



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA



Fábio Ribeiro von Glehn

Integração de telemática e *machine learning* para
determinação das condições de transporte de cargas

Goiânia
2024



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor e o orientador firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Fábio Ribeiro Von Glehn

3. Título do trabalho

Integração de telemática e *machine learning* para determinação das condições de transporte de cargas

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

- a) consulta ao(à) autor(a) e ao(à) orientador(a);
 - b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
- O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Marlipe Garcia Fagundes Neto, Professor do Magistério Superior**, em 26/02/2024, às 08:57, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fábio Ribeiro Von Glehn, Discente**, em 26/02/2024, às 09:00, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4405765** e o código CRC **96B57375**.

Referência: Processo nº 23070.068307/2023-84

SEI nº 4405765

Fábio Ribeiro von Glehn

Integração de telemática e *machine learning* para determinação
das condições de transporte de cargas

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Engenharia Mecânica da Universidade Federal de Goiás, como requisito parcial para obtenção do Título de Mestre em Engenharia Mecânica. Área de Concentração: Ciências Mecânicas

Orientador: Prof. Dr. João Paulo da Silva Fonseca

Coorientador: Prof. Dr. Marlipe Garcia Fagundes Neto

Universidade Federal de Goiás
Escola de Engenharia Elétrica, Mecânica e de Computação

Goiânia
2024

Glehn, Fábio Ribeiro von

Integração de telemática e machine learning para determinação das condições de transporte de cargas [manuscrito] / Fábio Ribeiro von Glehn. - 2024.

110 f.: il.

Orientador: Prof. João Paulo da Silva Fonseca; co-orientador Marlipe Garcia Fagundes Neto.

Dissertação (Mestrado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa de Pós-graduação em Engenharia Mecânica, Goiânia, 2024.

Bibliografia. Anexos. Apêndice.

1. telemática. 2. caminhões pesados. 3. economia de combustível. 4. análise de dados. 5. aprendizado de máquina. I. Fonseca, João Paulo da Silva, orient. II. Título

CDU 621



UNIVERSIDADE FEDERAL DE GOIÁS

ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº **01** da sessão de Defesa de Dissertação de **Fábio Ribeiro Von Glehn**, que confere o título de Mestre em **Engenharia Mecânica**, na área de concentração em **Ciências Mecânicas**.

Aos **dois dias do mês de fevereiro de dois mil e vinte quatro**, a partir das 08h30min, realizou-se a sessão pública de Defesa de Dissertação intitulada “**Integração de telemática e machine learning para determinação das condições de transporte de cargas**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **João Paulo da Silva Fonseca (EMC/UFG)** com a participação dos demais membros da Banca Examinadora: Professor Doutor **Marlize Garcia Fagundes Neto (EMC/UFG)** - coorientador, Professor Doutor **Anderson da Silva Soares (INF/UFG)** membro titular externo e Prof. Dr. **Max Mauro Dias Santos - (DAELE/UTFPR-PG)**, membro titular externo, **cujas participações ocorreram através de videoconferência pelo link: <https://meet.google.com/mka-bpmw-pqp>**. Durante a arguição os membros da banca **não fizeram** sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor **João Paulo da Silva Fonseca**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos **dois dias do mês de fevereiro de dois mil e vinte quatro**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Joao Paulo Da Silva Fonseca, Professor do Magistério Superior**, em 02/02/2024, às 10:56, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marlize Garcia Fagundes Neto, Professor do Magistério Superior**, em 02/02/2024, às 10:57, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Anderson Da Silva Soares, Professor do Magistério Superior**, em 02/02/2024, às 16:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Demostenes Ferreira Filho, Coordenador de Pós-Graduação**, em 27/02/2024, às 08:22, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4337763** e o código CRC **87799DF0**.

Referência: Processo nº 23070.068307/2023-84

SEI nº 4337763

AGRADECIMENTOS

Os meus agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Universidade Federal de Goiás (UFG), ao Programa de Pós-Graduação em Engenharia Mecânica (PPGMEC) e, em especial, ao Grupo Masut, ao srs. Anderson, Clóvis, Bruno Cortes e os professores João Paulo da Silva Fonseca e Marlipe Garcia Fagundes Neto.

RESUMO

Esse estudo explora a aplicação da telemática na gestão de frotas para controlar o consumo de combustível e as emissões poluentes. Iniciando com uma análise do impacto econômico e ambiental do uso de diesel, o trabalho identifica o desafio de reduzir tais efeitos e destaca a telemática e o *machine learning* como possíveis soluções. Revisões bibliográficas validam essa abordagem, mostrando que a combinação de telemática e *machine learning* otimizam o consumo de combustível, identifica padrões de direção e melhora a manutenção. A pesquisa aplicada utiliza dados veiculares disponíveis nas redes de comunicação dos caminhões e em dispositivos telemáticos conectados à rede SAEJ1939. Esses dados formaram 3 bases distintas: a primeira coletada de uma frota de 10 caminhões de uma transportadora de combustível e outras duas de equipamentos Geotab[®] instalados em dois caminhões de transportadoras distintas: uma de areia e outra de produtos alimentícios. O estudo desenvolveu um modelo com oito diferentes algoritmos de *machine learning* para determinar se um caminhão está carregado ou vazio. Após validar os dados, a análise final após 10 execuções do código revelou, com a primeira base de dados, uma acurácia superior a 85% para percursos acima de 1 km ou mais de 120 segundos de movimento. Com os dados Geotab[®] foi construído um *dashboard* que permitiu acompanhar o comportamento diários dos dois caminhões com a elaboração de mapas, gráficos, identificação de locais de interesse e seleção de trechos frequentes e indicadores de desempenho que incluem a carga útil transportada.

Palavras-chaves: telemática; caminhões pesados; economia de combustível; análise de dados, aprendizado de máquina

ABSTRACT

This study explores the application of telematics in fleet management to control fuel consumption and pollutant emissions. Beginning with analyzing the economic and environmental impact of diesel usage, the paper identifies the challenge of reducing these effects and highlights telematics and machine learning as possible solutions. Literature reviews validate this approach, demonstrating that combining telematics and machine learning optimizes fuel consumption, identifies driving patterns, and enhances maintenance. The applied research uses vehicle data available in the trucks' communication networks and telematics devices connected to the SAEJ1939 network. This data formed three separate databases: the first was collected from a fleet of 10 trucks belonging to a fuel transporter and the other two from Geotab[®] equipment installed in two trucks belonging to different transporters: one of sand and the other of food products. The study developed a model with eight different machine learning algorithms to determine whether a truck is loaded or empty. After validating the data, the final analysis after 10 runs of the code revealed, with the first database, an accuracy of over 85% for routes over 1 km or more than 120 seconds of movement. With the Geotab[®] data, a dashboard was built that made it possible to monitor the daily behavior of the two trucks with the creation of maps, graphs, identification of places of interest and selection of frequent stretches and performance indicators that include the payload transported.

Keywords: *telematics; heavy-duty vehicles; fuel saving; big data analytics; machine learning.*

LISTA DE FIGURAS

Figura 1.1: Participação das fontes energéticas nos Transportes e consumo final em 2022.....	22
Figura 1.2: Emissões totais (2022) em Mt CO ₂ -eq por setor da economia.....	23
Figura 1.3: Eficiência energética média de veículos novos vendidos (com carga) (km/L).....	24
Figura 1.4: Diagrama de sankey do transporte de carga no Brasil.....	25
Figura 1.5: Curvas de Torque e Potências dos Motores D13C e D13K.....	26
Figura 1.6: Diagrama de Ishikawa dos fatores que afetam o consumo de diesel.....	27
Figura 2.1: Como funciona a telemática.....	34
Figura 4.1: Estimativas do volume de combustível consumido.....	54
Figura 4.2: Erro absoluto (%) de cada abastecimento na estimativa do volume de combustível por diferentes métodos de cálculo.....	54
Figura 4.3: Comparativo entre as estimativas de volume abastecido e o relatório de abastecimento.....	55
Figura 4.4: Economia de combustível (km/L) em função da carga e do relevo.....	58
Figura 4.5: Metodologia aplicada.....	60
Figura 4.6: Acurácia (%) na classificação da condição de carga em distâncias curtas e longas com diferentes técnicas de <i>machine learning</i>	63
Figura 4.7: Método dos trechos integrais e de incorporação longa-curta.....	64
Figura 4.8: Acurácia (%) na classificação da condição de carga em função do método e do tamanho do percurso.....	64
Figura 4.9: Acurácia (%) na classificação da condição de carga em função do método e do tempo de aquisição do dado.....	66
Figura 4.10: <i>Dashboard</i> desenvolvido a partir das bases de dados 2 e 3.....	69
Figura 4.11: Quadro explicativo.....	69
Figura 4.12: Mapa de um dia de trabalho de um dos caminhões analisados.....	71
Figura 4.13: Economia diária em diferentes condições de carga.....	72
Figura 4.14: Economia de combustível acumulada em um período.....	73
Figura 4.15: Perfil de altitude de dois trechos frequentes.....	76

LISTA DE TABELAS

Tabela 1.1: Consumo energético (Mtep) por modo e fonte nos anos de 2021 e 2022	24
Tabela 1.2: Nomenclatura do controle de velocidade adaptativo por marca.....	30
Tabela 3.1. Palavras de busca para formação dos conjuntos.....	42
Tabela 3.2: Relatório após a unificação das bases.....	43
Tabela 3.3. Relatório de palavras chaves mais frequentes	44
Tabela 3.4: Artigos classificados no <i>cluster</i> consumo de combustível ou emissão de poluentes	46
Tabela 3.5: Artigos classificados no <i>cluster</i> estilo ou assistente de direção	47
Tabela 3.6: Artigos classificados no <i>cluster</i> utilização de dados para aprimoramento de planos de manutenção	47
Tabela 4.1: Bases de dados dos grupos de estudo	49
Tabela 4.2: Consumo estratificado em função do relevo	57
Tabela 4.3: Acurácia (%) na classificação da condição de carga em distâncias curtas.....	62
Tabela 4.4: Acurácia (%) na classificação da condição de carga em distâncias longas.....	62
Tabela 4.5: Quadro de distâncias percorridas após um período determinado	65
Tabela 4.6: Locais de interesse identificados	73
Tabela 4.7: <i>Status</i> de carga dos caminhões	74
Tabela 4.8: Acurácia (%) da condição de carga com dados dos caminhões 1 e 2 das bases 2 e 3 com diferentes métodos de classificação.....	74
Tabela 4.9: Tabela de <i>features</i> utilizadas para classificação com <i>machine learning</i>	75
Tabela 5.1: Prestação de contas.....	81

LISTA DE SÍMBOLOS

Siglas

ABS	<i>Anti-Lock Braking System</i>
AEA	Associação Brasileira de Engenharia Automotiva
ANFAVEA	Associação Nacional dos Fabricantes de Veículos Automotores
API	<i>Application Programming Interface</i>
ARLA-32	Agente Redutor Líquido Automotivo com nível de concentração da solução de ureia (32,5%) em água desmineralizada
BEN	Balanço Energético Nacional
CAN	<i>Controller Area Network</i>
cl	Volume em centilitro (cL)
CO ₂	Dióxido de carbono
CO ₂ -eq	Dióxido de carbono equivalente
CONAMA	Conselho Nacional do Meio Ambiente
CMT	Capacidade Máxima de Tração
DOI	<i>Digital Object Identifier</i>
DPF	Filtro de partículas diesel (<i>Diesel particulate filter</i>)
EGR	Recirculação dos gases de escape (<i>Exhaust gas recirculation</i>)
ELR	<i>European Low-Speed Cycle</i>
EPE	Empresa de Pesquisa Energética
ESC	<i>European Steady Cycle</i>
ETC	<i>European Transient Cycle</i>
GAV	Gasolina de Aviação
GEE	Gases de Efeito Estufa
GNC	Gás Natural Comprimido
GNL	Gás Natural Liquefeito

GNV	Gás Natural Veicular
GPS	<i>Global Positioning System</i>
HC	Hidrocarbonetos (<i>HydroCarbons</i>)
HEV	Veículo Híbrido (<i>Hybrid Eletric Vehicle</i>)
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais
IBGE	Instituto Brasileiro de Geografia e Estatística
IEA	Agência Internacional de Energia
INMETRO	Instituto Nacional de Metrologia, Qualidade e Tecnologia
IPT	Instituto Brasileiro de Petróleo e Gás
ISC	<i>In-service Conformity</i>
ISO	<i>International Organization for Standardization</i>
kmph	Velocidade em quilometro por hora (km/h)
kmpl	Consumo em quilometro por litro (km/L)
KNN	<i>k-nearest neighbors algorithm</i>
kpa	Pressão em quilopascal (kPa)
L	Litro
lge	Litro de gasolina equivalente, sendo a gasolina C, com 27% de etanol anidro, como referência (1.000lge = 0,706 tep = 29,57GJ)
L/tku	Unidade de Intensidade Energética medida em Litros por tonelada-quilometro útil
ISM	International Conference on Industry 4.0 and Smart Manufacturing
LGN	Líquido de Gás Natural
m	Distância em metros (m)
MME	Ministério de Minas e Energia
ML	<i>Machine learning</i>
MP	Material particulado
Mt	Milhões de toneladas

Mtep	Milhões de tonelada equivalente de petróleo
NMHC	Hidrocarbonetos Não Metano (<i>Non-Methane HydroCarbons</i>)
NOx	Óxidos de Nitrogênio
NP	Número de partículas
OBD	<i>On-Board Diagnostic.</i>
OBD II	<i>On-Board Diagnostic</i> fase II
OCE	<i>Off-Cycle Emissions</i>
perc	Percentual de algo
PIB	Produto Interno Bruto
PBT	Peso Bruto Total
PBTC	Peso Bruto Total Combinado
PCA	<i>Principal Component Analysis</i>
PDE	Plano Decenal de Expansão de Energia
PGN	<i>Parameter Group Numbers</i>
PPS	<i>Pulse Per Second</i>
PROCONVE	Programa de Controle de Emissões Veiculares
pkm	Passageiro-quilômetro
ppm	Partícula por milhão
rpm	Velocidade angular em rotação por minuto (rpm)
QAV	Querosene de aviação
SCR	Catalisadores seletivos de redução (<i>Selective catalytic reduction</i>)
SDK	<i>Software Development Kit</i>
seg	Tempo em segundos (s)
SUV	<i>Sport Utility Vehicle</i>
SPN	<i>Suspect Parameter Numbers</i>
SVM	<i>Support Vector Machine</i>

t	Toneladas
tempc	Temperatura em graus Celcius
tep	Tonelada equivalente de petróleo. (1 tep = 41,868 GJ = 1415,9 lge)
tu	Tonelada Útil Tracionada. É o total de carga movimentada, no transporte remunerado
THC	Hydrocarbonetos Totais (<i>Total HydroCarbons</i>)
tku	Tonelada-quilômetro útil - representa o somatório dos produtos das toneladas úteis (tu) tracionadas pelas distâncias de transporte. É a quantidade de toneladas úteis transportadas multiplicadas pela quilometragem percorrida pelas mesmas.
VECTO	<i>Vehicle Energy Consumption Calculation Tool</i>
WHSC	<i>World Harmonized Stationary Cycle</i>
WHTC	<i>World Harmonized Transient Cycle</i>
WNTE	<i>Worldwide harmonized Not to Exceed</i>
WoS	<i>Web of Science</i>

SUMÁRIO

1.	INTRODUÇÃO	21
1.1	Contextualização do Problema	21
1.2	Evolução dos motores de combustão	25
1.3	Fatores que afetam o consumo de diesel.....	26
1.4	Objetivos.....	27
1.5	Justificativa e Importância da Pesquisa	28
1.6	Linhas de pesquisas para redução de emissões	28
1.7	Definições importantes	30
2.	FERRAMENTAS PARA ANÁLISE	33
2.1	Telemática.....	33
2.1.1	Protocolo de comunicação veicular SAE J1939.....	34
2.2	Machine Learning	36
2.2.1	Categorias de Machine Learning	37
2.2.2	Exemplos de aplicação de <i>machine learning</i>	38
2.2.3	Algoritmos de <i>machine learning</i> clássicos.....	38
3.	REVISÃO BIBLIOGRÁFICA	41
3.1	Metodologia da revisão bibliográfica	41
3.1.1	A formação de cada conjunto de pesquisa	42
3.1.2	Análise bibliométrica, unificação dos dados e remoção de duplicatas.....	43
3.1.3	Identificação e remoção de artigos não elegíveis através de palavras-chave	44
3.1.4	Leitura e classificação da lista resultante	44
3.2	Resultados da Revisão Bibliográfica	45
4.	APLICAÇÃO PRÁTICA.....	49
4.1	Validação da base de dados 1	51
4.1.1	Descrição e análise do banco de dados 1	51
4.1.2	Análise e manipulação de dados da base 1	53
4.1.3	Inserção do relevo entre as características relevantes	56
4.1.4	Discussão sobre fontes de erro	59

4.2	Preparação para os algoritmos de <i>Machine Learning</i>	59
4.3	Resumo dos experimentos realizados com a base 1	60
4.3.1	Método dos trechos integrais divididos em curtos e longos.....	60
4.3.2	Método de incorporação longa-curta para enriquecimento estatístico	63
4.3.3	Método de segmentação temporal.....	65
4.4	Resumo dos experimentos realizados com as bases 2 e 3	66
4.4.1	Descrição e análise do banco de dados	66
4.4.2	Metodologia adotada com os dados das bases 2 e 3.....	67
4.4.3	Resultados com os dados das bases 2 e 3.....	69
4.4.4	Observações a respeito do trabalho com os dados das bases 2 e 3.....	73
4.4.5	Classificação com técnicas de <i>machine learning</i>	74
4.4.6	Trechos frequentes	76
5.	DISCUSSÃO DOS RESULTADOS	79
6.	CONCLUSÃO	83
7.	TRABALHOS FUTUROS	87
8.	REFERÊNCIAS BIBLIOGRÁFICAS	89
9.	APÊNDICE A	95
10.	APÊNDICE B	101
11.	APÊNDICE C	109

1. INTRODUÇÃO

A engenharia mecânica é um pilar da sociedade contemporânea, ao impulsionar o desenvolvimento de máquinas e sistemas que têm um impacto direto em nossa qualidade de vida. Dentro deste vasto conjunto de máquinas, os veículos automotores assumem um papel importante no transporte de pessoas e mercadorias. Nesse contexto, compreender os desafios enfrentados na operação e manutenção desses veículos torna-se essencial, especialmente diante do crescente foco em eficiência energética e confiabilidade.

Mnyakin (2023) destaca que a integração de tecnologias emergentes como a Inteligência Artificial (IA), a Internet das Coisas (IoT) e a Computação em Nuvem tem sido usadas gerenciar tráfego, assistir o motorista, gerenciar frota em tempo real, realizar diagnósticos remotos e otimizar processos de manutenção preditiva, entre outros.

Uma aplicação dessas tecnologias visa estudar o consumo de energia e eficiência, pois os veículos automotores são grandes consumidores, e otimizar seu desempenho não apenas reduz custos operacionais, como também minimiza o impacto ambiental, como salienta o trabalho de Walnum e Morten (2015).

Em outro exemplo, Zhang *et al.* (2022) trabalham com dados para manutenção, no entanto, destacam que diagnosticar falhas pode ser complexo, envolvendo correlações difíceis de serem avaliadas.

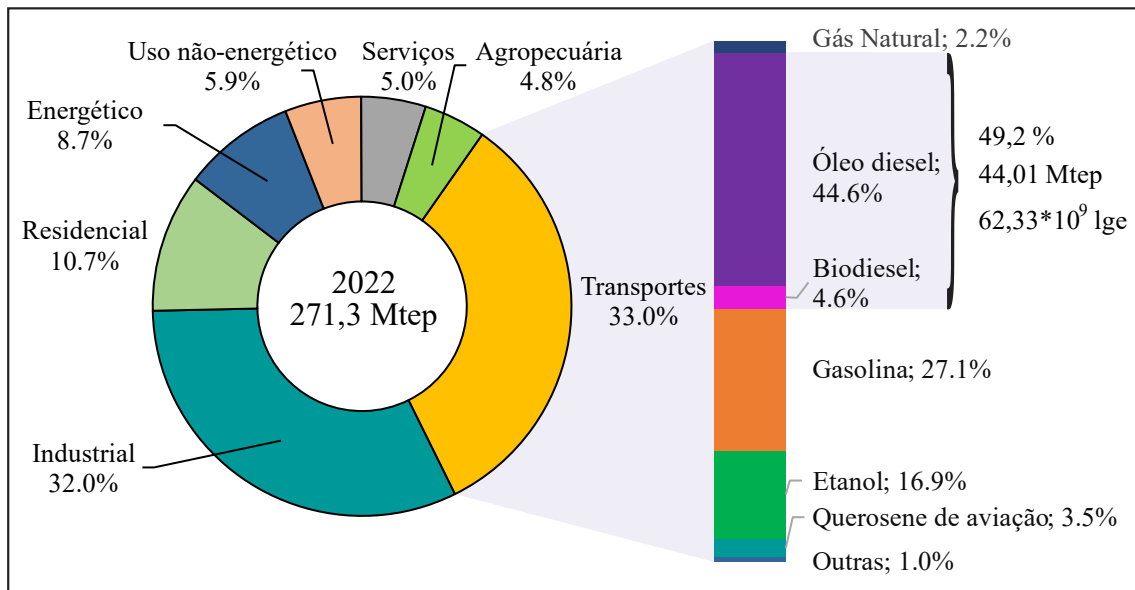
Ao adentrar no universo da pesquisa sobre o consumo de diesel em caminhões rodoviários, é preciso estabelecer uma fundamentação que contextualize a importância econômica e ambiental desse estudo. Inicialmente, exploraremos a participação das fontes energéticas nos transportes, delineando o panorama atual do consumo energético no país. Em seguida, ressaltaremos a necessidade de compreender o desenvolvimento tecnológico dos caminhões. Posteriormente, examinaremos fatores que influenciam o consumo de diesel, preparando para a apresentação dos objetivos e metodologias deste projeto. Tudo isso visa informar e proporcionar uma visão abrangente do cenário que envolve a investigação proposta.

1.1 Contextualização do Problema

O setor de transporte no Brasil consome 33% de toda a energia do país, conforme indicado no Balanço Energético Nacional (BEN 2023). Esse consumo é majoritariamente

impulsionado pela queima de combustíveis, destacando a relevância do estudo sobre o consumo de diesel em caminhões rodoviários. Esse dado ressalta a necessidade de compreender e otimizar o consumo energético nesse setor. A estratificação das fontes de energia, conforme exposto na Fig. 1.1, revela como esse consumo se dá.

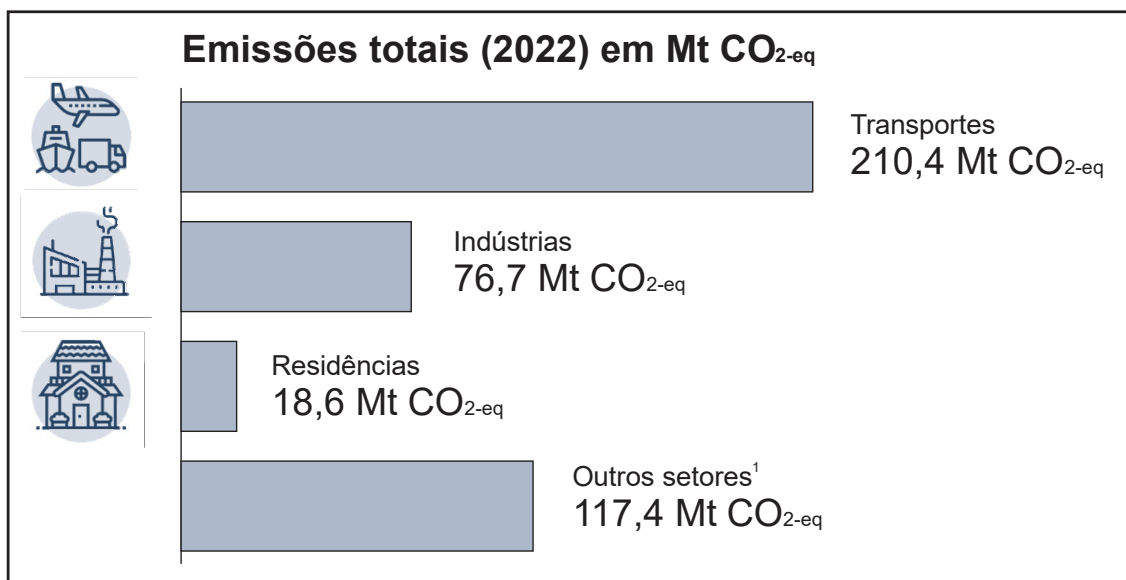
Figura 1.1: Participação das fontes energéticas nos Transportes e consumo final em 2022



Fonte: Balanço Energético Nacional – BEN 2023, elaboração: autor

O Balanço Energético Nacional (BEN 2023) afirma ainda que “Em 2022, o total de emissões antrópicas associadas a matriz energética brasileira atingiu 423 milhões de toneladas de dióxido de carbono equivalente ($Mt CO_{2-eq}$), sendo a maior parte (210,4 $Mt CO_{2-eq}$) gerada no setor de transportes”. A Fig. 1.2 ilustra o impacto desse modelo nas emissões totais em 2022. O setor de transportes responde por quase 50% do total de emissões entre todos os setores da economia do Brasil.

Figura 1.2: Emissões totais (2022) em Mt CO_{2-eq} por setor da economia



Fonte: Balanço Energético Nacional – BEN 2023

Para uma compreensão mais aprofundada sobre como essa energia é consumida no setor de transportes, é crucial analisar os relatórios da Matriz Energética Nacional (2023), Matriz Energética Nacional (2022) e do Atlas da Eficiência Energética Brasil (2022). Os dados compilados, apresentados na Tab. 1.1, destacam a expressiva participação do transporte rodoviário de cargas nesse processo.

Finalmente, os Estudos do Plano Decenal de Expansão de Energia 2032 - PDE 32 (2023) conclui que o diesel permanecerá como o combustível mais importante para o setor de transportes no Brasil e que a eletrificação na frota de veículos leves e pesados deverá avançar em nichos de aplicação, tais como frete de última milha em grandes cidades e junto aos consumidores de maior renda e que, apesar do avanço, a demanda por eletricidade no setor de transportes brasileiro continuará pouco representativa em 2032. Em resumo, o relatório afirma que o transporte rodoviário de cargas, ferroviário de cargas e rodoviário coletivo de passageiros continuará quase em 100% tendo o diesel como fonte energética.

