidades operacionais anteriores, foram realizados treinamentos, capacitações e gestão de mudança junto aos líderes de operação e operadores de caminhões fora de estrada. O objetivo era deixar claro a importância da participação deles para alcançar os resultados descritos na tabela 5, fazendo com que se sentissem parte da solução.

Os treinamentos visam conectá-los ao propósito de não apenas operar os equipamentos, mas operá-los de forma mais eficiente energeticamente, seguindo as indicações de velocidade conforme o trecho, indicadas pelo modelo de análise avançada.

Somente com a aderência em campo às velocidades recomendadas é possível atingir os níveis potenciais de redução de consumo de diesel e de emissão de gases

de efeito estufa dos caminhões fora de estrada. Dessa forma, a capacitação e o engajamento dos operadores são parte fundamental para o atingimento dos objetivos de aplicação de técnicas de Análise Avançada de Dados e Inteligência Artificial na busca da maior eficiência energética, redução do consumo de combustíveis fósseis e diminuição das emissões de gases de efeito estufa dos caminhões de grande porte.

Para exemplificar, durante as implantações do modelo de eficiência energética em unidades operacionais anteriores, foram treinados 268 operadores no Complexo de Itabira, 98 na Mina de Horizontes e 117 na Mina de Capão Xavier. Os treinamentos foram extremamente relevantes para o aumento da aderência às velocidades recomendadas e, conseqüentemente, para a maior eficiência energética dos caminhões fora de estrada.

Ou seja, para que haja o atendimento do benefício potencial, é necessário colocar as pessoas no centro e envolver a área que mais influencia nos resultados de eficiência energética desde o início dos trabalhos. Os resultados só acontecem combinando a evolução de processos, pessoas e sistemas.

4.5 Interpretabilidade dos modelos e resultados

4.5.1 XGBoost

4.5.1.1 Curva de aprendizagem

A curva de aprendizagem mostra a evolução do R^2 em relação ao aumento do conjunto de base de treinamento. Foi aplicado 80% da base de dados para treinamento e 20% para teste.

Pode-se observar que, para o conjunto da base de teste, o valor é 0,8934. Isso significa que o modelo explica 89,34% da variância da variável dependente (litros consumidos) a partir dos regressores (variáveis independentes: GVW, DMT e Velocidade do trecho) incluídos.

Detalhes do Modelo (*XGBoost*) para o conjunto de teste:

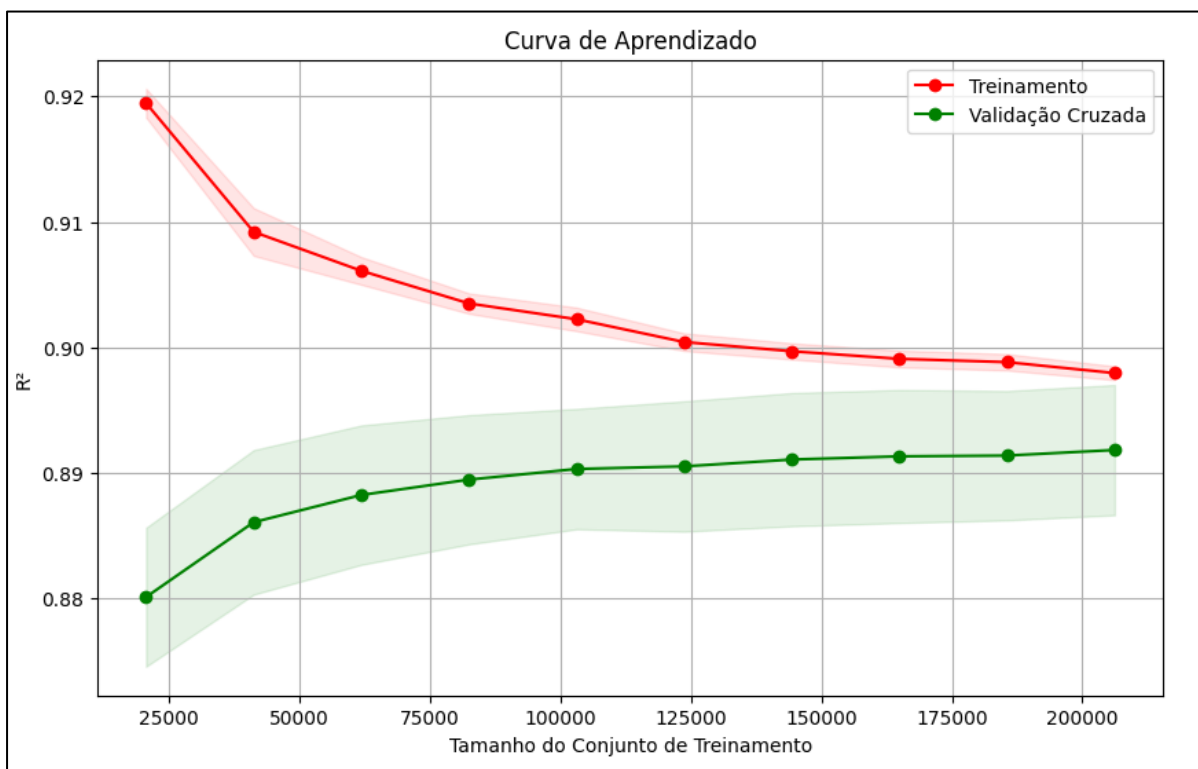
- a) Erro Médio Absoluto (MAE): 0,2462;
- b) Coeficiente de Determinação (R^2): 0,8934;
- c) Erro Médio Absoluto Percentual (MAPE): 3,0151%.

Detalhes do Modelo (*XGBoost*) para o conjunto de treino:

- a) Erro Médio Absoluto (MAE): 0,2401;
- b) Coeficiente de Determinação (R^2): 0,8975;
- c) Erro Médio Absoluto Percentual (MAPE): 2,9792%.

A Figura 11, curva de aprendizagem, mostra a evolução do R^2 em relação ao aumento do conjunto de base de treinamento. Pode-se notar uma aproximação e estabilização do R^2 a partir de 175 mil registros. O indicador ajuda a mitigar a existência de *overfitting* (um fenômeno em modelos de *machine learning* que ocorre quando um modelo aprende detalhes específicos dos dados de treinamento que não são representativos do padrão geral) e demonstra a capacidade da solução em explicar a variância da variável dependente (litros consumidos).

Figura 11 - Curva de aprendizado (XGboost)



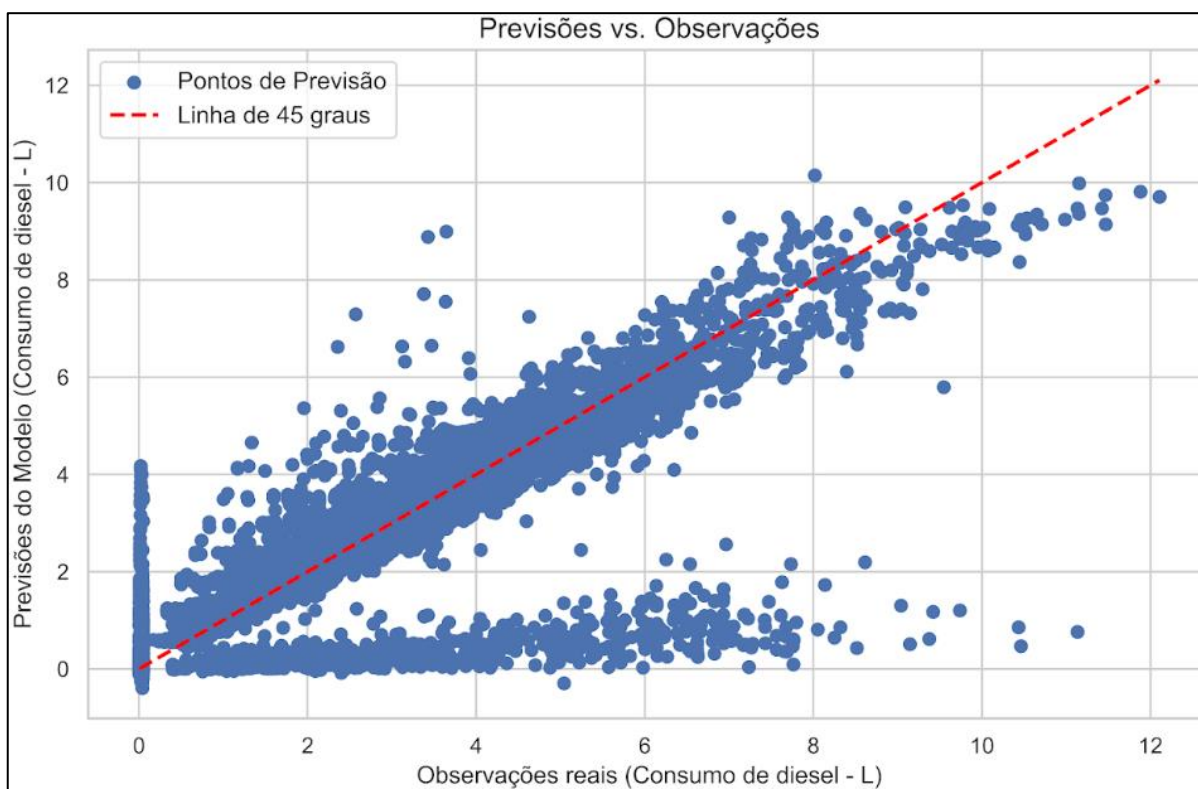
Fonte: Próprio Autor

4.5.1.2 Predições e observações

A Figura 12, o gráfico de previsões versus observações - *XGBoost*, apresenta as seguintes características:

