



UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
FACULDADE DE CIÊNCIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

IGOR SANTOS PORFÍRIO E SILVA

**Manutenção Preditiva de Superaquecimento em Motores de Ônibus Urbanos
com Aprendizado de Máquina**

APARECIDA DE GOIÂNIA - GO
2026



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE CIÊNCIAS E TECNOLOGIA

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(a) e o(a) orientador(a) 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

IGOR SANTOS PORFIRIO E SILVA

3. Título do trabalho

Manutenção Preditiva de Superaquecimento em Motores de Ônibus Urbanos com Aprendizado de Máquina

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 **Marcos Paulino Roriz Junior, Professor do Magistério Superior**, em 03/02/2026, às 19:29, 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 **Igor Santos Porfirio E Silva, Discente**, em 06/02/2026, às 11:36, 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 **5954468** e o código CRC **753D987F**.

IGOR SANTOS PORFÍRIO E SILVA

**Manutenção Preditiva de Superaquecimento em Motores de Ônibus Urbanos
com Aprendizado de Máquina**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Produção, da Faculdade de Ciências e Tecnologia da Universidade Federal de Goiás (UFG), como requisito para obtenção do título de Mestre em Engenharia de Produção.

Área de concentração: Gestão Estratégica e Operacional da Produção

Orientador(a): Dr. Marcos Paulino Roriz Junior

Coorientador(a): Dr. George Wilton Albuquerque Rangel

APARECIDA DE GOIÂNIA
2026

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Silva, Igor Santos Porfírio e
Manutenção Preditiva de Superaquecimento em Motores de Ônibus Urbanos com Aprendizado de Máquina [manuscrito] / Igor Santos Porfírio e Silva. - 2026.

91 f.: 2026

Orientador: Prof. Dr. Marcos Paulino Roriz Junior; co-orientador: Dr. George Wilton Albuquerque Rangel

Dissertação (Mestrado) - Universidade Federal de Goiás, Faculdade de Ciências e Tecnologia, Programa de Pós-graduação em Engenharia de Produção, Aparecida de Goiânia, 2026.

Apêndice.

Bibliografia.

Inclui: siglas, símbolos, tabelas, lista de figuras, lista de tabelas.

1. Manutenção Preditiva. 2. Aprendizado de Máquina. 3. Superaquecimento do Motor. 4. Dados de Telemetria. 5. Transporte Público Urbano.

I. Roriz Junior, Marcos Paulino, orient. II. Rangel, George Wilton Albuquerque, co-orient. III. Título.

CDU 658.5



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE CIÊNCIAS E TECNOLOGIA

ATA DE DEFESA DE DISSERTAÇÃO

Ata nº **64** da sessão de Defesa de Dissertação de **IGOR SANTOS PORFIRIO E SILVA**, que confere o título de Mestre em **Engenharia de Produção**, na área de concentração em **Gestão Estratégica e Operacional da Produção**.

Aos **sete dias do mês de janeiro de dois mil e vinte e seis**, a partir das **9h00min**, de forma virtual, através da plataforma **Google Meet**, realizou-se a sessão pública de Defesa de Dissertação intitulada “**Manutenção Preditiva em Ônibus Urbanos com Aprendizado de Máquina**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Marcos Paulino Roriz Junior (PPGEP/UFG)** com a participação dos demais membros da Banca Examinadora: Professor Doutor **Rafael Oliveira Vasconcelos (PROCC-UFS)**, membro titular externo, cuja participação ocorreu através de videoconferência; Professora Doutora **Symone Gomes Soares Alcalá (PPGEP/UFG)**, membro titular interno, cuja participação ocorreu através de videoconferência; e Professor Doutor **George Wilton Albuquerque Rangel (FCT/UFG)**, coorientador, cuja participação ocorreu através de videoconferência. Durante a arguição os membros da banca **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 **MARCOS PAULINO RORIZ JUNIOR**, 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 **sete dias do mês de janeiro de dois mil e vinte e seis**.

TÍTULO SUGERIDO PELA BANCA

Manutenção Preditiva de Superaquecimento em Motores de Ônibus Urbanos com Aprendizado de Máquina



Documento assinado eletronicamente por **Marcos Paulino Roriz Junior, Professor do Magistério Superior**, em 07/01/2026, às 11:15, 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 **Symone Gomes Soares Alcalá, Professora do Magistério Superior**, em 07/01/2026, às 11:16, 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 **George Wilton Albuquerque Rangel, Professor do Magistério Superior**, em 07/01/2026, às 11:16, 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 **Rafael Oliveira Vasconcelos, Usuário Externo**, em 09/01/2026, às 08:58, 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 **5895827** e o código CRC **1760BEFE**.

Referência: Processo nº 23070.064068/2025-55

SEI nº 5895827

Aos que vieram antes de mim, cuja história tornou possível que eu chegasse até aqui.

Agradecimentos

Agradeço, primeiramente, à minha família, em especial à minha mãe Paula Maria, ao meu pai Carlos Antônio e às minhas irmãs Fernanda Santos e Tamara Santos, pelo apoio, pela compreensão e pelo incentivo constantes ao longo de toda a minha vida, especialmente nos momentos difíceis, como aqueles vivenciados durante o período deste mestrado.

Ao meu orientador, agradeço pela orientação, pelas contribuições técnicas e pela disponibilidade ao longo de todo o percurso deste mestrado. Mesmo em um período marcado por desafios profissionais e pessoais, sua postura foi decisiva para a continuidade e a conclusão deste trabalho. As orientações firmes quando necessárias e, sobretudo, a confiança e a disposição em me apoiar, mesmo diante das dificuldades enfrentadas, foram fundamentais para que este trabalho se concretizasse.

Agradeço ao meu coorientador, que ao longo do processo contribuiu com observações relevantes, especialmente durante a etapa de qualificação. Sua participação foi importante para o amadurecimento conceitual e metodológico desta pesquisa.

Aos professores do programa de pós-graduação, agradeço pelas disciplinas oferecidas, pelas discussões acadêmicas e pelas contribuições que enriqueceram minha formação e o desenvolvimento deste trabalho.

À instituição de ensino e ao programa de pós-graduação, agradeço pela infraestrutura e pelo suporte acadêmico disponibilizados ao longo do curso.

À empresa parceira e às instituições envolvidas, agradeço por viabilizarem o acesso aos dados utilizados nesta pesquisa e pelo apoio institucional ao desenvolvimento do estudo.

Aos colegas de trabalho e de pesquisa, agradeço pelas trocas de experiências, pelas discussões técnicas e pelo apoio ao longo deste percurso.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“O pensamento crítico é sempre um ato de coragem.”

Lélia Gonzalez

Resumo

A manutenção adequada de veículos de transporte público urbano é essencial para garantir a confiabilidade do serviço, reduzir custos operacionais e mitigar impactos ambientais, especialmente em contextos de alta intensidade de operação. Nesse cenário, abordagens de manutenção preditiva baseadas em aprendizado de máquina têm sido amplamente propostas como alternativas às estratégias tradicionais, ao possibilitar a identificação antecipada de falhas a partir de dados operacionais. Este trabalho investigou a viabilidade da predição antecipada de eventos de superaquecimento do motor em ônibus urbanos utilizando dados reais de telemetria agregados por viagem, provenientes de um sistema comercial em operação real de frota. Foram avaliados modelos supervisionados, incluindo Floresta Aleatória, Máquinas de Vetores de Suporte e Redes Neurais Artificiais, combinados a diferentes estratégias de seleção de variáveis e janelas temporais retrospectivas definidas a partir da data e hora de abertura de ordens de serviço associadas a alertas térmicos, utilizadas para rotular as viagens como críticas ou normais. As variáveis explicativas foram obtidas por meio da agregação de eventos de telemetria e dados operacionais ao nível de cada viagem. Os modelos foram treinados com dados históricos e avaliados em um conjunto de teste temporalmente independente, utilizando métricas apropriadas a cenários desbalanceados, incluindo AUC, F1, precisão, recall e MCC. Os resultados indicaram desempenho global limitado e elevada variabilidade entre cenários. Embora, nos melhores experimentos, o AUC tenha atingido valores de até 0,875, as métricas associadas à detecção da classe crítica permaneceram baixas, com valores de F1 inferiores a 0,36, precisão abaixo de 0,25 e MCC inferiores a 0,34, refletindo colapsos de predição para uma única classe e baixa capacidade de discriminação entre viagens normais e críticas. As análises mostraram que tais limitações estão fortemente associadas à granularidade insuficiente da telemetria, à escassez e heterogeneidade dos eventos críticos e à qualidade dos rótulos derivados dos processos de manutenção. Conclui-se que, no estado atual da base de dados, a predição antecipada de superaquecimento do motor apresenta restrições significativas para uso operacional, sendo necessários avanços estruturais na instrumentação, na coleta, na integração e no registro dos dados para viabilizar aplicações futuras de manutenção preditiva em frotas de ônibus urbanos.

Palavras-chave: Manutenção preditiva; Aprendizado de máquina; Superaquecimento do motor; Dados de telemetria; Transporte público urbano.

Abstract

Proper maintenance of urban public transport vehicles is essential to ensure service reliability, reduce operational costs, and mitigate environmental impacts, particularly in high-intensity operating contexts. In this scenario, machine-learning-based predictive maintenance approaches have been widely proposed as alternatives to traditional strategies, enabling the early identification of failures from operational data. This study investigated the feasibility of early prediction of engine overheating events in urban buses using real telemetry data aggregated at the trip level, obtained from a commercial fleet management system in real operation. Supervised models, including Random Forest, Support Vector Machines, and Artificial Neural Networks, were evaluated in combination with different feature-selection strategies and retrospective temporal windows defined from the date and time of maintenance work order openings associated with thermal alerts, which were used to label trips as critical or normal. Explanatory variables were obtained by aggregating telemetry events and operational data at the trip level. The models were trained on historical data and evaluated on a temporally independent test set using metrics appropriate for imbalanced scenarios, including AUC, F1-score, precision, recall, and MCC. The results indicated globally limited performance and high variability across scenarios. Although, in the best experiments, the AUC reached values up to 0.875, the metrics associated with critical-class detection remained low, with F1-scores below 0.36, precision below 0.25, and MCC below 0.34, reflecting prediction collapses to a single class and low discrimination between normal and critical trips. The analyses showed that these limitations are strongly associated with insufficient telemetry granularity, the scarcity and heterogeneity of critical events, and the quality of labels derived from maintenance processes. It is concluded that, in the current state of the dataset, early prediction of engine overheating presents significant restrictions for operational use, and structural advances in instrumentation, data collection, integration, and recording are required to enable future predictive maintenance applications in urban bus fleets.

Keywords: Predictive maintenance; Machine learning; Engine overheating; Telemetry data; Public transport .

Lista de ilustrações

Figura 2.1 – Fluxograma PRISMA	36
Figura 3.1 – Etapas da pesquisa	42
Figura 3.2 – Modelos de ônibus que compõem a frota analisada. Fontes: RMTTC (2015), Diário da Manhã (2021), Meneses (2023).	44
Figura 3.3 – Histogramas de quilometragem acumulada para os principais modelos de ônibus da frota analisada.	46
Figura 3.4 – Fluxograma do processo de coleta, transmissão e extração de dados de telemetria.	51
Figura 3.5 – Fluxo de processamento e pré-processamento dos dados adotado na pesquisa	57
Figura 3.6 – Associação entre eventos de telemetria e viagens por meio de correspondência temporal.	60
Figura 3.7 – Agregação dos eventos por viagem e construção das variáveis analíticas.	62
Figura 3.8 – Fluxo de integração entre o banco de viagens e o banco de eventos, resultando no banco de viagens ampliado.	63
Figura 3.9 – Representação das viagens em escala temporal relativa à ordem de serviço, considerada como tempo zero. Os valores negativos indicam viagens anteriores à falha, enquanto os positivos indicam viagens posteriores.	65
Figura 3.10–Horizontes preditivos considerados na pesquisa, definidos em termos de janelas temporais (5, 10 e 15 dias) e número de viagens (20, 50 e 100 viagens anteriores à ordem de serviço).	66
Figura 3.11–Exemplo da aplicação de um horizonte temporal de 5 dias. As viagens destacadas em vermelho correspondem às que foram rotuladas como críticas por estarem dentro da janela anterior à ordem de serviço, enquanto as demais permanecem como normais.	66
Figura 3.12–Exemplo da aplicação de um horizonte baseado nas 20 viagens anteriores à ordem de serviço. As viagens destacadas em vermelho foram rotuladas como críticas, por estarem mais próximas da ocorrência do superaquecimento.	67
Figura 3.13–Distribuição original das viagens normais e críticas ao longo do tempo e entre diferentes veículos, evidenciando o desbalanceamento entre as classes no conjunto de dados.	69
Figura 3.14–Resultado do balanceamento das classes no conjunto de treinamento por meio de pareamento temporal, com a seleção de viagens normais realizadas no mesmo dia e em faixas horárias equivalentes às viagens críticas.	69