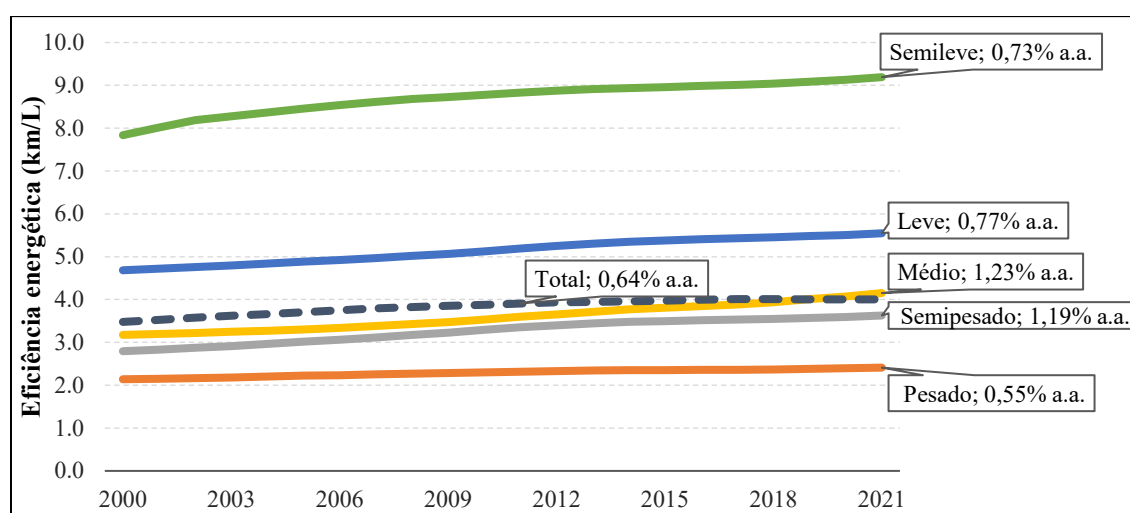
Tabela 1.1: Consumo energético (Mtep) por modo e fonte nos anos de 2021 e 2022

Fonte de energia	Modo	2022	2021	Fonte 2022 e 2021
Óleo Diesel + Biodiesel	Carga rodoviário	35,56	34,53	2022 - Estimado 2021 – Atlas (2022)
	Coletivo de passageiros	4,96	4,82	
	Individual de passageiros	2,08	2,02	
	Carga ferroviário	1,05	1,04	2022 – Matriz Energética Nacional (2023)
	Carga hidroviário	0,36	0,34	
Gasolina	Individual de passageiros	24,19	22,14	2021 – Matriz Energética Nacional (2022)
Etanol	Individual de passageiros	15,16	14,84	
Gás natural	Individual de passageiros	1,99	1,91	
Querosene	Aviação	3,16	2,55	
Eletricidade	Metroferroviário	0,17	0,17	
Óleo combustível	Carga hidroviário	0,74	0,80	
Total Transporte		89,42	85,14	

Fontes indicadas na própria tabela. Estimativa e organização: Autor

Atualmente os veículos pesados do Brasil não possuem metas de eficiência de combustível e as metas para 2032 possivelmente serão desenvolvidas até 2027, conforme afirma o Benchmarking Internacional (2021). Isso significa que o Brasil não possui um índice de CO_{2eq} para caminhões, e a evolução da eficiência energética é ilustrada pela Fig. 1.3 com dados do Atlas da Eficiência Energética Brasil (2022).

Figura 1.3: Eficiência energética média de veículos novos vendidos (com carga) (km/L)

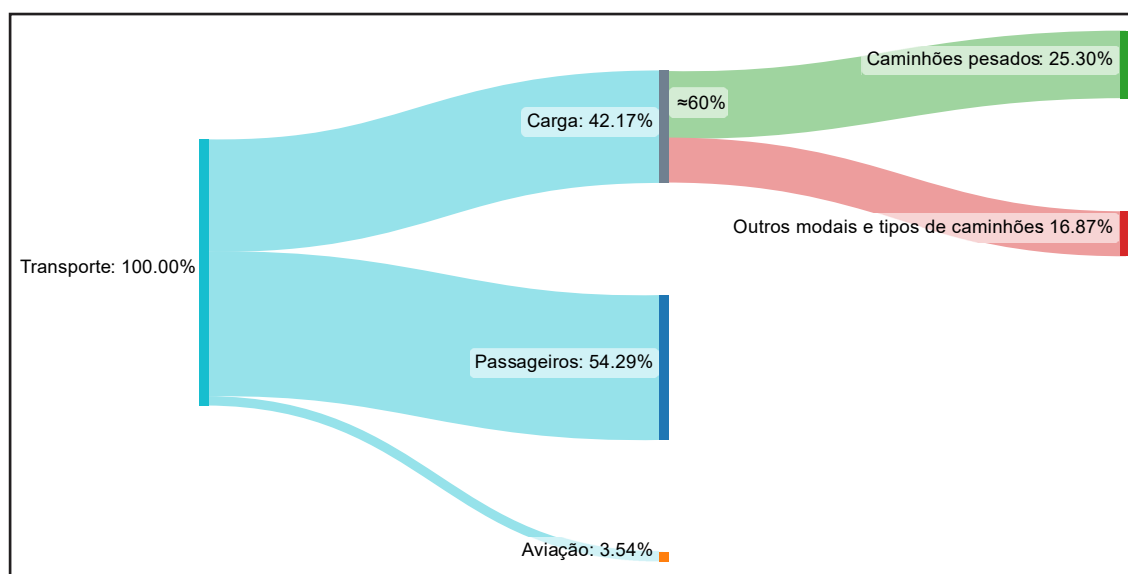


Fonte: Atlas da Eficiência Energética Brasil (2022). Elaboração: Autor

Observa-se pela Fig. 1.3 que os caminhões pesados saíram de uma eficiência energética de 2,14km/L em 2000 para 2,41km/L após 21 anos.

No ‘*Panorama Brasil*’ o Benchmarking Internacional (2021) destaca o papel significativo dos caminhões pesados no consumo de energia do setor de transporte e faz uma distinção entre os caminhões entre pesados, médios e leves. A Fig. 1.4 utiliza-se dos dados da Tab. 1.1 para distinguir o consumo entre carga, passageiro e aviação e os dados do Benchmarking Internacional (2021) para destacar os caminhões pesados de outros tipos de modais e de caminhões.

Figura 1.4: Diagrama de *sankey* do transporte de carga no Brasil



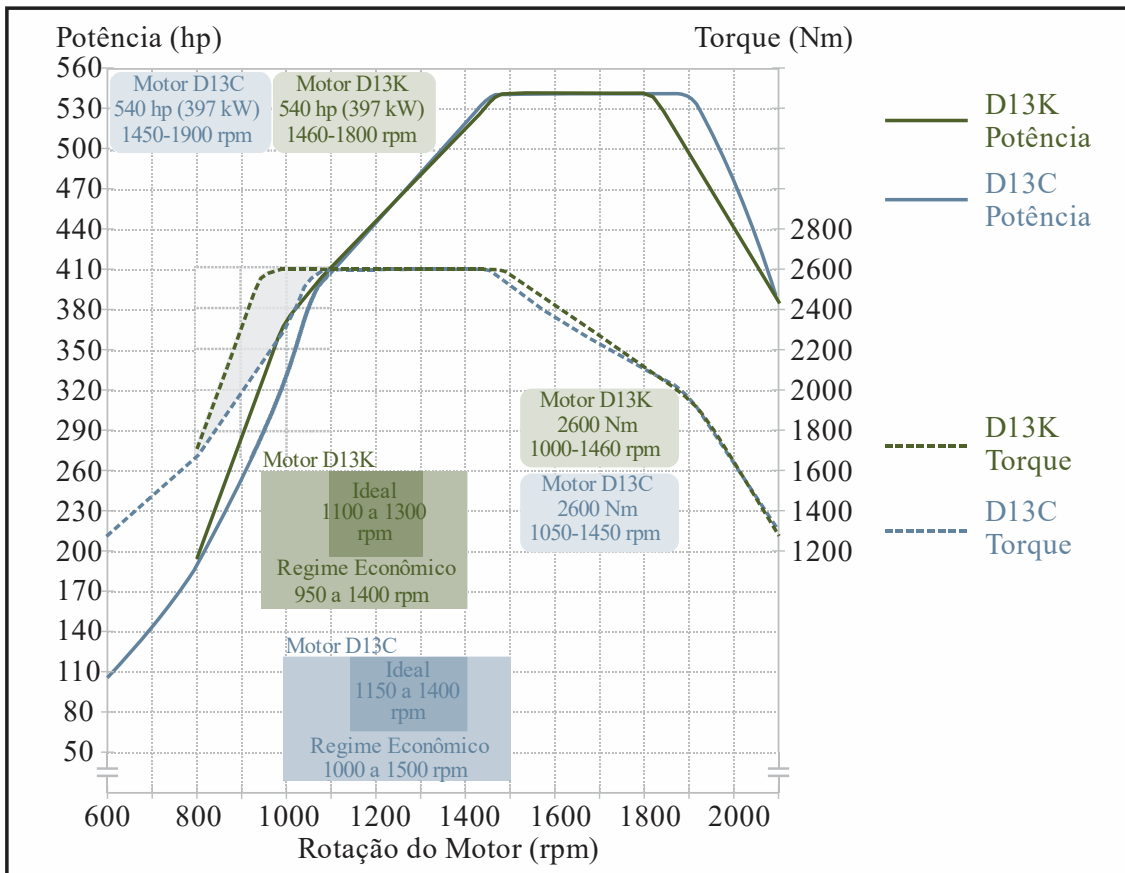
Fontes: Matriz Energética Nacional (2022) e (2023), Atlas da Eficiência Energética (2022) e Benchmarking Internacional (2021). Elaboração do gráfico: Autor

1.2 Evolução dos motores de combustão

Os principais fabricantes de caminhões possuem soluções e resultados similares, então a título de exemplificação utilizou-se a evolução dos motores Volvo da geração D13.

De acordo com Volvo Group (2018), em 2012, para atender a nova legislação de emissões à época, a Volvo trouxe para o Brasil o motor D13C “padrão Euro 5”. Segundo Volvo Group (2022b), entre as novidades para atender a nova regulamentação Euro 6/Proconve P8 a Volvo trouxe o motor D13K para a linha de caminhões 2023. As folhas de especificações dos motores D13C e D13K estão respectivamente disponíveis em Volvo Truck (2010) e Volvo Truck (2019). O gráfico apresentado na Fig. 1.5 é uma sobreposição das curvas de torque e potência dos dois motores.

Figura 1.5: Curvas de Torque e Potências dos Motores D13C e D13K



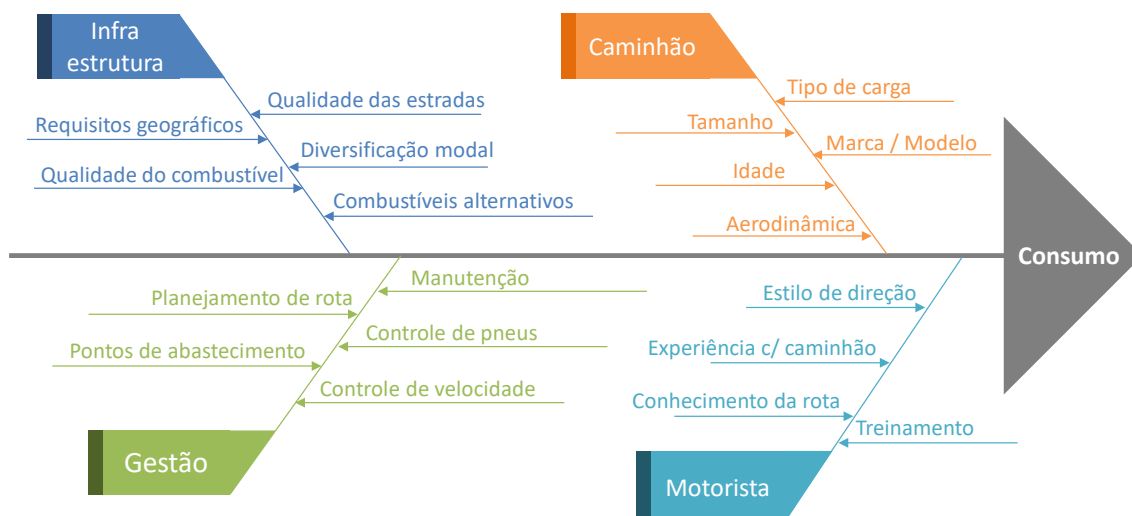
Fonte: Folhas de especificações Folha de Especificações FH Motor D13C540 e Folha de Especificações FH Motor D13K540. Organização: Autor

O que se observa na Fig. 1.5 é que a evolução de um motor para o outro não representou um ganho de potência máxima ou de torque máximo, e sim o deslocamento da faixa econômica para rotações menores. O processo de transição do regime de marcha-lenta para o de torque máximo se faz mais rapidamente nos novos motores e isso proporciona economia de combustível.

1.3 Fatores que afetam o consumo de diesel

Diversos fatores exercem influência no consumo energético no transporte de cargas, como abordado no Benchmarking Internacional (2021). A Fig. 1.6 condensa de maneira visual e eficaz os elementos-chave identificados no relatório conjunto do MME/EPE/IEA, concentrando-se especificamente na Infraestrutura e no Caminhão. As dimensões de Gestão e Motorista foram elaboradas por meio de entrevistas realizadas com profissionais da condução e gestores de frota.

Figura 1.6: Diagrama de Ishikawa dos fatores que afetam o consumo de diesel



Fonte: Benchmarking Internacional (2021) e entrevistas com motoristas e gestores de frota.

Organização: Autor

Sem dúvida, há uma variedade de fatores adicionais que não foram contemplados na representação gráfica da Fig. 1.6. No entanto, mesmo sendo incompleto, esse gráfico se configura como uma ferramenta visual valiosa, facilitando a compreensão do processo e a análise dos diversos elementos que impactam o consumo de diesel.

1.4 Objetivos

Este projeto tem como objetivo compreender os fatores determinantes no consumo de diesel e nas emissões de poluentes associadas a veículos ou frotas do modal rodoviário, buscando formas de mensurá-los de maneira eficaz. Essa abordagem visa não apenas identificar os fatores influenciadores, além de fornecer ferramentas práticas para a gestão eficiente da frota.

Como parte do processo para alcançar o objetivo geral deste trabalho, foram estabelecidos objetivos específicos, a serem progressivamente atingidos com base em dados automotivos e de georreferenciamento coletados *in loco* por meio da tomada OBD II do veículo e armazenados em nuvem. O foco é compreender os fatores determinantes no consumo de diesel e desenvolver métodos para mensurá-los. Os objetivos específicos abrangem:

- a) Identificar se o caminhão trafega com a carreta carregada ou vazia, utilizando apenas os dados de funcionamento do veículo;
- b) Identificar locais de interesse e estabelecer os instantes de carga e descarga da carreta;
- c) Identificar trechos de trânsito frequentes e criar um banco de dados correspondente a cada trecho para comparações posteriores;
- d) Criar um *dashboard* contendo indicadores relevantes e de fácil interpretação;
- e) Estabelecer uma base de dados contendo acíves frequentes de interesse.
- f) Comparar o desempenho dos caminhões em trechos semelhantes.

Estes objetivos específicos visam não apenas à obtenção de dados relevantes, mas também à criação de ferramentas práticas para uma gestão mais eficiente da frota, proporcionando *insights* valiosos para a manutenção preventiva e otimização do consumo de combustível.

1.5 Justificativa e Importância da Pesquisa

A análise detalhada dos dados telemáticos possibilitará uma gestão mais eficiente da frota de caminhões, permitindo a identificação precoce de falhas mecânicas. Além disso, a integração da topografia das estradas com a captura de dados telemáticos oferece uma abordagem avançada para monitorar e preservar a saúde da frota de caminhões em operação, impactando positivamente a produtividade e segurança das operações.

1.6 Linhas de pesquisas para redução de emissões

Os fabricantes de caminhões pesados estão investindo em várias linhas de pesquisa para reduzir as emissões de poluentes. Algumas dessas áreas de pesquisa incluem:

- a) Eletrificação e veículos híbridos: Isso envolve o uso de motores elétricos, baterias de alta capacidade e sistemas de gerenciamento de energia avançados. Os veículos elétricos e híbridos têm potencial para reduzir significativamente as emissões de poluentes, especialmente quando combinados com fontes de energia renovável. Exemplo: Mercedes-Benz (2022) “*IAA Transportation 2022: Daimler Truck revela*

caminhão eActros LongHaul elétrico à bateria e expande portfólio de mobilidade elétrica”.

- b) Motores de baixa emissão: Isso inclui o desenvolvimento de tecnologias de pós-tratamento, como filtros de partículas diesel (DPF), sistemas de recirculação de gases de escape (EGR), catalisadores seletivos de redução (SCR) e injeção de ureia. Essas tecnologias visam reduzir as emissões de óxidos de nitrogênio (NOx), material particulado (MP) e outros poluentes. Exemplo: DAF (2022a) “*DAF entrega o seu primeiro caminhão com motorização Euro 6 no Brasil*”.
- c) Combustíveis alternativos: Uso de combustíveis alternativos, como biocombustíveis, gás natural comprimido (GNC) e gás natural liquefeito (GNL). Esses combustíveis podem reduzir as emissões de gases de efeito estufa e melhorar a qualidade do ar. Exemplo 1: Scania (2023) “*Scania X-gas: o primeiro caminhão a gás com 900 km de autonomia*”. Exemplo 2: Volvo Group (2023b) “*Volvo começa a testar caminhões a hidrogênio em estradas na Europa*”.
- d) Otimização aerodinâmica: Os fabricantes estão projetando caminhões com formas mais aerodinâmicas, incluindo carenagens, saias laterais, defletores de ar e para-choques otimizados. Essas melhorias ajudam a reduzir o arrasto aerodinâmico e, conseqüentemente, o consumo de combustível e as emissões de CO₂. Exemplo: Volvo Group (2022a) “*Novo pacote aerodinâmico deixa Volvo FH ainda mais econômico*”.
- e) Sistemas avançados de controle de velocidade adaptativo: Esses sistemas utilizam dados de mapeamento, GPS e informações topográficas (elevações, descidas, curvas e declives) para otimizar a operação dos veículos pesados. Eles ajustam automaticamente a velocidade, as trocas de marcha e a frenagem, com o objetivo de melhorar a eficiência energética, economizar combustível e reduzir as emissões, ao adaptar a performance dos caminhões às condições da estrada em tempo real. Exemplo: Volvo Group (2023a) “*i-See pré mapeado da Volvo Caminhões*”. A Tab. 1.2 traz referências de vídeos explicativos a respeito desses sistemas.

Tabela 1.2: Nomenclatura do controle de velocidade adaptativo por marca

Fabricante	Nomenclatura	Vídeo explicativo
DAF	Predictive Cruise Control	https://youtu.be/nMVRdG5VTh8
MAN	EfficientCruise	https://youtu.be/PVxETVKneO8
Mercedes-Benz	Predictive Powertrain Control	https://youtu.be/lX0iOZVN_L0
Scania	Scania Active Prediction	https://youtu.be/zVCyvOePs4E
Volvo	I-See	https://youtu.be/YJMUtL0rXEA

Fontes: Sites dos fabricantes. Organização da tabela: Autor

1.7 Definições importantes

As seguintes definições foram retiradas da Demanda Energética do Setor de Transportes - Estudos do Plano Decenal de Expansão de Energia 2032 (2023) e servirão de orientação ao longo do texto.

Automóveis: Veículos destinados ao transporte de passageiros (ex.: sedã, *hatch*, minivan, conversível, *Sport Utility Vehicle* (SUV) e utilitários

Comerciais leves: Veículos comerciais para transporte de pessoas e/ou cargas, com PBT de até 3,5t (ex. picape, furgões e vans).

Categorias de caminhões em função do Peso Bruto Total (PBT), Capacidade Máxima de Tração (CMT) e Peso Bruto Total Combinado (PBTC):

- **Semileves:** $3,5t < PBT < 6t$
- **Leves:** $6,0t \leq PBT < 10t$
- **Médios:** $10,0t \leq PBT < 15t$
- **Semipesados:**
 - Caminhões-chassi com $PBT \geq 15t$ e $CMT \leq 45t$
 - Caminhões-trator com $PBT \geq 15t$ e $PBTC < 40t$
- **Pesados:**
 - Caminhões-chassi com $PBT \geq 15t$ e $CMT > 45t$
 - Caminhões-trator com $PBT \geq 15t$ e $PBTC \geq 40t$

Intensidade energética: Montante de energia necessário para produzir uma unidade de serviço de transporte (lge/tku ou lge/pkm). Segundo o Atlas (2022) a intensidade energética se

refere ao montante de energia necessário para produzir uma unidade de produto final ou de serviço. É a razão entre um indicador de energia (tonelada equivalente de petróleo [tep], Joule, calorias, Btu, entre outras) e um indicador de atividade (US\$, R\$, m², tonelada-quilômetro, passageiro-quilômetro, entre outros. Nesse trabalho, a intensidade energética será indicada pela razão entre o volume de diesel consumido (indicador de energia) e 1000 tonelada-quilômetro útil, tku (indicador de atividade).

Outras definições importantes à compreensão são feitas pela Norma NIE-DIMEL-100 (2011) do INMETRO.

Constante k do cronotacógrafo: É o fator característico que qualifica e quantifica a informação que o instrumento deve receber a cada quilômetro percorrido. A constante k deve ser expressa em rotações por quilômetro (rot/km) ou pulsos por quilômetro (pulsos/km).

Coefficiente w do veículo: É o fator característico que qualifica e quantifica a informação fornecida pelo veículo correspondente a uma distância de 1 km. O coeficiente w deve ser expresso em rotações por quilômetro (rot/km) ou pulsos por quilômetro (pulsos/km).

2. FERRAMENTAS PARA ANÁLISE

A crescente preocupação com as questões ambientais e a busca por eficiência energética têm impulsionado a adoção de tecnologias inovadoras na indústria automotiva. Nesse contexto, a telemática e o *machine learning* emergem como poderosas ferramentas para a análise e compreensão dos problemas relacionados ao consumo de combustível e emissão de poluentes em veículos.

A telemática, que envolve a coleta e transmissão de dados em tempo real, possibilita a obtenção de informações detalhadas sobre o desempenho dos veículos, padrões de condução e até mesmo o estado das vias. Já o *machine learning*, um subcampo da inteligência artificial, é capaz de analisar esses dados em larga escala, identificar padrões ocultos e desenvolver modelos preditivos mais precisos. Nesta abordagem, explora-se o potencial da combinação entre telemática e *machine learning* como uma poderosa estratégia para otimizar a eficiência do consumo de combustível e reduzir as emissões de poluentes, contribuindo para um transporte mais sustentável e alinhado com as demandas ambientais e sociais.

2.1 Telemática

Nos termos de SimonHu *et al.* (2022): “A telemática envolve enviar, receber e armazenar informações usando dispositivos de telecomunicação para controlar objetos remotos. Utilizando tecnologias sem fio atuais e sistemas computacionais, os dados de telemática veicular podem capturar várias características de um veículo em execução na estrada e geralmente podem ser usados não apenas como uma ferramenta de avaliação para veículos e motoristas individuais, mas também para avaliar os impactos ambientais do transporte urbano de carga em uma escala muito mais ampla”.

A telemática é, portanto, um campo da tecnologia que combina comunicação, computação e informações veiculares para fornecer recursos avançados de monitoramento e controle em tempo real. Ela envolve a coleta, transmissão e análise de dados de veículos para auxiliar na gestão de frotas, diagnóstico de falhas, otimização de desempenho e muito mais.

A Fig 2.1 fornece uma visão concisa da essência da telemática. Dentro desse cenário, um dispositivo de telemática é integrado à rede interna do veículo, desempenhando a função de capturar uma ampla variedade de dados, que incluem informações operacionais, registros de falhas e a localização atual. Posteriormente, esses dados são encaminhados através da rede de

telefonia ou via satélite para um servidor pertencente ao fornecedor dos serviços de telemática. Nesse servidor, os dados são recebidos e processados para serem convertidos em informações relevantes e úteis. Em um estágio adicional, essas informações são transmitidas a outros servidores por meio de *Application Programming Interface* (API), permitindo uma customização conforme o tipo específico de análise desejada. Por fim, esses dados são disponibilizados ao usuário final, completando o ciclo de obtenção, processamento e entrega de informações na área da telemática.

Figura 2.1: Como funciona a telemática

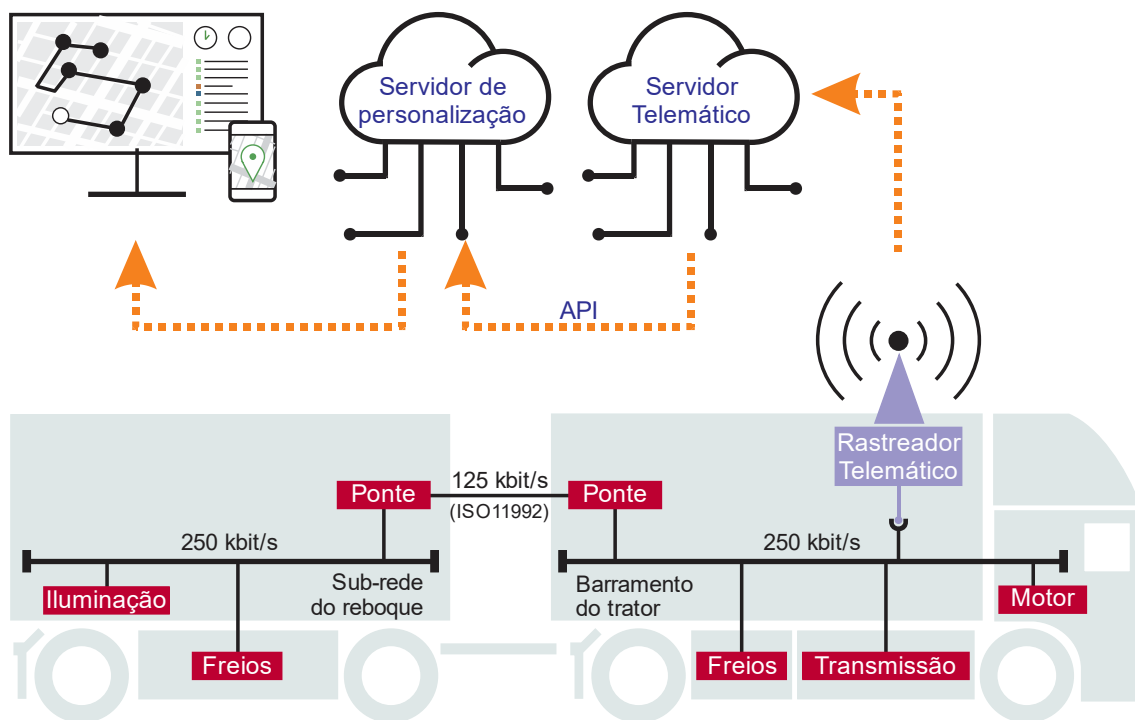


Imagem adaptada de Vector (2023) e Ware *et al.* (2022)

Nesse contexto as normas SAE J1939 desempenham um papel fundamental, uma vez que as estas definem um conjunto de protocolos de comunicação padronizados para a troca de informações entre os componentes eletrônicos em veículos pesados. A telemática aproveita essas normas para permitir a coleta de dados de sensores e sistemas de controle do veículo.

2.1.1 Protocolo de comunicação veicular SAE J1939

Em 1993 a International Organization for Standardization (ISO) publicou a primeira norma ISO 11898 e estabeleceu o protocolo de comunicação CAN (*Controller Area Network*) entre sistemas eletrônicos embarcados para permitir a troca de informações entre diferentes

subsistemas do veículo. O protocolo CAN é conhecido por sua confiabilidade, velocidade e capacidade de suportar a comunicação em tempo real (ISO 2023).

As mensagens que trafegam em uma rede CAN são compostas por um identificador numérico e até 8 bytes de carga útil, além de alguns bits de controle. Elas podem ser enviadas a partir de qualquer equipamento presente na rede, porém seguem uma ordem de prioridade de acesso ao barramento, que é definida pelo identificador da mensagem (SAE 2023a).