- o eixo horizontal representa os valores reais (observações) da variável de destino (litros de diesel consumidos);
- o eixo vertical representa os valores previstos pelo modelo *XGBoost* para a variável de destino (litros de diesel consumidos);
- a linha diagonal pontilhada representa a linha de ajuste perfeito, onde as previsões correspondem exatamente às observações reais;

Figura 12 - Previsões vs Observações (XGBoost)



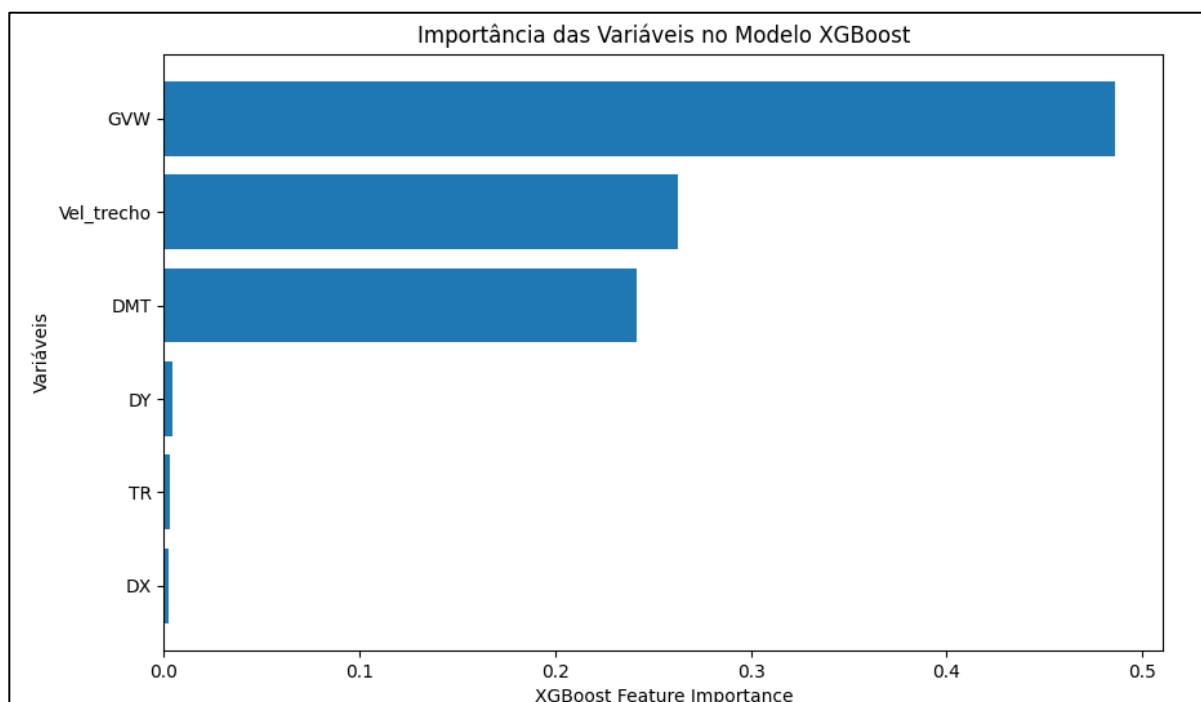
Fonte: Próprio Autor

A maioria dos pontos está próxima à linha de ajuste perfeito, indicando que o modelo *XGBoost* está fazendo previsões precisas para a maioria das instâncias. Há alguns pontos acima da linha de ajuste perfeito, o que significa que o modelo superestimou o valor da variável de destino.

Há também alguns pontos abaixo da linha de ajuste perfeito, o que significa que o modelo subestimou o valor da variável de destino para essas instâncias. Em resumo, o gráfico fornece uma visão geral do desempenho do modelo *XGBoost* em todo o conjunto de dados. O modelo apresenta previsões precisas para a maioria das instâncias, o que demonstra sua boa representatividade. A dispersão dos pontos ao redor da linha de ajuste perfeito indica a variabilidade das previsões, confirmando também a ausência de *overfitting*, corroborando com o gráfico anterior.

A Figura 13 demonstra a contribuição média dos valores de SHAP por módulo. Pode-se notar que o *GVW* contribui com cerca de 48% para a determinação da variável alvo, litros de combustível. Em sequência, a Velocidade por trecho contribui com cerca de 25%, e o *DMT* com aproximadamente 22%.

Figura 13 - Importância das variáveis (XGboost)



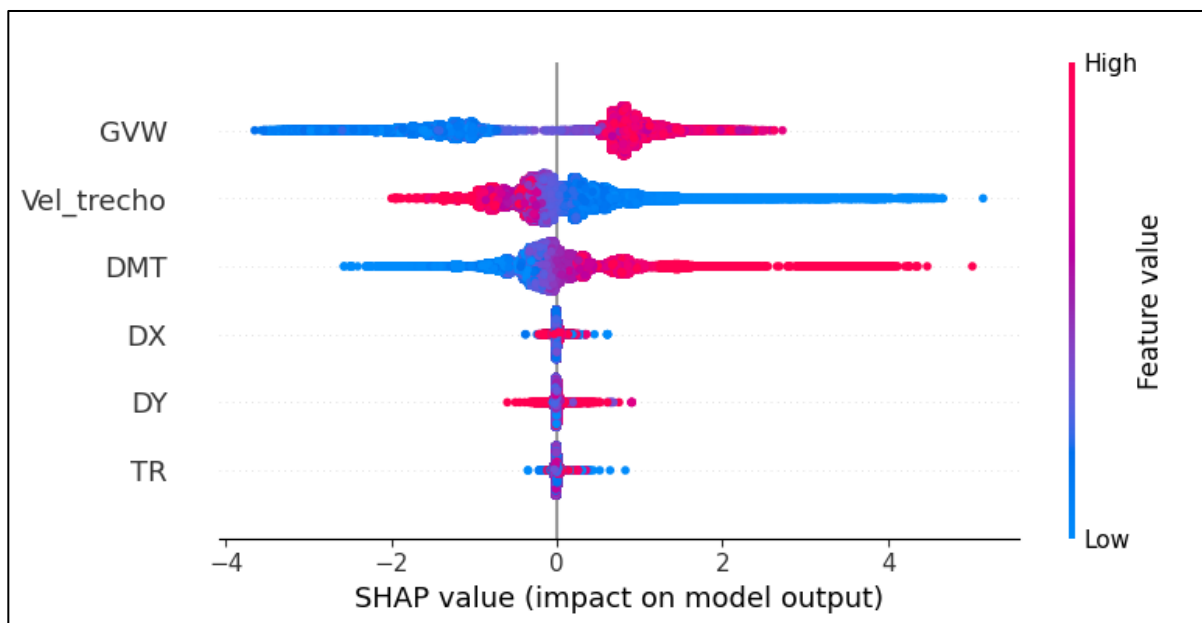
Fonte: Próprio Autor

4.5.1.3 Análise SHAP das variáveis

As Figuras 14 e 15 são resultados da função *Shap Summary Plot*, representando a base de dados de treinamento e teste, respectivamente.

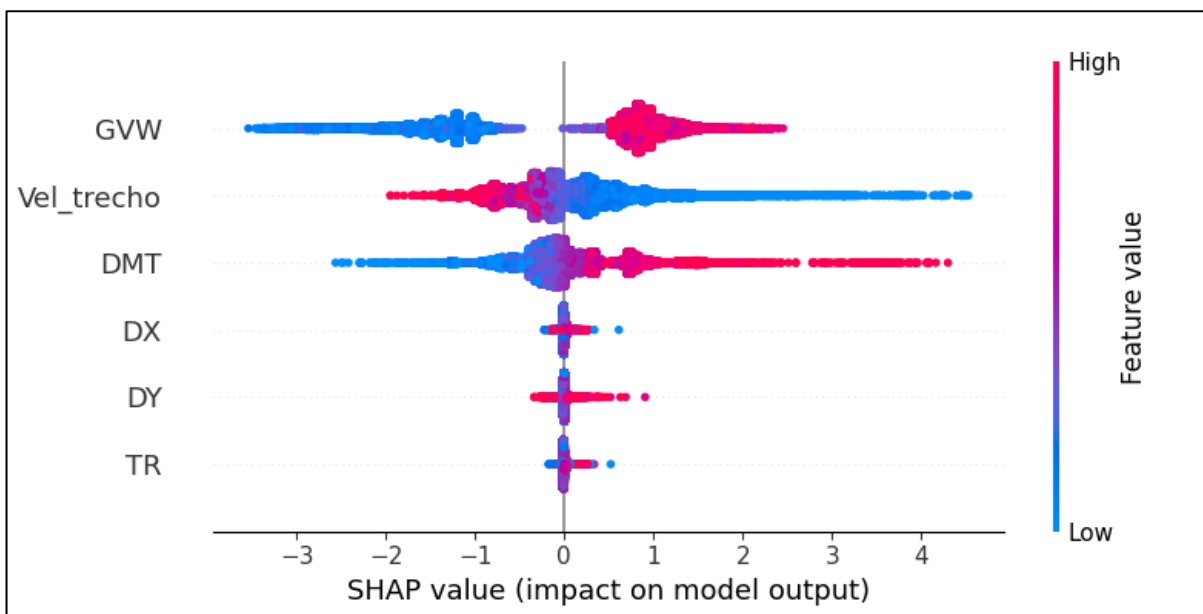
Observa-se um comportamento semelhante nas duas bases de dados, o que pode ser interpretado como um comportamento homogêneo e bem distribuído da base.