Figura 3.15– <i>Workflow</i> utilizado para treinamento e avaliação dos modelos no <i>Orange</i> <i>Data Mining</i>	74
Figura 3.16–Diagrama da etapa de avaliação dos modelos de aprendizado de máquina.	76

Lista de tabelas

Tabela 2.1 – Matriz de confusão genérica.	32
Tabela 2.2 – Resumo dos Estudos de Manutenção Preditiva em Veículos	38
Tabela 3.1 – Especificações técnicas dos principais modelos da frota.	47
Tabela 3.2 – Exemplo de registros de eventos de telemetria disponibilizados pela plataforma utilizada	53
Tabela 3.3 – Exemplo de registros de viagens (<i>trips</i>) disponibilizados pela plataforma de telemetria	54
Tabela 3.4 – Exemplo de registros de ordens de serviço utilizados no estudo	56
Tabela 3.5 – Exemplo de conjunto de dados agregado por viagem	68
Tabela 3.6 – Configurações de parâmetros para <i>Random Forest</i>	73
Tabela 3.7 – Configurações de parâmetros para Redes Neurais.	73
Tabela 3.8 – Configurações de parâmetros para SVM.	73
Tabela 4.1 – Top 20 atributos mais frequentes entre todos os filtros e estratégias de amostragem.	80
Tabela 4.2 – Distribuição de viagens normais e críticas nos conjuntos de treino e teste para cada estratégia de amostragem.	83
Tabela 4.3 – Desempenho médio dos modelos para diferentes estratégias de amostragem.	84
Tabela 4.4 – Estatísticas descritivas das métricas de desempenho obtidas para cada algoritmo e configuração	85
Tabela 4.5 – Matriz de confusão <i>Random Forest</i> (100 viagens, <i>Chi-Square</i>).	87
Tabela 4.6 – Matriz de confusão da RNA (2), horizonte de 10 dias, filtro <i>ReliefF</i>	88
Tabela 4.7 – Matriz de confusão da RNA (1), horizonte de 10 dias, filtro <i>ReliefF</i>	89
Tabela 4.8 – Matriz de confusão da SVM (2), 20 viagens, <i>Information Gain</i>	89
Tabela 4.9 – Matriz de confusão da SVM (2), 50 viagens, <i>ReliefF</i>	90
Tabela 4.10–Desempenho dos modelos por filtros de seleção de variáveis.	91
Tabela 4.11–Estatísticas descritivas das métricas dos 216 experimentos de classificação.	93
Tabela 4.12–Top 20 valores de AUC entre os 216 experimentos.	94
Tabela A.1–Variáveis térmicas do motor e do óleo.	106
Tabela A.2–Variáveis de condição mecânica e funcionamento do motor.	107
Tabela A.3–Variáveis de condução e comportamento operacional.	107
Tabela A.4–Variáveis de velocidade, deslocamento e operação da viagem.	108
Tabela A.5–Variáveis contextuais e temporais.	108

Lista de abreviaturas e siglas

ADASYN	<i>Adaptive Synthetic Sampling</i>
Adam	Algoritmo de otimização adaptativo para treinamento de redes neurais
ANTP	Associação Nacional de Transportes Públicos
AUC	Área sob a Curva ROC (<i>Area Under the Curve</i>)
Big Data	Conjunto de técnicas para armazenamento e análise de grandes volumes de dados
CA	Acurácia de Classificação (<i>Classification Accuracy</i>)
CAN	<i>Controller Area Network</i>
CNN-1D	<i>Convolutional Neural Network</i> unidimensional
COSMO	Método de detecção de anomalias baseado na comparação de modelos
COVID-19	Doença causada pelo coronavírus SARS-CoV-2
CSV	<i>Comma-Separated Values</i>
ENDF	Método de detecção de desvios baseado em histogramas
EventId	Identificador único de eventos de telemetria
EventTypeId	Identificador do tipo de evento de telemetria
F1-score	Média harmônica entre precisão e <i>recall</i>
FN	Falso Negativo (<i>False Negative</i>)
FP	Falso Positivo (<i>False Positive</i>)
GNL	Gás Natural Liquefeito
GPS	<i>Global Positioning System</i>
GRU	<i>Gated Recurrent Unit</i>
IA	Inteligência Artificial
IoT	Internet das Coisas (<i>Internet of Things</i>)
KPI	Indicador-Chave de Desempenho (<i>Key Performance Indicator</i>)

k-means	Algoritmo de agrupamento não supervisionado baseado em centróides
kNN	<i>k-Nearest Neighbors</i>
LR	Regressão Logística (<i>Logistic Regression</i>)
LSTM	<i>Long Short-Term Memory</i>
MCOD	Método de detecção de anomalias baseado em densidade
MCC	Coefficiente de Correlação de Matthews (<i>Matthews Correlation Coefficient</i>)
ML	Aprendizado de Máquina (<i>Machine Learning</i>)
MLP-ANN	<i>Multi-Layer Perceptron Artificial Neural Network</i>
NASA	<i>National Aeronautics and Space Administration</i>
NB	<i>Naive Bayes</i>
NO _x	Óxidos de Nitrogênio
NTU	Associação Nacional das Empresas de Transportes Urbanos
OBD-II	<i>On-Board Diagnostics</i>
OS	Ordem de Serviço
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i>
ReLU	<i>Rectified Linear Unit</i>
RF	<i>Random Forest</i>
RNA	Redes Neurais Artificiais
RPM	Rotações por Minuto
SGD	<i>Stochastic Gradient Descent</i>
SIMOB	Sistema de Informações da Mobilidade Urbana
SMOTE	<i>Synthetic Minority Over-sampling Technique</i>
SVM	<i>Support Vector Machine</i>
TN	Verdadeiro Negativo (<i>True Negative</i>)
TP	Verdadeiro Positivo (<i>True Positive</i>)

TripId	Identificador único de viagem
XGBoost	<i>Extreme Gradient Boosting</i>
χ^2	Estatística do teste Qui-quadrado

Lista de símbolos

°C	Temperatura em graus Celsius
km	Quilômetros percorridos
rpm	Rotações por minuto do motor
%	Percentual
Σ	Operador de soma
χ^2	Estatística do teste Qui-quadrado

Sumário

1	Introdução	18
1.1	Motivação	19
1.2	Justificativa	20
1.3	Objetivos	20
1.4	Metodologia	21
1.5	Contribuições	22
1.6	Estrutura da Dissertação	22
2	Referencial teórico	24
2.1	Fundamentação teórica	24
2.1.1	Manutenção de motores de ônibus	24
2.1.2	Sensores e dispositivos para coleta de dados em veículos	27
2.1.3	Métodos e algoritmos utilizados em manutenção preditiva	29
2.1.4	Seleção de Variáveis em Aprendizado de Máquina	30
2.1.5	Métricas de Avaliação de Modelos de Classificação	31
2.1.5.1	Exatidão	32
2.1.5.2	Precisão e Recall	32
2.1.5.3	F1-score	33
2.1.5.4	Área sob a Curva ROC (AUC)	33
2.1.5.5	Coefficiente de Correlação de Matthews	33
2.2	Revisão sistemática	33
2.3	Trabalhos relacionados	37
3	Metodologia	41
3.1	Delimitação da frota e do problema de estudo	42
3.2	Consulta e Extração de Dados	50
3.2.1	Dados de Telemetria	50
3.2.2	Dados de Manutenção (Ordens de Serviço)	55
3.3	Processamento dos Dados	56
3.3.1	Pré-processamento inicial e Filtragem de veículos	57
3.3.1.1	Filtragem de Viagens Válidas	58
3.3.2	Associação entre Viagens e Eventos	58
3.3.3	Agrupamento, cálculo de métricas e agregação dos eventos por viagem	61
3.3.4	Categorização de Variáveis	63
3.3.5	Definição de viagens críticas a partir das ordens de serviço	64
3.3.6	Conjunto de Dados Agregado por Viagem e Particionamento	67
3.3.7	Balanceamento das Classes no Conjunto de Treinamento	68
3.4	Identificação e Seleção de Variáveis	70

3.5	Escolha e Configuração dos Algoritmos de Aprendizado de Máquina	71
3.6	Avaliação e Validação do Modelo Preditivo	74
4	Resultados e Discussões	77
4.1	Caracterização das Variáveis e Seleção de Atributos	78
4.1.1	Critérios de seleção de variáveis	79
4.1.2	Atributos Discriminativos	79
4.1.3	Atributos Discriminativos	80
4.1.4	Síntese e Implicações da Seleção de Variáveis	81
4.2	Desempenho dos Modelos por Horizonte Preditivo	82
4.3	Comparação entre Algoritmos	85
4.3.1	Caso representativo: melhor configuração do Random Forest	87
4.3.2	Padrões extremos de predição nos modelos não lineares	88
4.3.2.1	RNA prevendo quase tudo como positivo.	88
4.3.2.2	RNA prevendo quase tudo como negativo.	88
4.3.2.3	SVM com predição massivamente positiva	89
4.3.2.4	SVM prevendo quase tudo como negativo	90
4.4	Impacto da Seleção de Variáveis no Desempenho dos Algoritmos de Aprendizado de Máquina	90
4.5	Panorama Geral do Desempenho	92
4.6	Melhores Desempenhos (Top 20 por AUC)	94
4.7	Discussão dos Resultados	95
5	Conclusão	97
5.1	Trabalhos Futuros	99
	Referências	101
	APÊNDICE A Descrição das Variáveis Utilizadas	105

1 Introdução

O transporte público por ônibus é um pilar fundamental para a mobilidade urbana no Brasil, sendo responsável por 86% do transporte coletivo no país, conforme dados do Sistema de Informações da Mobilidade Urbana [ANTP \(2020\)](#) e da Associação Nacional das Empresas de Transportes Urbanos [NTU \(2024\)](#). Esse modal transporta bilhões de passageiros anualmente, gerando um custo anual de R\$ 83 bilhões, com um patrimônio estimado em R\$ 122 bilhões em veículos e R\$ 376 bilhões em infraestrutura.

O Relatório Geral de 2018 do Sistema de Informações da Mobilidade Urbana da ANTP aponta que o tempo médio de viagem no transporte coletivo urbano foi de 44 minutos. O mesmo relatório indica que os ônibus respondem por aproximadamente 30% das emissões de poluentes do setor de transportes, contribuindo de forma significativa para a emissão de gases de efeito estufa ([ANTP, 2020](#)). Nesse contexto, a manutenção adequada da frota torna-se essencial não apenas para garantir a eficiência do sistema, mas também para reduzir impactos ambientais e preservar os investimentos e o patrimônio associados ao transporte público urbano.

Outro desafio relevante do setor é a pressão crescente por eficiência operacional em um contexto de restrições econômicas e envelhecimento da frota. Em 2018, o transporte coletivo urbano por ônibus no Brasil atingiu mais de 16 bilhões de passageiros transportados e 9 bilhões de quilômetros percorridos por uma frota de aproximadamente 117 mil veículos, com o suporte de cerca de 620 mil profissionais ([ANTP, 2020](#)). Nos anos seguintes, o setor passou por um período de retração na demanda, segundo a [NTU \(2024\)](#), e, em 2023, a idade média da frota atingiu aproximadamente 6,5 anos. Independentemente de flutuações conjunturais de demanda, esses indicadores evidenciam um cenário de frota envelhecida, margens operacionais pressionadas e necessidade crescente de otimização dos processos de manutenção, de forma a garantir confiabilidade, segurança e sustentabilidade do serviço.

Nesse contexto, para reverter o declínio observado, é essencial priorizar a melhoria no tempo operacional dos sistemas de transporte, na pontualidade dos veículos em suas rotas e na manutenção dos ônibus, aumentando a confiabilidade do transporte público. De acordo com [Guo, Xu e Xiao \(2018\)](#), a confiabilidade é um indicador crítico para a satisfação dos passageiros e para o aumento da taxa de participação no transporte público, que contribuem para o sucesso de políticas de mobilidade urbana sustentáveis.

1.1 Motivação

Nas últimas décadas, o aumento da complexidade operacional do transporte público urbano, aliado à crescente demanda por eficiência e confiabilidade, tem impulsionado a adoção de tecnologias capazes de aprimorar o monitoramento e a gestão de frotas. Essas tecnologias, baseadas em sistemas de telemetria, Internet das Coisas (IoT), *Big Data*, computação em nuvem, entre outros, permitem acompanhar em tempo real parâmetros operacionais críticos dos veículos, como temperatura, pressão e comportamento dinâmico durante as viagens. Essa infraestrutura digital viabiliza a coleta massiva de dados e possibilita o desenvolvimento de aplicações avançadas para planejamento de rotas, segurança, redução de custos e melhoria na experiência do usuário (ZHAI, 2024).