A coleção de normas SAE J1939 surgiu em 1994 e define uma rede de comunicação CAN, de alta velocidade à época, baseada na norma ISO 11898-1. Essa rede suporta funções de controle em tempo real, trocas simples de informações e trocas de dados de diagnóstico entre unidades de controle eletrônico em todo o veículo. A SAE J1939 é reconhecida como a solução preferencial do protocolo CAN para veículos aplicados nos setores de construção, incêndio/salvamento, silvicultura, manuseio de materiais e rodovias. Os veículos de interesse incluem, mas não se limitam a caminhões rodoviários e fora de estrada e seus reboques, equipamentos de construção e equipamentos e implementos agrícolas (SAE 2023e).

Desde sua criação, a coleção de normas SAE J1939 passa por revisões para incorporar avanços tecnológicos e atender às necessidades em constante evolução da indústria automotiva. Essas revisões visam melhorar a eficiência da comunicação, adicionar funcionalidades e solucionar possíveis problemas identificados nas versões anteriores (SAE 2023d).

A SAE J1939 possui pelo menos 16 padrões essenciais que compõe o protocolo (SAE 2023c). O documento SAE J1939-21, é um desses padrões essenciais e define o esquema de interpretação do identificador CAN de 29 bits. Dentre esses 29 bits, 8 bits definem um código para origem da mensagem, 3 bits definem a prioridade da mensagem e um conjunto de 18 bits é definido como *Parameter Group Numbers* (PGN). Vector (2023) explica que os PGNs agrupam informações relacionadas em uma única mensagem e são usados para identificar e categorizar diferentes tipos de dados transmitidos pela rede SAE J1939. Por exemplo, o PGN 61444 está relacionado ao controle eletrônico do motor e o PGN 61441 ao controle eletrônico dos freios. Existem cerca de 450 PGNs categorizados.

Por outro lado, os *Suspect Parameter Numbers* (SPN) são identificadores específicos atribuídos a parâmetros de dados individuais, tais como temperatura do motor, velocidade do veículo, pressão do óleo, entre outros. Os SPNs são usados para acessar dados específicos dentro dos PGNs e permitem que os sistemas interpretem e utilizem esses dados de maneira adequada (Vector 2023). Existem mais de 3500 SPNs listados.

Como uma amostra, o PGN 65269 refere-se às condições do ambiente e a carga útil da mensagem de 8 bytes são divididos da seguinte forma: o byte 1 corresponde ao SPN 108, que revela a pressão barométrica do meio; os bytes 2 e 3 formam o SPN 170, que diz respeito à temperatura no interior da cabine; os bytes 4 e 5 representam o SPN 171, relacionado à temperatura do ar ambiente; o byte 6 corresponde ao SPN 172, que representa a temperatura do ar de admissão no motor, e, por fim os bytes 7 e 8 são do SPN 79, que indica a temperatura de superfície da estrada.

Em muitos casos, o SPN pode ser representado por bits de um único byte do PGN. Para ilustrar, o SPN 563 é utilizado para informar se o sistema de freios com ABS (*Anti-Lock Braking System*) está ativo e esse dado pode ser capturado nos bits 5 e 6 do primeiro byte que compõe o PGN 61441.

O anexo digital SAE J1939DA disponibiliza todos os SPNs, PGNs e demais dados da coleção de normas SAE J1939 em um formato de planilha, o que possibilita pesquisas, classificações e conversões para outros formatos de maneira ágil e eficiente (SAE 2023b).

A série SAE J1939 trata esse conjunto de definições como ‘Práticas Recomendadas’. Isso significa que cabe ao fabricante do veículo adotá-la em seu todo ou em parte. Portanto, normalmente a lista de SPNs disponíveis no protocolo de comunicação do veículo é muito inferior às possibilidades listadas na J1939, uma vez que os SPNs foram definidos de maneira a atender uma enorme gama de funcionalidades previstas para veículos, com tipos de combustíveis, acessórios, sensores, entre outros itens, diferentes (SAE 2011).

No Apêndice A, encontra-se uma lista com os principais fornecedores de telemática aos frotistas de caminhões rodoviários no Brasil

2.2 Machine Learning

Conforme mencionado por Zhang *et al.* (2022), existe uma dificuldade em construir um conjunto de dados ideal para o treinamento de modelos de diagnóstico inteligentes. Os autores destacam que, no contexto da engenharia, as máquinas geralmente operam em condições normais, com falhas raras, o que dificulta a obtenção de dados de falhas suficientes para treinar modelos de diagnóstico. Além disso, experimentos em laboratórios podem ser caros e a simulação computacional pode não ser prática o suficiente para obter dados de falhas.

Uma solução proposta é o uso de telemática veicular, que permite a captura de dados de funcionamento dos veículos em tempo real e em grande quantidade. Isso implica um desafio

para a aplicação de conhecimentos científicos na análise desses dados. As abordagens relacionadas a *big data*, inteligência artificial e *machine learning* têm sido cada vez mais utilizadas, conforme demonstrado pela crescente quantidade de publicações nos últimos anos, especialmente com *machine learning*.

Segundo Alpaydin (2014) o aprendizado de máquina, ou *machine learning*, é um programa de computador utilizado para otimizar um critério de desempenho usando dados de exemplo ou experiência anterior. O *machine learning* é um nicho de estudos que surge quando não é possível criar uma sequência de instruções que devem ser executadas para transformar a entrada em saída (algoritmo), mas que é possível construir uma aproximação boa e útil capaz de explicar parte dos dados, detectar padrões e regularidades para fazer previsões, ajudar a entender o processo e ter a capacidade de aprender com um ambiente em constante mudança.

Ao contrário da programação tradicional, onde as regras são explicitamente codificadas, no *machine learning* os algoritmos são treinados em dados para aprenderem automaticamente a realizar tarefas específicas. Segundo Alpaydin (2014), a base é a teoria da estatística na criação de modelos matemáticos, onde o primeiro papel é o treinamento do modelo e o segundo é a eficiência do modelo aprendido.

2.2.1 Categorias de Machine Learning

Alpaydin (2014) destaca três categorias de aprendizado de máquina:

- a) **Aprendizado Supervisionado:** Nesse tipo, o algoritmo é treinado em um conjunto de dados rotulados, ou seja, dados onde já se conhece a resposta correta. O objetivo é que o modelo aprenda a mapear entradas para saídas corretas, de forma que possa fazer previsões precisas com novos dados. Exemplos de aplicação incluem classificação de e-mails como spam ou não spam e previsão do preço de imóveis com base em características específicas.
- b) **Aprendizado Não Supervisionado:** Nessa categoria, o algoritmo é treinado em dados não rotulados e deve encontrar padrões e estruturas ocultas nos dados. O objetivo é descobrir informações relevantes ou agrupamentos sem a necessidade de rótulos predefinidos. Um exemplo clássico é a segmentação de clientes para análise de comportamento de consumo.
- c) **Aprendizado por Reforço:** Aqui, o algoritmo aprende através de tentativa e erro, recebendo *feedback* em forma de recompensas ou penalidades. O objetivo é que o

modelo descubra a melhor maneira de realizar uma tarefa específica interagindo com um ambiente em constante mudança. Jogos, como xadrez ou Go, são exemplos onde o aprendizado por reforço tem sido amplamente aplicado.

2.2.2 Exemplos de aplicação de *machine learning*

- a) Classificação – objetivo: identificar a qual categoria um objeto pertence. Exemplo: reconhecimento de dígito escrito à mão.
- b) Regressão – objetivo: prever um atributo de valor contínuo associado a um objeto. Exemplos: Resposta a medicamentos, preços de ações.
- c) *Clustering* – objetivo: Agrupamento automático de objetos semelhantes em conjuntos. Exemplo: Segmentação de clientes
- d) Redução de dimensionalidade – objetivo: Reduzir o número de variáveis aleatórias a serem consideradas.

2.2.3 Algoritmos de *machine learning* clássicos

Os algoritmos clássicos de aprendizado de máquina têm sido há muito tempo a base da análise de dados e do reconhecimento de padrões. Esses algoritmos formam o alicerce sobre o qual o campo do aprendizado de máquina evoluiu, fornecendo técnicas fundamentais para resolver uma ampla gama de problemas complexos. Apesar de terem sido desenvolvidos há alguns anos, são muito eficazes hoje em dia e são frequentemente usados em aplicações práticas, como reconhecimento de padrões, classificação e regressão.

Nesse trabalho serão utilizados os seguintes algoritmos de *machine learning*:

- Naive Bayes: É um classificador probabilístico. Ele baseia-se em modelos de probabilidade que incorporam suposições de independência forte. Muitas vezes, as suposições de independência não têm impacto na realidade. Portanto, elas são consideradas ingênuas, IBM (2023).
- Árvore de Decisão: É um modelo de aprendizado supervisionado que utiliza regras de decisão com base nos valores dos atributos. Ela consiste em *nós*, sendo um deles a *raiz*, estabelecendo uma hierarquia entre eles. Essa abordagem é empregada para resolver problemas de classificação e regressão. As decisões são tomadas com base nas características mais relevantes, Lauretto (2010).

- *Random Forest*: É um algoritmo baseado em conjuntos (*ensemble*) que combina várias árvores de decisão para melhorar a precisão e a generalização. Cada árvore é treinada em uma amostra aleatória dos dados e suas previsões são combinadas para obter um resultado mais robusto, Scikit-Learn (2023).
- *AdaBoost (Adaptive Boosting)*: É outro algoritmo de conjunto que melhora o desempenho de modelos fracos, dando mais peso aos exemplos classificados incorretamente em cada iteração. A combinação desses modelos ponderados resulta em um modelo forte e preciso, Scikit-Learn (2023).
- *KNN (K-Nearest Neighbors)*: É um algoritmo de aprendizado supervisionado usado para classificação e regressão. Ele calcula a proximidade entre os pontos de dados e classifica novos pontos com base na maioria dos k vizinhos mais próximos, Scikit-Learn (2023).
- *Regressão Logística*: É um algoritmo de aprendizado supervisionado usado para problemas de classificação binária. Apesar do nome, é usado para prever probabilidades de pertencer a uma classe específica e não para tarefas de regressão, Scikit-Learn (2023).
- *SVM (Support Vector Machines)*: É um algoritmo de aprendizado supervisionado usado tanto para problemas de classificação quanto de regressão. O SVM busca encontrar um hiperplano que melhor separe os dados em classes distintas, Scikit-Learn (2023).
- *Redes Neurais Artificiais*: São modelos inspirados na estrutura do cérebro humano, compostos por camadas de neurônios interconectados. As redes neurais artificiais são capazes de aprender representações complexas dos dados e são amplamente utilizadas em problemas de classificação, regressão e processamento de imagens, fala e texto, Scikit-Learn (2023).
- *PCA*: Embora não seja um algoritmo de aprendizado supervisionado, o PCA é uma técnica de redução de dimensionalidade amplamente utilizada. Ele transforma um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas (componentes principais), permitindo uma representação mais compacta dos dados, Scikit-Learn (2023).

3. REVISÃO BIBLIOGRÁFICA

Essa revisão bibliográfica teve como objetivo buscar artigos que poderiam ajudar a solucionar problemas de consumo de combustível e emissões de poluentes em caminhões rodoviários, por meio da análise de dados obtidos por telemática veicular. A pesquisa realizada nas bases Scopus e Web of Science (WoS), usou a intersecção de quatro grandes conjuntos: economia de combustível, caminhões, telemática e aprendizado de máquina. Os resultados foram analisados por programas bibliométricos para selecionar trinta artigos relevantes.

Para reduzir o escopo do estudo e lidar com a grande diversidade de máquinas na Engenharia Mecânica, optou-se por focar na revisão sistemática que envolve a aquisição de dados pela telemática veicular em caminhões rodoviários, com o objetivo de reduzir o consumo de combustível e realizar diagnósticos inteligentes de falhas. Essa escolha se baseia no relatório do Benchmarking Internacional (2021), que aponta que os caminhões pesados são responsáveis por cerca de 60% da energia consumida no transporte de cargas no Brasil (Fig. 1.4).

A telemática fornecerá um enorme volume de dados de funcionamento dos caminhões em que o equipamento telemático estiver instalado. Esses equipamentos estão disponíveis para aquisição a um preço acessível, o que se traduz em vantagem da não necessidade de um laboratório físico nas instalações da universidade, visto que o dado é coletado em campo e transmitido em armazenado em nuvem pelo fabricante do equipamento. O trabalho consistirá em construir uma *Application Programming Interface* (API) e buscar nos dados coletados as respostas aos problemas que se deseja estudar.

3.1 Metodologia da revisão bibliográfica

A metodologia proposta consistiu em responder a seguinte pergunta: 'Como a utilização de dados da telemática tem sido usada em pesquisas relacionadas à Engenharia Mecânica, em especial em caminhões rodoviários?' Para coletar os dados necessários, foram utilizadas as bases Scopus - Elsevier (2022) e WoS - Clarivate (2022), e o tratamento dos dados foram realizados por meio das ferramentas Biblioshiny de Aria and Cuccurullo (2017) e pybibx de Pereira (2022).

Inicialmente foram criados quatro grupos para organizar as palavras-chave utilizadas na busca, sendo eles:

- a) a região limitante do problema;
- b) o problema que se deseja estudar;
- c) como os dados seriam adquiridos; e
- d) as técnicas de análise dos dados.

Definidos os grupos de palavras, a busca resultou em alguns artigos em duplicidade, devido a inconsistências na indexação, tais como a identificação do *Digital Object Identifier* (DOI) em apenas uma das bases. Essas inconsistências foram identificadas e removidas. Além disso, foi identificado e contabilizado apenas uma vez os artigos que apareceram em mais de uma base.

Após essa organização, os artigos foram classificados como elegíveis ou não. Entradas do tipo *Conference Review* em bases de dados bibliográficas como Scopus e WoS são índices que listam os artigos publicados em uma determinada conferência e, muitas vezes, trazem uma breve descrição do evento em si. Ou seja, essas entradas não se referem a um artigo específico, mas sim a uma lista de artigos publicados em um determinado evento. Resultados listados com essa classificação, *Conference Review*, foram criteriosamente verificados para evitar confusões na interpretação dos resultados.

3.1.1 A formação de cada conjunto de pesquisa

Foram definidas palavras-chave para cada grupo, e realizadas buscas nas bases Scopus e WoS, agrupadas por *Article title*, *Abstract*, *Keywords*. Os resultados obtidos em 03/01/2023 são apresentados na Tab. 3.1. O Scopus retornou uma lista com 100 resultados e a WoS 45.

Tabela 3.1. Palavras de busca para formação dos conjuntos

Grupo	Palavras-chave	Scopus	WoS
A	truck* OR heavy-duty OR vehicle* OR fleet	1.090.198	601.203
B	"fuel consumption" OR maintenance OR "mass detection" OR "fault detection" OR "intelligent fault diagnosis"	950.798	583.678
C	telematic* OR obd OR few-shot OR (sample AND (small OR few OR limited OR unbalance OR imbalance*))	780.436	1.004.776
D	"machine learning" OR "deep learning" OR "neural network" OR "artificial intelligence"	1.532.583	702.675
Lista	$A \cap B \cap C \cap D$	100	45

Fonte: Autor

3.1.2 Análise bibliométrica, unificação dos dados e remoção de duplicatas

Para realizar a análise bibliométrica, utilizamos o Biblioshiny, uma ferramenta de código aberto para pesquisa quantitativa em cienciometria e bibliometria. Com o auxílio dessa ferramenta, foram obtidos resultados do Scopus e WoS individualmente.

Com a ajuda do Biblioshiny foi possível perceber que palavras como “*fault detection*”, “*failure analysis*” e “*deep learning*” formam um agrupamento de palavras chaves que levam a interpretação de que existe um conjunto de artigos que utilizam de métodos de “*deep learning*” para detecção e análise de falhas.

Ainda com o auxílio de relatórios do Biblioshiny é possível notar que “*fuel consumption*” é uma palavra que aparece com frequência e está conectada a técnicas de *machine learning* como “*support vector*”, “*random forest*” e “*neural network*”, ou que as pesquisas por detecção de falhas utilizam-se mais da técnica de redes neurais.

O trabalho de unificação das bases dos dados foi realizado com o pybibx e o conjunto se solidificou conforme a Tab. 3.2.

Tabela 3.2: Relatório após a unificação das bases

Tipo de documento	Exclusivo Scopus	Comum a Scopus e WoS	Exclusivo WoS	Total
Article	24	21	6	51
Article; Early Access		1	1	2
Book Chapter	2			2
Conference Paper / Proceedings Paper	32	11	4	47
Conference Review	8			8
Editorial			1	1
Review	1			1
Duplicado	0	0	0	0
Total	67	33	12	112

Fonte: Autor

Foram identificados, como mostra a Tab. 3.2, 33 artigos comuns às duas bases, ou seja, aquele grupo inicial de 100 artigos Scopus e 45 WoS, reduziu-se a 112.

Tanto Biblioshiny quanto o pybibx fornecem uma lista das palavras mais frequentes, as quais foram organizadas e apresentadas no formato da Tab. 3.3.

Tabela 3.3. Relatório de palavras chaves mais frequentes

Local de pesquisa	Uma palavra - ngram = 1	Par de palavras - ngram = 2	Trio de palavras - ngram = 3
resumo	<i>data</i> (285)	fuel consumption (89)	support vector machine (12)
título	learning (34)	machine learning (81)	vehicle fuel consumption (4)
palavras-chave	systems (91)	machine learning (42)	support vector machine (10)
palavras-chave do autor	learning (31)	fuel consumption (16)	eco driving fuel (6)

* Entre parêntesis a frequência de repetição das palavras

Fonte: Autor

Os resultados apresentados na Tab. 3.3 reforçam a conexão de “*fuel consumption*” com “*machine learning*” e “*support vector machine*”.

3.1.3 Identificação e remoção de artigos não elegíveis através de palavras-chave

Os documentos do tipo *Conference Review* foram avaliados cuidadosamente para evitar que um artigo fosse selecionado pelo fato das palavras de busca que originaram o resultado na base Scopus ou WoS aparecerem no resumo da Conferência, porém em artigos distintos. Após essa análise, a lista inicial de 112 artigos foi reduzida a 78 documentos elegíveis.

3.1.4 Leitura e classificação da lista resultante

Após a leitura dos 78 resumos, e com a ajuda das palavras-chave mais frequentes relacionadas pelo Biblioshiny (expansão da Tab. 3.3), iniciou-se um processo de classificação destes resumos e chegou-se em quatro grupos:

- Pesquisas relacionadas à busca de causas ou estimativas para explicar o consumo de combustível ou a emissão de poluentes, com as seguintes palavras em destaque: *fleet operations* e *fuel consumption*;
- Pesquisas relacionadas à determinação ou medição da influência do estilo de direção no desempenho do caminhão ou atuação como assistente de direção, onde as palavras de destaque são: *eco-driving*, *fuels*, *driving styles* e *fuel economy*;
- Pesquisas relacionadas à utilização de dados para aprimoramento de planos de manutenção, cujas palavras *predictive maintenance*, *maintenance* e *commercial vehicles* são os destaques; e

- Pesquisas com interesse em engenharia elétrica, sistemas de comunicação, diagnóstico por imagem, pesagem de veículo em movimento e outras. A análise das palavras mais frequentes permitiu identificar aqueles artigos que não eram elegíveis para este estudo. No caso, apareceram palavras como *autonomous vehicle*, *battery*, *child*, *peach*, *photovoltaic*, *railway*, *road*, *smartphone*, *UAV*, *vegetation* e *weather*, as quais foram utilizadas para indicar a não elegibilidade de um artigo.

3.2 Resultados da Revisão Bibliográfica

As Tabs. 3.4, 3.5 e 3.6 identificam, respectivamente, os títulos relacionados aos *clusters* 1, 2 e 3, ou seja, pesquisas relacionadas à busca ou estimativas para explicar o consumo de combustível ou a emissão de poluentes; pesquisas relacionadas a determinar ou medir a influência do estilo de direção no desempenho do caminhão ou atuar com assistente de direção; e pesquisas relacionadas à utilização de dados para aprimoramento de planos de manutenção. Os títulos foram ordenados do mais antigo para o mais recente e em ordem alfabética.

Na produção científica por país, sobressai-se a China com 15 autores, seguido da Espanha com 9. O artigo com maior número de citações é *Application of machine learning for fuel consumption modeling of trucks* por Perrotta *et al.* (2017) com 59 citações totais.

Perrotta *et al.* (2017) utilizaram SVM, Random Forest e Redes Neurais Artificiais para modelar o consumo de combustível em caminhões articulados. Os dados dos caminhões foram coletados da rede SAE J1939 e cruzados com dados de relatórios referentes a características das estradas da Inglaterra.

Destacam-se outros artigos muito relacionados ao trabalho aqui proposto:

Rimpas *et al.* (2020) avaliaram dados recuperados através do protocolo de diagnóstico On-Board Diagnostic II (OBD-II) para monitoramento da operação e do consumo do veículo.

Pavlovic *et al.* (2021) afirmam que a análise estatística mostrou que a precisão do OBD para medir o consumo de combustível é impactada pela velocidade média do veículo, aceleração e dinamicidade geral da viagem.

Walnum e Morten (2015) analisaram a influência do comportamento do motorista no consumo de combustível e ao final descobriram que as variáveis usadas, que dependem principalmente da infraestrutura e do terreno, têm efeitos que são 10 a 12 vezes maiores do que os efeitos de variáveis que atribuíram ao comportamento de dirigir.

Por fim, Eagon *et al.* (2022) afirmam que prever a massa do veículo é fundamental para estimar, com precisão, o uso de energia e as emissões dos caminhões comerciais, e propõem um modelo de detecção de massa.

Entre os 30 artigos selecionados, a produção científica anual oscila entre 1 e 3 artigos entre 2013 e 2020 e salta a 9 em 2021 e 2022. A *SAE Technical Papers* é a fonte mais relevante - 3 artigos publicados. A maioria das fontes locais citadas, 21, é a *Transportation Research Part D: Transport and Environment*, seguida da *Applied Energy* e *Sensors* ambas com 10 citações cada uma. Os autores mais relevantes foram CHEN e YEN, ambos com 3 artigos. Dentre as afiliações, a *National Chin-Yi University of Technology* de Taiwan se destacou com 4 artigos e a *University of Nottingham* da Inglaterra com 3 artigos.

Tabela 3.4: Artigos classificados no *cluster* consumo de combustível ou emissão de poluentes

Título	Ano
<i>Intelligent prediction model of vehicle fuel consumption</i>	2014
<i>Application of machine learning for fuel consumption modelling of truck</i>	2017
<i>Accessing and constructing driving data to develop fuel consumption forecast model</i>	2018
<i>Data driven analysis and forecasting of medium and heavy truck fuel consumption</i>	2019
<i>Methodology to recognize vehicle loading condition - an indirect method using telematics and machine learning</i>	2019
<i>Route-sensitive fuel consumption models for heavy-duty vehicles</i>	2020
<i>A deep learning method for monitoring vehicle energy consumption with GPS data</i>	2021
<i>Combining a universal OBD-ii module with deep learning to develop an eco-driving analysis system</i>	2021
<i>Estimation of pollutant emissions in real driving conditions based on data from OBD and machine learning</i>	2021
<i>Neural models to quantify the determinants of truck fuel consumption</i>	2021
<i>Remote NO_x emission prediction model based on LSTM neural network</i>	2021
<i>Estimating truck fuel consumption with machine learning using telematics, topology, and weather data</i>	2022
<i>Integration of second-generation on-board diagnostics data via deep learning to develop eco-driving analysis system applicable to large and small car</i>	2022
<i>Interpretable machine learning models for predicting and explaining vehicle fuel consumption anomalies</i>	2022
<i>Mass detection for heavy-duty vehicles using gaussian belief propagation</i>	2022
<i>Predicting gasoline vehicle fuel consumption in energy and environmental impact based on machine learning and multidimensional big data</i>	2022
<i>Traffic emissions clustering using OBD-ii dataset based on machine learning algorithms</i>	2022

Fonte: Scopus e WoS; Organização: Autor

Tabela 3.5: Artigos classificados no *cluster* estilo ou assistente de direção

Título	Ano
<i>The impact of using gamification on the eco-driving learning</i>	2014
<i>Assessing the impact of driving behavior on instantaneous fuel consumption</i>	2015
<i>Artemisa: A personal driving assistant for fuel saving</i>	2016
<i>Drivingstyles: A mobile platform for driving styles and fuel consumption characterization</i>	2017
<i>Lot sensing for reality-enhanced serious games, a fuel-efficient drive use case</i>	2021
<i>Truck driving assistance system</i>	2021
<i>Driving style analysis by studying pid's signals for determination of its influence on pollutant emission</i>	2022

Fonte: Scopus e WoS; Organização: Autor

Tabela 3.6: Artigos classificados no *cluster* utilização de dados para aprimoramento de planos de manutenção

Título	Ano
<i>Overview of telematics-based prognostics and health management systems for commercial vehicles</i>	2013
<i>On-board predictive maintenance with machine learning</i>	2019
<i>Preventive maintenance for heterogeneous industrial vehicles with incomplete usage data</i>	2021
<i>Scenario uncertainty modeling for predictive maintenance with recurrent neural adaptive processes (RNAPs)</i>	2021
<i>Handling imbalanced data for aircraft predictive maintenance using the bache algorithm</i>	2022
<i>Predictive maintenance of vehicle fleets using LSTM autoencoders for industrial IoT datasets</i>	2022

Fonte: Scopus e WoS; Organização: Autor

4. APLICAÇÃO PRÁTICA

Este estudo engloba a análise de três bases de dados provenientes de diferentes contextos operacionais no setor de transporte rodoviário. Cada conjunto de dados foi obtido a partir de experiências distintas, proporcionando uma visão abrangente e contextualizada sobre o consumo de diesel e emissões associadas a veículos comerciais.

A Tab 4.1 apresenta um resumo das características de cada base de dados, destacando detalhes cruciais sobre a frota, equipamento telemático e a disponibilidade dos dados para análise.

Tabela 4.1: Bases de dados dos grupos de estudo

Base de Dados	Transportadora	Frota	Equipamento Telemático	Período de Coleta	Dados Adicionais
1	Combustível	10 caminhões (2 Volvo, 8 DAF)	Desenvolvido internamente	62 dias	Relatórios de abastecimento, carregamento/d Descarregamento
2	Alimentícios	1 caminhão (Iveco)	Geotab®	Março/2023 a Janeiro/2024	Dados externos obtidos somente a partir do final de julho/2023
3	Areia	1 caminhão (Volvo)	Geotab®	Agosto/2023 a Janeiro/2024	Dados de abastecimento e carga fornecidos pela transportadora, com algumas falhas na comunicação

Elaboração: Autor

Esse capítulo apresenta formas de extrair métricas importantes do banco de dados e validará se essas informações são confiáveis para aplicação em outros estudos. Além disso, são apresentados valores de interesse sobre o consumo de diesel dos caminhões analisados. Dessa forma, pretende-se avaliar se os dados de telemática podem ser aplicados confiavelmente em estudos que visam analisar o consumo de combustível.

Na avaliação proposta, o objetivo será o de utilizar técnicas de aprendizado supervisionado para determinar o *status* de carregamento dos caminhões, tendo como base de dados as informações já presentes na rede de comunicação desses caminhões, ou seja, sem adicionar qualquer hardware extra, como sensores específicos para essa finalidade.