Figura 14 - Importância das variáveis XGBoost (treinamento)



Fonte: Próprio Autor

Figura 15 - Importância das variáveis XGBoost (teste)



Fonte: Próprio Autor

Seguem algumas premissas:

- a) eixo X: impacto do valor da variável regressora no resultado da previsão do modelo (litros de Diesel);
- b) eixo Y: variáveis regressoras. A função ordena, por padrão, as características de maior relevância para o modelo;
- c) pontos azuis: indicam valores baixos das variáveis regressoras;
- d) pontos vermelhos indicam valores altos das variáveis regressoras.

Análise das principais variáveis regressoras:

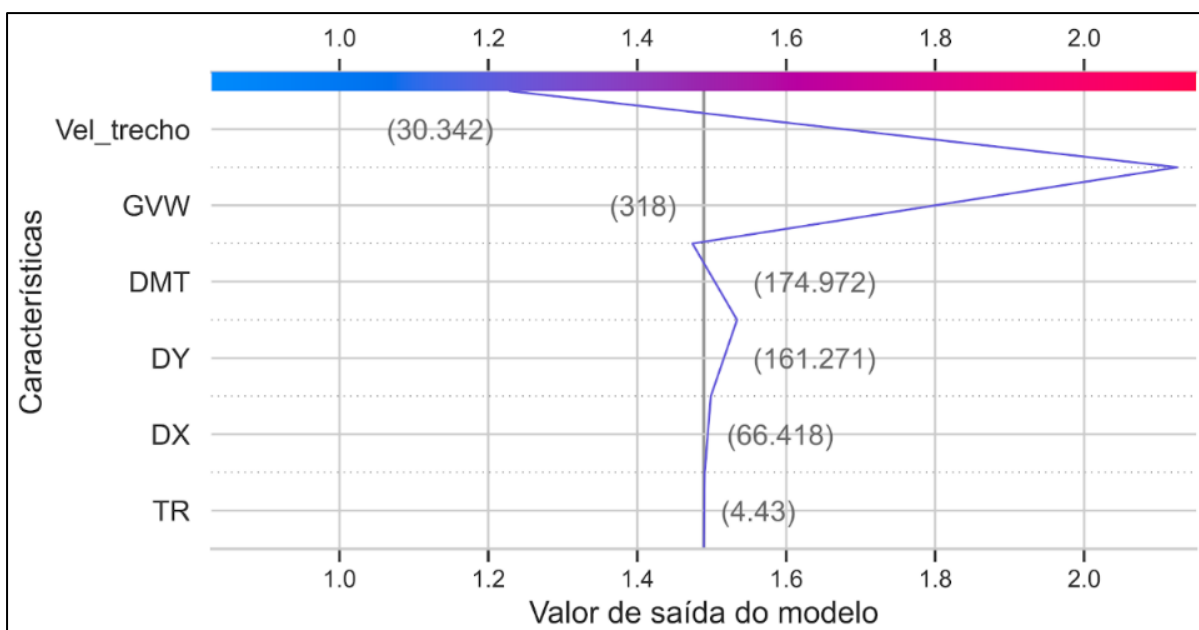
- a) *GVW*: observa-se uma concentração de pontos vermelhos, influenciando positivamente o modelo. Ou seja, valores altos de *GVW* contribuem para o aumento do consumo de Diesel;
- b) Velocidade: há uma clara separação, com concentração de valores azuis no lado positivo e valores vermelhos no lado negativo (o inverso pode ser observado em *GVW*). Isso indica que velocidades mais altas estão associadas a menor consumo de Diesel, enquanto velocidades mais baixas estão relacionadas a maior consumo de Diesel;
- c) *DMT*: Quanto maior a distância média de transporte, maior o consumo de Diesel;
- d) as variáveis de localização (*DX* e *DY*) e *TR* (resistência total de rolamento) não influenciam na determinação do consumo de Diesel.

Interpretação Geral: as características Velocidade e *GVW* são inversamente proporcionais, sendo o cenário mais eficiente em termos de consumo de Diesel aquele com velocidades mais altas, menor *GVW* e *DMT* abaixo da média.

A Figura 16 apresenta a implementação da função *Shap Decision Plot*. Com a função, é possível avaliar o "percurso" da previsão de litros de Diesel por amostra. A solução permite o "diagnóstico" e a interpretabilidade do modelo, exemplificando o comportamento dos dados por amostra. Através do *Shap Decision Plot*, é possível demonstrar como o modelo chegou à previsão e qual foi a influência das variáveis regressoras nesse processo.

Como mencionado anteriormente, que *TR*, *DX* e *DY* tiveram pouca influência no resultado, assim como o *DMT* (trecho de 174 metros). No entanto, o valor de *GVW* em 318 toneladas influenciou para um aumento no consumo, enquanto a velocidade de aproximadamente 30 km/h ajudou a reduzir o consumo de Diesel.

Figura 16 - Shap Decision Plot (XGBoost)



Fonte: Próprio Autor

O modelo prevê um consumo de 1.22 litros de Diesel naquele trecho específico. A Figura 17 implementa a função *Shap Dependence Plot*, semelhante aos demais gráficos.

A função também utiliza como base os valores de *Shapley* para a construção da nuvem de pontos. Com essa função, é possível criar diversos gráficos que demonstram a correlação entre as variáveis e suas influências nos resultados dos modelos.

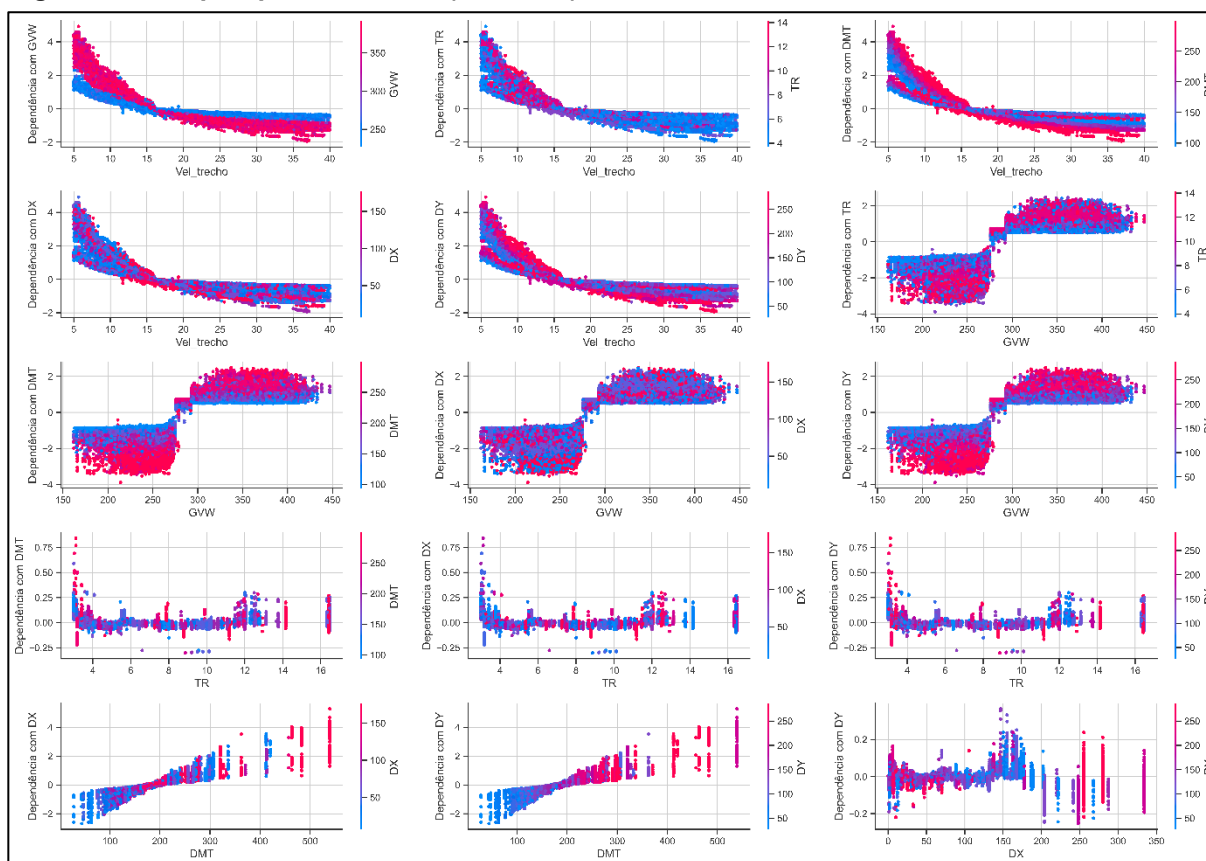
A inclinação da nuvem de pontos indica a relação entre a variável e a previsão do modelo. Se a inclinação for positiva, significa que valores maiores da variável tendem a aumentar o valor de previsão do modelo. Se a inclinação for negativa, significa que valores maiores da variável tendem a diminuir o valor da previsão do modelo. Se a inclinação for nula, significa que a variável não tem influência na previsão do modelo.