Apesar desses avanços, a manutenção dos veículos ainda é predominantemente realizada por meio de estratégias tradicionais, como a manutenção corretiva, que se baseia na identificação de falhas somente após sua ocorrência, e a manutenção preventiva, executada em intervalos de tempo predefinidos, sem considerar a condição real do veículo. Embora sejam práticas amplamente utilizadas devido à facilidade de implementação, essas abordagens frequentemente resultam em indisponibilidade operacional, atrasos na intervenção e custos elevados com reparos emergenciais e substituição de peças (CHEN, 2020). No contexto específico dos sistemas de transporte coletivo urbano, falhas inesperadas podem ainda provocar interrupções no serviço, engarrafamentos, impactos no cumprimento contratual e efeitos indiretos sobre o consumo de combustível e o desgaste de componentes.

Diante desse cenário, a manutenção preditiva surge como uma abordagem mais alinhada às atuais capacidades tecnológicas, uma vez que pode utilizar dados dos sensores embarcados para identificar padrões de comportamento atípicos e antecipar falhas antes que estas se manifestem de forma crítica (ERSÖZ *et al.*, 2022). Ao contrário da manutenção preventiva, que atua de forma sistemática e programada, a preditiva possibilita intervenções orientadas por indicadores de condição e degradação do sistema, extraídos dos dados operacionais, o que contribui para a redução de custos de manutenção, aumento da disponibilidade da frota e maior confiabilidade do serviço prestado.

Entretanto, a interpretação manual dos dados produzidos pelas diversas tecnologias implementadas, sobretudo em ambientes com alta frequência de operação e grande diversidade de veículos, torna-se progressivamente mais onerosa, sujeita a inconsistências e pouco escalável em função do volume, da variabilidade e da complexidade dos registros gerados. Nesse contexto, a utilização de técnicas baseadas em inteligência artificial, particularmente o aprendizado de máquina, surge como uma alternativa promissora para automatizar a detecção de anomalias e apoiar a antecipação de falhas mecânicas, permitindo que intervenções sejam planejadas de forma mais precisa (GÜVEN; ŞAHİN, 2022).

1.2 Justificativa

Diversos estudos recentes têm demonstrado o potencial da manutenção preditiva aplicada a sistemas de transporte público urbano, utilizando dados provenientes de sensores veiculares para identificação antecipada de falhas, como discutido em (GIANNOULIDIS; GOUNARIS, 2022; GÜVEN; ŞAHİN, 2022; ERSÖZ *et al.*, 2022; ZHAI, 2024; CORAZZA *et al.*, 2018; CHEN, 2020). Entretanto, apesar dos avanços tecnológicos observados nos últimos anos, ainda persistem lacunas relevantes na aplicação dessa abordagem em frotas de ônibus de transporte coletivo urbano. Para que os modelos preditivos baseados em aprendizado de máquina alcancem desempenho satisfatório nesse contexto, é necessário que sejam capazes de lidar com a natureza dinâmica e volátil das operações, considerando variações entre os veículos, condições operacionais diversas e a disponibilidade de dados coletados por sistemas de telemetria.

Além desses desafios operacionais, a implementação de estratégias de manutenção preditiva exige ferramentas computacionais capazes de processar grandes volumes de dados com confiabilidade. Como observado por Rögnvaldsson *et al.* (2018), a robustez dos modelos depende de sua capacidade de contornar limitações associadas à confiabilidade das medições e à ocorrência de ruídos operacionais. Em grande parte das empresas do setor, a ausência de infraestrutura adequada leva à dependência de estratégias tradicionais de manutenção que, embora sejam de fácil adoção, são, em condições ideais, menos eficientes do que a manutenção preditiva. Carvalho *et al.* (2019) destacam que essa preferência pela manutenção corretiva e preventiva está frequentemente associada a custos iniciais mais baixos e menor complexidade operacional, o que, contudo, tende a resultar em maior incidência de falhas não previstas e maiores despesas com reparos emergenciais.

1.3 Objetivos

O objetivo geral deste trabalho é investigar a viabilidade do uso de algoritmos de aprendizado de máquina, aplicados a dados de telemetria agregados por viagem, para identificar padrões operacionais associados à ocorrência de eventos de superaquecimento do motor em ônibus urbanos.

Este trabalho visa contribuir para a área de diagnóstico e manutenção preditiva em frotas de ônibus urbanos, tendo como objetivos específicos:

- a) Levantar e sistematizar o estado da arte em manutenção preditiva aplicada a ônibus urbanos, com ênfase em abordagens baseadas em aprendizado de máquina;
- b) Integrar e organizar bases de dados heterogêneas de telemetria, eventos e ordens de serviço, estruturando uma base unificada ao nível de viagem;

- c) Definir e implementar um procedimento de rotulagem das viagens com base em janelas temporais retrospectivas em relação às aberturas de ordens de serviço associadas a alertas térmicos;
- d) Realizar o pré-processamento, a seleção de variáveis e a construção dos conjuntos de treinamento e teste;
- e) Analisar a relevância e a recorrência das variáveis de telemetria e operação por meio de filtros de seleção de atributos, identificando aquelas mais associadas aos padrões pré-falha de superaquecimento;
- f) Avaliar o desempenho de diferentes algoritmos de aprendizado de máquina na identificação de viagens associadas a padrões pré-falha de superaquecimento do motor;
- g) Analisar criticamente os limites e as fontes de incerteza associadas aos dados, aos rótulos e aos modelos empregados.

1.4 Metodologia

A metodologia adotada neste trabalho foi estruturada a partir de uma abordagem experimental baseada na análise de dados de telemetria veicular, com o objetivo de avaliar o potencial de técnicas de aprendizado de máquina na predição de superaquecimento do motor em ônibus do transporte coletivo urbano. Inicialmente, realizou-se um levantamento sobre métodos de manutenção aplicados ao transporte público e sobre o uso de algoritmos preditivos em sistemas operacionais de frota, o que permitiu identificar os requisitos técnicos e os desafios associados ao processamento de dados provenientes de sensores embarcados.

A etapa subsequente consistiu na coleta e preparação dos dados utilizados nos experimentos. Os registros de telemetria, obtidos em ambiente operacional, passaram por procedimentos de limpeza, padronização, tratamento de valores ausentes e integração com informações operacionais, de forma a compor o conjunto de dados utilizado nos testes. Em seguida, foram conduzidas etapas de seleção de variáveis, nas quais foram aplicadas técnicas voltadas à identificação dos atributos mais relevantes para a predição do evento estudado.

Com base no conjunto de dados tratados, foram testados algoritmos de aprendizado de máquina em diferentes configurações parametrizadas, considerando múltiplos horizontes de predição. O processo experimental envolveu o treinamento e avaliação dos modelos sob diferentes cenários, permitindo comparar seu desempenho e identificar configurações mais adequadas ao problema em análise.

O ambiente, a caracterização do conjunto de dados e o problema técnico considerado serão descritos em detalhe posteriormente, no Capítulo ??.

1.5 Contribuições

As contribuições deste trabalho decorrem dos objetivos estabelecidos na Seção 1.3 e podem ser sintetizadas nos seguintes pontos:

- a) Construção de uma base de dados integrada ao nível de viagem, a partir da fusão de dados de telemetria, eventos e ordens de serviço provenientes de sistemas não originalmente comunicados;
- b) Proposição e implementação de um procedimento de rotulagem retrospectiva de viagens com base em janelas temporais definidas em relação às aberturas de ordens de serviço associadas a alertas térmicos;
- c) Análise sistemática da relevância e da recorrência de variáveis operacionais e de telemetria na identificação de padrões pré-falha de superaquecimento do motor;
- d) Avaliação comparativa do desempenho de diferentes algoritmos de aprendizado de máquina na identificação de viagens associadas a eventos de superaquecimento, sob validação temporal realista;
- e) Evidenciação empírica dos limites e das fontes de incerteza associados ao uso de dados agregados por viagem e rótulos operacionais na predição antecipada de falhas térmicas;
- f) Geração de subsídios técnicos para a tomada de decisão em empresas de transporte público urbano, no que se refere à viabilidade e às condições necessárias para a adoção de estratégias de manutenção preditiva.

1.6 Estrutura da Dissertação

O restante desta dissertação está organizado em cinco capítulos conforme descrito a seguir. O Capítulo 2 apresenta o referencial teórico que fundamenta este trabalho, abordando conceitos relacionados à manutenção de frotas de ônibus, bem como tecnologias utilizadas em veículos e técnicas de aprendizado de máquina aplicadas à detecção e predição de falhas.

No Capítulo 3 são apresentados os dados, as ferramentas e os métodos utilizados no desenvolvimento deste trabalho. Inicialmente, é realizada a caracterização do problema investigado, com foco nas ocorrências de superaquecimento do motor em veículos da frota

analisada. Em seguida, descreve-se o estudo de caso, contemplando as particularidades da empresa de transporte público urbano, sua estrutura operacional e os procedimentos de manutenção atualmente empregados.

Ainda nesse capítulo, são detalhadas as etapas de coleta, tratamento e preparação dos dados de telemetria utilizados na modelagem, incluindo critérios de seleção de viagens, padronização das variáveis, remoção de identificadores que possam induzir viés e definição dos horizontes de predição. Posteriormente, é apresentado o método de construção dos modelos preditivos, os algoritmos de aprendizado de máquina aplicados, o processo de seleção de variáveis relevantes, a definição das configurações dos algoritmos e os critérios estabelecidos para avaliação e validação dos modelos.

O Capítulo 4 apresenta os resultados da análise da relação entre os eventos de telemetria e a ocorrência de superaquecimento do motor, bem como da aplicação dos modelos de aprendizado de máquina. São discutidos o desempenho das diferentes configurações avaliadas e os principais achados referentes à identificação de padrões associados ao risco de superaquecimento.

Por fim, o Capítulo 5 apresenta as conclusões deste trabalho, destacando as contribuições alcançadas e as limitações identificadas, além de sugerir direções para trabalhos futuros que possam aprofundar ou ampliar os resultados aqui obtidos.

2 Referencial teórico

Para investigar a viabilidade de identificar problemas mecânicos em veículos a partir de dados de telemetria, este capítulo apresenta os fundamentos teóricos que sustentam a pesquisa. São abordados três tópicos principais: manutenção preditiva, telemetria veicular e aprendizado de máquina. O objetivo é compreender como o uso de dados coletados durante a operação dos veículos pode apoiar a construção de modelos preditivos voltados à detecção antecipada de falhas, com ênfase no superaquecimento do motor. Assim, este capítulo revisa brevemente os conceitos essenciais e os estudos relacionados, destacando avanços, limitações e lacunas presentes na literatura.

2.1 Fundamentação teórica

2.1.1 Manutenção de motores de ônibus

A manutenção de frotas de ônibus equipados com motores a diesel constitui um fator crítico para garantir a disponibilidade operacional, a confiabilidade e a segurança do transporte público urbano. Esses motores, por serem submetidos a ciclos intensivos de operação, caracterizados por várias viagens por dia, variações térmicas do ambiente e condições severas de uso típicas do tráfego urbano, apresentam tendência ao desgaste acelerado de suas peças e componentes. Nesse contexto, a adoção de estratégias de manutenção eficazes torna-se essencial para prevenir falhas mecânicas e prolongar a vida útil dos sistemas veiculares.

Segundo a [ABNT NBR 5462 \(1994\)](#) manutenção é definida como “a combinação de todas as ações técnicas e administrativas, incluindo as de supervisão, destinadas a manter ou recolocar um item em um estado no qual possa desempenhar uma função requerida”. Assim, a manutenção abrange todas as atividades relacionadas a finalidade de manter a funcionalidade do item, e podem incluir desde ações corretivas até modificações do equipamento.

Para manter a continuidade de operação, é importante planejar e executar a manutenção de forma estratégica. Essas estratégias de manutenção ou políticas de manutenção são atividades, como renovação, reparo ou substituição de peças, realizadas para garantir a continuidade do estado de saúde dos equipamentos, visando prolongar sua vida útil e capacidade de cumprir suas funções ([ERSÖZ et al., 2022](#))

Segundo [Carvalho et al. \(2019\)](#) a manutenção é um fator que afeta a eficiência e operação dos equipamentos e tem três categorias de estratégias principais:

- a) Manutenção corretiva ou *Run-to-Failure* (R2F): se caracteriza por esperar o equipamento parar de funcionar e então realizar seu reparo.
- b) Manutenção preventiva ou *Preventive maintenance* (PvM): é a estratégia que utiliza técnicas de manutenção periodicamente, com um cronograma baseado no tempo ou quantidade de iterações de processo.
- c) Manutenção preditiva ou *Predictive Maintenance* (PdM): usa ferramentas para monitorar os parâmetros da máquina ou a integridade do processo, e assim indicar quando uma ação de manutenção é necessária, antes da falha acontecer.