Foram criados dois grupos distintos de desenvolvimento: O primeiro recebeu os dados em forma de arquivos ‘.csv’ em coleta realizada durante 62 dias consecutivos por equipamentos

de telemática instalados em 10 caminhões a diesel de uma transportadora de combustível (base 1). No segundo foi instalado um equipamento Geotab GO9 em dois caminhões de duas transportadoras diferentes (bases 2 e 3) com acesso aos dados através da plataforma my.geotab.com ou através do conjunto de ferramentas do *Software Development Kit* (SDK) que a Geotab® disponibiliza em <https://geotab.github.io/sdk/>

Algumas variáveis obtidas por meio da telemática, como a rotação do motor e a velocidade do veículo, serão enriquecidas com suas médias, desvios padrão e medianas correspondentes. Embora essa abordagem aumente o número de variáveis, optou-se por descartar a redução de dimensionalidade com a *Principal Component Analysis* (PCA). Essa decisão foi tomada devido à natureza tabular de todas as variáveis e ao fato de que a quantidade existente não compromete o desempenho computacional do código. Eagon *et al.* (2022) justificaram a mesma decisão ao descobrirem que oferecia pouco valor para os dados de treinamento fornecidos.

Para completar as informações oriundas da telemática, deve-se buscar informações das especificações técnicas de cada caminhão; relatórios de abastecimento de combustível e relatórios de carregamento desses caminhões.

As especificações técnicas devem conter, no mínimo:

- a) Faixa de rotação de marcha lenta;
- b) Faixa de rotação de torque máximo;
- c) Faixa de rotação de potência máxima;
- d) Limite máximo de rotação do motor;
- e) Rotação na aplicação do freio motor;
- f) Volume de cada tanque de combustível;
- g) Fase da legislação à qual está submetido o caminhão;
- h) Outros dados, tais como número de cilindros e cilindrada total.

Os relatórios de abastecimento servirão para confrontar os registros de volume de combustível abastecido com as estimativas de volume de combustível consumido, calculadas a partir dos dados oriundos da telemática.

Finalmente, o relatório de carregamento servirá de base para o aprendizado supervisionado, cujo objetivo é identificar se o caminhão se encontra carregado ou vazio.

4.1 Validação da base de dados 1

A aplicação prática se inicia com dados já coletados em uma transportadora de combustível. Tais dados não continham informações de identificação dos veículos, condutores e cargas transportadas. Sabe-se, no entanto, que a transportadora não costuma realizar transporte de cargas fracionadas, e que as cargas possuem o mesmo volume, com variações inferiores a 1%. Esses dados foram anonimizados pela transportadora para manter a privacidade, tanto em relação à operação, quanto ao condutor. Para esse estudo, isso também é relevante de forma a garantir a imparcialidade no tratamento dos dados.

No contexto dos caminhões de carga, algumas características relevantes para a eficiência energética, tais como velocidade de deslocamento, rotação do motor, distância percorrida, volume de combustível consumido e variações de altitude na estrada podem ser mensuradas pelo sistema eletrônico do veículo e extraídas com o auxílio de equipamentos de aquisição apropriados.

4.1.1 Descrição e análise do banco de dados 1

Esse estudo considera um banco de dados obtido a partir da coleta realizada por equipamentos de telemática instalados em 10 caminhões a diesel de uma transportadora. São 8 caminhões DAF XF105 FTT, equipados com motorização PACCAR MX 13 de 375 kW, e 2 caminhões Volvo FH-540, equipados com motorização D13C e potência máxima de 397 kW.

Os equipamentos de telemática foram desenvolvidos pela equipe de Pesquisa e Desenvolvimento da transportadora. O equipamento concatena dados coletados no GPS (*Global Positioning System*), com informações internas de execução, conexões com redes de dados e sensores internos, e dados dos protocolos SAE J1939, ISO 15765. Como regra de aquisição telemática, foi definida a coleta de 1 em 1 segundo para se ajustar com o PPS (*Pulse Per Second*) do GPS.

Ao se analisar os dados, nota-se que nem todas as variáveis são preenchidas para todos os veículos. Isso acontece pelo fato de que alguns veículos não reportam determinados SPNs na rede de dados, ou os dados são lidos a partir de um protocolo proprietário.

A transportadora disponibilizou 620 arquivos '.csv', os quais representaram 62 dias consecutivos de operação dos 10 caminhões. Nesse período, foram percorridos cerca de 165 mil quilômetros e consumidos pouco mais de 95 mil litros de diesel. A frequência de aquisição característica do equipamento utilizado é de uma amostra de 174 variáveis por segundo (1 Hz).

Em condições ideais, um arquivo deve conter 86.400 linhas, no entanto, observou-se uma ausência global de 168.560 amostras, o que corresponde a uma perda de 0,31% dos registros temporais. Desse modo, compôs-se uma planilha geral com um total de 53.399.440 linhas e 174 colunas. A lista completa de variáveis analisadas encontra-se no Apêndice B.

Durante a análise dos arquivos, identificaram-se 123 variáveis com seus respectivos SPNs. No entanto, é importante ressaltar que alguns dados identificados com SPNs não foram provenientes diretamente da rede CAN dos caminhões, e sim do sistema GPS, tais como latitude, longitude e altitude, e que possuem SPNs relatados na coleção de normas SAE J1939. Além disso, outras 51 variáveis não possuíam SPNs correlacionados.

Por se tratar de uma quantidade massiva de dados, optou-se por extrair medidas descritivas de algumas variáveis, como valores máximo, mínimo, média, moda, mediana e desvio padrão. Tal abordagem foi inspirada pelo trabalho de Wolberg *et al.* (1993), no qual os autores transformaram 10 características relevantes para o diagnóstico do câncer de mama em 30 variáveis de entrada, ao considerarem a média, o desvio padrão e a média dos três maiores valores, para alimentar um sistema de aprendizado de máquina.

As estatísticas foram extraídas agrupando os dados de acordo com a condição de operação dos caminhões ao longo do dia, que poderia ser desligado, parado com o motor ligado ou em movimento. Admitiu-se motor desligado quando a rotação era inferior à rotação mínima de marcha lenta e o caminhão parado quando a velocidade era igual a zero km/h. Desse modo, criou-se um arquivo composto por 39.266 linhas (99,926% a menos que a base original).

Durante a análise, encontraram-se dados que indicavam o caminhão com a rotação inferior à rotação mínima de marcha lenta e com velocidade maior do que zero, o que representa um veículo desligado e em movimento. Esses dados representaram apenas 0,98% do total de dados e foram desprezados para fins de análise.

Para cada estado de funcionamento, foi calculada a duração do intervalo, o tamanho do percurso percorrido e as estatísticas relacionadas à velocidade (SPN 84), rotação (SPN 190), percentual de torque (SPN 513), percentual de carga (SPN 82) e informações de condições iniciais e finais de latitude (SPN 584), longitude (SPN 585), hodômetro (SPN 917), consumo total de combustível (SPN 250 para caminhões DAF ou 5054 para Volvo) e percentual de nível de combustível (SPN 96). Foram incluídas outras variáveis relevantes disponíveis, como altitude (SPN 580), vazão instantânea de combustível (SPN 183), estado do pedal de freio (SPN 521) e número de marcha atual (SPN 523).

Os dados obtidos pela telemática foram complementados com informações técnicas dos caminhões em Volvo Truck (2010) – Folha de Especificações FH Motor D13C, ANFAVEA (2015) – Tabelas de Emissões Diesel Volvo, DAF (2019) – Paccar MX-13 Engines, DAF (2017) – Ficha Técnica XF105 FTT e DAF (2022b) – Índices de Emissão, relatórios de abastecimento e relatório das condições de carga do caminhão. É de fundamental importância destacar que esses relatórios de carga dependem de informações do motorista do caminhão e estão sujeitas a divergência, apesar dos esforços da gerência da transportadora. Tais informações foram utilizadas na etapa subsequente, com o intuito de validar a nova base de dados para estimativa do volume de combustível consumido, bem como do percurso percorrido.

4.1.2 Análise e manipulação de dados da base 1

Para validação dos dados relacionados ao volume de combustível, consideraram-se três métodos de cálculo:

- O consumo pela variação do nível do tanque (NT) de combustível, que se dá pelo produto entre a capacidade nominal do tanque (CN), especificada pelo fabricante (em litros), e a variação do percentual do nível (%PN) de combustível (SPN 96), anotada durante o processo de abastecimento. Eq. (4.1):

$$Vol_{k=NT} = CN * (\%PN_{MAX} - \%PN_{MIN}) \quad (4.1)$$

- O valor acumulado (VA) de consumo de diesel (em litros), registrado pelos SPN 250 (para caminhões DAF) e SPN 5054 (para caminhões Volvo). O consumo estimado é dado pela diferença dos valores indicados (TA) pelo respectivo SPN entre o final e o início do estado de funcionamento. Eq. (4.2):

$$Vol_{k=VA} = TA_{MAX} - TA_{MIN} \quad (4.2)$$

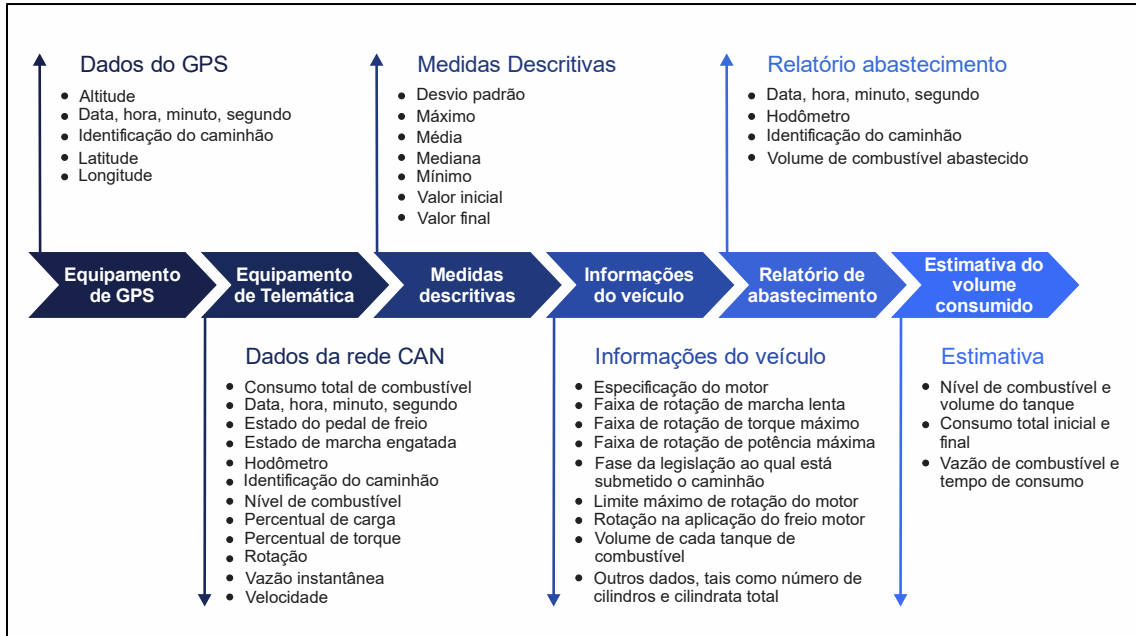
- A vazão instantânea (VI) de combustível. O valor de consumo é obtido pelo somatório do produto entre a vazão instantânea (Q) de combustível (dada pelo SPN 183), convertida para litros por segundo, e o intervalo de tempo (t) em segundos. Eq. (4.3):

$$Vol_{k=VI} = \sum_{i=1}^n Q_{ij}(t_i - t_{i-1}) \quad (4.3)$$

Basicamente, o primeiro método fornece uma estimativa aproximada do consumo de combustível entre abastecimentos, o segundo entre trechos e o terceiro a cada dado capturado e transmitido.

A Fig. 4.1 ilustra o procedimento aplicado na análise dos dados. O objetivo para validação dos dados foi comparar os valores registrados no relatório de abastecimento com as estimativas de consumo resultantes dos três métodos anteriormente descritos.

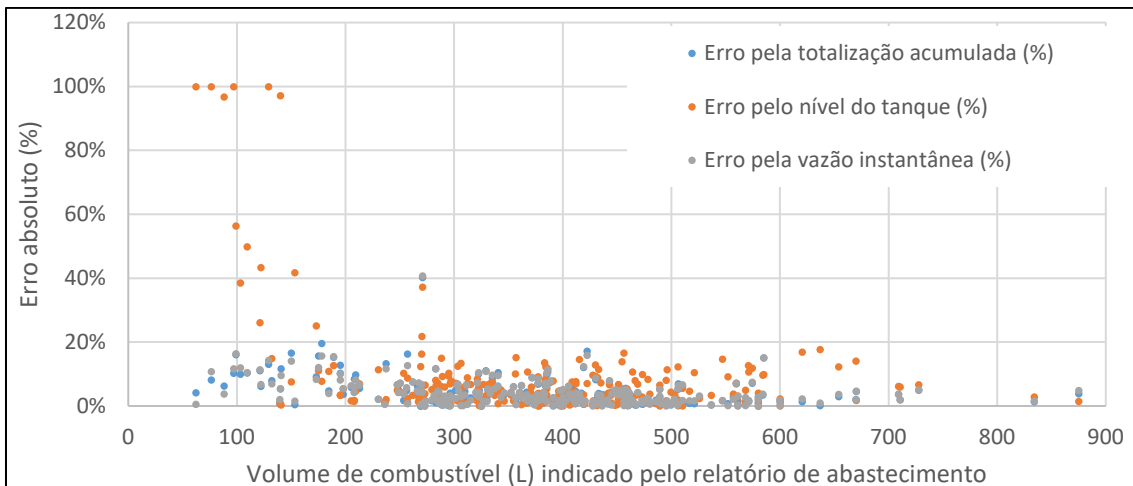
Figura 4.1: Estimativas do volume de combustível consumido



Elaboração: Autor

Ao se analisar os dados de maneira individual, verificou-se que episódios de abastecimentos pequenos (inferiores a 300 litros) estão suscetíveis a maiores divergências entre o valor registrado no relatório de abastecimento e a variação de nível de combustível (SPN 96).

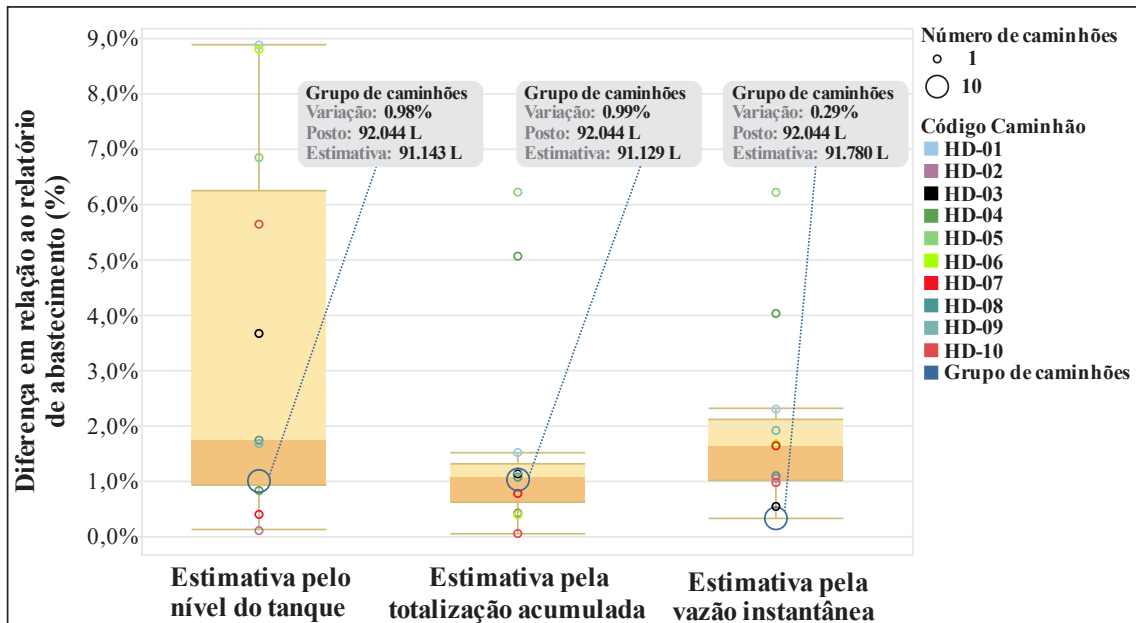
Figura 4.2: Erro absoluto (%) de cada abastecimento na estimativa do volume de combustível por diferentes métodos de cálculo



Elaboração: Autor

Em pelo menos 4 episódios com abastecimentos inferiores a 150 litros o nível de combustível não teve sensibilidade suficiente para detectar o acréscimo de volume. No entanto, ao acumular os abastecimentos ao longo de todo o tempo analisado, os resultados indicam que, no pior caso, a diferença máxima fica abaixo de 9%, como ilustra a Fig. 4.3.

Figura 4.3: Comparativo entre as estimativas de volume abastecido e o relatório de abastecimento



Fonte de dados: Transportadora. Organização: Autor

Os registros de abastecimentos disponibilizados pela transportadora informam que os 10 caminhões abasteceram 92.044 litros de diesel no período avaliado pelo estudo, descontado o primeiro abastecimento de cada caminhão. De modo geral, o terceiro método (estimativa do consumo pela vazão instantânea, utilizando o SPN 183) foi o que mais se aproximou do relatório de abastecimento, apresentando uma diferença de apenas 264 litros (0,29%) em relação aos 92.044 litros registrados. A estimativa pelo segundo método (consumo acumulado) apresentou diferença de 0,99%, enquanto a estimativa pelo primeiro método (pela variação de nível), apresentou uma diferença de 0,98%.

De forma semelhante, a comparação entre o valor da distância percorrida, verificada pela variação do hodômetro (HO) (SPN 917) – Eq. 4.4, e a estimativa de deslocamento, dada pelo produto entre a velocidade (V) registrada (SPN 84) e o intervalo de tempo (t) admitido – Eq. 4.5, possui boa aproximação – Eq. 4.6.

$$\Delta O_{k=hodometro} = HO_{MAX} - HO_{MIN} \quad (4.4)$$

$$D_{k=velocimetro} = \sum_{i=1}^n V_{ij}(t_i - t_{i-1}) \quad (4.5)$$

$$\Delta O_{k=hodometro} = HO_{MAX} - HO_{MIN} \approx D_{k=velocimetro} = \sum_{i=1}^n V_{ij}(t_i - t_{i-1}) \quad (4.6)$$

Em percursos abaixo de 300 m, é possível encontrar diferenças superiores a 5% entre o valor estimado e o registrado pelo hodômetro. Já para percursos acima de 10 km, essa diferença foi inferior a 0,59%, e com tendência de redução com aumento da distância percorrida.

Esses resultados, obtidos ao calcular o volume usando a vazão instantânea (SPN 183 – Eq.4.3) e o percurso usando a velocidade (SPN 84 – Eq.4.5), forneceram credibilidade ao processamento de dados e flexibilidade ao processo.

4.1.3 Inserção do relevo entre as características relevantes

Como alertado por Walnum e Morten (2015) a respeito das implicações do relevo, buscou-se incluir essas características na lista de características disponíveis para futura análise e, uma vez que se torna possível calcular o volume de combustível consumido e o percurso percorrido em intervalos de tempo de um segundo, então, no estado de "veículo em movimento", a diferença de altitude revela as condições do relevo da estrada e é possível calcular o consumo e o percurso dentro da seguinte classificação:

- 'active' – a altitude atual é maior do que a altitude no instante anterior;
- 'declive' – a altitude atual é menor do que a altitude no instante anterior
- 'plano' – a altitude atual é igual a altitude no instante anterior

É importante evidenciar que a medida de altitude reportada pelo GPS é influenciada por diversos fatores, além da inclinação da via, como, por exemplo, a qualidade do sinal GPS e trepidações. Para efeito desse estudo, os possíveis erros foram considerados desprezíveis, uma vez que seriam necessários outros sensores, como giroscópio e acelerômetros, para permitir melhor representação das condições ambientais.

Com essas definições, além das medidas descritivas mencionadas anteriormente, a utilização da vazão instantânea (SPN 183) permitiu estimar não apenas o volume de combustível consumido durante um percurso, mas também obter informações sobre o consumo em diferentes condições de relevo. Na Tab. 4.2 têm-se os parâmetros de distância percorrida, volume consumido e economia de combustível em diferentes condições de carga e relevo. Por fim, foram agregadas variáveis relacionadas à quantidade de trocas de marcha e ao uso do pedal de freio em diferentes condições de relevo.

É importante destacar que os volumes de combustível apresentados na Tab. 4.2. são diferentes dos volumes destacados na Fig. 4.3, porque, para que fosse possível a comparação de resultados entre o relatório de abastecimento e os diferentes métodos de cálculo, foi preciso ignorar o primeiro abastecimento de cada caminhão, uma vez que este carrega consigo consumos anteriores ao relatório de dados.

Tabela 4.2: Consumo estratificado em função do relevo

Grupo	Medida	Método	Status de relevo ou funcionamento	Carregado ou Vazio	Carregado	Vazio
10 caminhões	Volume de combustível consumido (L)	Eq. 4.2	Independente de relevo	95.122,90	78.224,04	16.898,86
			Independente de relevo	95.798,01	78.614,71	17.183,30
		Eq. 4.3	Parado com motor ligado	799,40	473,40	326,00
			Aclive	65.835,96	54.185,97	11.649,99
			Declive	17.232,32	14.012,68	3.219,63
			Plano	11.566,92	9.643,13	1.923,79
			Dados não categorizados	363,42	299,54	63,88
	Eq. 4.4	Independente de relevo	165.118,44	119.469,82	45.648,62	
		Independente de relevo	165.090,65	119.461,14	45.629,51	
		Eq. 4.5	Aclive	71.188,17	51.195,55	19.992,61
			Declive	72.925,08	52.731,53	20.193,54
			Plano	20.324,36	15.048,57	5.275,79
			Dados não categorizados	653,05	485,48	167,56
	Duração de funcionamento (h)	Σ intervalo	Independente de relevo	3.548,52	2.575,00	973,52
			Parado com motor ligado	308,56	181,81	126,75
	Consumo (L/h)	Volume / Duração	Parado com motor ligado	2,59	2,60	2,57
	Economia (km/L)	Distância / Volume	Independente de relevo	1,72	1,52	2,66
			Aclive	1,08	0,94	1,72
			Declive	4,23	3,76	6,27
			Plano	1,76	1,56	2,74

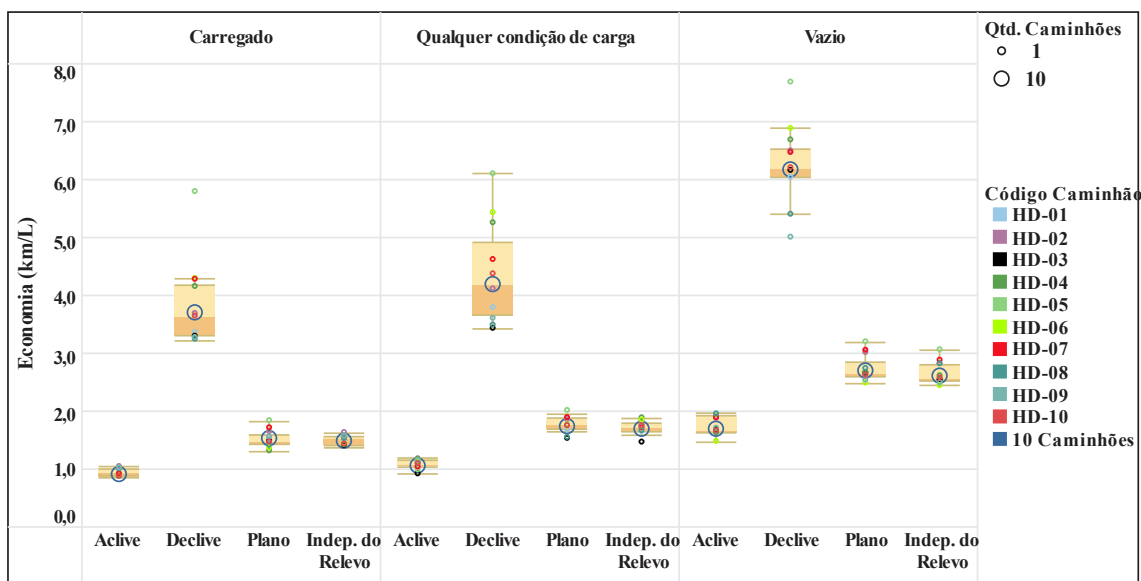
Fonte de dados: Transportadora. Cálculos e organização: Autor

As Eq. 4.2 e 4.4 não proporcionam uma estratificação dos resultados com base no relevo, ao contrário das Eq. 4.3 e 4.5 (índice 'j' representativo da categoria), que permitem o cálculo em diversas condições de relevo, ou consumo durante períodos de veículo parado com o motor ligado. Para assegurar que a soma dos valores em cada condição fosse equivalente ao total, uma linha, denominada 'dados não categorizados', foi introduzida na tabela.

Observa-se ainda que o consumo médio do conjunto de caminhões, desconsiderando as condições do relevo, movimentando em vazio, tem uma economia média de 2,66 km/L, enquanto na condição carregado essa média cai para 1,52 km/L, uma redução de quase 43% na autonomia desses caminhões.

Na Fig. 4.4. apresenta-se o resultado de economia de combustível (km/L) em função do relevo e da condição de carga para os 10 caminhões analisados. Importante mencionar que o cálculo da economia de combustível considerou a estimativa de distância percorrida indicada pelo hodômetro e o volume consumido pela estimativa de vazão instantânea.

Figura 4.4: Economia de combustível (km/L) em função da carga e do relevo



Fonte de dados: Transportadora. Cálculos e Organização: Autor

Na Fig. 4.4, merece destaque a capacidade de discernir 12 médias distintas de economia de combustível, levando em consideração tanto a carga quanto o relevo. Sendo os dados registrados a cada segundo, é viável calcular essas 12 médias, juntamente com seus respectivos desvios, em intervalos personalizáveis. Atualmente, de maneira geral, um gestor de frotas possui apenas o resultado mensal de economia (km/L), sem incorporar o *status* da carga e, ainda menos, as nuances das condições de relevo.

4.1.4 Discussão sobre fontes de erro

Existem diversas outras fontes de erros em processos de medição. Cada SPN especifica uma resolução, porém esse tópico traz um destaque a respeito de fontes de erro que nesse momento estão desconsideradas do processo e tomados como verdadeiro. A título de exemplo:

- O volume do reservatório especificado pelo fabricante;
- As curvas de torque e potência especificados pelo fabricante apenas por pares conjugados de torque máximo e rotação ou de potência máxima e rotação;
- O volume indicado pela bomba de abastecimento;
- Abastecimento acima ou abaixo do nível máximo especificado para o veículo;
- Os relatórios de carga são preenchidos manualmente;
- O efeito da temperatura sobre a densidade do diesel;
- Foi desconsiderada a hipótese de transporte de cargas parciais. Ou o caminhão desloca-se totalmente carregado ou totalmente vazio;
- Considerar que a vazão instantânea e a velocidade não sofrem alteração de valor ao longo dos intervalos de tempo superiores a 1 segundo, ou seja, nos momentos em que foram observadas perda de dados; e
- A altitude revelada pelo GPS.

Apesar da consciência dessas possíveis fontes de erro, e antes de buscar mecanismos que pudessem minimizá-los, optou-se por avaliar a fiabilidade dos dados e do procedimento aplicado.