A dispersão dos pontos indica a variabilidade dos valores *SHAP* para cada valor da variável. Se os pontos estiverem próximos de uma linha, significa que a variável tem uma influência consistente na previsão do modelo.

Se os pontos estiverem espalhados, significa que a variável tem uma influência variável na previsão do modelo, dependendo de outras variáveis.

A cor dos pontos indica o valor de outra variável que pode interagir com a variável principal.

Figura 17 - Shap Dependence Plot (XGBoost)



Fonte: Próprio Autor

Se os pontos de cores diferentes tiverem valores SHAP diferentes para o mesmo valor da variável principal, significa que há uma interação entre as duas variáveis. Se os pontos de cores diferentes tiverem valores SHAP similares para o mesmo valor da variável principal, significa que não há uma interação entre as duas variáveis. Essa função é uma análise complementar às demais, onde é possível, por exemplo, observar que quanto maior a velocidade e menor o GVW, menor é o consumo de diesel.

4.5.2 Random Forest

O comportamento do modelo *Random Forest* é extremamente semelhante ao modelo *XGBoost*. Dessa forma, para evitar repetições na análise dos gráficos, irei me ater às diferenças encontradas entre os modelos.

4.5.2.1 Curva de aprendizagem

Para o modelo *Random Forest*, pode-se observar que, para o conjunto da base de teste, o R^2 está em 0,8842. Isso significa que o modelo explica 88,42% da variância da variável dependente (litros consumidos) a partir dos regressores (variáveis independentes: GVW, DMT e Velocidade do trecho) incluídos.

Detalhes do Modelo (*Random Forest*) para o conjunto de teste:

- a) Erro Médio Absoluto (MAE): 0,2499.
- b) Coeficiente de Determinação (R^2): 0,8842;
- c) Erro Médio Absoluto Percentual (MAPE): 2,9565%.

Detalhes do Modelo (*Random Forest*) para o conjunto de treino:

- a) Erro Médio Absoluto (MAE): 0,2442;
- b) Coeficiente de Determinação (R^2): 0,8952;
- c) Erro Médio Absoluto Percentual (MAPE): 2,9115%.

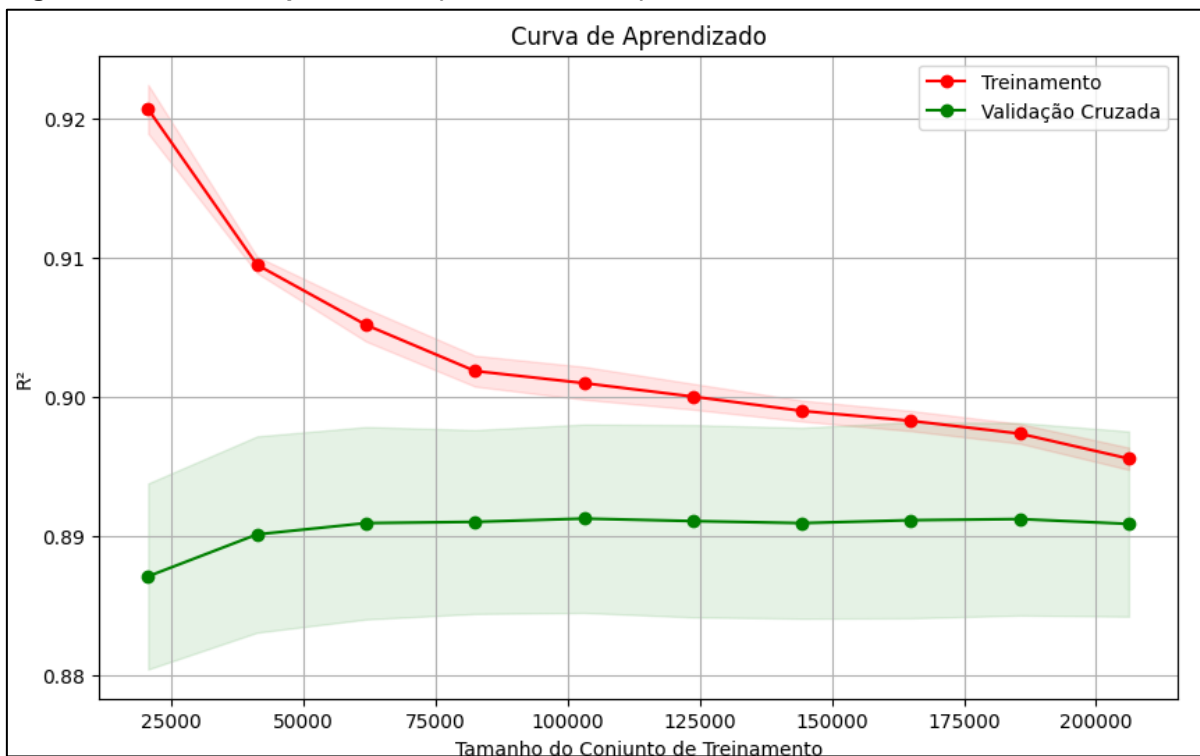
A Figura 18, pode-se notar uma aproximação e estabilização do R^2 a partir de 175 mil registros. O indicador ajuda a mitigar a existência de *overfitting* e demonstra a capacidade da solução em explicar a variância da variável dependente (litros consumidos).

4.5.2.2 Predições e observações

A Figura 19, o gráfico de previsões versus observações (*Random Forest*), apresenta as seguintes características:

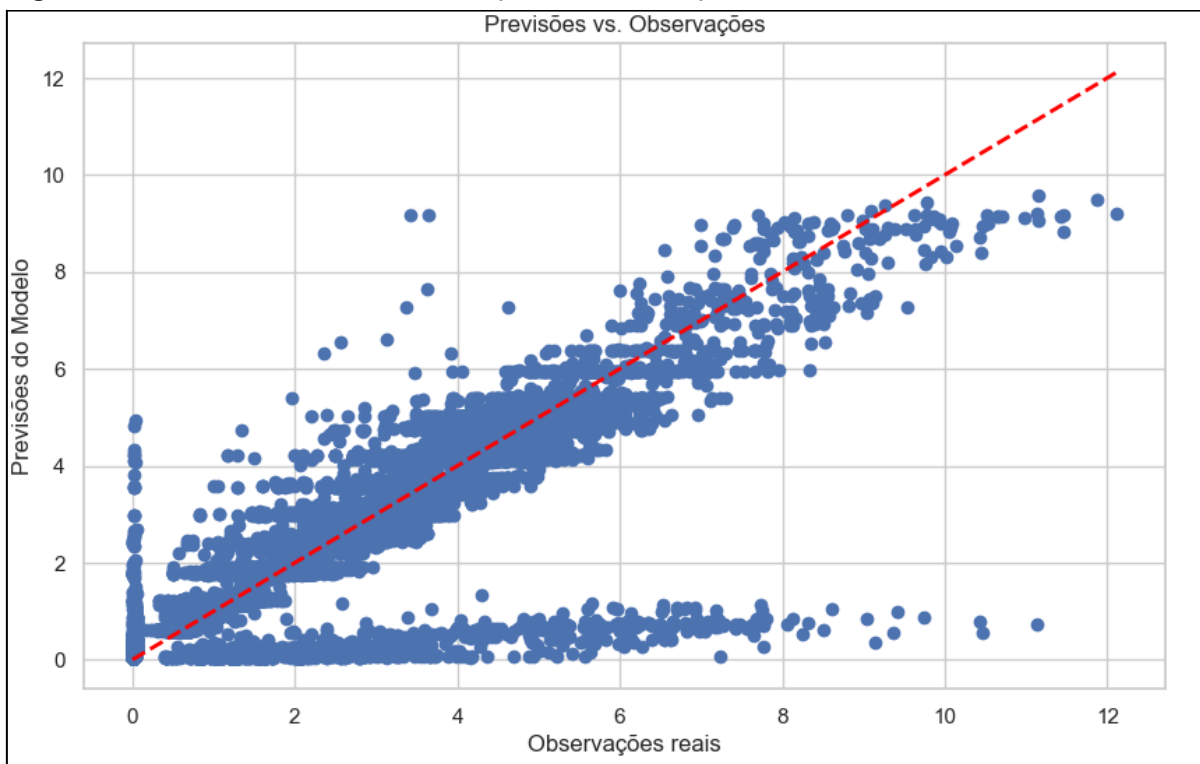
- a) O eixo horizontal representa os valores reais (observações) da variável de destino (litros de diesel consumidos);
- b) O eixo vertical representa os valores previstos pelo modelo para a variável de destino (litros de diesel consumidos);
- c) A linha diagonal pontilhada representa a linha de ajuste perfeito, onde as previsões correspondem exatamente às observações reais.