Por ser uma estratégia mais eficiente, a manutenção preditiva vem ganhando destaque nos setores de transporte, especialmente na aeronáutica e na indústria automotiva, como evidenciado por [Ersöz et al. \(2022\)](#). Essa revisão de literatura, composta por 52 trabalhos publicados entre 2017 e 2022, indica um crescimento acentuado no volume de estudos sobre PdM, que passou de 1 publicação em 2017 para 17 em 2021, com picos de 15 e 17 estudos em 2020 e 2021, respectivamente, sendo o número reduzido em 2022 explicado pelo encerramento da coleta em junho daquele ano. Os autores atribuem esse crescimento ao rápido avanço das tecnologias de sensores e de comunicação, bem como à consolidação de *Big Data*, inteligência artificial e unidades de processamento gráfico, que viabilizam o processamento de sinais como pressão, temperatura e vibração para a estimação da vida útil remanescente dos componentes. Além disso, a crescente preocupação com segurança e confiabilidade em sistemas de transporte, nos quais falhas em subcomponentes como motores e engrenagens podem levar à imobilização do veículo ou a acidentes graves, reforça o papel da PdM como ferramenta essencial para a operação segura e eficiente das frotas.

Quando se trata de motores, [Cai et al. \(2020\)](#) e [Cao et al. \(2015\)](#) apontam que diversas falhas estão associadas aos sistemas de injeção, lubrificação e arrefecimento, bem como à mecânica interna do motor. Entre os problemas relatados destacam-se variações anômalas nas folgas das válvulas, empenamento da haste de válvulas de exaustão, desgaste da sede e da mola de válvulas de admissão, falhas no sistema de injeção de combustível, incluindo baixa pressão no trilho de injetores, falhas na unidade de injeção de fluido de exaustão e degradação de componentes estruturais, como a junta do cabeçote. [Corazza et al. \(2018\)](#) indicam que aproximadamente 80% das falhas em motores estão relacionadas à contaminação do óleo lubrificante, decorrente do acúmulo de partículas sólidas, como fuligem e metais provenientes do desgaste, bem como de contaminantes líquidos e gasosos, incluindo água e combustível, que não são removidos por filtros convencionais de fluxo total. Esses contaminantes promovem tanto desgaste físico das superfícies metálicas quanto degradação química do óleo, reduzindo sua lubrificidade e acelerando a oxidação, o que contribui diretamente para o aumento da taxa de falhas internas do motor.

Determinar as causas das falhas exige análises detalhadas. Alguns estudos indicam que a ocorrência de falhas no motor está diretamente relacionada a fatores como características operacionais, condições externas e o estilo de condução dos motoristas. [Rosero *et al.* \(2020\)](#) destacam que operar o motor frequentemente em marcha lenta pode gerar acúmulo de desgaste e diminuir a eficiência do sistema. Os trabalhos de [Guo, Xu e Xiao \(2018\)](#) e [Massaro, Selicato e Galiano \(2020\)](#) ressaltam que algumas práticas de motoristas na condução dos veículos, como acelerações e desacelerações bruscas, frenagens constantes e sobrecarga na operação, aumentam o esforço sobre os componentes do motor e reduzem sua vida útil. Além disso, fatores contextuais também contribuem para aumentar a probabilidade de falhas, como condições climáticas adversas (chuva, calor ou frio extremo), estradas em condições ruins (terra, buracos ou subidas íngremes) e oscilações de velocidade, conforme descrito por [Giannoulidis e Gounaris \(2022\)](#), também influenciam negativamente o desempenho do motor e podem acelerar o aparecimento de anomalias no funcionamento do veículo.

O superaquecimento de motores é amplamente reconhecido na literatura como uma das ocorrências mais críticas em frotas de ônibus, em função de sua elevada frequência, do impacto direto sobre a confiabilidade operacional e das consequências para a segurança e a disponibilidade da frota ([LI *et al.*, 2019](#); [MASSARO; SELICATO; GALIANO, 2020](#); [RÖGNVALDSSON *et al.*, 2018](#)). Em geral, o superaquecimento resulta da combinação de fatores como falhas no sistema de arrefecimento, incluindo vazamentos, falhas de bombas e mau funcionamento de ventiladores, lubrificação inadequada, sobrecarga operacional, condução inadequada e condições extremas de temperatura ambiente, conforme discutido por ([CORAZZA *et al.*, 2018](#); [GUO; XU; XIAO, 2018](#); [MASSARO; SELICATO; GALIANO, 2020](#); [RÖGNVALDSSON *et al.*, 2018](#); [CHEN, 2020](#)).

Como consequência do aumento excessivo da temperatura do motor, podem ocorrer falhas críticas, incluindo o travamento do conjunto pistão–cilindro, empenamento de componentes estruturais, como o cabeçote, e degradação das propriedades mecânicas das superfícies metálicas em decorrência da perda do filme lubrificante e do aumento da oxidação do óleo ([RÖGNVALDSSON *et al.*, 2018](#); [CORAZZA *et al.*, 2018](#); [TURNER; AUSTIN, 2000](#)). O superaquecimento também compromete a integridade das juntas e dos sistemas de vedação, favorecendo vazamentos de óleo e do líquido de arrefecimento, frequentemente associados a falhas em juntas do cabeçote, bombas e tanques de expansão ([CAI *et al.*, 2020](#); [RÖGNVALDSSON *et al.*, 2018](#)). Esses mecanismos resultam em perda de desempenho, aumento do desgaste químico e físico dos componentes e redução significativa da vida útil do motor ([CORAZZA *et al.*, 2018](#); [TURNER; AUSTIN, 2000](#)).

Em situações menos severas, intervenções localizadas, como a substituição de juntas, a retífica do cabeçote ou a manutenção do sistema de arrefecimento, podem ser suficientes para restabelecer o funcionamento adequado do motor ([CAI *et al.*, 2020](#); [RÖGNVALDSSON](#)

et al., 2018). Contudo, episódios graves de superaquecimento podem evoluir para danos estruturais, como o travamento de cilindros, exigindo a retificação completa ou até a substituição do motor (RÖGNVALDSSON *et al.*, 2018). Tais falhas implicam não apenas em elevação significativa dos custos de manutenção, incluindo despesas com inspeções, reboque e penalidades contratuais, mas também em paralisações prolongadas que comprometem a disponibilidade da frota e o cumprimento da programação operacional (FAN; NOWACZYK; RÖGNVALDSSON, 2014; CHEN, 2020; GIANNOULIDIS; GOUNARIS, 2022).

2.1.2 Sensores e dispositivos para coleta de dados em veículos

Segundo Fraden (2015), um sensor é um dispositivo que recebe e responde a um estímulo, como movimento, som, força ou luz, convertendo-o em um sinal de saída, geralmente elétrico, mas que também pode ser de outra forma de energia. Um sensor pode ser constituído por um transdutor, que converte o estímulo em sinal elétrico, amplificadores, que aumentam a intensidade do sinal para torná-lo mais fácil de medir, e circuitos de processamento de sinal, que ajustam e filtram o sinal para facilitar sua interpretação e análise, além de outros componentes necessários para otimizar o funcionamento e a precisão do sensor. De forma geral, os sensores desempenham o papel de conectar o mundo físico a dispositivos eletrônicos, possibilitando o processamento de dados provenientes do ambiente.

No contexto automotivo, os sensores desempenham um papel crucial ao transformar grandezas físicas, como pressão do óleo, temperatura do motor e aceleração do veículo, em sinais elétricos (FLEMING, 2001). Esses dados constituem a base para diversas aplicações contemporâneas. Segundo Siegel, Erb e Sarma (2018), eles são fundamentais não apenas para a operação em tempo real, mas também para estratégias de manutenção proativa. Além de viabilizarem sistemas de controle de movimento e de segurança, esses sinais permitem análises preditivas, otimizando o tempo de serviço e possibilitando a estimativa da vida útil dos componentes.

Nos veículos modernos, há uma ampla variedade de sensores distribuídos estrategicamente, monitorando variáveis como temperatura, pressão, posição, velocidade, fluxo de massa de ar, torque, viscosidade, entre outras, conforme evidenciado por Mohankumar *et al.* (2019). Essa crescente complexidade dos sistemas embarcados em veículos automotivos e o aumento do número de Unidades de Controle Eletrônico (ECUs) demandaram a criação de soluções eficientes para a troca de informações e integração de dados entre os subsistemas. Para atender a esses requisitos, a comunicação interna entre sensores, atuadores e ECUs é realizada predominantemente por meio da *Controller Area Network* (CAN), tecnologia desenvolvida pela Bosch na década de 1980 para suportar aplicações automotivas em tempo real (FARSI; Schiffmann, 1999). O uso da rede CAN possibilitou uma expressiva redução da complexidade da fiação, permitindo a interconexão de múltiplos dispositivos

utilizando apenas um par de fios, além de oferecer alta confiabilidade e mecanismos de detecção e correção automática de erros por *hardware*.

A comunicação via CAN segue um modelo baseado em mensagens identificadas por rótulos que definem seu conteúdo e prioridade, em vez de endereçamento direto entre nós. Quando um dispositivo transmite dados, o controlador CAN encapsula esses valores em um frame, que é então recebido por todos os nós da rede. Cada módulo realiza um teste de aceitação e processa apenas as mensagens relevantes ao seu funcionamento (FARSI; Schiffmann, 1999). Essa arquitetura permite uma comunicação multimestre e tolerante a falhas, características essenciais para sistemas automotivos de alto desempenho.

Nos veículos pesados, como ônibus urbanos, o protocolo CAN serve como base para padrões de comunicação de nível superior, destacando-se o SAE J1939, amplamente adotado para interconectar ECUs e padronizar a transmissão de dados operacionais (CHEN, 2020). Esse protocolo utiliza quadros CAN estendidos de 29 bits e organiza as mensagens em *Protocol Data Units* (PDU), que incluem identificadores e até oito bytes de dados. Durante a operação do veículo, sensores embarcados monitoram variáveis físicas, como temperatura do motor, pressão do óleo, velocidade e nível de combustível, e transmitem essas informações por meio da rede CAN, sendo encapsuladas como fluxos J1939 para análise posterior (CHEN, 2020).

Adicionalmente, fatores externos ao veículo, como condições climáticas da região onde o veículo opera, características das vias (velocidade da via, buracos, inclinação) e padrões de condução do motorista, também influenciam o desempenho e nas abordagens de manutenção dos veículos. Estudos como os de Li, Guo e Xue (2018), Massaro, Selicato e Galiano (2020), Mckinley *et al.* (2020) destacam a relevância de dados complementares, como temperatura ambiente, precipitação e tráfego, para a elaboração de estratégias preditivas mais robustas. Dessa forma, a integração de sensores internos e externos aos veículos, aliados a tecnologias de análise preditiva, representa um avanço significativo para a eficiência e confiabilidade da manutenção veicular.

De acordo com Turner e Austin (2000), a integração de sensores automotivos com sistemas de telemetria otimiza o desempenho dos veículos, pois permite que os dados sejam analisados mesmo com o veículo ainda operando, possibilitando identificar possíveis falhas antes que se tornem problemas críticos. Essa dinamicidade fornecida pela telemetria ajuda nas estratégias de manutenção, pois fornece informações sobre o veículo, mesmo sem precisar tirá-lo de operação, o que faz reduzir o tempo de inatividade do veículo para manutenções.

2.1.3 Métodos e algoritmos utilizados em manutenção preditiva

A inteligência artificial (IA) é definida como o campo que busca desenvolver métodos, sistemas e algoritmos capazes de realizar tarefas que exigem inteligência, para pensar ou agir de forma humana ou racional (RUSSELL; NORVIG, 2021). Entre os métodos de IA, destaca-se o aprendizado de máquina (em inglês, *Machine Learning* - ML), que visa construir sistemas computacionais capazes de aprimorar seu desempenho automaticamente por meio da experiência (JORDAN; MITCHELL, 2015).

A manutenção preditiva faz uso de diversos métodos para processar dados oriundos de sensores ou registros históricos, aplicando abordagens estatísticas, estocásticas e de inteligência artificial (ERSÖZ *et al.*, 2022). O avanço de tecnologias emergentes, como a IA e a IoT, tem transformado as práticas tradicionais de manutenção, permitindo não apenas a predição de falhas, mas também a coleta e análise de dados em tempo real, baseando as predições em análises de dados históricos e operacionais.

A inteligência artificial tem aberto novas oportunidades de melhoria em diversos contextos, incluindo o transporte público e a gestão de frotas de veículos. Jevinger *et al.* (2023) apresenta uma revisão das principais abordagens de IA, com destaque para o aprendizado de máquina, aplicadas para otimizar tarefas como agendamento, roteamento, definição de caminhos mais curtos em redes de transporte público e alocação de veículos. Os autores também reforçam a importância da qualidade e da disponibilidade dos dados para o sucesso dessas aplicações, que são igualmente relevantes para a manutenção preditiva em frotas de veículos.

A classificação dos algoritmos no contexto de aprendizado de máquina é dividida em duas categorias principais: supervisionados e não supervisionados. Algoritmos supervisionados são treinados com dados rotulados, ou seja, dados para os quais as respostas ou resultados desejados já são conhecidos. Por outro lado, algoritmos não supervisionados não utilizam rótulos nos dados de entrada e buscam identificar padrões ou estruturas por conta própria, como agrupamentos ou detecção de anomalias (SARKER, 2021).