4.2 Preparação para os algoritmos de *Machine Learning*

De forma resumida, os algoritmos seguiram a sequência:

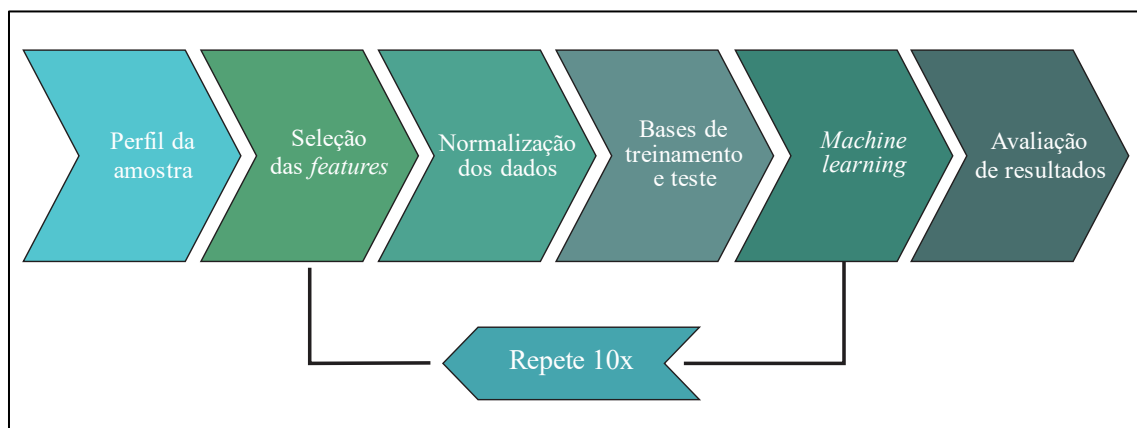
- a) Definir o perfil da amostra conforme metodologia específica;
- b) Selecionar o conjunto de variáveis (características);
- c) Normalizar os dados. Ou seja, garantir que os dados estejam na mesma escala ou intervalo, o que pode melhorar o desempenho de muitos algoritmos de *machine learning*. Isso também pode ajudar a evitar que atributos com valores maiores

dominem os resultados em relação aos atributos com valores menores, o que poderia causar distorções nos resultados do modelo;

- d) Dividir as bases em treinamento (75%) e teste (25%);
- e) Executar a rotina de cada método de *machine learning* considerado. Nesse caso, optou-se pela biblioteca *scikit-learn*, de Pedregosa *et al.* (2011), por ser uma biblioteca de aprendizado de máquina de código aberto que inclui vários algoritmos de classificação, dentre os quais foram escolhidos: *AdaBoost*, *Decision Tree*, *KNeighbors* (KNN), *Logistic Regression*, *Naive Bayes*, *Neural Network*, *Random Forest* e *Support Vector Machine* (SVM). O apêndice C traz um código resumido da aplicação de cada uma das técnicas;
- f) Repetir as etapas “b” a “e” até que o número de repetições seja igual a 10;
- g) Calcular o valor médio de resultado para cada método;

A Fig 4.5 resume graficamente a metodologia aplicada após a validação dos dados até a avaliação dos resultados.

Figura 4.5: Metodologia aplicada



Fonte: Autor

4.3 Resumo dos experimentos realizados com a base 1

Com a base 1 foram realizadas algumas configurações de experimentos.

4.3.1 Método dos trechos integrais divididos em curtos e longos

Rimpas *et al.* (2020) fizeram experimentações em viagens de 5 km com tráfego baixo e pesado. Inspirado nisso e a partir da base de dados disponibilizada, admitiu-se pequenas

distâncias trechos entre 0,5 e 5 km, e longas distâncias trechos entre 5 e 500 km. Destaca-se que na base disponível não existem distâncias superiores a 500 km.

A primeira separou os dados em trechos integrais de 0,5 a 5 km com 1.018 amostras, chamados de trechos curtos e outro grupo de 5 a 500km com 2.387 amostras, chamados de trechos longos. Nessa configuração, o melhor resultado para classificação do estado de carregamento do caminhão foi de 81,96% para trechos curtos e de 95,81% para trechos longos.

A partir de sete variáveis provenientes diretamente da comunicação SAE J1939 – *Engine Fuel Rate* (SPN 183), *Wheel-Based Vehicle Speed* (SPN 84), *Engine Fuel Delivery Pressure* (SPN 94), *Engine Percent Load At Current Speed* (SPN 92), *Accelerator Pedal Position I* (SPN 91), *Actual Engine – Percent Torque* (SPN 513), e *Engine Speed* (SPN 190) – denominadas, respectivamente, como por `clph_vazao_instantanea_combustivel`, `kmph_velocidade`, `kpa_pressao_combustivel`, `perc_carga`, `perc_posicao_acelerador`, `perc_torque` e `rpm_rotacao_motor`, foram criados três conjuntos de características. O primeiro, intitulado V7média, compreende as médias de cada variável; o segundo, V7mediana, inclui as medianas de cada variável; e o terceiro, V7desvio, abrange os desvios padrão correspondente a cada variável.

Além dessas, foram incorporadas mais oito variáveis (V8), originadas do cálculo da economia de combustível em km/L, dada pela Eq. 4.7, e da diferença desse valor em relação à média geral (Eq. 4.8), considerando quatro cenários distintos: aplane, declive, plano e qualquer tipo de relevo.

$$EC_j = \frac{\text{Distância percorrida}_j}{\text{Volume consumido}_j} \quad (4.7)$$

$$DEC_j = EC_j - \overline{EC_j} \quad (4.8)$$

Onde EC significa Economia de Combustível, DEC é a diferença entre economia de combustível e à sua média, e o índice ‘j’ refere-se às condições categóricas de relevo, tais como aplane, declive, plano e qualquer tipo de relevo.

A Tab. 4.3 apresenta os resultados obtidos com diferentes conjuntos de *features*, considerando a possibilidade de combinações entre os conjuntos criados, tais como V15 pela composição de V7média e V8; V21 pela composição de V7média, V7mediana e V7desvio e V29 pela utilização de todos os conjuntos.

Tabela 4.3: Acurácia (%) na classificação da condição de carga em distâncias curtas

km min	km max	Núm. amost	Técnica de classificação	V7 média	V8	V15	V21	V29
0,5	5	1018	AdaBoost	74,90	69,41	76,47	75,69	79,61
			Árvore de Decisão	66,27	62,75	70,59	69,80	71,37
			KNN	79,22	72,94	80,39	80,78	80,39
			Naive Bayes	70,59	64,71	72,55	70,98	70,59
			Random Forest	79,22	71,37	78,04	81,18	81,18
			Redes Neurais	78,12± 1,58	74,43± 1,15	79,18± 1,82	75,76± 1,78	78,04± 1,89
			Regressão Logística	78,04	71,76	79,22	79,61	78,04
			SVM	78,43	74,12	79,22	80,00	81,96

Fonte: Autor

Tabela 4.4: Acurácia (%) na classificação da condição de carga em distâncias longas

km min	km max	Núm. amost	Técnica de classificação	V7 média	V8	V15	V21	V29
5	500	2387	AdaBoost	94,14	94,14	93,47	94,64	93,97
			Árvore de Decisão	93,30	92,13	92,46	91,29	92,13
			KNN	94,30	94,64	94,64	93,80	94,97
			Naive Bayes	90,62	93,97	92,46	91,12	91,96
			Random Forest	93,80	94,30	94,30	93,97	94,47
			Redes Neurais	94,19± 0,34	95,43± 0,30	94,39± 0,60	92,93± 0,65	93,52± 0,76
			Regressão Logística	93,80	94,47	95,31	93,97	94,97
			SVM	94,30	94,30	95,81	93,80	95,14

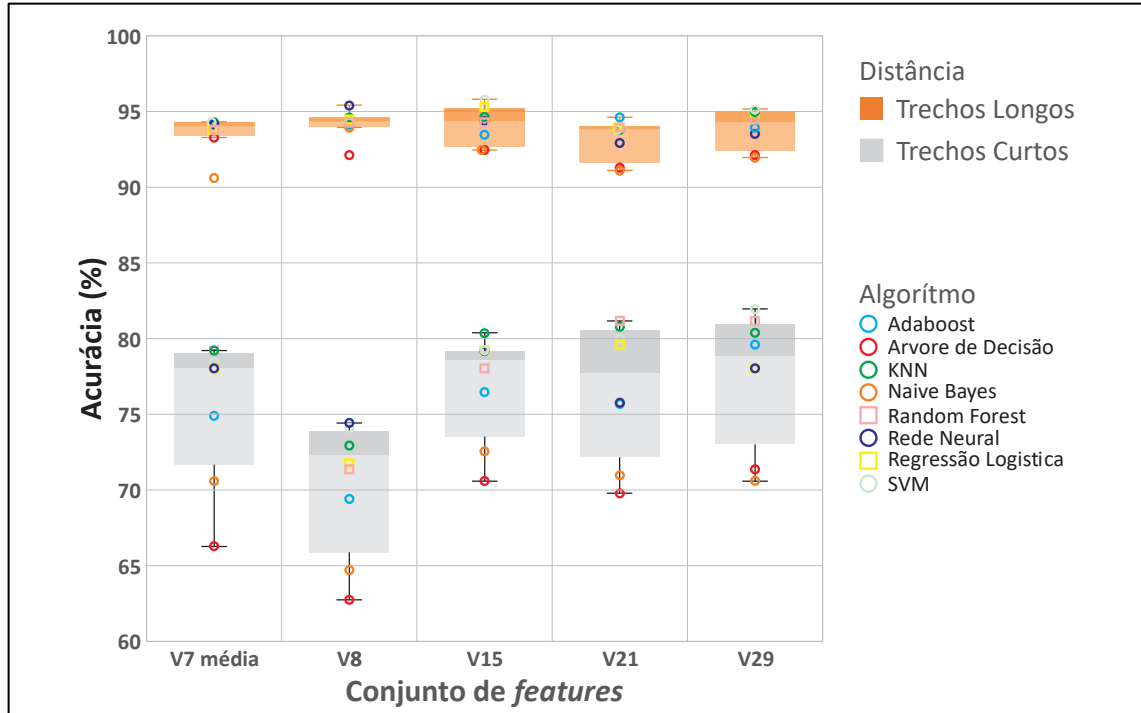
Fonte: Autor

O processo de repetição do código em 10 vezes resultou em um desvio padrão registrado para unicamente para o método de Redes Neurais, tanto nos trechos curtos quanto nos longos.

A Fig. 4.6 condensa os resultados das Tabs. 4.4 e 4.5 em um único gráfico, e oferece uma representação visual das informações. Neste gráfico, torna-se evidente não apenas o aumento positivo na precisão da classificação da condição de carga ao passar de trechos curtos para trechos longos, mas também é observável que a disparidade entre os resultados devido aos

algoritmos de *machine learning* utilizados é menor nos trechos mais extensos em comparação com os trechos mais curtos.

Figura 4.6: Acurácia (%) na classificação da condição de carga em distâncias curtas e longas com diferentes técnicas de *machine learning*



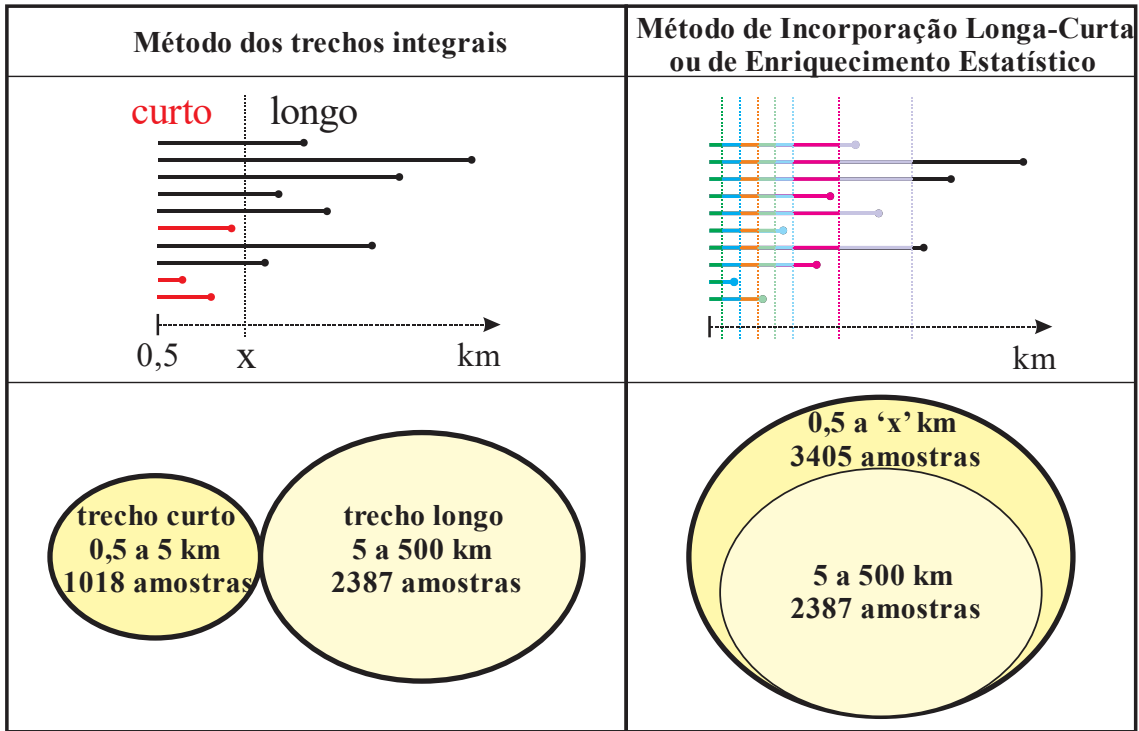
Fonte: Autor

4.3.2 Método de incorporação longa-curta para enriquecimento estatístico

Uma segunda configuração foi criada com a finalidade de aumentar o número de amostras nos trechos curtos, tendo em vista que para percorrer um percurso longo necessariamente se precisa percorrer o trecho curto. Para facilitar a compreensão a respeito do impacto sobre o número de amostras comparado à estratégia anterior, elaborou-se a Fig. 4.7.

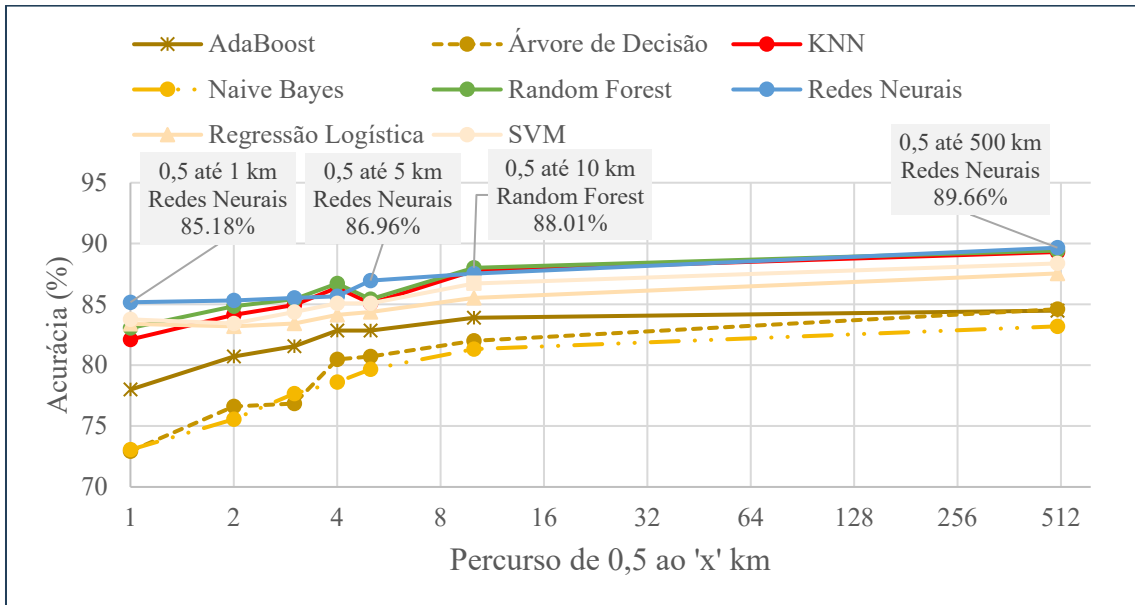
Com a mudança de metodologia, o tamanho do percurso foi fixado entre 0,5 km e um limite de percurso máximo para análise. Criou-se sete grupos de aquisição estatística: até 1 km, até 2 km, até 3 km, até 4 km, até 5 km, até 10 km e até 500 km. Os filtros para formação dos conjuntos de amostra foram: *status* igual a 'Ligado em Movimento', percurso mínimo de 0,5 km e percurso máximo igual ao limite do grupo. Todos os grupos contiveram com 3405 amostras e optou-se pelo conjunto de 21 variáveis, pela facilidade de programação. Os resultados podem ser mais bem avaliados com o gráfico da Fig. 4.8.

Figura 4.7: Método dos trechos integrais e de incorporação longa-curta



Organização: Autor

Figura 4.8: Acurácia (%) na classificação da condição de carga em função do método e do tamanho do percurso



Fonte: Autor

Essa modelagem dos dados resultou em 85,17% a acurácia da classificação a partir do primeiro quilometro de tráfego do caminhão com esse valor se elevando a 89,65% na análise

até 500 km. Ambos os resultados explicitados utilizaram como técnica de classificação a Rede Neural.

4.3.3 Método de segmentação temporal

A terceira configuração definiu o percurso mínimo de 0,5 km e agrupou os dados em função do tempo de movimento do caminhão, e formatou novamente sete grupos com um limite mínimo de tempo de aquisição: 10 s, 60 s, 120 s, 300 s, 600 s, 1800 s e sem limite temporal. Novamente foi escolhido o conjunto de 21 variáveis formadas pelas médias, medianas e desvios padrão das sete variáveis originais. O número de amostras e as distâncias percorridas variaram e são mostrados na Tab.4.5.

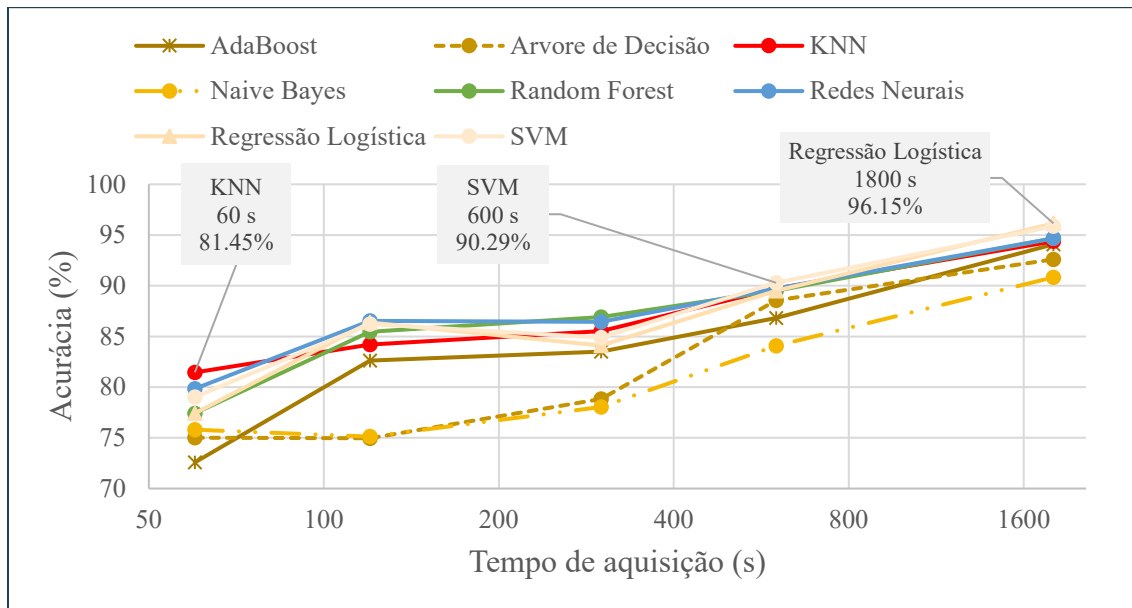
Tabela 4.5: Quadro de distâncias percorridas após um período determinado

Período	Amostras	Distância mínima (km)	Distância média (km)	Distância Máxima (km)
10s	0	0	0	0
60s	496	0,5	0,6	0,97
120s	2555	0,5	1,04	2,07
300s	2568	0,5	3,23	5,58
600s	2059	1,33	7,73	11,62
1800s	1350	11,69	26,26	35,80
Período completo	3405	0,5	48,16	252,68

Fonte: Autor

Esse tipo de abordagem tenderia a um valor de amostras decrescente de 10 s a 1800 s, no entanto esse fato não ocorre devido ao fato de que a amostra se torna inelegível caso não se confirme a distância mínima percorrida dentro do limite de tempo, ou seja, para efeitos de composição estatística da amostra não importa a duração total do percurso e sim a distância percorrida dentro do limite de tempo selecionado. A Fig. 4.9 traça um gráfico do resultado.

Figura 4.9: Acurácia (%) na classificação da condição de carga em função do método e do tempo de aquisição do dado



Fonte: Autor

Com 10 s não se obteve amostras, porém a partir de 60 s de movimento do caminhão conseguiu-se uma acurácia de 81,45% no processo de classificação com esse resultado crescendo a 96,15% com 1800 s de dados de funcionamento do caminhão em movimento.

4.4 Resumo dos experimentos realizados com as bases 2 e 3

O equipamento modelo GO9 de fabricação da Geotab® foi instalado em 15 de março de 2023 num caminhão Iveco Hi Way 480 equipado com um motor Cursor 13 de 353 kW de uma empresa que transporta produtos alimentícios. E em 29 de agosto de 2023 num caminhão Volvo FH460 e motor D13C que transporta areia, calcário ou brita. Bases 2 e 3 segundo a Tab. 4.1.

4.4.1 Descrição e análise do banco de dados

Em seu site, a Geotab® afirma ter, em janeiro de 2024, mais de 4 milhões de dispositivos telemáticos em funcionamento. Geotab (2024). Uma das principais preocupações de um fornecedor de dados telemáticos é a de tentar manter apenas os pontos necessários para fornecer uma representação precisa dos dados. A Geotab® (2022) patenteou o método de Registro de Curvas (do inglês *Curve Logging*) e afirma que essa lógica permite a eles reterem os dados telemáticos mais detalhados com a menor quantidade de sobrecarga de dados.

Na prática, esse método de registro de curvas faz com que os dados possuam espaçamentos de tempo irregulares e o dado somente é transmitido, armazenado e disponibilizado segundo as regras do algoritmo do registro de curvas para cada variável.

Eagon *et al.* (2022) também utilizaram equipamentos Geotab[®] em seus experimentos e explicam que o algoritmo de registro de curva verifica os pontos de erro máximo, determinando se a diferença entre a posição prevista e a real é pequena ou grande. Se a diferença for pequena, os pontos não serão registrados no servidor. O algoritmo é aplicado a todos os parâmetros mensuráveis do sistema, o que resulta em conjunto de dados de amostragem não uniforme e com muitos pontos de dados ausentes ou incompletos, ou seja, um desafio adicional a quaisquer métodos de estimativa orientado por dados.

Entre os dados disponíveis aos quais se obteve acesso estão: latitude, longitude, total horas de funcionamento do motor, velocidade do veículo, rotação do motor, temperatura do motor, pressão de óleo, hodômetro, combustível total usado, combustível total usado em marcha lenta e outros de menor relevância para esse estudo. Ainda não se obteve acesso aos dados de vazão instantânea.

4.4.2 Metodologia adotada com os dados das bases 2 e 3

Existe uma grande diferença entre a formatação dos dados da base 1 recebidos da transportadora de combustível para a base de dados coletada da Geotab[®] que forma as bases 2 e 3, dessa forma elaborou-se uma sequência de passos desde a aquisição do dado até as estratégias de *machine learning* e publicação dos resultados em um *website*.

Passo 1: Conectar à base Geotab[®] e identificar locais de interesse.

Conectar-se à base de dados da Geotab[®] e extrair duas bases distintas: uma denominada '*StatusData*', contendo dados operacionais do caminhão, e outra chamada '*LogRecord*', que fornece informações de localização geográfica. Ambos os conjuntos de dados são organizados em um único *dataframe*. Para lidar com a lacuna nos dados, natural do método de registro de curvas adotado pela Geotab[®], optou-se por realizar interpolações lineares para cada variável operacional, preenchendo assim os dados intermediários entre as linhas existentes. O *dataframe* resultante é então reduzido, mantendo apenas as linhas que contêm valores de latitude e longitude.

Um mapa que delinea a trajetória percorrida é gerado, com ênfase nos pontos de parada que ultrapassam 600 segundos (equivalente a 10 minutos). O processo resulta na criação de dois

arquivos: um arquivo *CSV* contendo dados específicos e outro arquivo que destaca a trajetória no mapa, ressaltando os mencionados pontos de parada. A intenção por trás dessa abordagem é possibilitar a identificação visual de potenciais locais de interesse.

Com o apoio das ferramentas Google Maps e Google Earth, à medida que novos locais foram identificados, procedeu-se a atualização de um arquivo Excel denominado 'Locais_Interesse_*.xlsx', em que '*' representa o nome da empresa. As variáveis incorporadas neste arquivo incluem o nome do local, latitude, longitude, raio de influência (km) e o motivo associado (carga/descarga, abastecimento, manutenção). Essa abordagem implica a dar o nome do local idêntico ao do local de interesse sempre que o caminhão estiver localizado dentro de um círculo com raio de influência desse local de interesse.

Passo 2: Identificar momentos de carga e descarga.

Ler o arquivo salvo na etapa anterior e fazer a fusão com 'Locais_Interesse_*.xlsx' atualizado, para que a coluna 'nome do local' no *dataframe* de estudo traga os nomes com base nos locais de interesse. O objetivo agora é mostrar no *dataframe* os pontos de interesse para que se possa identificar os momentos em que foi realizada a carga ou a descarga de modo a atualizar o arquivo 'Status_Carga_*.xlsx', onde '*' representa a placa do caminhão.

A ideia por trás dessa estratégia é alterar de Carregado para Vazio ou de Vazio para Carregado a partir do último instante de velocidade igual a zero, antes que o caminhão saia do raio de influência de um local de interesse.

Passo 3: Salvar os dados diários e fazer um mapa do trajeto

Ler o arquivo salvo na etapa anterior e atualizar nomes de locais e *status* de carga. Fragmentar o *dataframe* e salvar arquivos diários em formato *csv* e *html* (mapa).

Passo 4: Prepara os dados para apresentação e para o *machine learning*

Ler os arquivos *csvs* diários e criar um único relacionado à placa com as principais medidas de avaliação separado por *status* carregado ou vazio como preparatório para uma lógica de *machine learning* para determinar se o caminhão está carregado ou vazio a partir dos dados de funcionamento.

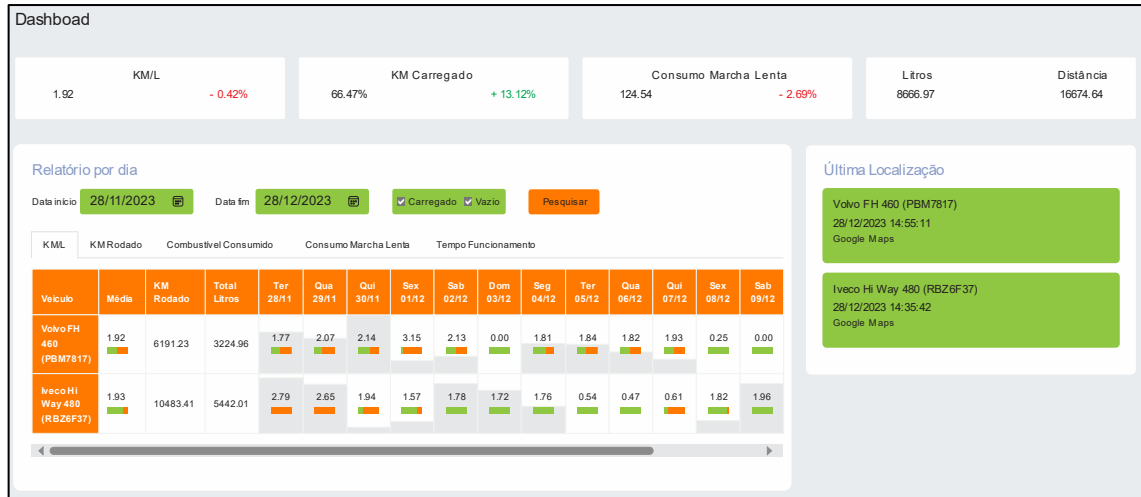
Passo 5: *Dashboard* diário

Ler o arquivo *csv* da etapa anterior, calcular as métricas diárias e categorizar as variáveis em função do *status* de carga, ou seja, cada variável passa a ser calculada na condição de carga do caminhão (carregado, vazio ou global).