Figura 18 - Curva de aprendizado (Random Forest)



Fonte: Próprio Autor

Figura 19 - Previsões vs Observações (Random Forest)



Fonte: Próprio Autor

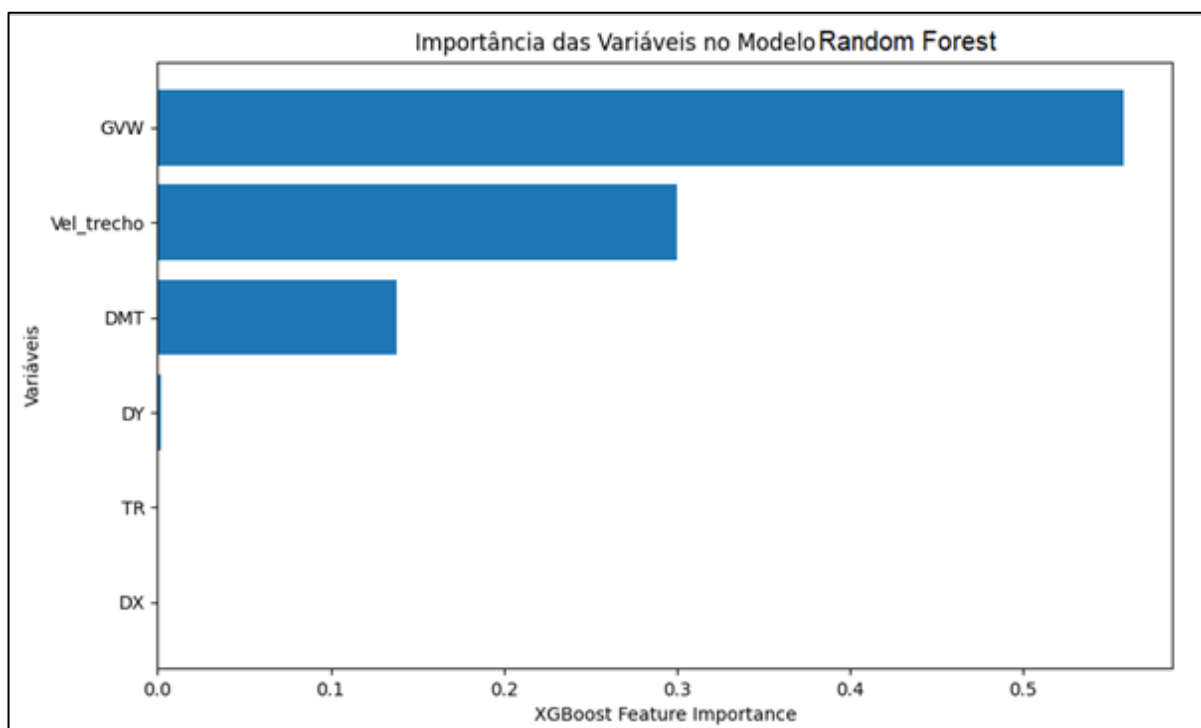
Semelhante ao gráfico do modelo *XGBoost*, a maioria dos pontos está próxima à linha de ajuste perfeito, indicando que previsões precisas para a maioria das instâncias.

Há alguns pontos acima da linha de ajuste perfeito, o que significa que o modelo superestimou o valor da variável de destino (litros de diesel consumidos) para essas instâncias. Há também alguns pontos abaixo da linha de ajuste perfeito, o que significa que o modelo subestimou o valor da variável de destino para essas instâncias.

Em resumo, o gráfico fornece uma visão geral do desempenho do modelo *Random Forest* em todo o conjunto de dados. O modelo também apresenta previsões precisas para a maioria das instâncias, o que demonstra sua boa representatividade. A dispersão dos pontos ao redor da linha de ajuste perfeito indica a variabilidade das previsões, confirmando a ausência de *overfitting*, corroborando com o gráfico anterior.

A Figura 20 demonstra a contribuição média dos valores de *SHAP* por módulo. Pode-se notar que o *GVW* (Peso do veículo vazio mais a carga útil) contribui com cerca de 55% para a determinação da variável alvo, litros de combustível. Em sequência, a Velocidade por trecho contribui com cerca de 30%, e o *DMT* (Distância média de transporte) com aproximadamente 13%

Figura 20 - Importância das variáveis (Random Forest)

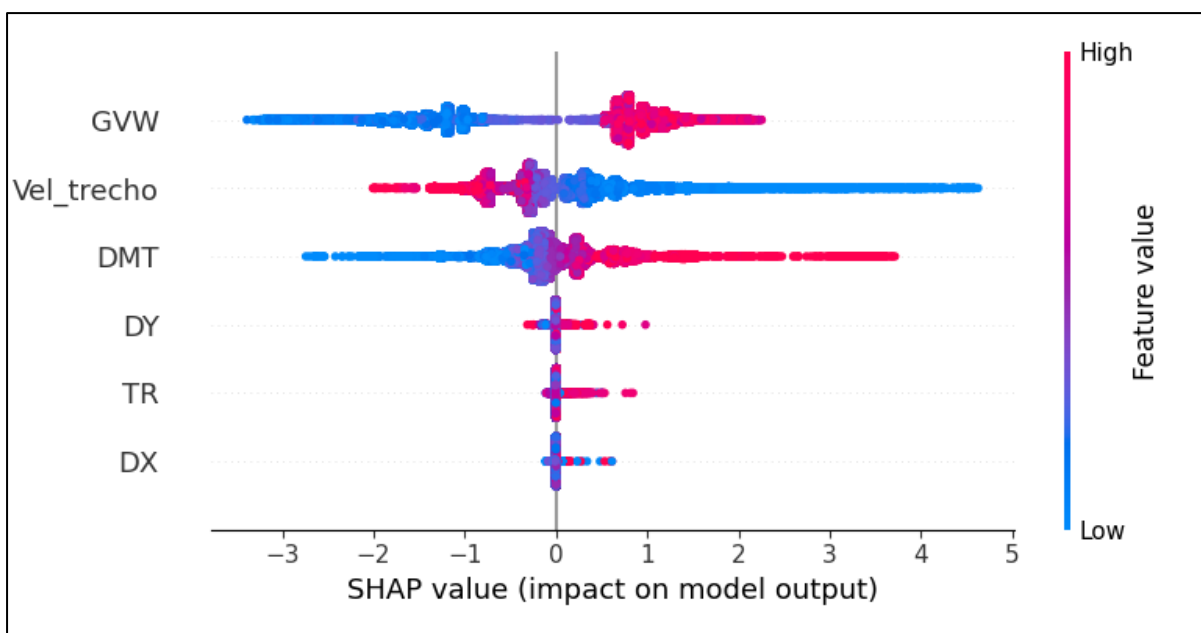


Fonte: Próprio Autor

4.5.2.3 Análise SHAP das variáveis

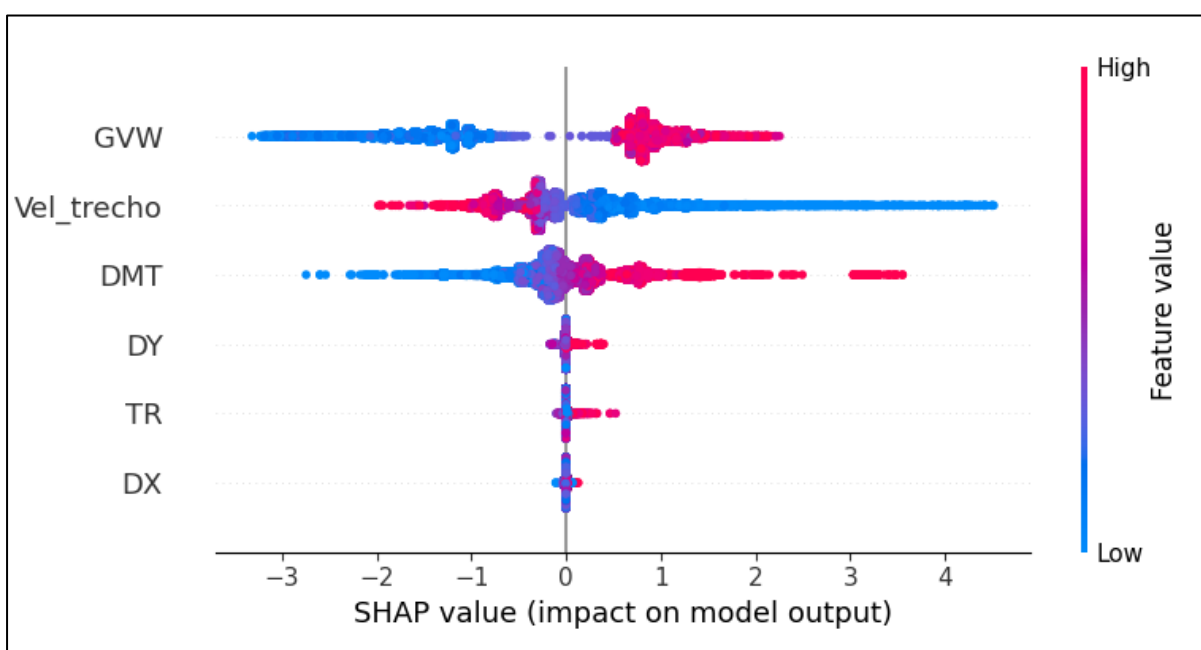
As Figuras 21 e 22 são resultados da função *Shap Summary Plot*, representando a base de dados de treinamento e teste, respectivamente.