Modelos de aprendizado de máquina, particularmente algoritmos de classificação, são amplamente utilizados na manutenção preditiva para antecipar falhas em equipamentos e otimizar os cronogramas de manutenção. O objetivo principal é diferenciar condições saudáveis de condições críticas do sistema com base nos dados observados do veículo. Enquanto modelos de regressão podem estimar a vida útil restante, modelos de classificação são empregados para a determinação direta do “estado de saúde” do equipamento (SUSTO *et al.*, 2015).

Carvalho *et al.* (2019) apresentam os métodos de aprendizado de máquina mais utilizados em aplicações de manutenção preditiva, como Floresta Aleatória (*Random Forest*, RF), Redes Neurais Artificiais (RNA), Máquinas de Vetores de Suporte (*Support*

Vector Machine, SVM) e *k*-means, com base em artigos publicados entre 2009 e 2018. Os autores também ressaltam que a escolha do método de aprendizado de máquina mais adequado depende do ambiente de aplicação, dos equipamentos utilizados e dos dados disponíveis. Por outro lado, Davari *et al.* (2021) revisam métodos de manutenção preditiva aplicados a dados da indústria ferroviária, cobrindo o período de 2002 a 2021, e observam um crescimento no uso de Redes Neurais, além de destacar a importância das técnicas de pré-processamento de dados para melhorar o desempenho dos algoritmos.

O método Floresta Aleatória é uma técnica de aprendizado de máquina baseada na combinação de múltiplas árvores de decisão construídas a partir de amostras aleatórias, o que reduz o risco de sobreajuste e melhora a capacidade de generalização. Seu desempenho depende da força preditiva das árvores individuais e do baixo grau de correlação entre elas, características que conferem robustez frente a ruídos e *outliers*, além da capacidade de estimar a importância das variáveis (BREIMAN, 2001).

De acordo com Aggarwal (2018) as Redes Neurais Artificiais constituem uma técnica de aprendizado de máquina inspirada no funcionamento do sistema nervoso humano, em que unidades computacionais denominadas “neurônios” são interligadas de forma semelhante às conexões sinápticas biológicas. Essas conexões são representadas por pesos ajustáveis, os quais são modificados durante o processo de treinamento a partir dos dados disponíveis, possibilitando à rede a capacidade de aprender padrões complexos.

As Máquinas de Vetores de Suporte são algoritmos de classificação avançados que buscam criar hiperplanos ótimos para separar classes em tarefas de classificação. O objetivo é maximizar a margem entre as classes, ou seja, a distância entre os exemplos mais próximos de cada classe, conhecidos como vetores de suporte. Segundo Ben-Hur e Weston (2010), as SVM são amplamente utilizadas em áreas como bioinformática devido à sua elevada precisão, capacidade de lidar com dados de alta dimensionalidade e flexibilidade para modelar fontes de dados heterogêneas. Trata-se de métodos baseados em kernel, permitindo que os produtos internos do algoritmo sejam substituídos por funções kernel, o que possibilita a construção de fronteiras de decisão não lineares mantendo a estrutura de algoritmos originalmente lineares. A utilização eficaz das SVMs requer atenção à escolha de parâmetros, do tipo de kernel e ao pré-processamento adequado dos dados.

2.1.4 Seleção de Variáveis em Aprendizado de Máquina

A seleção de variáveis, ou *feature selection*, desempenha papel central na construção de modelos de aprendizado de máquina eficientes, interpretáveis e com boa capacidade de generalização. Conforme discutido por Kohavi e John (1997), a escolha adequada dos atributos pode reduzir a dimensionalidade do problema, mitigar redundâncias, diminuir o risco de sobreajuste e reduzir o custo computacional do treinamento, especialmente em conjuntos de dados com grande número de variáveis.

De modo geral, a literatura classifica as técnicas de seleção de variáveis em três grandes categorias [Cai et al. \(2018\)](#): métodos do tipo filtro, métodos do tipo *wrapper* e métodos embutidos. Os métodos do tipo filtro avaliam a relevância dos atributos a partir de medidas estatísticas de associação com a variável-alvo, de forma independente do algoritmo de aprendizado. Os métodos *wrapper* utilizam o desempenho de um modelo preditivo específico para avaliar diferentes subconjuntos de variáveis, enquanto os métodos embutidos realizam a seleção durante o próprio processo de treinamento, como ocorre em árvores de decisão ou modelos regularizados.

Neste trabalho, optou-se pelo uso de métodos do tipo filtro, uma vez que essa abordagem apresenta baixo custo computacional, maior transparência na interpretação dos resultados e independência em relação ao algoritmo de classificação adotado. Essa característica é particularmente relevante em cenários exploratórios e comparativos, nos quais se busca avaliar o impacto da seleção de variáveis em diferentes modelos e horizontes preditivos, sem introduzir viés associado a um classificador específico.

Entre as métricas de seleção do tipo filtro mais utilizadas na literatura estão o índice de Gini (*Gini index*), o teste qui-quadrado (*Chi-square*), o ganho de informação (*Info Gain*), a razão de ganho (*Gain Ratio*) e o algoritmo ReliefF. O índice de Gini e o ganho de informação são baseados em conceitos de impureza e entropia, avaliando o quanto um atributo contribui para reduzir a incerteza na classificação da variável-alvo. O teste qui-quadrado mede o grau de dependência estatística entre um atributo e a classe, sendo especialmente adequado para variáveis categóricas ou discretizadas. A razão de ganho é uma extensão do ganho de informação, proposta para reduzir o viés em favor de atributos com grande número de valores distintos. Já o algoritmo ReliefF avalia a relevância dos atributos com base na capacidade de distinguir instâncias próximas pertencentes a classes diferentes, considerando interações locais entre os dados.

A utilização conjunta de múltiplos critérios de seleção permite avaliar a consistência da relevância das variáveis sob diferentes perspectivas estatísticas. Dessa forma, reduz-se a dependência de um único método de avaliação e aumenta-se a robustez do processo de seleção, aspecto particularmente importante em conjuntos de dados reais, nos quais as variáveis podem apresentar diferentes distribuições, escalas e níveis de correlação.

2.1.5 Métricas de Avaliação de Modelos de Classificação

A avaliação do desempenho de modelos de aprendizado de máquina em problemas de classificação requer o uso de métricas adequadas às características do conjunto de dados e ao objetivo da aplicação. Em cenários com classes desbalanceadas, métricas tradicionais como a exatidão podem fornecer uma visão limitada do desempenho do modelo, tornando necessária a adoção de medidas complementares.

Considere uma matriz de confusão, como a da Tabela 2.1, composta por verdadeiros positivos (TP), verdadeiros negativos (TN), falsos positivos (FP) e falsos negativos (FN). A partir dessa matriz, podem ser definidas as métricas utilizadas neste estudo.

Tabela 2.1 – Matriz de confusão genérica.

Real	Predito		Σ
	Não	Sim	
Não (normal)	TN	FP	TN + FP
Sim (crítica)	FN	TP	FN + TP
Σ	TN + FN	FP + TP	TN + FN + FP + TP

2.1.5.1 Exatidão

A exatidão (*accuracy*) representa a proporção total de classificações corretas realizadas pelo modelo, sendo definida por:

$$\text{Exatidão} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

Embora amplamente utilizada, a exatidão pode ser inadequada em conjuntos de dados desbalanceados, pois modelos que classificam predominantemente a classe majoritária podem apresentar valores elevados dessa métrica sem, necessariamente, identificar corretamente a classe de interesse.

2.1.5.2 Precisão e Recall

A precisão (*precision*) e o *recall* são métricas que avaliam, respectivamente, a confiabilidade das previsões positivas e a capacidade do modelo em identificar corretamente os exemplos da classe positiva.

A precisão é definida como:

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (2.2)$$

Já o *recall*, também denominado sensibilidade, é dado por:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.3)$$

No contexto deste trabalho, o *recall* é particularmente relevante, pois mede a capacidade do modelo em identificar viagens críticas, reduzindo a ocorrência de falsos negativos, que correspondem a falhas não antecipadas.

2.1.5.3 F1-score

O *F1-score* consiste na média harmônica entre precisão e *recall*, sendo definido por:

$$\text{F1-score} = 2 \cdot \frac{\text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (2.4)$$

Essa métrica é especialmente útil quando se deseja um equilíbrio entre a capacidade de identificar corretamente a classe positiva e a confiabilidade das previsões realizadas, sendo amplamente empregada em problemas com desbalanceamento entre classes.

2.1.5.4 Área sob a Curva ROC (AUC)

A métrica AUC (*Area Under the Curve*) corresponde à área sob a curva ROC (*Receiver Operating Characteristic*), que relaciona a taxa de verdadeiros positivos (*true positive rate*) e a taxa de falsos positivos (*false positive rate*) para diferentes limiares de decisão.

Diferentemente das métricas baseadas em um único ponto de corte, a AUC avalia a capacidade discriminativa global do modelo, isto é, sua habilidade em separar corretamente as classes positiva e negativa ao longo de diferentes limiares. Valores de AUC próximos a 0,5 indicam desempenho equivalente ao acaso, enquanto valores próximos a 1,0 indicam elevada capacidade de discriminação.

2.1.5.5 Coeficiente de Correlação de Matthews

O coeficiente de correlação de Matthews (*Matthews Correlation Coefficient*, MCC) é uma métrica que considera simultaneamente todos os elementos da matriz de confusão, sendo definida por:

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (2.5)$$

O MCC assume valores no intervalo $[-1, 1]$, onde 1 indica classificação perfeita, 0 corresponde a uma classificação aleatória e valores negativos indicam desempenho inferior ao acaso. Essa métrica é reconhecida por sua robustez em cenários com classes desbalanceadas, sendo particularmente adequada para avaliar modelos quando há grande assimetria entre as classes, como no problema abordado neste trabalho.

2.2 Revisão sistemática

Com o intuito de levantar informações e identificar as tecnologias de inteligência artificial aplicadas à manutenção preditiva de motores de ônibus, foi realizada uma revisão

sistemática sobre o assunto. Para tanto, foram definidas questões de pesquisa a serem respondidas, sendo elas:

- a) Quais tecnologias, sensores e ferramentas podem ser usadas para monitorar dados de telemetria em ônibus urbanos?
- b) Quais as vantagens e desvantagens dos métodos de aprendizado de máquina atualmente utilizados para analisar dados de telemetria e prever problemas de superaquecimento do motor?
- c) Os sistemas de telemetria e manutenção preditiva conseguem identificar padrões que antecedem ou indicam a probabilidade de ocorrência de superaquecimento do motor?
- d) As soluções atuais enfrentam problemas em termos de custo e escalabilidade para frotas urbanas grandes ou que necessitem de monitoramento constante?
- e) Quais modelos matemáticos e algoritmos de aprendizado de máquina são usados para transformar dados de sensores em previsões de falhas relacionadas ao superaquecimento do motor?

Com base nas questões levantadas e visando garantir a padronização e a reprodutibilidade desta pesquisa, optou-se pela adoção do protocolo *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA), conforme apresentado por [Moher \(2009\)](#). Este protocolo abrange as fases de identificação, triagem e inclusão dos estudos, sendo que a Figura 2.1 ilustra cada uma dessas etapas e o número de artigos selecionados em cada fase.

A primeira fase do processo consistiu na identificação dos artigos na literatura. Para isso, foram selecionadas as bases de dados *Web of Science* e Periódicos da CAPES, com o objetivo de reunir publicações relevantes sobre o tema.

Em seguida, foi realizado o refinamento da string de busca. A elaboração da string baseou-se na aplicação de testes nas bases de dados e na avaliação dos resultados obtidos. As strings desenvolvidas foram compostas pelas seguintes palavras-chave:

- a) (*Public transportation OR Transit bus OR Urban bus OR Fleet management OR Diesel engine*) **AND**
- b) (*Maintenance OR Predictive maintenance OR Failure prediction OR Engine failure OR Engine overheating*) **AND**
- c) (*Sensor OR IoT OR Telemetry OR Temperature data OR Thermal data OR Machine learning OR Predictive algorithms*)

Essas palavras-chave ajudaram a direcionar a busca para trabalhos que abordam a manutenção preditiva de ônibus de transporte público urbano, com ênfase em soluções que utilizam sensores, telemetria e algoritmos de aprendizado de máquina para prever problemas de superaquecimento do motor, no período de 2010 a 2024. A busca bibliográfica foi realizada entre dezembro de 2024 e fevereiro de 2025 e resultou em 189 trabalhos, dos quais 31 eram duplicados.

Além disso, foram incluídos 8 trabalhos recomendados por especialistas na área, que foram considerados relevantes para o tema, mesmo não tendo sido identificados na busca sistemática. Dessa forma, 166 trabalhos avançaram para a triagem.