4.4.3 Resultados com os dados das bases 2 e 3

Os resultados obtidos nos passos 3 e 5 alimentaram um *website* conforme ilustra a Fig. 4.10:

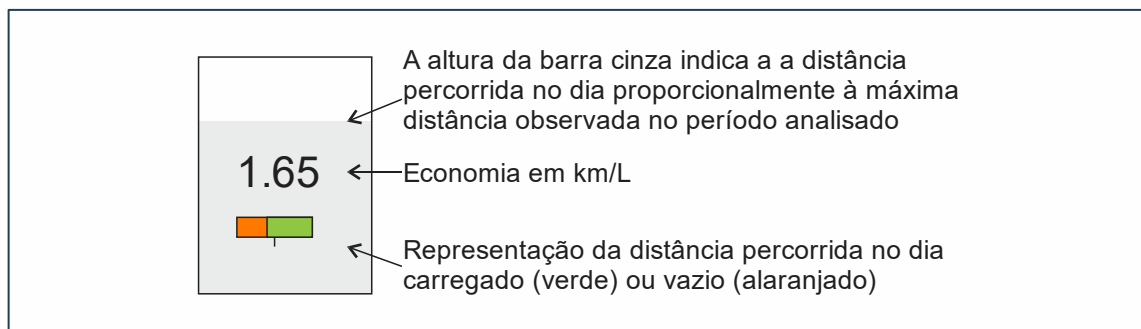
Figura 4.10: *Dashboard* desenvolvido a partir das bases de dados 2 e 3



Elaboração: Autor e Bruno Cortes Siqueira

A Fig. 4.11 é uma legenda do que se apresenta na Fig. 4.10 como resultado de um dia de trabalho de cada caminhão. O valor indica a economia de combustível em km/L conforme a escolha do usuário: somente carregado, somente vazio ou ambos. A barra abaixo do valor mostra a proporção da distância diária percorrida pelo caminhão na condição vazia (parte alaranjada da barra) e na condição carregado (parte verde da barra). A barra cinza ao fundo possui uma altura que representa a relação da distância percorrida no dia com a distância máxima percorrida em um dia por este caminhão dentro do período analisado.

Figura 4.11: Quadro explicativo



Elaboração: Autor

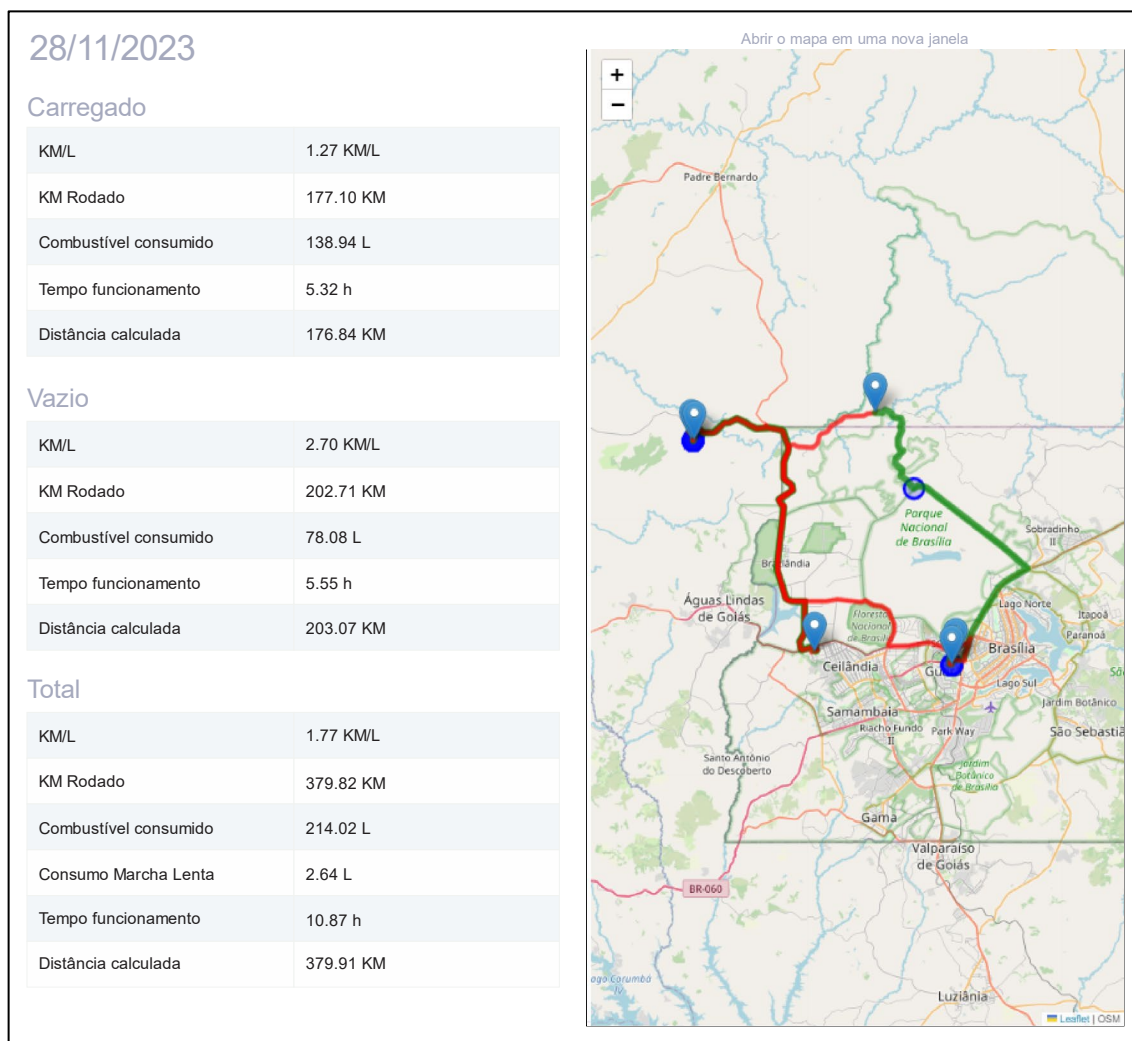
A ideia da barra cinza ao fundo é sinalizar visualmente ao gestor que valores anômalos talvez devam ser desprezados em função da baixa quilometragem percorrida naquele dia, e a barra alaranjada/verde pode mostrar rapidamente que valores baixos de economia de combustível, na maioria das vezes, estão relacionados à distância percorrida na condição de carreta carregada.

O *dashboard* permite ao usuário selecionar o período a ser analisado, e especificar se deseja os resultados globais, condição em que as opções carregado e vazio estão selecionadas, somente carregado ou somente vazio. Os resultados são apresentados em 5 abas:

- a) Economia de combustível em quilometro por litro (km/L);
- b) Distância percorrida em quilômetros (km);
- c) Combustível consumido em litros (L);
- d) Combustível consumido na condição de marcha-lenta em litros (L);
- e) Tempo de funcionamento em horas (h)

O *dashboard* ainda mostra um mapa da trajetória diária com um resumo dos resultados carregado, vazio e global, como mostra a Fig. 4.12.

Figura 4.12: Mapa de um dia de trabalho de um dos caminhões analisados



Elaboração: Autor e Bruno Cortes Siqueira

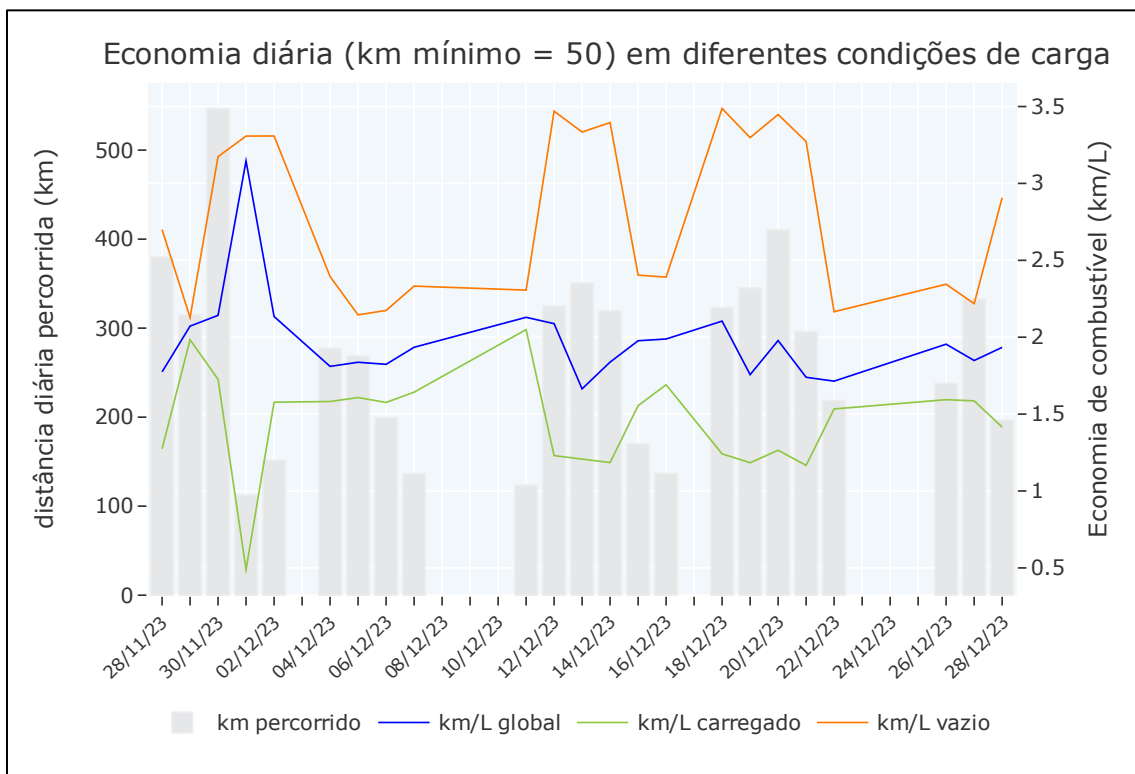
A opção pelo alaranjado para representar a carreta vazia foi substituída pela cor vermelha para dar maior destaque no mapa. O verde se manteve para representar a carreta carregada. Os pontos azuis em forma de gota invertida representam os locais de interesse visitados naquele dia e os círculos azuis indicam os locais onde foram registradas paradas superiores a 600 s (10 minutos).

Os dados de hodômetro são comparados com as estimativas de distância percorrida a partir do deslocamento de uma coordenada geográfica a outra. Quando a diferença é maior do que 10% surge um alerta em forma de exclamação. O objetivo é conferir constantemente o coeficiente "w" do veículo e certificar-se da qualidade dos dados oriundos do hodômetro, e algumas situações em particular: veículo sendo transportado por balsa – fato que ocorreu com o caminhão de transporte de produtos alimentícios quando esteve em Macapá – AP, veículo

sendo rebocado por outro, falha no sistema de aquisição de dados – outra condição que ocorreu, desta vez no caminhão que transporta areia após uma pane elétrica que afetou o tacógrafo. Neste dia o hodômetro registrou apenas 30 km enquanto os cálculos pelo GPS registraram corretamente 65 km.

O gráfico da Fig. 4.13 foi construído com base nos dados diários de um dos caminhões. Nele é possível verificar a distância percorrida no dia (barras) com leitura pelo eixo vertical primário à esquerda do gráfico, e a economia de combustível global (linha azul), carregado (linha verde), vazio (linha alaranjada) com leitura e pelo eixo vertical secundário à direita do gráfico.

Figura 4.13: Economia diária em diferentes condições de carga



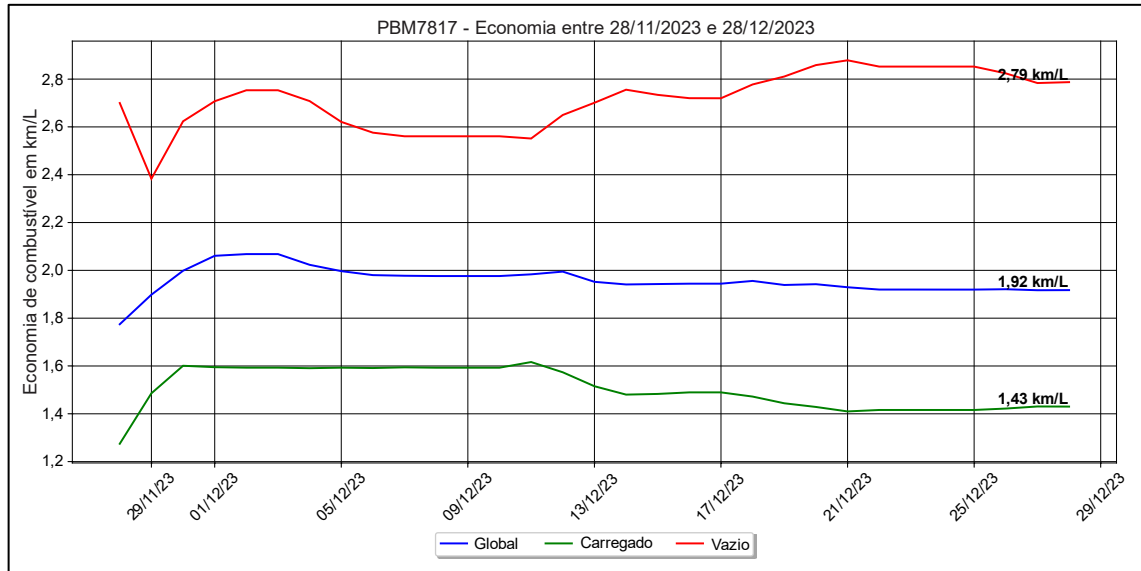
Elaboração: Autor

Distâncias globais percorridas inferiores a 50km foram excluídas para mitigar as distorções provocadas pelas pequenas distâncias. Mesmo assim é possível verificar na Fig. 4.13 que o consumo no dia 01/12/2023 na condição carregado foi muito abaixo do normal. Nesse dia o caminhão rodou 112,5 km e apenas 1,0 km registrado como carregado.

Uma outra possibilidade de análise é apresentada no gráfico da Fig. 4.14. A estratégia foi calcular a economia acumulada após cada dia de trabalho. Dessa forma, o enésimo dia de

trabalho leva em consideração o somatório da distância percorrida em todos os dias anteriores, bem como todo o volume de combustível consumido no mesmo período.

Figura 4.14: Economia de combustível acumulada em um período



Elaboração: Autor

4.4.4 Observações a respeito do trabalho com os dados das bases 2 e 3

Os locais de interesse foram identificados no mapa manualmente por dificuldade de comunicação com o escritório das empresas. O mais produtivo seria existir uma integração que possibilitasse trazer os dados de gerenciamento de transporte, de abastecimento de combustível e de manutenção. Dessa maneira, o sistema poderia alimentar a base de dados de locais de interesse sempre que uma nova origem ou um novo destino surgisse para que fosse questionado sobre as coordenadas geográficas desse novo ponto. A Tab. 4.6 resume o conteúdo dos dois arquivos de locais de interesse gerados.

Tabela 4.6: Locais de interesse identificados

Motivo	Empresa 1	Empresa 2	Observação
Carga / Descarga	38	48	
Abastecimento		66	O caminhão da empresa 1 abastece na própria empresa
Manutenção	2	7	
Fiscalização		1	
Diverso	3	3	Serve a vários motivos
TOTAL	43	125	

Fonte: Autor

O peso transportado, foi conseguido através de comunicação direta por aplicativo de mensagens com o motorista do caminhão, quase na totalidade recebida por áudio. Esse dado poderia facilmente ser obtido caso houvesse uma integração com o ERP da transportadora. A Tab. 4.7 apresenta números extraídos dos arquivos de ‘Status_Carga’.

Tabela 4.7: *Status* de carga dos caminhões

Descrição	Caminhão 1	Caminhão 2
Registros	583	147
Registros Carregado	246	55
Registros de peso transportado	187	23

Fonte: Autor

Com a possibilidade do equipamento estar armazenando dados sem transmiti-los por estar em um ‘ponto cego’ sem conexão de telefonia, adotou-se a estratégia de sempre buscar e trabalhar com os últimos 9 dias de dados Geotab[®]. Isso significa que um único dia teve seus dados recalculados 9 vezes até que tenha ficado fora da faixa de busca de dados no servidor da Geotab[®].

4.4.5 Classificação com técnicas de *machine learning*

Como exercício comparativo, executou-se a mesma rotina de classificação utilizada com os dados da base 1. O resultado é apresentado na Tab. 4.8.

Tabela 4.8: Acurácia (%) da condição de carga com dados dos caminhões 1 e 2 das bases 2 e 3 com diferentes métodos de classificação

Cami nhão	Trecho (km)	Amost.	Ada boost	Árv. Dec.	KNN	Naive Bayes	Random Forest	Regr. Logist.	Rede Neural	SVM
1	0,5 a 5	61	79.04	63.41	79.04	84.38	84.38	84.38	75.98± 3.22	78.90
1	5 a 500	503	99.21	99.21	99.21	100.00	100.00	100.00	99.90± 0.33	98.41
2	0,5 a 5	37	40.00	40.00	40.00	40.00	40.00	40.00	44.00± 8.43	40.00
2	5 a 500	112	82.14	82.14	78.57	85.71	75.00	78.57	80.36± 5.12	82.14

Fonte: Autor

Na condição de trechos curtos para o caminhão 1, os métodos com maior acurácia (%) para classificação da carreta carregada ou vazia foram 3: Naive Bayes, Random Forest e

Regressão Logística com 84,38% com um número de amostras de 61. Já em trechos longos, os mesmos métodos obtiveram uma acurácia de 100% com um número de amostras igual a 503.

Para o caminhão 2 a maior acurácia para trechos curtos foi obtida pela Rede Neural com um valor médio de 44% e um desvio padrão de 8,43% em 37 amostras. Nos trechos longos o índice de maior acerto foi o de Naive Bayes com 85,71% em 112 amostras.

Apenas a técnica de classificação pela Rede Neural apresentou desvio no processo de repetição, similar ao que já havia ocorrido nas experiências com a base 1.

Foram utilizadas 54 *features* relacionadas na Tab. 4.9.

Tabela 4.9: Tabela de *features* utilizadas para classificação com *machine learning*

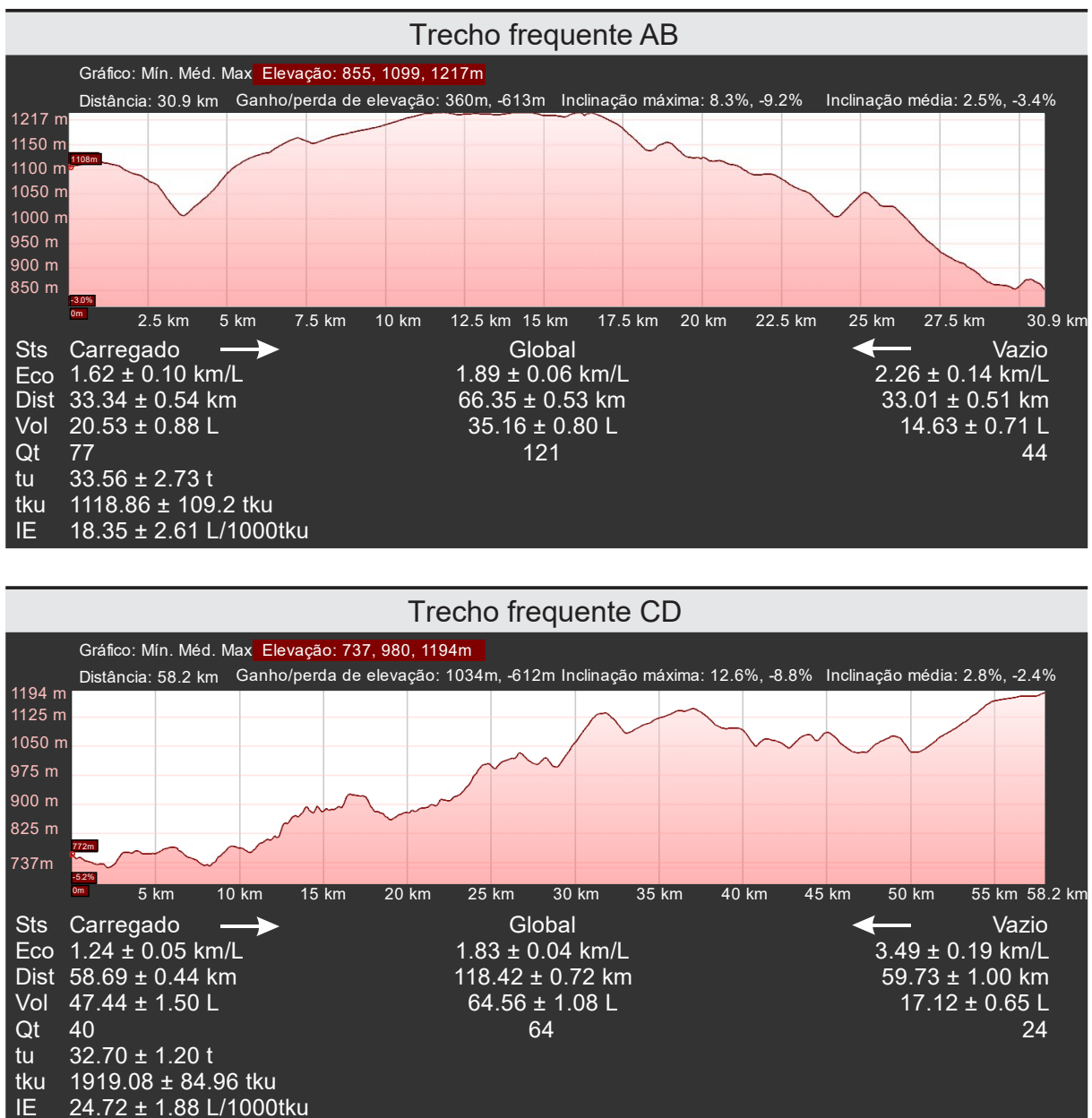
Data	Feature	Data	feature
0	dateTime_BR_duracao	27	kmph_speed_avg_desvio
1	dateTime_BR_parado	28	kmph_velocimetro_min
2	gC_coolant_temperature_min	29	kmph_velocimetro_max
3	gC_coolant_temperature_max	30	kmph_velocimetro_media
4	gC_coolant_temperature_media	31	kmph_velocimetro_desvio
5	gC_coolant_temperature_desvio	32	kmpL_consumo_odometro
6	gC_oil_temperature_min	33	L_fuel_used_inicio
7	gC_oil_temperature_max	34	L_fuel_used_fim
8	gC_oil_temperature_media	35	L_fuel_used_percurso
9	gC_oil_temperature_desvio	36	L_fuel_used_idle_inicio
10	gC_outside_temperature_min	37	L_fuel_used_idle_fim
11	gC_outside_temperature_max	38	L_fuel_used_idle_percurso
12	gC_outside_temperature_media	39	rpm_motor_min
13	gC_outside_temperature_desvio	40	rpm_motor_max
14	km_odometro_inicio	41	rpm_motor_media
15	km_odometro_fim	42	rpm_motor_desvio
16	km_odometro_percurso	43	s_funcionamento_inicio
17	km_gps_inicio	44	s_funcionamento_fim
18	km_gps_fim	45	s_funcionamento_duracao
19	km_gps_percurso	46	V_cranking_min
20	kmph_speed_min	47	V_cranking_max
21	kmph_speed_max	48	V_cranking_media
22	kmph_speed_media	49	V_cranking_desvio
23	kmph_speed_desvio	50	V_device_min
24	kmph_speed_avg_min	51	V_device_max
25	kmph_speed_avg_max	52	V_device_media
26	kmph_speed_avg_media	53	V_device_desvio

Fonte: Autor

4.4.6 Trechos frequentes

A partir da análise dos locais de interesse é fácil perceber que os caminhões realizam trajetos frequentes. Essa análise do trecho frequente permite uma melhor comparação entre caminhões que transportam cargas similares num mesmo trecho, seja este de mesma origem e destino como normalmente ocorre com a transportadora de areia, ou de parte de trechos de rodovias como geralmente ocorre com a transportadora de produtos alimentícios. A Fig. 4.15 traz um resumo de dois trechos frequentes.

Figura 4.15: Perfil de altitude de dois trechos frequentes



Dados do perfil de altitude: Google Earth. Elaboração: Autor

Os dois trechos mostrados na Fig. 4.14 são do mesmo caminhão, dirigido pelo mesmo motorista. O centro mostra o perfil de altitude obtido através do Google Earth. O Google Earth mostra a distância percorrida, os ganhos/perdas de elevação, a inclinação máxima, a inclinação média e as altitudes mínima, máxima e média. Da esquerda para a direita segue o caminhão carregado, e da direita para a esquerda o caminhão vazio.

No trecho AB o caminhão carregado segue um caminho que ganha 360m de elevação e tem boa parte do trecho em descida. No trecho CD o caminhão carregado tem um ganho de elevação de 1034m, sai de um ponto com pouco mais de 737m e chega ao destino a uma altitude de 1194m, é um trecho onde o caminhão carregado está em trecho de subida. No campo oposto, o caminhão vazio mais sobe pelo trecho BA e tem a seu favor uma longa descida no trecho DC.

Os resultados são mostrados abaixo do perfil de altitude em três condições: carregado, global – vai carregado e volta vazio, e vazio. Apresenta nas três condições a economia de combustível em km/L, a distância percorrida em km, o volume de diesel consumido em L, a quantidade de repetições de viagens nas condições especificadas.

Na condição carregado apresenta mais três informações: a tonelada útil transportada, a tonelada-quilometro útil em tku e a intensidade energética (1.7 - definições importantes) em L / 1000tku. A tonelada útil transportada destaca a capacidade real de carga, e reflete diretamente na otimização do transporte. A tonelada-quilômetro útil (tku) oferece uma medida do esforço necessário para transportar uma tonelada de carga por um quilômetro, e é essencial para avaliar a eficácia do deslocamento em termos de energia consumida. Além disso, a intensidade energética em litros por 1000 tku fornece uma perspectiva detalhada sobre o consumo de combustível, e contribui para a compreensão holística da eficiência do sistema do ponto de vista energético. Ao discutir esses indicadores, buscou-se destacar a relevância de cada um na avaliação do desempenho energético do sistema.

Na comparação pela intensidade energética, o transporte de carga no trecho CD consome 34,71% mais energia do que no trecho AB.

5. DISCUSSÃO DOS RESULTADOS

A realização prática que objetivou analisar dados de 10 veículos pesados para extrair características relevantes a serem utilizadas como variáveis de entrada para um sistema de aprendizado de máquina com foco no estudo da economia de energia no transporte rodoviário, mostrou-se muito promissora.