Figura 21 - Importância das variáveis Random Forest (treinamento)



Fonte: Próprio Autor

Figura 22 - Importância das variáveis Random Forest (teste)



Fonte: Próprio Autor

Também se observa um comportamento semelhante nas duas bases de dados, o que pode ser interpretado como um comportamento homogêneo e bem distribuído da base. Seguem algumas premissas:

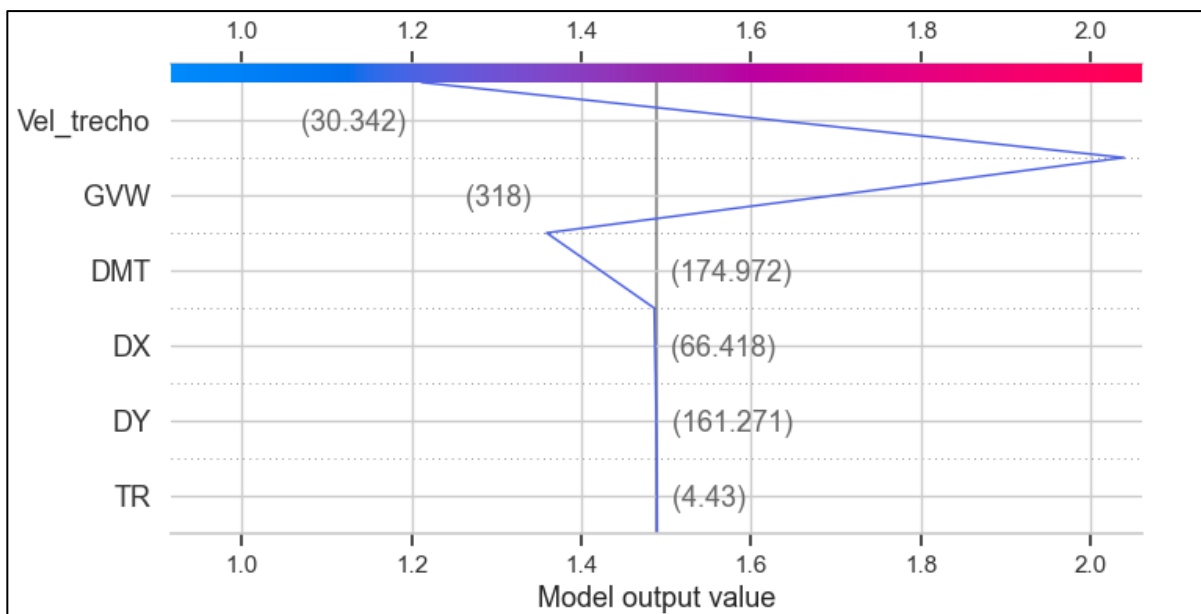
- a) eixo X: impacto do valor da variável regressora no resultado da previsão do modelo (litros de Diesel);
- b) eixo Y: variáveis regressoras. A função ordena, por padrão, as características de maior relevância para o modelo;
- c) pontos azuis: indicam valores baixos das variáveis regressoras;
- d) pontos vermelhos indicam valores altos das variáveis regressoras.

Análise das principais variáveis regressoras:

- a) *GVW*: observa-se uma concentração de pontos vermelhos, influenciando positivamente o modelo. Ou seja, valores altos de *GVW* contribuem para o aumento do consumo de Diesel;
- b) Velocidade: há uma clara separação, com concentração de valores azuis no lado positivo e valores vermelhos no lado negativo (o inverso pode ser observado em *GVW*). Isso indica que velocidades mais altas estão associadas a menor consumo de Diesel, enquanto velocidades mais baixas estão relacionadas a maior consumo de Diesel;
- c) *DMT*: quanto maior a distância média de transporte, maior o consumo de Diesel;
- d) as variáveis de localização (*DX* e *DY*) e *TR* (resistência total de rolamento) não influenciam na determinação do consumo de Diesel.

As características Velocidade e *GVW* também são inversamente proporcionais, sendo o cenário mais eficiente em termos de consumo de Diesel aquele com velocidades mais altas, menor *GVW* e *DMT* abaixo da média. A Figura 23 apresenta a implementação da função *Shap Decision Plot*. A solução permite o "diagnóstico" e a interpretabilidade do modelo e mostra como o modelo chegou à previsão e qual foi a influência das variáveis regressoras. Com a mesma amostra, é possível notar, que *TR*, *DX* e *DY* tiveram pouca influência no resultado, assim como o *DMT* (trecho de 174 metros). No entanto, o valor de *GVW* em 318 toneladas influenciou um aumento no consumo. Na velocidade de aproximadamente 30 km/h o modelo ajudou a reduzir o consumo de Diesel, e previu um consumo de 1.20 litros naquele trecho.

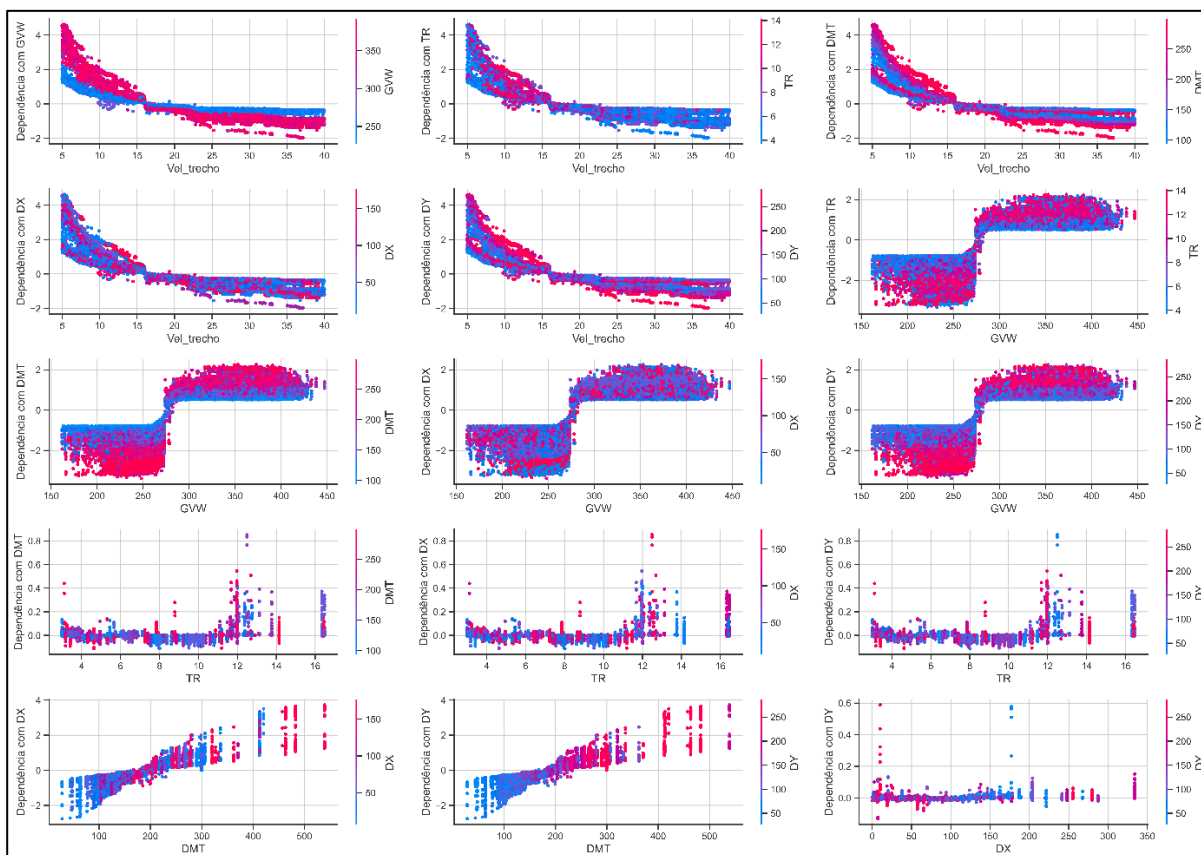
Figura 23 - Shap Decision Plot (Random Forest)



Fonte: Próprio Autor

A Figura 24 implementa a função *Shap Dependence Plot*. A função também utiliza como base os valores de *Shapley* para a construção da nuvem de pontos.

Figura 24 - Shap Dependence Plot (Random Forest)



Fonte: Próprio Autor

Semelhante ao *XGBoost* a função confirma as correlações e análises anteriores, onde também é possível, por exemplo, observar que quanto maior a velocidade e menor o *GVW*, menor é o consumo de diesel

4.5.3 Quadro Resumo

Apesar de serem modelo semelhantes, baseados em árvore de decisão, podem ser usados tanto para classificação quanto regressão.

Os modelos *Random Forest* e *XGBoost* possuem singularidades que os diferem. Conforme algumas características dos modelos levantadas nos tópicos anteriores, bem como os tipos dos dados, segue Tabela 6, com o resumo das características e comportamentos dos dois modelos de predição analisados (*Random Forest* e *XGBoost*).