A etapa de triagem inicial envolveu uma leitura exploratória dos títulos, resumos e palavras-chave dos artigos para verificar sua relevância. Foram aplicados os seguintes critérios de exclusão:

1. Descartar trabalhos que não abordam especificamente a predição de falhas do motor em veículos urbanos, mas apenas aspectos gerais de manutenção veicular;
2. Descartar artigos que se concentram em tecnologias voltadas exclusivamente para outros tipos de transporte ou não aplicáveis a ônibus de transporte público urbano;
3. Descartar trabalhos que não abordam especificamente o uso de sensores, telemetria, IoT, ou algoritmos de aprendizado de máquina;
4. Descartar trabalhos que não abordem o uso, desenvolvimento ou aplicação de algoritmos de aprendizado de máquina para transformar dados de sensores em predições de falhas relacionadas ao motor.

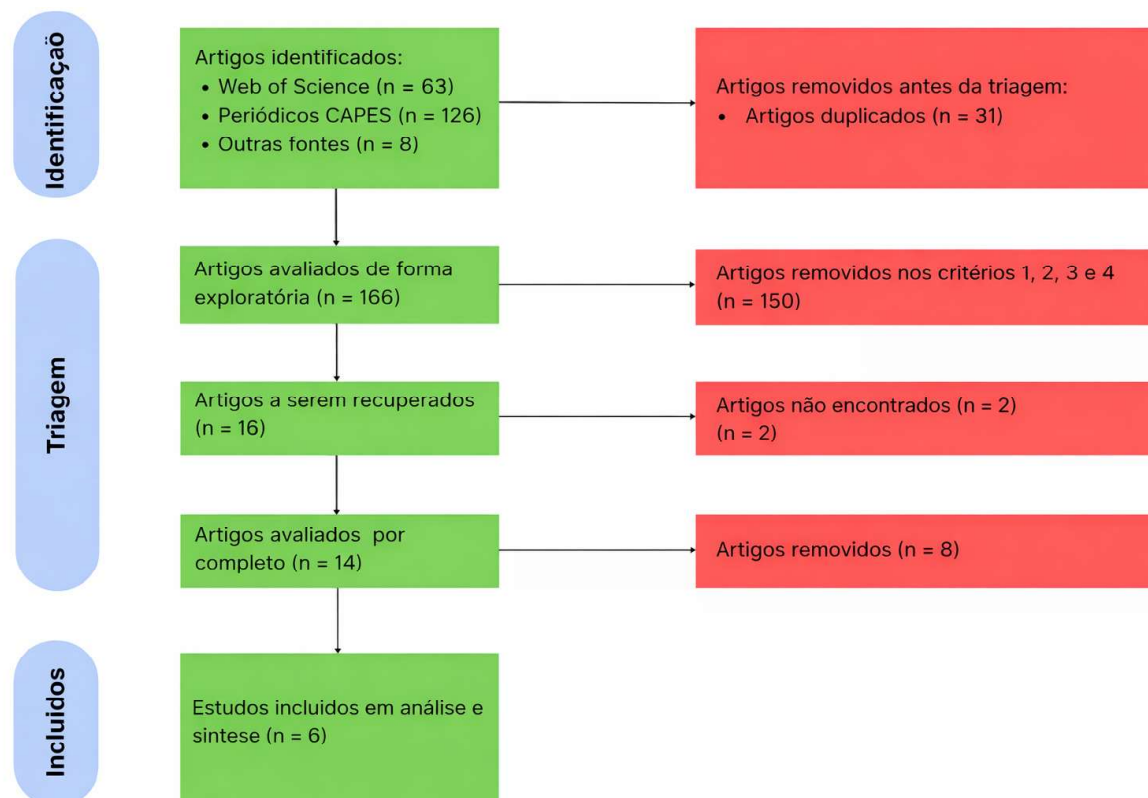


Figura 2.1 – Fluxograma PRISMA

Com base nesses critérios, 150 trabalhos foram descartados na etapa de triagem, restando 16 estudos que atenderam simultaneamente aos quatro critérios e avançaram para a fase de recuperação dos artigos. Dois trabalhos não puderam ser recuperados, pois não estavam disponíveis em acesso aberto nem em periódicos com os quais a UFG mantém convênio, resultando em 14 artigos para a etapa de leitura completa. Nessa fase, também foram excluídos oito estudos que apenas mencionavam manutenção preditiva ou falhas relacionadas ao motor, mas não apresentavam descrição dos métodos, tecnologias ou algoritmos de aprendizado de máquina efetivamente utilizados.

Ao final, foram selecionados seis trabalhos diretamente relacionados ao escopo desta pesquisa. Embora esse número possa parecer reduzido, ele reflete a aplicação de critérios de elegibilidade estritos, alinhados ao foco específico em ônibus urbanos, dados reais de telemetria, aprendizado de máquina e predição de falhas térmicas em motores. Como em toda revisão sistemática, os resultados dependem das questões de pesquisa e dos critérios adotados, de modo que variações nesses parâmetros poderiam levar à inclusão de outros estudos. Assim, o conjunto final deve ser interpretado como o subconjunto mais aderente ao problema investigado, e não como uma exaustão completa da literatura. As próximas seções exploram os conceitos e métodos utilizados nesses trabalhos, destacando tanto as soluções propostas quanto as lacunas que motivam a presente investigação.

2.3 Trabalhos relacionados

A revisão sistemática apresentada na Seção 2.2 identificou um conjunto de seis estudos que abordam a aplicação de aprendizado de máquina à manutenção preditiva de veículos, particularmente no contexto de frotas de ônibus. A partir desse conjunto, esta seção discute criticamente os métodos empregados, os tipos de dados utilizados, as estratégias de coleta e rotulagem, bem como as limitações estruturais observadas nesses trabalhos. A Tabela 2.2 sintetiza os estudos analisados, enquanto a discussão a seguir posiciona o presente trabalho em relação à literatura identificada.

Uma parte relevante dos trabalhos relacionados adota abordagens não supervisionadas para detecção de anomalias em frotas, explorando a capacidade desses métodos de identificar desvios operacionais sem a necessidade de dados rotulados. O estudo de Rögnavaldsson *et al.* (2018) propôs o método COSMO (*Consensus Self-organising Models*), baseado em detecção de *outliers* não supervisionada e na comparação entre modelos construídos a partir de sinais contínuos de sensores embarcados. Essa estratégia apresenta como principal vantagem a possibilidade de operar em ambientes com escassez ou ausência de rótulos de falha, o que é recorrente em frotas reais. A abordagem foi validada em um estudo de campo de longo prazo com uma frota de 19 ônibus urbanos, no qual foi capaz de identificar diferentes classes de desvios e anomalias, incluindo falhas no sistema de arrefecimento, como ventiladores de refrigeração descontrolados, vazamentos e obstruções em radiadores, falhas em compressores e sensores, aquecimento excessivo de rolamentos, além de anomalias térmicas e de pressão que antecederam falhas graves de motor, como o travamento de um cilindro. O método também identificou processos de degradação lenta, como a perda de desempenho de sensores de emissões, e desvios associados a condições operacionais e ambientais, como permanência prolongada em oficinas, diferenças de rota e estilo de condução. Apesar desses resultados, os autores destacam dificuldades em associar automaticamente as anomalias detectadas às causas exatas de falha, em função da baixa qualidade, inconsistência temporal e natureza não estruturada dos registros de manutenção, o que limita sua aplicação direta como ferramenta de apoio à decisão.

De forma semelhante, Giannoulidis e Gounaris (2022) investigaram e propuseram uma variedade de métodos não supervisionados para detecção de anomalias em frotas de ônibus urbanos, abrangendo abordagens baseadas em proximidade, métodos híbridos e técnicas de aprendizado profundo. Entre os métodos avaliados estão o Grand, uma técnica híbrida originalmente proposta por Rögnavaldsson *et al.* (2018), o MCOB (*Micro-cluster-based Continuous Outlier Detection*), baseado em distância para dados em fluxo, e uma abordagem de agrupamento que trata cada veículo como um grupo individual. Além disso, os autores propuseram um algoritmo em dois estágios, que adiciona uma verificação de proximidade para reduzir falsos positivos, e o modelo TranAD (*Transformer-based Anomaly Detection*), baseado em redes *Transformer* com consciência de contexto. Essas

Tabela 2.2 – Resumo dos Estudos de Manutenção Preditiva em Veículos

Autor/Ano	Entradas	Classificação e Algoritmos Usados	Contexto Operacional	Saídas
Giannoulidis (2022)	Dados de séries temporais de sensores de veículos (simulados - NASA Turbofan)	Sem Supervisão: Grand, MCOB, Cluster-based, TranAD e 2-stage	Gestão de frotas de veículos	<ul style="list-style-type: none"> • Predição da necessidade de manutenção com base na detecção de anomalias. • Comparação da relação custo-benefício de diferentes métodos.
Guyen (2022)	Dados de sensores de IoT de veículos	Supervisão: SVM, RF, LR, NB e kNN	Gestão de frotas de veículos	<ul style="list-style-type: none"> • Classificação de veículos como normal/defeituoso. • Probabilidade de falha do veículo. • Recomendação da velocidade de manutenção usando lógica fuzzy.
Li (2018)	Dados de sensores de motores de ônibus a GNL: Velocidade do instrumento, RPM, Pressão do óleo, Temperatura do líquido de arrefecimento, Temperatura externa e Precipitação.	Supervisão: RF e Rough Set	Predição de falhas em motores de ônibus a GNL	<ul style="list-style-type: none"> • Predição da ocorrência de falhas no motor com base em indicadores chave.
Massaro (2020)	Dados OBD-II, GPS e de sensores inerciais de ônibus	Supervisão para comportamento do condutor: MLP-ANN Sem Supervisão para detecção de anomalias: K-means	Gestão de frotas de ônibus	<ul style="list-style-type: none"> • KPIs do condutor com base na análise de agrupamento. • Predição de desgaste do veículo e necessidade de manutenção usando MLP-ANN. • Dashboards de eficiência da frota.
McKinley (2020)	Dados de garantia e instantâneos do motor de ônibus de trânsito	Supervisão: XGBoost (Extreme Gradient Boosting)	Predição de falhas em sensores NOx	<ul style="list-style-type: none"> • Probabilidade de falha do sensor NOx. • Análise de importância de variáveis usando Shapley Values. • Recomendações para ajuste do uso do veículo e melhoria do produto.
Rögnvaldsson (2018)	Dados de sensores de frotas de ônibus	Sem Supervisão: Histogramas, Autoencoders e ENDF	Gestão de frotas de veículos	<ul style="list-style-type: none"> • Detecção de desvios em frotas de veículos. • Identificação de sinais relevantes para monitoramento. • Associação de desvios com registros de reparo.

abordagens apresentam como vantagem a capacidade de detectar desvios sem rótulos explícitos, porém sua eficácia depende fortemente da disponibilidade de sinais contínuos, da correta segmentação dos regimes operacionais e da escolha da representação dos dados. Os resultados indicam que estratégias como segmentação operacional, agrupamento por condições de operação e redução de dimensionalidade são decisivas para viabilizar esses

modelos em ambientes reais, afetando diretamente o desempenho, a estabilidade e o custo computacional.

Em contraste, outros trabalhos adotam abordagens supervisionadas para a predição de falhas específicas, explorando a disponibilidade de dados rotulados para treinar classificadores diretamente sobre estados normais e de falha. [Li, Guo e Xue \(2018\)](#) propuseram um modelo para motores de ônibus a Gás Natural Liquefeito (GNL), combinando Teoria de Conjuntos Aproximados e Floresta Aleatória, alcançando F1-score de 0,86, com precisão de 0,81 e revocação de 0,80. Para a classe de falha, o modelo atingiu revocação igual a 1,00, indicando a detecção de todos os casos de falha no conjunto de teste, ainda que com ocorrência de falsos positivos. Esses resultados refletem a capacidade do modelo em aprender padrões fortemente associados às falhas quando rótulos confiáveis e janelas temporais bem definidas estão disponíveis. Contudo, o estudo apresenta limitações relacionadas à definição simplificada dos rótulos e ao uso de uma janela temporal restrita, baseada em médias de apenas 10 minutos antes da falha, o que compromete a generalização dos resultados para contextos operacionais mais heterogêneos. Desempenhos ainda mais elevados são relatados por [Güven e Şahin \(2022\)](#), com exatidão entre 0,9960 e 0,9999 e AUC superior a 0,996, porém os próprios autores reconhecem restrições associadas ao tamanho do conjunto de dados e à necessidade de validação adicional em tempo real, indicando que tais resultados estão condicionados a ambientes de teste controlados.