Ao avaliar o consumo de combustível e as distâncias percorridas, pretendeu-se validar a base de dados, tomando como referência dados externos àqueles disponíveis no banco de dados obtidos via telemática, neste caso, do relatório de abastecimento e do hodômetro do caminhão. Foram apresentadas três estimativas para o volume de diesel consumido e duas para a distância percorrida. Entendeu-se que o fato de os valores estimados apresentarem pequenas diferenças quando comparados aos dados externos (abaixo de 5% para o consumo e 1% para a distância percorrida), valida o método aplicado para estimativas de nível percentual de combustível (SPN 96), consumo total de combustível (SPN 250 ou SPN 5054), vazão instantânea de combustível (SPN 183), hodômetro (SPN 917) e velocidade do caminhão (SPN 84). Essa interpretação foi importante para ampliar o entendimento de que toda a base de dados era confiável.

A partir disso, a validação da metodologia de estimar o volume pelo fluxo instantâneo de combustível e o percurso pela velocidade do veículo permitiu o cálculo do consumo de energia em trechos fragmentados, classificados com o auxílio da altitude revelada pelo GPS em aclive, declive e plano.

Ao se observar a Fig. 4.4, referente aos dados da base 1 (transportadora de combustíveis), é possível observar que os caminhões possuem desempenho semelhante no aclive e no plano, e que existe uma grande dispersão no declive, o que pode apontar aos gestores a necessidade de reforçar o treinamento dos motoristas para melhor aproveitamento dos efeitos da gravidade na economia de combustível.

A busca por novas características para estimar o *status* de carga do caminhão revelou-se dispendiosa e, talvez, desnecessária, uma vez que as técnicas de *machine learning* demonstraram ser capazes de determinar essa informação de forma precisa usando apenas as variáveis disponíveis na rede SAE J1939. Os resultados mostraram que, ao avaliar distâncias percorridas entre 5 e 500 km, as técnicas de *machine learning* alcançaram precisões superiores a 93%, considerando apenas as variáveis originais disponíveis na rede SAE J1939. Ao se

utilizarem as variáveis adicionais propostas, as precisões alcançadas ganharam menos de 1 ponto percentual, tornando as características adicionais pouco relevantes para a classificação.

O estudo realizado para avaliar pequenas distâncias revelou a importância de um melhor preparo do conjunto de amostras, buscando aumentá-las quantitativamente. Esse aprimoramento teve um impacto significativo nos resultados, proporcionando uma determinação mais promissora do estado de carga do caminhão. Os índices de acerto foram superiores a 85% para trechos maiores que 1 km e a mais de 86% após 120 segundos de movimento do caminhão.

Esses resultados demonstram o valor de um conjunto de amostras mais robusto e representativo para a análise do *status* de carga, permitindo que as técnicas de *machine learning* obtenham resultados mais precisos e confiáveis. Ao considerar trechos de maior extensão e um intervalo de tempo mais abrangente, o modelo de classificação se torna mais eficiente na identificação das características distintas entre caminhões carregados e vazios

A melhoria na determinação do estado de carga é de suma importância para a eficiência e segurança das operações de transporte rodoviário, pois torna esse processo independente da informação humana para ser obtida.

O segundo grupo de dados, coletados pelos equipamentos Geotab[®], representou alguns desafios adicionais, como por exemplo o fato das duas bases que se conseguiu conexão, um referente aos dados funcionamento e a outra às posições geográficas do equipamento, terem coletas não homogêneas. Uma terceira base de dados foi encontrada na API da Geotab[®], uma que mostra os códigos de falha de funcionamento, entretanto, foi uma descoberta tardia que ainda necessita de estudos, interpretações e conversões.

Com os dados da Geotab[®] e anuência das empresas proprietárias dos caminhões, foi possível construir um *dashboard* com a possibilidade de visualização da trajetória percorrida a cada dia pelos caminhões.

Foi incluído o cálculo da distância com base nas coordenadas de latitude e longitude, a fim de compará-la com a distância indicada pelo hodômetro. Essa análise busca estabelecer um limite de divergência entre uma diferença natural, decorrente dos métodos de cálculo, e um erro causado por possíveis modificações no coeficiente "w" do veículo.

O *dashboard* mostra diariamente o consumo, distância percorrida, volume de combustível consumido, volume de combustível consumido em marcha-lenta e tempo de funcionamento, e permite ao usuário filtrar os resultados na condição carregado, vazio ou global

dentro de um período selecionável. O uso de cores e planos de fundo aumentam a carga de informação em um único quadro sem confundir o leitor e facilita o processo de análise.

Os gráficos de consumo revelam que existe uma considerável diferença no consumo de combustível ao se trafegar transportando uma carreta carregada ou uma vazia.

A integração com sistemas de gerenciamento de carga, manutenção e abastecimento podem potencializar os resultados e novos *insights*, entretanto, caso isso não seja possível é viável a implementação de códigos de *machine learning* para identificação de locais de interesse, trechos frequentes e, possivelmente ter uma previsão de carga útil transportada. Essa afirmação se baseia no fato de que nesse trabalho, os locais de interesse e trechos frequentes foram determinados manualmente sem a informação dos proprietários dos caminhões. As cargas foram obtidas depois de um demorado processo de aproximação com os motoristas dos caminhões e o trabalho de inclusão da predição de carga com regressão teve que ser adiado.

Para orientação orçamentária, apresenta-se a Tab 5.1 com os principais custos deste projeto:

Tabela 5.1: Prestação de contas

Rubrica	Quant.	Valor unit	Total	Participação
Inscrição ISM 2023 (nov/2023) ⁽¹⁾	1	R\$1.712,00	R\$1.712,00	17,68%
Despesas não reembolsáveis ISM 2023			R\$123,32	1,27%
Aquisição de equipamentos Geotab® ⁽²⁾	2	R\$1.024,00	R\$2.048,00	21,15%
Aquisição de chicotes de adaptação ⁽³⁾⁽⁴⁾	2	R\$515,00	R\$1.030,00	10,64%
Transmissão e armazenamento (2 equip. 16 meses) ⁽⁴⁾	32	R\$149,00	R\$4.768,00	49,25%
TOTAL			R\$ 9.681,32	100,00%

(1) Reembolsado pelo Programa de Apoio à Pós-Graduação (PROAP) da CAPES

(2) O equipamento foi adquirido com desconto de 50% - valor unitário normal: R\$2.048,00

(3) O chicote Y de adaptação foi adquirido com 15% de desconto – valor unitário normal: R\$606,00

(4) Instalação própria, então nesse caso sem custos de instalação.

(5) A mensalidade foi negociada com 15% de desconto – mensalidade unitária normal: R\$170,00

Fonte: Autor

Os valores discriminados na Tab. 5.1 são verdadeiros na data de aquisição de cada item. Deslocamentos até as transportadoras e instalação de equipamentos não foram contabilizados. A base de dados (1) da transportadora de combustível foi obtida em cortesia. O desenvolvimento Web não teve custo.

6. CONCLUSÃO

O transporte rodoviário é um dos principais consumidores de diesel, representando mais de 13% da energia total consumida pelos diversos setores da economia. Além disso, é responsável por 40% das emissões de gases do efeito estufa de todos os setores, tornando-se um ponto crucial para pesquisa e desenvolvimento de estratégias mais sustentáveis. Embora haja foco na limitação de emissões em veículos novos, é igualmente importante abordar a eficiência energética dos veículos em circulação. Nesse aspecto, a falta de um procedimento claro para medir a eficiência energética de veículos pesados representa um desafio.

A telemática fornece informações valiosas aos gestores de frota e aos fabricantes para melhorar a eficiência operacional, o monitoramento de saúde do veículo e a tomada de decisões. Através da combinação da telemática com as normas SAE J1939, é possível obter uma visão abrangente e em tempo real do desempenho e condição dos veículos, contribuindo para uma operação mais eficiente e segura da frota.

Embora as normas SAE J1939 ofereçam muitos dados para análise, as empresas de telemática fornecem principalmente informações relevantes para a gestão de frotas, não atendendo necessariamente às demandas necessárias para a pesquisa científica focada em eficiência energética e redução na emissão de poluentes.

A redução do escopo do estudo na análise de consumo de diesel em caminhões pesados, não limita a generalização dos resultados para outros tipos de máquinas e pode contribuir para o entendimento do consumo de energia de forma mais ampla.

Os resultados mostram que é possível identificar o estado de carregamento da carreta transportada com os dados de funcionamento obtidos pela telemática e analisados por técnicas de *machine learning* com alto percentual de acerto - acima de 90% com mais de 10 minutos de movimento do caminhão.

Pelos pontos de tempos de parada foi relativamente fácil identificar os locais de interesse onde ocorrem carga/descarga, abastecimento e manutenção.

A identificação dos trechos frequentes foi uma parte importante porque facilitou o processo de comparação de desempenho entre caminhões, visto que alguns fatores importantes como relevo e qualidade da pavimentação passam a ser irrelevantes na comparação de desempenho em trechos idênticos.

A inclusão de índices como tku (tonelada-quilômetro útil) e L/1000tku (litros de diesel por mil tku) fragmentado por trechos frequentes introduz a carga útil como variável e retira relevo e infraestrutura como variáveis ocultas. É comum entre as transportadoras premiarem os motoristas com melhor resultado de economia (km/L), entretanto esse resultado pode estar contaminado com transporte de cargas mais leves, maiores trechos em vazio ou percursos com menor variação de relevo. Criar índices que proporcione uma comparação de desempenho mais justa entre caminhões e motoristas deve ajudar a criar um ambiente mais saudável e que envolva cada vez mais a participação do motorista no processo de economia de energia no transporte e com isso produzir resultados cada vez mais relevantes.

O *dashboard* criado permite avaliar rapidamente o consumo diário nas condições carregado, vazio ou global. Não apenas o consumo diário, mas outros fatores de interesse como por exemplo a distância percorrida, o tempo de funcionamento, o consumo em marcha lenta e, a comparação de desempenho em trechos frequentes permite, além da comparação de desempenho de caminhões e motoristas, a comparação de custos entre trechos frequentes diferentes.

Durante o desenvolvimento deste trabalho, uma das principais dificuldades encontradas foi a obtenção de informações precisas sobre o estado de carga dos caminhões. Essa questão também foi uma das principais reclamações dos gestores de frota. Mesmo a transportadora que nos forneceu os principais dados de análise, apesar de possuírem tecnologia e anos de experiência no controle de suas operações, enfrenta dificuldades para afirmar em que momento exato um caminhão carregou ou descarregou.

A dependência da informação humana nesse processo pode gerar imprecisões e incertezas, o que impacta diretamente nas decisões estratégicas e operacionais das empresas de transporte. Ao utilizar técnicas de *machine learning* e análise de dados telemáticos, é possível superar essas dificuldades, tornando a determinação do estado de carga mais objetiva e precisa. E, com a integração de novos elementos e abordagens tecnológicas esse trabalho contribuirá para uma gestão mais eficaz da frota, redução de custos operacionais e um menor impacto ambiental.

Este trabalho se destaca dos artigos científicos selecionados na revisão bibliográfica pelos seguintes aspectos:

- Transição metodológica. Enquanto muitos artigos traçam um caminho do laboratório para a realidade, com ambientes controlados, tais como trajetos específicos, tipos de

estradas, velocidades e testes em motores de bancada, como por exemplo: Perrotta *et al.* (2017): “o estudo se centrou apenas em registos de caminhões realizados em autoestradas a uma velocidade constante”. Por outro lado, esse trabalho segue a trajetória que parte da realidade capturada em dados para condições que se vem a conhecer. Essa transição metodológica destaca uma abordagem, na qual a investigação se baseia na análise de dados do mundo real para, posteriormente, identificar, isolar e estudar fatores específicos.

- Isolamento de fatores. Em ambientes controlados, as pesquisas geralmente buscam alterar um ou mais fatores para observar mudanças nos dados. Por exemplo, Rei-Jo Yamashita *et al.* (2018) desenvolveram um modelo de previsão para estimar o consumo de combustível com base no comportamento de direção, no qual os veículos e as rotas são conhecidas. Enquanto esse trabalho adota a abordagem de isolar os fatores à medida que se tornam conhecidos;
- Indicador de desempenho. Embora a carga seja, sem dúvida, um dos principais fatores que afetam o consumo de energia no transporte, o volume de combustível consumido por milhares de toneladas-quilômetro útil não é comumente utilizado como indicador de desempenho pelos gestores de transporte rodoviário, e, nem nos artigos científicos previamente selecionados. Isso difere do transporte ferroviário e das práticas adotadas pelos institutos governamentais de pesquisa energética.

7. TRABALHOS FUTUROS

Para trabalhos futuros sugere-se a continuidade dos trabalhos a fim de atingir o objetivo geral proposto, pois a hipótese inicial de criar um grande banco de dados para avaliar a saúde mecânica do caminhão em aclives íngremes parece ser viável, especialmente considerando os esforços dos fabricantes de caminhões em direção ao desenvolvimento de sistemas avançados de controle de velocidade adaptativo, capaz de prever antecipadamente as marchas a serem utilizadas com base na topografia da estrada.

Uma hipótese inicial foi formulada considerando a viabilidade de avaliação frequente do estado geral da manutenção dos caminhões, especialmente quando trafegam carregados em grandes aclives. A comparação dos dados de um caminhão, operando em condições de carga máxima e torque máximo, com as estatísticas próprias do veículo ou de caminhões semelhantes naquele trecho pode servir como indicativo da 'saúde' do conjunto ou apontar a necessidade de intervenção na manutenção. Para tanto é necessário:

- a) Estimar a carga transportada utilizando técnicas de regressão;
- b) Estabelecer uma base de dados contendo aclives frequentes de interesse.
- c) Analisar o desempenho de caminhões em aclives íngremes.
- d) Identificar problemas no motor, transmissão, suspensão ou pneus com base em comportamentos anormais de consumo de combustível e uso de marchas.
- e) Comparar o desempenho dos caminhões em trechos semelhantes para detectar variações indicativas de problemas em veículos específicos.
- f) Implementar a visualização dos códigos de falha identificados no *dashboard*.

Nesse trabalho, a identificação de locais frequentes foi feita manualmente através de pesquisas a cada ponto novo que surgia no mapa. Um desenvolvimento com *machine learning* deverá potencializar o processo de identificação.

O trabalho de identificação de trechos frequentes no caso do caminhão de transporte de areia foi simples pela característica de repetição de origem e destino. No caso do caminhão de transporte alimentício a realidade é diferente, pois o trecho frequente não está relacionado às mesmas origens e destino e sim à repetição de trajeto em partes de uma rodovia, por exemplo: uma carga de Goiânia – GO com destino a São Luis – MA usará parte de uma mesma rodovia

de uma outra carga que sai da Anápolis – GO com destino a Belém – PA. Essa segunda realidade é a mais comum e, independente da integração de *softwares*, a identificação dos trechos frequentes terá que ser executada por técnicas de *machine learning*.

A inclusão dos trechos frequentes permite a inserção de uma base de dados relacionada ao perfil de altitude do trajeto e à qualidade da infraestrutura, o que provavelmente refinará ainda mais o resultado de uma predição de carga.

Um aspecto importante a ser considerado é a classificação de cargas fracionadas, permitindo uma análise mais detalhada dos diferentes tipos de carga transportada e sua influência no desempenho energético dos caminhões.

Usando as frequências de variação da velocidade como referência, provavelmente será possível estabelecer critérios de classificação para distinguir se um caminhão está operando em vias urbanas ou transitando no pátio para carregamento/descarregamento, trafegando em uma rodovia de pista simples ou numa autoestrada. Ou seja, há a perspectiva de ampliar esse conceito, expandindo as categorias das vias de teste para abranger não somente as classificações convencionais de urbano, rural e expresso, mas também considerar uma gama mais representativa das situações enfrentadas pelos caminhões nas estradas brasileiras.

A busca por informações de manutenção será um passo crucial para compreender como a substituição de filtros, fluidos e peças, juntamente com a limpeza e calibração de injetores, além do processo de substituição, rodízio e monitoramento do desgaste dos pneus, têm um impacto significativo no consumo de combustível.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ANFAVEA. **Tabela de Emissões Diesel Volvo.** (2015). Disponível em <http://anfavea.com.br/Emissoes/diesel/TABELAEMISSOESDIESELVOLVO.pdf>. Citado na página 53.

ARIA, M. AND CUCCURULLO, C. **Bibliometrix: An r-tool for comprehensive science mapping analysis.** Journal of Informetrics, 2017. Citado na página 39.

BRASIL. Ministério de Minas e Energia - MME, Empresa de Pesquisa Energética - EPE, Agência Internacional de Energia (IEA), **Transporte rodoviário de cargas – Brasil 2021 – benchmarking internacional.** Disponível em [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-626/IEA-EPE_Brazilian_Road_Freight_Transport_Benchmarking-2021.09.09%20\[PT\].pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-626/IEA-EPE_Brazilian_Road_Freight_Transport_Benchmarking-2021.09.09%20[PT].pdf). Citado nas páginas 23, 24, 25 e 39.

BRASIL. Ministério de Minas e Energia - MME, Empresa de Pesquisa Energética - EPE, **Balço Energético Nacional (BEN) - Relatório Síntese 2023, Ano base 2022.** Disponível em https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-748/topico-681/BEN_S%C3%ADntese_2023_PT.pdf. Citado nas páginas 21 e 22.

DAF CAMINHÕES. **DAF entrega o seu primeiro caminhão com motorização Euro 6 no Brasil.** Publicado em: 08/12/2022. Disponível em: www.dafcaminhoes.com.br/pt-br/noticias-e-midia/news-articles/br/2022/q4/08-12-2022-daf-entrega-o-seu-primeiro-caminhao-euro-6 - acesso em 12/07/2023. Citado na página 28.

DAF. **DAF Caminhões - Índices de Emissão.** Publicado em: 08 dez. 2022 Disponível em: <https://www.dafcaminhoes.com.br/-/media/files/countries/br/brochures/07-02-2022-indices-de-emissao.pdf>. Acesso em 12 jul. 2023. Citado na página 32.

DAF. **Ficha Técnica XF105 FTT.** Publicado em: 08 ago. 2017. Disponível em: https://www.dafcaminhoes.com.br/-/media/files/countries/br/brochures/daf-xf105/2708-ficha_tecnica_xf105_ftt_s-corte_af.pdf. Acesso em 12 jul. 2023. Citado na página 32.

DAF. **Paccar MX-13 engines.** (2019). Disponível em: https://paccarpowertrain.com/wp-content/uploads/2022/03/MX13_2022Brochure_DigitalLoRes.pdf. Acesso em 12 jul. 2023. Citado na página 32.

DIMITRIOS RIMPAS, ANDREAS PAPADAKIS, MARIA SAMARAKOU. **OBD-II sensor diagnostics for monitoring vehicle operation and consumption**. Energy Reports, Volume 6, Supplement 3, 2020, Pages 55-63, ISSN 2352-4847, <https://doi.org/10.1016/j.egy.2019.10.018>. Citado nas páginas 25 e 39.

EMPRESA DE PESQUISA ENERGÉTICA - EPE, **Atlas da Eficiência Energética Brasil | 2022 – Workbook em português**. Disponível em: https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-741/Atlas_Eficiencia_Energetica_Brasil_2022.xlsb - acesso em 07/07/2023. Citado nas páginas 23, 24, 25 e 30.

EMPRESA DE PESQUISA ENERGÉTICA - EPE, **Matriz ab2021**. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-675/topico-633/Matriz%20ab2021.xlsx> - acesso em 10/07/2023. Citado nas páginas 23, 24 e 25.

EMPRESA DE PESQUISA ENERGÉTICA - EPE, **Matriz ab2022**. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-748/topico-678/Matriz%20ab2022%20site.xlsx> - acesso em 07/07/2023. Citado nas páginas 23, 24 e 25.

ETHEM ALPAYDIN. **Introduction to Machine Learning**. 3th Edition. Publicado por: Massachusetts Institute of Technology. United States of America, 2014. ISBN 978-0-262-02818-9. Citado na página 37.

F. PERROTTA, T. PARRY, L. C. NEVES. **Application of machine learning for fuel consumption modelling of trucks**. 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 2017, pp. 3810-3815, <https://doi.org/10.1109/BigData.2017.8258382>. Citado nas páginas 45 e 85.

GEOTAB. **Como funciona o algoritmo de curva patentado da Geotab**. Publicado em: 14 de julho de 2022. Disponível em: <https://www.geotab.com/pt-br/blog/algoritmo-de-curva/> - acesso em 26/07/2023. Citado na página 66.

GEOTAB. **Geotab**. Disponível em: <https://www.geotab.com/pt-br/> - acesso em 10/01/2024. Citado na página 66.

HANS JAKOB WALNUM, MORTEN SIMONSEN. **Does driving behavior matter? An analysis of fuel consumption data from heavy-duty trucks.** Transportation Research Part D: Transport and Environment, Volume 36, 2015, Pages 107-120, ISSN 1361-9209, <https://doi.org/10.1016/j.trd.2015.02.016>. Citado nas páginas 21, 45 e 56.

IBM, **Documentação Db2 – versão 11.5 Naive Bayes.** Disponível em <https://www.ibm.com/docs/pt-br/db2/11.5?topic=building-naive-bayes> – acesso em 10/07/2023. Citado na página 37.

INSTITUTO NACIONAL DE METROLOGIA, QUALIDADE E TECNONOLOGIA - INMETRO. **Norma N° NIE-DIMEL-100 - Verificação Subsequente de Cronotacógrafos.** Aprovada em set/2011. Disponível em <https://cronotacografo.rbmlq.gov.br/legislacoes/downloadArquivo/MjcwMzUxMQ%3D%3D>. Citada na página 31.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **Search for 11898.** Disponível em: https://www.iso.org/search.html?q=11898&hPP=10&idx=all_en&p=1. Acesso em 28/05/2023. Citado na página 34

J. PAVLOVIC, G. FONTARAS, S. BROEKAERT, B. CIUFFO, M.A. KTISTAKIS, T. GRIGORATOS. **How accurately can we measure vehicle fuel consumption in real world operation?** Transportation Research Part D: Transport and Environment, Volume 90, 2021, 102666, ISSN 1361-9209, <https://doi.org/10.1016/j.trd.2020.102666>. Citado na página 45.

LAURETTO, MARCELO S. **Árvores de Decisão.** EACH-USP 2010. Disponível em https://edisciplinas.usp.br/pluginfile.php/4469825/mod_resource/content/1/ArvoresDecisao_normalsize.pdf. Acesso em 15/12/2023. Citado na página 38.

MERCEDES-BENZ. **IAA Transportation 2022: Daimler Truck revela caminhão eActros LongHaul elétrico à bateria e expande portfólio de mobilidade elétrica.** Publicado em 19 de setembro de 2022. Disponível <https://www.bancomercedes-benz.com.br/sala-de-imprensa/noticias/caminhoes/iaa-transportation-2022-daimler-truck-revela-caminhao-eactros-longhaul-eletrico-a-bateria-e-expande-portfolio-de-mobilidade-eletrica> - acesso em 25/07/2023. Citado na página 28.

M. MNYAKIN. **Applications of AI, IoT, and Cloud Computing in Smart Transportation: A Review,** Artificial Intelligence in Society, vol. 3, no. 1, pp. 9–27, Feb. 2023, [Online]. Available: <https://researchberg.com/index.php/ai/article/view/108>. Citado na página 21

M. EAGON, S. FAKHIMI, A. PERNSTEINER E W. F. NORTHROP. **Mass Detection for Heavy-Duty Vehicles using Gaussian Belief Propagation** 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Alemanha, 2022, pp. 1655-1661, doi: 10.1109/IV51971.2022.9827370. Citado nas páginas 46, 50 e 67.

PEDREGOSA, F. - VAROQUAUX, G. - GRAMFORT, A. - MICHEL, V. - THIRION, B. - GRISEL, O. - BLONDEL, M. - PRETTENHOFER, P. - WEISS, R. - DUBOURG, V. - VANDERPLAS, J. - PASSOS, A. - COURNAPEAU, D. - BRUCHER, M. - PERROT, M. - DUCHESNAY, E. **Scikit-learn: Machine Learning in Python**. Journal of Machine Learning Research, volume 12, páginas 2825-2830, ano 2011. Citado na página 60.

PEREIRA, V. **Project pybibx**. 2022, <https://github.com/Valdecy/pyBibX>. Citado na página 41.

REI-JO YAMASHITA, HSIU-HSEN YAO, SHIH-WEI HUNG, ACQUAH HACKMAN, **Accessing and constructing driving data to develop fuel consumption forecast model**, IOP Conference Series: Earth and Environmental Science, volume 113, number 1, Feb/2018, <https://dx.doi.org/10.1088/1755-1315/113/1/012217>. Citado na página 85.

SAE INTERNATIONAL. **Data Link Layer J1939/21_199407**. Disponível em https://www.sae.org/standards/content/j1939/21_202205/ - acesso em 28/05/2023. Citado na página 35.

SAE INTERNATIONAL. **J1939 Digital Annex J1939DA_202301**. Disponível em: https://www.sae.org/standards/content/j1939da_202301/ - acesso em 28/05/2023. Citado na página 36.

SAE INTERNATIONAL. **Items included in the J1939 Standards Collection**. Disponível em: <https://www.sae.org/standards/development/ground-vehicle/sae-j1939-standards-collection-on-the-web> - acesso em 28/05/2023. Citado na página 35.

SAE INTERNATIONAL. **SAE J1939 Standards Collection**. Disponível em: https://www.sae.org/publications/collections/content/j1939_dl/ - acesso em 28/05/2023. Citado na página 35.

SAE INTERNATIONAL. **Surface Vehicle Recommended Practice**. J1939_71, Mar/2011. Citado na página 35.

SAE INTERNATIONAL. **Vehicle Application Layer J1939/71_202208**. Disponível em: https://www.sae.org/standards/content/j1939/71_202208/ - acesso em 28/05/2023. Citado na página 35.

SCANIA. **Scania X-gas: o primeiro caminhão a gás com 900 km de autonomia.** Publicado em: 12/07/2023. Disponível em: <https://www.scania.com/br/pt/home/about-scania/newsroom/news/2021/news-article-template-simple26.html> - acesso em 23/07/2023. Citado na página 29.

SCIKIT-LEARN. Disponível em <https://scikit-learn.org/> . Acesso em 25/07/2023. Citado na página 39.

SIMON HU, SIQI SHU, JUSTIN BISHOP, XIAOXIANG NA, MARC STETTLER. **Vehicle telematics data for urban freight environmental impact analysis.** Transportation Research Part D: Transport and Environment Volume 102, January 2022 - <https://doi.org/10.1016/j.trd.2021.103121>. Citado na página 33.

VECTOR. **SAE J1939 – Know-how.** Disponível em: <https://www.vector.com/br/pt/know-how/protocols/sae-j1939/#> - acesso em 28/05/2023. Citado nas páginas 34 e 35.

VOLVO GROUP. **25 anos inovando sem parar.** Publicado em: 23/11/2018. Disponível em: <https://www.volvogroup.com/br/news-and-media/news/2018/nov/25-anos-inovando-sem-parar.html> - acesso em 23/07/2023. Citado na página 25.

VOLVO GROUP. **I-See pré-mapeado da Volvo Caminhões.** Disponível em: <https://www.volvotrucks.com.br/pt-br/trucks/features/i-see.html> - acesso em 12/07/2023. Citado na página 29.