Tabela 6 - Características dos modelos XBoost e Random Forest

Características	Modelo de Predição	
	<i>XGBoost</i>	<i>Random Forest</i>
Treinamento	Sequencial, com ajuste dos erros anteriores	Paralelo e independente
Overfitting	Menos propenso, por usar árvores rasas	Mais propenso, por usar árvores profundas
Complexidade	Maior, com mais parâmetros a otimizar	Menor, com menos parâmetros a ajustar
Interpretabilidade	Menor, por ser um modelo mais complexo	Maior, por produzir árvores que podem ser visualizadas
Dados desbalanceados	Melhor desempenho, por usar pesos nas instâncias	Pior desempenho, por usar amostragem aleatória
Implementação	Pacote próprio em Python	Scikit-learn em Python

Fonte: Adaptado de Mljar (2023)

Abaixo, a Tabela 7 apresenta a *performance* dos modelos considerando o estudo de caso apresentado.

Tabela 7 - Performance dos modelos XBoost e Random Forest

Características	Modelo de Predição	
	<i>XGBoost</i>	<i>Random Forest</i>
Coeficiente de Determinação (R ²) base de teste	0,8934	0,8842
Benefício Potencial, redução de consumo real (%)	-22,55%	-24,64%
Tempo de processamento (modelo completo)	1 hora	36 horas
Importância das variáveis (%)	GVW (48%), Velocidade (25%) e DMT (22%)	GVW (55%), Velocidade (30%) e DMT (13%)

Fonte: Próprio Autor

5 CONSIDERAÇÕES FINAIS

O futuro da mineração está intrinsecamente ligado à necessidade de equilibrar a demanda crescente por recursos minerais com os imperativos da sustentabilidade ambiental, social e econômica.

Conforme citado por Vasconcelos (2022), os recursos minerais representam um diferencial competitivo para o Brasil, visando um futuro pautado em ciência, tecnologia e sustentabilidade. O país detém cerca de 9% de todas as reservas minerais do mundo. Os principais minerais explorados no Brasil incluem alumínio, cobre, estanho, ferro, manganês, nióbio, níquel e ouro, correspondendo a 98,6% do valor de toda a produção mineral brasileira comercializada. Adicionalmente, o Brasil possui consideráveis reservas de petróleo e gás.

No âmbito dos minerais estratégicos, destaca-se que o Brasil é o terceiro maior fornecedor mundial de grafita e possui a segunda maior reserva global desse material, crucial para a produção de grafeno. O país também possui grandes reservas de lítio e terras-raras, minerais estratégicos vinculados a muitos dos novos desenvolvimentos tecnológicos em curso. Além disso, merece destaque o fato de o Brasil ter recursos significativos de urânio e ser um dos países que domina o uso pacífico da energia nuclear.

A mineração emerge como um elemento vital no contexto de um futuro sustentável, desempenhando um papel crucial na transição para fontes de energia limpa, na redução de emissões de carbono e no impulsionamento da revolução tecnológica. O crescente interesse em baterias de lítio, energias renováveis e a busca por práticas industriais de baixo carbono convergem para colocar a mineração no centro das discussões sobre sustentabilidade.

Contudo, a sociedade caminha para um futuro em que não haverá espaço para organizações ou setores que não mitigam seus impactos ambientais, não trabalham na redução do consumo de recursos naturais, emitem cada vez menos GEE e não fortalecem a economia circular.

Nesse sentido, a ciência e a tecnologia desempenham um papel fundamental na promoção de estudos e soluções, como evidenciado neste trabalho, que visam a redução do impacto socioambiental na indústria, utilizando dados e informações já disponíveis no processo. Como demonstrado, de maneira ágil e com baixo custo, é

possível reduzir drasticamente o impacto que uma frota de transporte traz com emissão de GEE e consumo de Diesel.

Certamente, o estudo para aumento da eficiência energética não substitui a necessidade de mudança na matriz energética das frotas de transporte e/ou a implementação de veículos autônomos (mais eficientes operacionalmente), por exemplo. Entretanto, pode ser aplicado de maneira iterativa e paralela aos grandes movimentos de descarbonização.

Observa-se no estudo que tanto o modelo *Random Forest* quanto o modelo *XGBoost* conseguiram gerar um benefício potencial de mais de 20% no consumo real de combustível.

É importante ressaltar que esse valor traz ainda incertezas operacionais, tecnológicas ou de aderência às recomendações. Contudo, representa um primeiro passo para conscientizar a indústria, parceiros, empregados e a sociedade em geral sobre a urgência de abordar o tema.

Dessa forma, conclui-se que os modelos de *Random Forest* e *XGBoost* são aptos para a predição de consumo de combustível, sem exaurir a possibilidade de estudo com outros modelos.

Também há espaço para otimização local das variáveis para encontrar valores ainda mais precisos de recomendação, visto que o Algoritmo Genético busca uma otimização global dos parâmetros.

Importante ressaltar que a qualidade dos dados, assim como sua categorização e refinamento, são essenciais para gerar recomendações de qualidade. Existe também uma grande necessidade de engajamento e gestão de mudança dos operadores, para que sigam em campo as recomendações de velocidade conforme o trecho.

Também importante ressaltar que há uma grande oportunidade de recomendar a carga útil do caminhão e/ou construir modelos embarcados que consigam gerar velocidade em tempo real, conforme a carga do caminhão, o que traria ainda mais eficiência ao transporte, dado a importância do *GVW* (peso do caminhão vazio mais a carga útil) para o modelo.

A possibilidade de vinculação direta do método a sistemas especialistas (otimizadores de frota da mina) também é um diferencial e torna o modelo altamente produtivo e de fácil operacionalização.

5.1 Limitações do trabalho

Foi premissa do trabalho verificar a aplicação de técnicas de Análise Avançada de Dados e Inteligência Artificial (*Random Forest* e *XGBoost*) na busca da maior eficiência energética, redução do consumo de combustível fóssil e diminuição das emissões de gases de efeito estufa dos caminhões de grande porte. Há a possibilidade de utilizar outros algoritmos ou técnicas de previsão.

O trabalho também se limitou a um estudo de caso em operações de mina de céu aberto no Brasil, mais especificamente na empresa Vale S/A, e em equipamentos dos modelos Catterpillar 793D e 797F. Pode-se estender o estudo para outras mineradoras, assim como para outros modelos de equipamentos.

O estudo não é exaustivo ao analisar todas as variáveis que podem influenciar no consumo de Diesel de uma frota de transporte na mina. Da mesma forma, não foi intenção do trabalho otimizar outros parâmetros que não a velocidade do trecho no momento.

Por questões de capacidade de armazenamento e processamento de dados, o estudo limitou-se a um período de aproximadamente um mês de dados (01/Ago/2023 a 09/Set/2023).

Para compreender toda a variabilidade sazonal dos dados, pode-se estender o período de estudo para seis meses ou um ano, para entender melhor o comportamento dos dados durante um período maior de observação.

5.2 Trabalhos Futuros

Explorar o impacto do *GVW* (peso do veículo vazio mais a carga útil) na eficiência energética, utilizando métodos adicionais de análise de dados e fornecendo informações aos operadores.

Investigar como as tendências tecnológicas emergentes, como a IA Generativa, Frota de Veículos Autônomos e/ou *Video Analytics*, podem ser integradas para aprimorar a interatividade com o operador e resultar em maior eficiência energética.

Realizar um estudo longitudinal para compreender a evolução ao longo do tempo das variáveis analisadas neste trabalho (velocidade por trecho, resistência total ao rolamento, GVW, DMT), assim como integrar os resultados de eficiência energética encontrados no modelo ao planejamento futuro de outras minas a céu aberto.

Ampliar a amostra para incluir outras categorias e/ou grupos, como, por exemplo, questões ambientais (temperatura, pluviometria, etc.), condição de manutenção e saúde do ativo (motor, transmissão, comando), entre outras variáveis de confiabilidade do ativo.

Explorar a aplicação de outras técnicas ou modelos, como Redes Neurais, ou outras técnicas de regressão e otimização local/global, como alternativa aos métodos de previsão utilizados neste estudo, comparando os resultados para validar as conclusões.