No contexto de ônibus urbanos, [Massaro, Selicato e Galiano \(2020\)](#) e [Mckinley et al. \(2020\)](#) exploraram sistemas integrados de coleta e análise de dados baseados em sensores embarcados e sinais contínuos da rede CAN (*Controller Area Network*). O estudo de [Massaro, Selicato e Galiano \(2020\)](#) utiliza uma infraestrutura composta por uma unidade embarcada conectada à porta OBD-II (*On-Board Diagnostics*), sensores inerciais, GPS e câmeras, com transmissão dos dados via rede móvel para uma plataforma em nuvem, na qual técnicas como *k-means* e Redes Neurais Multicamadas (*Multi-Layer Perceptron*, MLP) são empregadas para analisar o estresse do motor e o comportamento do motorista. Já [Mckinley et al. \(2020\)](#) combinam dados multimodais provenientes de registros de garantia, *snapshots* do módulo de controle do motor e variáveis ambientais, utilizando engenharia de atributos e modelos baseados em XGBoost (*Extreme Gradient Boosting*) com explicabilidade via SHAP (*Shapley Additive Explanations*). Embora essas arquiteturas ofereçam alta granularidade e forte capacidade de diagnóstico, elas dependem de infraestruturas dedicadas, integração de múltiplas fontes e condições específicas de amostragem, o que pode introduzir vieses e limitar sua escalabilidade e replicabilidade em ambientes operacionais reais.

De modo geral, os trabalhos analisados indicam que resultados positivos em manutenção preditiva estão fortemente associados à disponibilidade de dados contínuos e de alta granularidade, isto é, medições de alta frequência coletadas diretamente nos sensores

do veículo, como séries temporais de temperatura, pressão, vibração e rotação do motor. Esses dados são tipicamente obtidos em infraestruturas controladas ou integradas a fabricantes e fornecedores de tecnologia veicular. Em contraste, ambientes operacionais reais frequentemente dispõem apenas de dados agregados ou orientados a eventos, com menor resolução temporal, maior heterogeneidade e perda de informações dinâmicas relevantes para a antecipação de falhas.

Diante desse cenário, o presente estudo diferencia-se ao investigar a viabilidade do uso de dados de telemetria consolidados e eventos operacionais provenientes de plataformas comerciais de monitoramento de frotas amplamente utilizadas em operações de transporte urbano, para a identificação antecipada de risco de superaquecimento do motor. Ao contrário dos trabalhos que utilizam sinais contínuos de sensores ou infraestruturas experimentais dedicadas, esta pesquisa baseia-se em dados operacionais efetivamente disponíveis no ambiente de gestão de frotas, refletindo as condições reais de uso desses sistemas no mercado. Esse enfoque permite avaliar de forma mais crítica os limites, as incertezas e a robustez dos modelos de manutenção preditiva quando aplicados a cenários operacionais reais.

3 Metodologia

Este capítulo descreve o delineamento metodológico adotado neste estudo, detalhando o tipo de pesquisa, os dados utilizados, os procedimentos de pré-processamento, a construção dos modelos de aprendizado de máquina e os critérios de avaliação empregados. Para a classificação da pesquisa, utiliza-se a tipologia proposta por [Turrioni e Mello \(2012\)](#), que organiza estudos segundo sua natureza, objetivo, abordagem e método.

Quanto à natureza, a pesquisa é classificada como aplicada, uma vez que utiliza dados reais de operação para investigar a possibilidade de apoiar a gestão da manutenção de frotas por meio de modelos preditivos. Em relação aos objetivos, o estudo possui caráter exploratório, pois avalia a viabilidade técnica dessa abordagem no contexto específico analisado. A abordagem é predominantemente quantitativa, baseada na análise estatística do desempenho dos modelos, complementada por interpretação qualitativa dos resultados. Do ponto de vista metodológico, trata-se de um estudo experimental, conduzido por meio de simulações computacionais com diferentes configurações de dados, modelos e horizontes de previsão.

Seguindo o modelo de classificação proposto pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico ([CNPQ, 2025](#)), a pesquisa enquadra-se na área de Engenharia de Produção, mais especificamente nos subcampos de “Garantia de Controle de Qualidade” e “Planejamento, Projeto e Controle de Sistemas de Produção”, os quais se alinham diretamente à linha de pesquisa Gerenciamento de Sistemas Produtivos do Programa de Pós-Graduação. Esse enquadramento reflete o foco do estudo na análise da confiabilidade, da disponibilidade e do desempenho de sistemas operacionais complexos, por meio do uso de dados e modelos preditivos para apoiar a gestão da manutenção em frotas de transporte público.

A Figura [3.1](#) apresenta o fluxograma das etapas metodológicas adotadas neste estudo, desde a organização das bases de dados até a avaliação dos modelos. Esse encadeamento explicita a lógica de execução da pesquisa e a dependência entre as etapas, que serão detalhadas nas seções subsequentes.

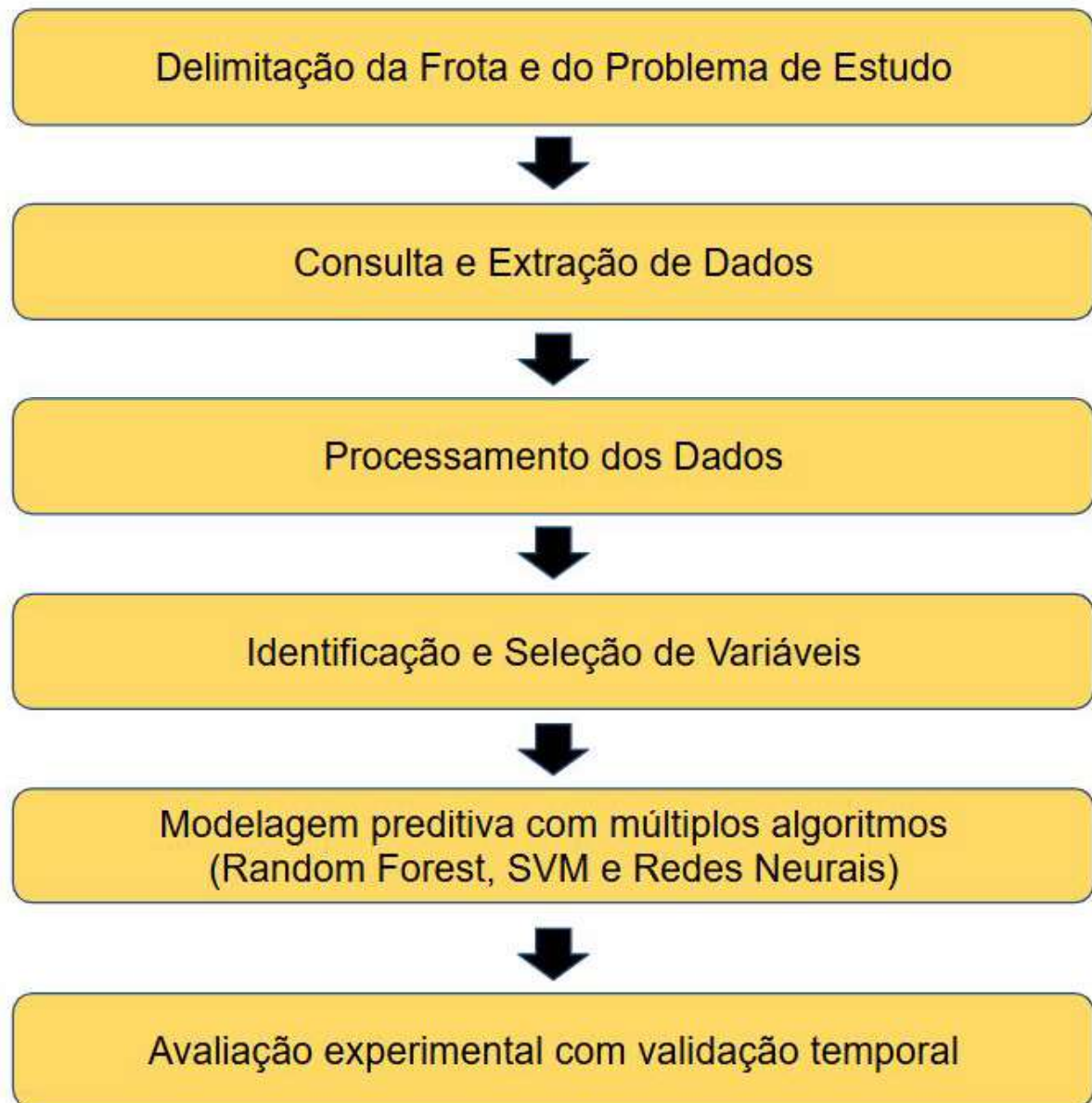


Figura 3.1 – Etapas da pesquisa

3.1 Delimitação da frota e do problema de estudo

Como forma de contextualizar a escala operacional do sistema como um todo, o boletim diário de fluxo e demanda da RMTC indica que, no dia 17 de abril de 2024, o transporte coletivo por ônibus de Goiânia atendeu aproximadamente 486.083 passageiros, utilizando 1.110 veículos, os quais realizaram cerca de 7.784 viagens ao longo do dia (RMTC, 2025).

A frota analisada pertence a uma das cinco empresas que compõem o consórcio responsável pela operação do transporte coletivo urbano por ônibus na região metropolitana de Goiânia. Em outubro de 2024, essa empresa operava 554 veículos, distribuídos conforme

apresentado a seguir:

a) **Volkswagen 17.230 Apache VIP-SC:**

- 449 ônibus;
- 81,0% da frota;
- Entraram em operação em 2008.

b) **Iveco/Masca 170S28:**

- 60 veículos;
- 10,8% da frota;
- Integrados em 2019.

c) **Mercedes-Benz O500 (articulados):**

- 32 ônibus;
- 5,8% da frota;
- Operacionais desde 2013.

d) **Outros modelos:**

- 13 veículos;
- 2,4% da frota.

Trata-se de uma frota urbana de grande porte, o que proporciona um contexto operacional amplo e a existência de dados históricos de utilização ao longo do período analisado. A frota apresenta variabilidade entre os modelos veiculares, com diferenças em idade, configuração mecânica e intensidade de uso, refletindo a diversidade operacional típica de sistemas de transporte público urbano. Essas características são relevantes para a contextualização do estudo e ajudam a delimitar o universo a partir do qual os recortes analíticos posteriores foram definidos.

A presença majoritária do modelo Volkswagen, em operação desde 2008, indica que parte significativa da frota possui maior tempo de uso, o que pode estar associado a diferentes demandas de manutenção ao longo de sua vida útil. Por outro lado, os modelos Iveco/Masca e Mercedes-Benz apresentam períodos de incorporação mais recentes (2019 e 2013, respectivamente), compondo um cenário que combina veículos com distintas faixas etárias e características operacionais. Essa contextualização fornece uma visão geral do ambiente em que os dados de telemetria foram obtidos e auxilia na compreensão das particularidades envolvidas na análise dos eventos registrados.

As Figuras 3.2 apresentam os três principais modelos de ônibus que compõem a frota analisada, permitindo sua identificação visual: Volkswagen 17.230, Iveco/Masca 170S28 e Mercedes-Benz O500 articulado.



(a) Volkswagen 17.230



(b) Iveco/Masca 170S28



(c) Mercedes-Benz O500

Figura 3.2 – Modelos de ônibus que compõem a frota analisada. Fontes: [RMTC \(2015\)](#), [Diário da Manhã \(2021\)](#), [Meneses \(2023\)](#).

Além da idade média, analisou-se a quilometragem acumulada por modelo de veículo com base no último valor de hodômetro registrado para cada ônibus que operou durante o mês de setembro de 2025, garantindo que cada veículo contribuísse com uma única observação ao cálculo. Essa estratégia evita viés decorrente de múltiplos registros por veículo e fornece uma estimativa mais representativa do uso acumulado da frota no período analisado. Observam-se diferenças significativas na distribuição e na consistência dos registros de odômetro entre os modelos, conforme apresentado a seguir:

- a) **Volkswagen 17.230 Apache VIP-SC**, com 312 veículos considerados: quilometragem média de 646.828 km, desvio padrão de 445.704 km e coeficiente de variação de 0,69;

- b) **Iveco/Masca 170S28**, com 54 veículos: quilometragem média de 475.031 km, desvio padrão de 123.888 km e coeficiente de variação de 0,26;
- c) **Mercedes-Benz O500 articulado**, com 33 veículos: quilometragem média de 644.828 km, desvio padrão de 63.304 km e coeficiente de variação de 0,10.

Os coeficientes de variação evidenciam diferenças marcantes na consistência dos registros de quilometragem entre os modelos. O valor elevado observado nos veículos Volkswagen indica forte dispersão relativa, sugerindo a coexistência de padrões muito distintos nos registros de hodômetro, enquanto os modelos Iveco/Masca e Mercedes-Benz apresentam distribuições significativamente mais concentradas, com maior estabilidade estatística.