VOLVO GROUP. **Novo pacote aerodinâmico deixa Volvo FH ainda mais econômico.** Publicado em: 27/04/2022. Disponível em: www.volvogroup.com/br/news-and-media/news/2022/apr/novo-pacote-aerodinamico-deixa-volvo-fh-ainda-mais-economico.html - acesso em 12/07/2023. Citado na página 29.

VOLVO GROUP. **Novo Volvo FH Euro 6 concilia redução de emissões e alto desempenho com mais economia.** Publicado em: 10/05/2022. Disponível em: <https://www.volvogroup.com/br/news-and-media/news/2022/oct/novo-volvo-fh-euro-6-concilia-reducao-de-emissoes-e-alto-desempenho-com-mais-economia.html> - acesso em 23/07/2023. Citado na página 29.

VOLVO GROUP. **Volvo começa a testar caminhões a hidrogênio em estradas na Europa.** Publicado em: 08/05/2023. Disponível em: www.volvogroup.com/br/news-and-media/news/2023/may/Volvo-comeca-a-testar-caminhoes-a-hidrogenio-em-estradas-na-Europa.html - acesso em 12/07/2023. Citado na página 29.

VOLVO TRUCK CORPORATION. **Folha de Especificações FH Motor D13C540, EU5SCR-M. POR – Version 02”** (2010). Disponível em: https://stpi.it.volvo.com/STPIFiles/Volvo/FactSheet/D13C540,%20EU5SCR-M_Por_02_1144750.pdf – acesso em 12/07/2023. Citado nas páginas 25 e 53.

VOLVO TRUCK CORPORATION. **Folha de Especificações FH Motor Motor D13K540, EU6SCR. POR – Version 07”** (2019). Disponível em: https://stpi.it.volvo.com/STPIFiles/Volvo/FactSheet/D13K540,%20EU6SCR_Por_07_310999624.pdf – acesso em 12/07/2023. Citado na página 25.

W. NICK STREET, W. H. WOLBERG, O. L. MANGASARIAN. **Nuclear feature extraction for breast tumor diagnosis.** Proc. SPIE 1905, Biomedical Image Processing and Biomedical Visualization, (29 July 1993); <https://doi.org/10.1117/12.148698>. Citado na página 52.

WARE, S. R, JAIN, M. M., BHARATI S. S. **GPS Data Logger and Data Visualizer.** International Research Journal of Engineering and Technology (IRJET), Volume: 09, Issue: 07, July 2022, e-ISSN: 2395-0056. Citado na página 34.

ZHANG, T., CHEN, J., LI, F., ZHANG, K., LV, H., HE, S. AND XU, E. **Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions.** ISA Transactions, Vol. 119, pp. 152–171. Citado nas páginas 21 e 36.

9. APÊNDICE A

Principais fornecedores de serviços de telemática no Brasil em ordem alfabética. Sites visitados em 19 de julho de 2023

(01) Autotrac

<https://www.autotrac.com.br/software/supervisor-telemetria/>

“Permite o controle eletrônico das condições de uso do veículo, colaborando para a redução dos custos de operação e manutenção, ajudando a prevenir acidentes. O Supervisor Telemetria emite relatórios sobre velocidade, RPM, freadas e acelerações bruscas, e outros importantes indicadores de performance da frota.”

(02) Bosch -

<https://www.bosch.com.br/produtos-e-servicos/mobilidade/telemetria-bosch-para-caminhoes/>

“Desenvolvido por especialistas da Bosch, o serviço de telemetria permite ao gestor de frota otimizar a operação e reduzir custos, através da apresentação de dados de maneira simples e intuitiva.

Utilizando a rede móvel de celulares e uma nuvem de dados, as informações são disponibilizadas praticamente em tempo real, aumentando a transparência e gerando recomendações precisas para a tomada de decisão.”

(03) Cobli

<https://www.cobli.co/telemetria-veicular>

“Com a telemetria de frota você consegue saber informações importantes sobre como os veículos estão sendo conduzidos, como acelerações e frenagens bruscas, curvas perigosas e velocidades excedidas. Com essas informações à mão, você terá maior certeza de quais decisões tomar no dia a dia e terá mais facilidade para fazer o gerenciamento de transportes da sua empresa.”

(04) DAF

<https://www.dafcaminhoes.com.br/pt-br/a-daf/sustainability/intelligent-logistics/daf-big-data>

“DAF Connect é um sistema de gerenciamento de frota online que monitora, compara e otimiza o desempenho de sua frota e motoristas em tempo real. Para máxima eficiência, os menores custos de combustível possíveis e emissões de CO2 muito favoráveis.

O progresso está sendo feito em um ritmo rápido. Com o DAF Connect já foi possível planejar as melhores rotas e monitorar o consumo de combustível de caminhões e motoristas individuais. E recentemente também se tornou possível ler a memória de cartões de motoristas e tacômetros online com o DAF Connect, o que economiza muito tempo e esforço.”

(05) Delta Global

<https://www.deltaglobal.com.br/fleet>

“Tenha informações precisas para fazer análises de desempenho da sua frota, resultando em economia e otimização da sua operação.”

(06) Geotab

<https://www.geotab.com/pt-br/>

“Gerencie, monitore e potencialize a sua frota. Soluções telemáticas baseadas em dados.

Mais de 55 bilhões de pontos de dados processados diariamente que geram insights valiosos para a gestão da sua frota. Siga com inteligência.”

(07) Gobrax

<https://gobrax.com.br/>

“Entregamos resultados financeiros para o transportador através da gamificação da gestão do motorista.

Conectividade oficial da DAF no Brasil”

(08) Intelbrás

<https://www.intelbras.com/pt-br/seguranca-eletronica/monitoramento-veicular>

“São equipamentos instalados nos mais variados tipos de veículos, cuja função está no monitoramento e armazenamento de gravações, alarmes e outras informações importantes sobre o veículo e seus passageiros.”

(09) Iveco

<https://www.iveco.com/brasil/servicos-e-solucoes/Servicos-Mais-productividade-e-menos-custo-operacional-para-a-frota>

“TELEMETRIA

Serviços conectados e inteligentes para te ajudar na gestão da sua frota e maximizar a rentabilidade do seu negócio.”

(10) Mercedes-Benz

<https://www.mercedes-benz-trucks.com.br/caminhoes/servicos-e-pecas/fleetboard/>

“Produto da Mercedes-Benz, o FleetBoard é uma ferramenta que auxilia o cliente na gestão de sua frota. Entre diversas vantagens, está a otimização dos veículos e sua disponibilidade, além de possibilitar uma redução nos custos operacionais em até 15%. Além disso, o novo módulo FleetBoard, também é um rastreador e bloqueador inteligente de fábrica. Além dos serviços de telemetria, também oferece serviços de auxílio à segurança do veículo e da carga.”

(11) Omnilink

<https://www.omnilink.com.br/omnitelemetria/>

“Os produtos da família Omnitelemetria permitem acesso a diversas informações fundamentais para o gerenciamento completo da sua frota com uma análise detalhada e inteligente do desempenho dos motoristas, pautadas pela segurança e economia financeira na sua operação.

O sistema preciso e detalhado de análise das soluções Omnitelemetria identifica o comportamento do motorista (direção agressiva e o mau uso do veículo) para mapeamento de locais de risco e alertas em tempo real.”

(12) SasCar

<https://www.sascar.com.br/telemetria-can>

“Corrija hábitos ruins que afetam diretamente o seu resultado!

Com a solução de Telemetria CAN Sascar você identifica os comportamentos ao volante que afetam os custos de sua operação ou que podem gerar riscos de acidente.”

(13) Scania

<https://www.scania.com/br/pt/home/services/data-driven-services.html>

“Em cada Scania que fabricamos, você encontrará tecnologia inteligente, sensores avançados e conectividade sem fio. Isso significa que temos centenas de milhares de veículos e motores constantemente conectados sendo usados em todo o mundo atualmente, fornecendo dados que

não apenas melhoram nossa engenharia, mas também nos permitem criar serviços com os quais você obtém valor comercial direto.

Desde a redução do consumo de combustível até a redução do desgaste e das necessidades de manutenção, até a automação do trabalho administrativo tedioso. Não importa se você opera uma frota mista ou uma frota somente de veículos Scania. Em resumo, nós nos concentramos nos dados, para que você possa se concentrar na gestão de seus negócios.”

(14) Texa

<https://www.texabrasil.com.br/products/etruck/>

“A evolução do diagnóstico remoto para caminhões pesados, veículos agrícolas e de construção”

(15) Trimble

<https://tl.trimble.com/solucoes/telemetria/>

“Embora 70% das frotas utilizem telemetria, 24% dos gestores não buscam melhorias constantes em tecnologia ou estão atrasados em relação aos concorrentes.

A falta de indicadores impossibilita uma visão correta do desempenho da frota. Empresas crescem menos e perdem espaço no mercado quando os indicadores de telemetria não entregam as informações que realmente precisam.”

(16) VDO

<https://www.fleet.vdo.com.br/solucoes/vdo-fts-fleet-tracking-service/>

“O VDO Fleet Tracking Service foi desenvolvido para proporcionar diversas soluções, pensando nas principais dificuldades encontradas no dia a dia de sua operação. O cliente terá acesso a um serviço completo de gestão de frota, com todas as informações disponíveis em tempo real, com a tranquilidade de ter uma plataforma criada por uma empresa pioneira em desenvolvimento de tecnologia automotiva.”

(17) Volkswagen

<https://www.vwco.com.br/servicos/Solu%C3%A7%C3%B5es%20de%20Servi%C3%A7o%20ao%20Cliente/Plataforma%20RIO?id=15&productid=113>

“Assim como um rio, que flui e muda, cada vez que encontra um afluente, estamos constantemente em movimento e trazendo novas soluções para os desafios diários do setor de transporte e logística. Acompanhe sua frota de onde estiver, com pacotes pensados sob medida

para a sua operação. Monitore localização, consumo, desempenho, perfil de condução e muito mais com as soluções RIO.”

(18) Volvo

<https://www.volvotrucks.com.br/pt-br/services/optimized-business/conducao-inteligente-volvo.html>

“Condução Inteligente Volvo

É o novo serviço de gestão de combustível do Volvo Connect. Compare o desempenho da sua frota com outros veículos que trabalham em operações similares e veja o quanto você pode economizar no mês. Condução Inteligente Volvo (CIV) é mais inteligência para economizar combustível.”

(19) - WebTrac

<https://www.webtrac.com.br/>

“São mais de 14 anos de experiência em telemetria e rastreamento, aplicados para lhe ajudar a tomar as decisões mais importantes do seu negócio.

Para isso, adaptamos nossas soluções à sua realidade de mercado, fornecendo informações precisas sobre o dia a dia da sua operação. E você acompanha tudo que está acontecendo, de onde estiver.”

(20) ZF

https://press.zf.com/press/pt/releases/release_46336.html

“As soluções de conectividade estão presentes nos desenvolvimentos voltados aos fabricantes de implementos. O Trailer Pulse é um dos destaques e, em conjunto com o iEBS, oferece serviços completos de telemetria da carreta, valendo-se, também, das informações geradas pelo sistema de controle de estabilidade. Trata-se de um sistema aberto, pronto para ser utilizado em todo o Brasil e região América do Sul. Considerando a dinâmica do transporte rodoviário de carga no Brasil, monitorar unitariamente cada implemento (geralmente rebocado por diferentes veículos) é fundamental para frotistas que buscam maior rentabilidade nas operações logísticas.”

10. APÊNDICE B

Lista das variáveis identificadas na base fornecida pela transportadora

Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
1	binTimestamp			
2	cl_total_consumo_arla			
3	cl_total_consumo_combustivel_hr	5054	High Resolution Engine Total Fuel Used	64777 / 5-8
4	cl_total_consumo_combustivel_lr	250	Engine Total Fuel Used	65257 / 5-8
5	cl_total_consumo_marcha_lenta	236	Engine Total Idle Fuel Used	65244 / 1-4
6	cl_total_consumo_rotacao_nao_economico			
7	clph_vazao_instantanea_combustivel	183	Engine Fuel Rate	65266 / 1-2
8	data_vdo			
9	data_veiculo			
10	estado_abs_ativo	563	Anti-Lock Braking (ABS) Active	61441 / 1.5
11	estado_abs_carreta	1836	Trailer ABS Status	61441 / 8.5
12	estado_abs_lampada_ambar	1438	ABS/EBS Amber Warning Signal (Powered Vehicle)	61441 / 6.5
13	estado_abs_offhoad	575	ABS Off-road Switch	61441 / 3.1
14	estado_abs_totalmente_operacional	1243	ABS Fully Operational	61441 / 6.1
15	estado_accelerator_pedal_kickdown	559	Accelerator Pedal Kickdown Switch	61443 / 1.3
16	estado_alerta_vermelho_ebs	1439	EBS Red Warning Signal	61441 / 6.3
17	estado_ar_condicionado_ativo			
18	estado_asr_ativo	562	ASR Brake Control Active	61441 / 1.3
19	estado_asr_motor_ativo	561	ASR Engine Control Active	61441 / 1.1
20	estado_atc_asr_info_signal	1793	ATC/ASR Information Signal	61441 / 6.7
21	estado_botao_accelerator_remoto_ativo	969	Remote Accelerator Enable Switch	61441 / 4.7
22	estado_botao_asr_hill_holder	577	ASR "Hill Holder"	61441 /

Var	Descrição Transportadora	SPN	Switch	3.5
			Descrição SAE J1939	PGN / Start Position
23	estado_botao_asr_offroad	576	ASR Off-road Switch	61441 / 3.3
24	estado_botao_auxiliar_desligamento_motor	970	Engine Auxiliary Shutdown Switch	61441 / 4.5
25	estado_botao_derate_motor	971	Engine Derate Switch	61441 / 4.3
26	estado_botao_ebs	1121	EBS Brake Switch	61441 / 1.7
27	estado_botao_sobrescrita_desligando_motor	1237	Engine Shutdown Override Switch	65265 / 8.7
28	estado_cartao_motorista_1	1615	Driver card, driver 1	65132 / 2.5
29	estado_cartao_motorista_2	1616	Driver card, driver 2	65132 / 3.5
30	estado_chave_ignicao			
31	estado_cinto_seguranca_motorista	1856	Seat Belt Switch	57344 / 4.7
32	estado_desligamento_tempo_ocioso_ativo			
33	estado_desligamento_tempo_ocioso_desligado			
34	estado_evento_tacografo	1622	System event	65132 / 4.1
35	estado_farol	2350	Low Beam Head Light Data	65088 / 1.5
36	estado_farol_luz_alta	2348	High Beam Head Light Data	65088 / 1.7
37	estado_farol_neblina_dianteiro	2388	Tractor Front Fog Lights	65088 / 2.1
38	estado_farol_neblina_traseiro	2390	Rear Fog Lights	65088 / 5.1
39	estado_faroleta_ligado			
40	estado_faroleta_selecionado			
41	estado_freio_carreta			
42	estado_freio_estacionamento	70	Parking Brake Switch	65265 / 1.3
43	estado_indicador_agua_combustivel	97	Water In Fuel Indicator	65279 / 1.1
44	estado_lampada_atencao_abs_carreta	1792	Tractor-Mounted Trailer ABS Warning Signal	61441 / 8.7
45	estado_limpador_parabrisa_ajuste_velocidade			
46	estado_limpador_parabrisa_ativo	2866	Front Operator Washer Switch	64973 / 6.6
47	estado_limpador_parabrisa_esguicho_agua			

48	estado_luzes_internas			
Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
49	estado_modos_economia_potencia			
50	estado_modos_partida_motor	1675	Engine Starter Mode	61444 / 7.1
51	estado_modos_retarder_motor	900	Retarder Torque Mode	61440 / 1.1
52	estado_modos_selecao_freio_motor			
53	estado_modos_torque	899	Engine Torque Mode	61444 / 1.1
54	estado_pedal_embreagem	598	Clutch Switch	65265 / 4.7
55	estado_pedal_freio_acionado	597	Brake Switch	65265 / 4.5
56	estado_performance_tacografo	1620	Tachograph performance	65132 / 4.5
57	estado_piloto_automatizado	527	Cruise Control States	65265 / 7.6
58	estado_piloto_automatizado_ativo	595	Cruise Control Active	65265 / 4.1
59	estado_porta_motorista_aberta	3413	Open Status of Door 1	64933 / 1.3
60	estado_porta_motorista_travada	3415	Lock Status of Door 2	64933 / 1.7
61	estado_porta_passageiro_aberta	3416	Open Status of Door 2	64933 / 2.1
62	estado_porta_passageiro_travada	3418	Lock Status of Door 3	64933 / 2.5
63	estado_pto	976	PTO Governor State	65265 / 7.1
64	estado_setas_direita_acionadas	2370	Right Turn Signal Lights	65088 / 2.5
65	estado_setas_direita_acionadas_travadas			
66	estado_setas_esquerda_acionadas	2368	Left Turn Signal Lights	65088 / 2.7
67	estado_setas_esquerda_acionadas_travadas			
68	estado_sobrescrita_controle_tracao	1238	Traction Control Override Switch	61441 / 3.7
69	estado_suporte_diagnostico	2804	FMS-standard Diagnostics Supported	64977 / 1.1
70	estado_suporte_request	2805	FMS-standard Requests Supported	64977 / 1.3
71	estado_tacografo_processando_informacao	1621	Handling information	65132 / 4.3
72	estado_trabalho_motorista_1	1612	Driver 1 working	65132 /

Var	Descrição Transportadora	SPN	state Descrição SAE J1939	1.1 PGN / Start Position
73	estado_trabalho_motorista_2	1613	Driver 2 working state	65132 / 1.4
74	f_razao_marcha_selecionada	526	Transmission Actual Gear Ratio	61445 / 2-3
75	gps_deg_course			
76	gps_deg_latitude	584	Latitude	65267 / 1-4
77	gps_deg_longitude	585	Longitude	65267 / 5-8
78	gps_dop_h			
79	gps_dop_p			
80	gps_dop_v			
81	gps_estado_fix			
82	gps_estado_fix_mode			
83	gps_m_altitude	580	Altitude	65256 / 7-8
84	gps_mag_variation			
85	gps_mps_speed			
86	gps_num_sats_in_use			
87	gps_num_sats_in_view			
88	gps_timestamp			
89	hardware_version			
90	idPlaca			
91	kmph_velocidade_def_piloto_automatico	86	Cruise Control Set Speed	65265 / 6
92	kmph_velocidade_primeiro_eixo	904	Front Axle Speed	65215 / 1-2
93	kmph_velocidade_relativa_eixo1_roda_direita	906	Relative Speed; Front Axle, Right Wheel	65215 / 4
94	kmph_velocidade_relativa_eixo1_roda_esquerda	905	Relative Speed; Front Axle, Left Wheel	65215 / 3
95	kmph_velocidade_relativa_eixo2_roda_direita	908	Relative Speed; Rear Axle #1, Right Wheel	65215 / 6
96	kmph_velocidade_relativa_eixo2_roda_esquerda	907	Relative Speed; Rear Axle #1, Left Wheel	65215 / 5
97	kmph_velocidade_relativa_eixo3_roda_direita	910	Relative Speed; Rear Axle #2, Right Wheel	65215 / 8
98	kmph_velocidade_relativa_eixo3_roda_esquerda	909	Relative Speed; Rear	65215 /

			Axle #2, Left Wheel	7
Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
99	kmph_velocidade_tacogrado	1624	Tachograph vehicle speed	65132 / 7-8
100	kmph_velocidade_veiculo	84	Wheel-Based Vehicle Speed	65265 / 2-3
101	kmpl_consumo_instantaneo_combustivel	184	Engine Instantaneous Fuel Economy	65266 / 3-4
102	kmpl_economia_media_combustivel	185	Engine Average Fuel Economy	65266 / 5-6
103	kpa_pressao_ar_inlet	106	Engine Air Intake Pressure	65270 / 4
104	kpa_pressao_barometrica	108	Barometric Pressure	65269 / 1
105	kpa_pressao_boost	1127	Engine Turbocharger 1 Boost Pressure	65190 / 1-2
106	kpa_pressao_carter	22	Engine Extended Crankcase Blow-by Pressure	65263 / 2
107	kpa_pressao_carter_estendido	101	Engine Crankcase Pressure	65263 / 5-6
108	kpa_pressao_combustivel	94	Engine Fuel Delivery Pressure	65263 / 1
109	kpa_pressao_freio_1	1087	Service Brake Circuit 1 Air Pressure	65198 / 3
110	kpa_pressao_freio_2	1088	Service Brake Circuit 2 Air Pressure	65198 / 4
111	kpa_pressao_lig_arrefecimento	109	Engine Coolant Pressure	65263 / 7
112	kpa_pressao_oleo_motor	100	Engine Oil Pressure	65263 / 4
113	m_hodometro_vdo			
114	m_hodometro_veiculo	917	High Resolution Total Vehicle Distance	65217 / 1-4
115	msUpTime			
116	num_marcha_atual	523	Transmission Current Gear	61445 / 4
117	num_marcha_selecionada	524	Transmission Selected Gear	61445 / 1
118	perc_abertura_termostato_intercooler_motor	1134	Engine Intercooler Thermostat Opening	65262 / 8

119	perc_carga_motor	92	Engine Percent Load At Current Speed	61443 / 3
Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
120	perc_demanda_torque_motor	2432	Engine Demand – Percent Torque	61444 / 8
121	perc_demanda_torque_motorista	512	Driver's Demand Engine - Percent Torque	61444 / 2
122	perc_desgaste_disco_transmissao			
123	perc_nivel_arla	1761	Aftertreatment 1 Diesel Exhaust Fluid Tank 1 Level	65110 / 1
124	perc_nivel_combustivel			
125	perc_nivel_combustivel_tanque_primario	96	Fuel Level 1	65276 / 2
126	perc_nivel_combustivel_tanque_secundario	38	Fuel Level 2	65276 / 7
127	perc_nivel_lig_arrefecimento	111	Engine Coolant Level	65263 / 8
128	perc_nivel_oleo_motor	98	Engine Oil Level	65263 / 3
129	perc_posicao_acelerador	91	Accelerator Pedal Position 1	61443 / 2
130	perc_posicao_acelerador_remoto	974	Remote Accelerator Pedal Position	61443 / 4
131	perc_posicao_pedal_freio	521	Brake Pedal Position	61441 / 2
132	perc_retarder_motor	520	Actual Retarder - Percent Torque	61440 / 2
133	perc_selecao_retarder_motor	973	Engine Retarder Selection	61441 / 5
134	perc_selecao_retarder_nao_motor	1716	Retarder Selection, non-engine	61440 / 7
135	perc_torque_atual	513	Actual Engine - Percent Torque	61444 / 3
136	primary_errors			
137	primary_firmware_version			
138	rpm_rotacao_marcha_lenta	188	Engine Speed At Idle, Point 1 (Engine Configuration)	65251 / 1-2
139	rpm_rotacao_motor	190	Engine Speed	61444 / 4-5
140	rpm_total_rpm			
141	rpm_velocidade_rotacao_eixo_tacografo	1623	Tachograph output shaft speed	65132 / 5-6

142	secondary_errors			
143	secondary_firmware_version			
Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
144	seg_tempo_motorista_1	1617	Driver 1 Time Related States	65132 / 2.1
145	seg_tempo_motorista_2	1618	Driver 2 Time Related States	65132 / 3.1
146	seg_total_horas_marcha_lenta	235	Engine Total Idle Hours	65244 / 5-8
147	seg_total_horas_motor	247	Engine Total Hours of Operation	65253 / 1-4
148	seg_total_horas_rotacao_nao_economica			
149	seg_total_horas_velocidade_controlada			
150	sys_serial			
151	tempc_temperatura_ambiente	171	Ambient Air Temperature	65269 / 4-5
152	tempc_temperatura_combustivel	174	Engine Fuel Temperature 1	65262 / 2
153	tempc_temperatura_inlet	172	Engine Air Intake Temperature	65269 / 6
154	tempc_temperatura_intake	105	Engine Intake Manifold 1 Temperature	65270 / 3
155	tempc_temperatura_intercooler_motor	52	Engine Intercooler Temperature	65262 / 7
156	tempc_temperatura_interior_cabine	170	Cab Interior Temperature	65269 / 2-3
157	tempc_temperatura_lig_arrefecimento	110	Engine Coolant Temperature	65262 / 1
158	tempc_temperatura_oleo_motor	175	Engine Oil Temperature 1	65262 / 3-4
159	tempc_temperatura_oleo_transmissao	177	Transmission Oil Temperature	65272 / 5-6
160	tempc_temperatura_oleo_turbo	176	Engine Turbocharger Oil Temperature	65262 / 5-6
161	tempc_temperatura_pista	79	Road Surface Temperature	65269 / 7-8
162	timestamp			
163	timestampHash			
164	txt_iccid			
165	txt_identificacao_motorista_1	1625	Driver 1 identification	65131 / a
166	txt_identificacao_motorista_2	1626	Driver 2 identification	65131 / b

167	txt_imei			
168	txt_nome_marcha_engatada	163	Transmission Current Range	61445 / 7-8
Var	Descrição Transportadora	SPN	Descrição SAE J1939	PGN / Start Position
169	txt_nome_marcha_selecionada	162	Transmission Requested Range	61445 / 5-6
170	txt_operadora			
171	txt_vin	237	Vehicle Identification Number	65260 / 1
172	u16_peso_carreta_1			
173	u16_peso_carreta_2			
174	v_tensao_bateria	168	Battery Potential / Power Input 1	65271 / 5-6

11. APÊNDICE C

Códigos utilizados para cada método de classificação utilizado:

Naive Bayes

```
naive_consum = GaussianNB()
naive_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesNB = naive_consum.predict(X_consum_teste)
acuraciaNB = accuracy_score(y_consum_teste, previsoesNB)*100
```

Árvore de Decisão

```
arvore_consum = DecisionTreeClassifier(criterion='entropy', random_state=0)
arvore_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesAD = arvore_consum.predict(X_consum_teste)
acuraciaAD = accuracy_score(y_consum_teste, previsoesAD)*100
```

Random Forest

```
random_forest_consum = RandomForestClassifier(n_estimators=100,
                                              criterion='entropy', random_state = 0)
random_forest_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesRF = random_forest_consum.predict(X_consum_teste)
acuraciaRF = accuracy_score(y_consum_teste, previsoesRF)*100
```

Adaboost

```
ada_consum = AdaBoostClassifier(n_estimators=100, algorithm='SAMME',
                                random_state = 0)
adaboost_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesAB = adaboost_consum.predict(X_consum_teste)
acuraciaAB = accuracy_score(y_consum_teste, previsoesAB)*100
```

KNN

```
knn_consum = KNeighborsClassifier(n_neighbors=10)
```

```
knn_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesKNN = knn_consum.predict(X_consum_teste)
acuraciaKNN = accuracy_score(y_consum_teste, previsoesKNN)*100
```

Regressão Logística

```
logistic_consum = LogisticRegression(random_state = 1)
logistic_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesRL = logistic_consum.predict(X_consum_teste)
acuraciaRL = accuracy_score(y_consum_teste, previsoesRL)*100
```

SVM

```
svm_consum = SVC(kernel='linear', random_state=1)
svm_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesSVM = svm_consum.predict(X_consum_teste)
acuraciaSVM = accuracy_score(y_consum_teste, previsoesSVM)*100
```

Redes Neurais Artificiais

```
rede_neural_consum = MLPClassifier(verbose=True, max_iter = 1000,
                                   tol=0.0000010, hidden_layer_sizes = (15,15))
rede_neural_consum.fit(X_consum_treinamento, y_consum_treinamento)
previsoesRN = rede_neural_consum.predict(X_consum_teste)
acuraciaRN = accuracy_score(y_consum_teste, previsoesRN)*100
```