REFERÊNCIAS

- Agência Nacional de Mineração (ANM). Relatório do Informe Mineral – 2022 4º trimestre. Disponível em: <<https://www.gov.br/anm/pt-br/centrais-de-conteudo/publicacoes/serie-estatisticas-e-economia-mineral/informe-mineral/publicacoes-nacionais/informe-mineral-2022-4o-trimestre>>. Acesso em: 18 dez. 2023. Ano de publicação não disponível.
- Barbieri, J. C. et al. (2010). Inovação e Sustentabilidade: Novos Modelos e Proposições. RAE-Revista de Administração de Empresas, São Paulo, v. 50, n. 2, p. 146-154, abr.-jun. Disponível em: <<http://dx.doi.org/10.1590/S0034-75902010000200002>>.
- BRASIL. Ministério do Meio Ambiente. Relatório: Plano Setorial de Mitigação e Adaptação à Mudança do Clima para a Consolidação de uma Economia de Baixa Emissão de Carbono na Indústria de Transformação. 2023. Disponível em: <<https://antigo.mma.gov.br/clima/politica-nacional-sobre-mudanca-do-clima/planos-setoriais-de-mitigacao-e-adaptacao.html>>. Acesso em: 26 dez. 2023.
- Broniatowsky, D. Professor Associado, Departamento de Gestão de Engenharia e Engenharia de Sistemas, Universidade George Washington, 2021.
- BULLIVANT, D. A. Current surface mining techniques. 1987. Disponível em: <<https://www.osti.gov/etdeweb/biblio/5664362>>.
- CARMIGNANO, Ottavio Raul Domenico Riberti. Inovação no Setor de Mineração de Ferro em Minas Gerais com Foco na Destinação de Rejeitos. UFMG, 2021. p. 23-24.
- CAT. Especificação técnica equipamentos 793D e 797F Caterpillar. Disponível em: <https://www.cat.com/pt_BR/products/new/equipment/off-highway-trucks/mining-trucks/18396313.html>. Acesso em: 06 jan. 2024.
- CESAR, Júnior da Silva Muniz. TCC: Dimensionamento e Operação de Caminhões Fora de Estrada em Mina de Céu Aberto. Faculdade Pitágoras, 2019.
- CHEN, Tianqi; GUESTRIN, Carlos. XGBoost: A Scalable Tree Boosting System. In: SIGKDD, 2016.
- CLIMA, Observatório. Relatório: Acordo de Paris: Um Guia para os Perplexos. 2023. Disponível em: <<https://oc.eco.br/acordo-de-paris-um-guia-para-os-perplexos-2/>>. Acesso em: 01 set. 2023.
- CLIMA, Observatório. Relatório: Mineração responde por 5% das emissões do Brasil. 2022. Disponível em: <<https://oc.eco.br/mineracao-responde-por-5-das-emissoes-do-brasil/>>. Acesso em: 18 ago. 2022.
- CONSELHO FEDERAL DE ENGENHARIA E AGRONOMIA (CONFEA). Mineração: Importância dos Minerais para a Sociedade Moderna. Acesso em: 11 nov. 2022. Ano

de publicação não disponível. Disponível em:
<https://www.confea.org.br/midias/web_cartilha_mineracao_170x240.pdf>.

COP26. Brasil se compromete a reduzir emissões de carbono em 50%, até 2030. Disponível em: <<https://www.gov.br/casacivil/pt-br/assuntos/noticias/2021/novembro/brasil-se-compromete-a-reduzir-emissoes-de-carbono-em-50-ate-2030>>. Acesso em: 07 jan. 2024.

DETERMINAÇÃO. Coeficiente de determinação. In: WIKIPÉDIA: a enciclopédia livre. Disponível em: <https://pt.wikipedia.org/wiki/Coeficiente_de_determinação>. Acesso em: 09 jan. 2024.

DIJKSTRA, E.W. A Short Introduction to the Art of programming. Technological University Endhoven. 1971

ELKINGTON, John. Cannibals with Forks: The Triple Bottom Line of 21st Century Business. 1997.

ELKINGTON, John. Canibais com garfo e faca. São Paulo: Makron Books, 2001.

Gall, R. Machine Learning explainability vs interpretability: two concepts that could restore trust in AI, KDnuggets, 2018

GARTNER. Gartner Says Advanced Analytics Is a Top Business Priority. 2014. Disponível em: <<https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>>. Acesso em: 07 jan. 2024.

GENÉTICO. Algoritmo. In: WIKIPÉDIA: a enciclopédia livre. Disponível em: <https://pt.wikipedia.org/wiki/Algoritmo_genético>. Acesso em: 07 jan. 2024.

GLAVIC, P.; LUKMAN, R. Review of sustainability terms and their definitions. Journal of Cleaner Production, v.15, p.1875-1885, 2007.

IBRAM, Instituto Brasileiro de Mineração : Relatório Mineração em Números 2022. Disponível em: <<https://ibram.org.br/mineracao-em-numeros/> acessado em 09/dez/23, 2022>.

IPEA. Item 7. Energia Acessível e Limpa; 13. Ação Contra a Mudança Global do Clima. 2023. Disponível em: <<https://www.ipea.gov.br/ods/index.html>>. Acesso em: 29 dez. 2023.

JAVA. What Is Java. Disponível em: <<https://www.java.com/pt-BR/>>. Acesso em: 09 jan. 2024.

MENEZES, Messias Gilmar. Curso de Introdução à Mineração, Geologia Geral e Pesquisa Geológica. Ouro Preto, MG: DEGEO-EM-UFOP, NUPEC-Fundação Gorceix, 2008.

METSO. Caminhão fora de estrada: conheça o seu custo operacional. 2019. Disponível em: <<https://www.metso.com/pt/insights/blog/mineracao-e-metais/caminhoes-fora-de-estrada-sao-os-peso-pesados-da-mineracao-mas-tem-um-custo-operacional-igualmente-gigante/>>. Acesso em: 29 dez. 2023.

"Mineração a Céu Aberto". In: WIKIPÉDIA: a enciclopédia livre. 2023. Disponível em: <https://pt.wikipedia.org/wiki/Minera%C3%A7%C3%A3o_a_c%C3%A9u_aberto>. Acesso em: 29 dez. 2023.

MLJAR, Título: Random Forest vs Xgboost, 2023. Disponível em: <<https://mljar.com/machine-learning/random-forest-vs-xgboost/>>. Acesso em 10 jan 2024,.

ZIVIANI, Nivio. Projeto De Algoritmos Com Implementações Em Pascal E C. 3ª ed. rev. e ampl. Capítulo 1.1, 2010.

OECD. The Oslo Manual: The Measurement of Scientific and Technical Activities. Paris: OECD; Eurostat, 1997.

OLIVEIRA, João Leandro Cássio de. Valoração de Tecnologias no Cenário de Transferência de Tecnologia entre Universidade e Empresa no Brasil: Uma Metodologia Proposta. UFMG, 2020. p. 32.

ONU. Agenda 2030 e os Objetivos de Desenvolvimento Sustentável. 2023. Disponível em: <<https://brasil.un.org/pt-br/sdgs>>. Acesso em: 29 dez. 2023.

PYTHON. About. Acesso em: 09 jan. 2024. Disponível em: <<https://www.python.org/about/>>.

RADOS, G. J. V.; DIAS, P. M. Introdução à gestão da inovação. 2015. Disponível em: <http://tvled.egc.ufsc.br/biblioteca/biblioteca/enova_abimaq/Livros%20M%C3%B3dulo%20B%C3%A1sico/introducao_a_gestao_da_inovacao.pdf>. Acesso em: 11 jun. 2018.

RIO+20. Relatório Rio+20. 2012. Acesso em: 15 dez. 2021. Disponível em: <http://www.rio20.gov.br/sobre_a_rio_mais_20/desenvolvimento-sustentavel.html>.

R-PROJECT. Disponível em:< <https://www.r-project.org/>>. Acesso em: 09 jan. 2024.

SARTORI, Simone; LATRÔNICO, Fernanda; CAMPOS, Lucila M. S. Sustentabilidade e desenvolvimento sustentável: uma taxonomia no campo da literatura. Universidade Federal de Santa Catarina - UFSC, 2014.

SKLEARN. Cross Validation Score. Documentação: `sklearn.model_selection.cross_val_score` — scikit-learn 1.3.2 documentation. Acesso em: 09 jan. 2024.

SOLUTIONS, Management S. Explainable artificial intelligence - Desafios na interpretabilidade de modelos. Disponível em:

<<https://www.managementsolutions.com/pt-br/microsites/whitepapers/explainable-artificial-intelligence>>. Acesso em: 03 jan. 2024.

SOOFASTAEI, Tese: Development of an Advanced Data Analytics Model to Improve the Energy Efficiency of Haul Trucks in Surface Mines. 2016. Disponível em: <https://www.researchgate.net/publication/309176186_Development_of_an_Advanced_Data_Analytics_Model_to_Improve_the_Energy_Efficiency_of_Haul_Trucks_in_Surface_Mines>.

SOUSA, Wilson Trigueiro, Geologia de Mina e Operações de Lavra, Departamento de Engenharia de Minas, UFOP, 2008.

SQL. Tutorial. Disponível em: <<https://www.sqltutorial.org/>>. Acesso em: 09 jan. 2024.

TABLEAU. Aprenda tudo sobre o conceito de Inteligência Artificial (IA). 2023. Disponível em: <<https://www.tableau.com/pt-br/learn/articles/ai#:~:text=A%20Intelig%C3%Aancia%20Artificial%20%C3%A9%20a,que%20recebem%20de%20seus%20usu%C3%A1rios>>. Acesso em: 26 dez. 2023.

TECH. O que é e como funciona o algoritmo RandomForest. Disponível em: <<https://didatica.tech/o-que-e-e-como-funciona-o-algoritmo-randomforest>>. Acesso em: 03 jan. 2024.

VALE. Relato Integrado 2022. Disponível em: <https://vale.com/documents/44618/3893145/VALE_RelatoIntegrado2022-BR-final.pdf/34101961-a68e-8259-f169-d465e182074d?version=4.0&t=1696882895347&download=true>. Acesso em: 08 mai. 2023.

VASCONCELOS, Ado Jorio de et al. Ciência para prosperidade: sustentável e socialmente justa. Brasília: Empresa Brasileira de Pesquisa e Inovação Industrial - EMBRAPPII, 2022. Disponível em: <<https://www.abc.org.br/wp-content/uploads/2022/10/CI%C3%80NCIA-PARA-PROSPERIDADE-compactado.pdf>>. Acesso em: 18 jan. 2024.

YIN, Robert K. Estudo de caso: planejamento e métodos. 2. ed. Tradução Daniel Grassi. Porto Alegre: Bookman, 2001.