No histograma do modelo Volkswagen (Figura 3.3c), observa-se uma distribuição claramente multimodal dos valores de hodômetro, caracterizada pela presença simultânea de veículos com quilometragens acumuladas elevadas e de outros com registros muito baixos. Considerando que todos os veículos desse modelo entraram em operação em 2008, esse padrão não é compatível com uma evolução contínua e homogênea da quilometragem ao longo do tempo e indica a ocorrência de reinicializações ou reconfigurações dos registros de hodômetro no sistema de gestão e de telemetria da frota. Essas reinicializações podem estar associadas a intervenções de oficina, como substituição de painéis, unidades eletrônicas, reconfiguração de módulos ou procedimentos de cadastro de veículos, cujos detalhes não estão disponíveis na base de dados analisada.

Como não há informações que permitam identificar quando, por qual motivo ou quantas vezes esses resets ocorreram para cada veículo, não é possível reconstruir de forma confiável a quilometragem real acumulada, o que compromete a utilização do hodômetro como indicador de desgaste para o modelo Volkswagen. Em contraste, os modelos Iveco/Masca (Figura 3.3b) e Mercedes-Benz (Figura 3.3a) apresentam distribuições mais concentradas e unimodais, compatíveis com o tempo de operação e com menor influência de problemas de registro. Essas diferenças foram consideradas nas etapas seguintes da pesquisa, em especial na definição do modelo veicular selecionado para a análise preditiva.

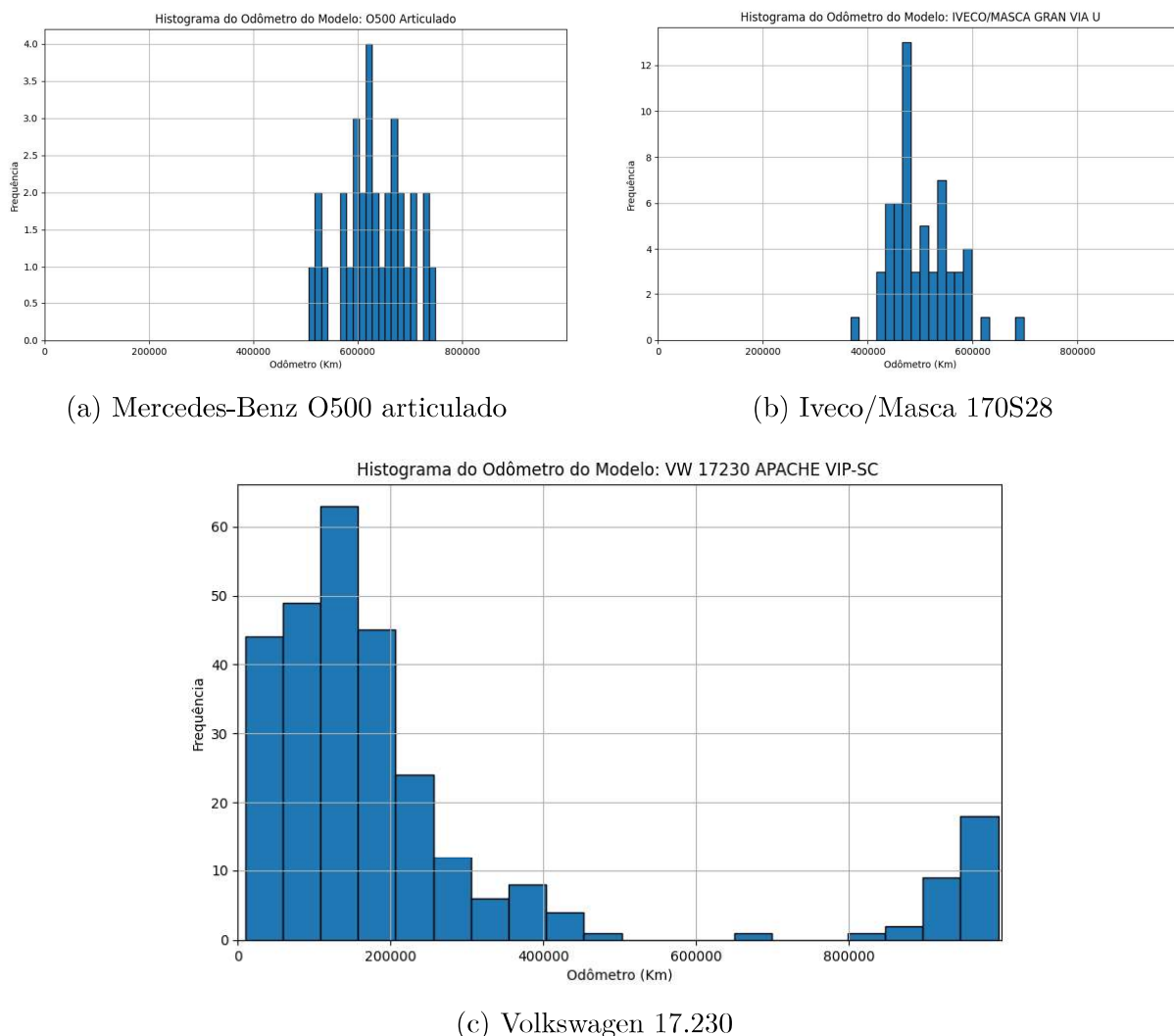


Figura 3.3 – Histogramas de quilometragem acumulada para os principais modelos de ônibus da frota analisada.

A Tabela 3.1 apresenta as principais especificações técnicas dos três modelos predominantes na frota analisada. Observa-se que há diferenças relevantes entre os veículos. Esses parâmetros estão diretamente relacionados à capacidade de tração e ao desempenho operacional, sobretudo em condições de tráfego intenso e em trechos com variações altimétricas.

Outro aspecto relevante é a relação entre peso bruto total e potência do motor, que influencia diretamente o esforço mecânico e térmico ao qual o conjunto motriz é submetido ao longo da operação. Veículos com maior razão peso/potência tendem a operar com cargas mais elevadas por unidade de potência disponível, especialmente em condições de tráfego urbano com paradas frequentes, acelerações sucessivas e variações altimétricas, o que pode aumentar a demanda térmica sobre o sistema de arrefecimento e acelerar processos de desgaste.

Nesse cenário, observa-se que o modelo Volkswagen 17.230 apresenta menor potência

Tabela 3.1 – Especificações técnicas dos principais modelos da frota.

Especificação	VW 17.230 Apache VIP-SC	Iveco / Masca 170S28	MB O500 Articulado
Potência (cv)	225	280	350
Torque (Nm)	830	950	1600
Transmissão	Manual, 6 marchas à frente + 1 ré	Manual, 6 marchas à frente + 1 ré	Automática, 4 marchas + retardador
PBT (kg)	16.000	16.000	29.500
Norma ambiental	Euro III	Euro V	Euro V
Tipo de suspensão	Molas semielípticas (dianteira e traseira)	Molas semielípticas (dianteira e traseira)	Pneumática (dianteira / interm. / traseira)
Distância entre eixos (mm)	5.180	5.950	5.250 + 6.700 (articulado)

e torque em relação aos demais, enquanto opera com um peso bruto total semelhante ao Iveco/Masca e inferior ao do modelo articulado, o que pode resultar em maior solicitação do motor em determinados regimes de operação. Por outro lado, o Mercedes-Benz O500, apesar de possuir maior PBT, dispõe também de potência e torque significativamente superiores, além de transmissão automática e suspensão pneumática, o que contribui para uma distribuição mais eficiente das cargas dinâmicas durante a operação. Essas diferenças reforçam a importância de considerar as características técnicas dos veículos na análise do comportamento térmico e na interpretação dos eventos de superaquecimento observados na frota.

No caso dos veículos *Volkswagen 17.230*, incorporados à frota a partir de 2008 e responsáveis pela maior parcela dos ônibus analisados, as limitações observadas não se relacionam à ausência de sensores ou à irrelevância do modelo para o fenômeno investigado. Pelo contrário, trata-se de um veículo com menor relação peso/potência em comparação aos demais, o que, do ponto de vista físico, o torna mais suscetível a regimes de maior esforço térmico do motor em condições típicas de operação urbana, como tráfego intenso, paradas frequentes e trechos com aclives. Isso faz desse modelo, em princípio, um candidato natural para o estudo de superaquecimento.

Entretanto, na prática, a utilização desse modelo como base para análises preditivas

foi limitada pela qualidade e pela consistência dos dados disponíveis. Segundo gestores da frota e técnicos da empresa de telemetria, há dificuldades recorrentes de ajuste e calibração de sensores em parte desses veículos, especialmente no sensor de pressão do óleo, o que compromete a confiabilidade de variáveis críticas para a caracterização do estado térmico e mecânico do motor. Soma-se a isso o fato de que os registros de hodômetro apresentam forte dispersão e múltiplos padrões de reinicialização ao longo do ciclo de vida da frota, decorrentes de intervenções operacionais realizadas em processos de manutenção, substituição de painéis ou reconfiguração de sistemas eletrônicos, sem que essas alterações sejam devidamente documentadas de forma estruturada.

Adicionalmente, a substituição progressiva dos veículos *Volkswagen 17.230* por modelos mais recentes a partir de 2025, associada à reutilização de códigos de identificação no sistema de Ordens de Serviço, introduziu descontinuidades e ambiguidades nos históricos de manutenção, dificultando a rastreabilidade temporal de falhas, reparos e eventos associados a veículos específicos. Essas inconsistências inviabilizam a construção de séries históricas confiáveis, requisito fundamental para o treinamento e a validação de modelos de aprendizado de máquina.

Por outro lado, os veículos *IVECO/MASCA 170S28*, incorporados à frota em 2019 e correspondentes ao segundo maior grupo amostral, apresentaram registros de telemetria significativamente mais consistentes, além de uma distribuição de quilometragem acumulada mais homogênea e menor evidência de reinicializações ou artefatos de medição. Esse equilíbrio entre relevância operacional, estabilidade estatística e qualidade dos dados torna esse modelo mais adequado para a investigação empírica proposta.

Diante desse compromisso entre sensibilidade ao fenômeno de superaquecimento e confiabilidade dos dados disponíveis, o modelo *IVECO/MASCA 170S28* foi selecionado como referência para o estudo de caso e para o desenvolvimento dos modelos preditivos voltados à identificação de falhas térmicas do motor.

O sistema de telemetria empregado registra eventos associados a variáveis operacionais, como velocidade, rotação do motor, temperatura do líquido de arrefecimento, pressão do óleo, consumo de combustível e localização via GPS. Durante a execução desta pesquisa, não houve acesso ao detalhamento técnico da instrumentação, restringindo-se a análise aos dados disponibilizados pela plataforma da empresa de telemetria, situação coerente com a realidade operacional de gestores de frotas.

Os registros de ordens de serviço constituem uma fonte complementar de informação, pois permitem relacionar eventos operacionais a intervenções de manutenção. No entanto, constatou-se que parte desses registros não corresponde a falhas efetivamente diagnosticadas, apresentando inconsistências, diagnósticos inconclusivos e ocorrências de retrabalho. Essas limitações afetam a confiabilidade dos rótulos de falha e exigem cautela na sua utilização em análises preditivas, aspecto considerado ao longo deste estudo.

A empresa adota predominantemente práticas de manutenção preventiva e corretiva, com controle informatizado das ordens de serviço, o que permite o registro histórico das intervenções realizadas nos veículos e, em princípio, seu cruzamento com os dados de telemetria. Ainda assim, observou-se que parte das ordens registradas não resulta em intervenção técnica efetiva, seja pela ausência de confirmação do problema relatado, seja por limitações no diagnóstico realizado pelas equipes de manutenção. Como consequência, podem ocorrer liberações de veículos sem correção definitiva da anomalia e a abertura recorrente de novas ordens de serviço para o mesmo sintoma.

Essas situações comprometem a consistência dos registros de manutenção e introduzem incertezas quanto à correspondência entre as ordens de serviço e a condição real dos veículos. Apesar dessas limitações, o histórico de intervenções permanece como uma fonte potencialmente relevante para a análise de padrões de falhas e de processos de degradação, desde que interpretado de forma crítica e contextualizada. Essas restrições são retomadas ao longo da dissertação, especialmente na discussão dos limites dos dados utilizados.

Diante desse cenário, optou-se por concentrar a investigação em uma falha cuja manifestação pudesse ser confirmada diretamente por meio dos dados de telemetria, reduzindo a influência das incertezas associadas às ordens de serviço. A ocorrência selecionada foi o superaquecimento do motor, definida como foco deste estudo preditivo por sua relevância operacional e por sua aderência às possibilidades analíticas disponíveis.

Sob a perspectiva técnica, o superaquecimento está associado à degradação de componentes críticos do sistema de propulsão, podendo resultar em falhas severas, indisponibilidade prolongada dos veículos e interrupções da operação em via. Além disso, esse tipo de falha apresenta elevada disponibilidade de variáveis sensoriais potencialmente informativas, como a temperatura do líquido de arrefecimento e a pressão do óleo, quando comparado a outras ocorrências registradas nas ordens de serviço.

No levantamento realizado, identificaram-se duas categorias principais de ordens de serviço relacionadas ao superaquecimento. A primeira corresponde a registros baseados no relato do motorista, classificados como “motor esquentando”. A segunda compreende ordens abertas pelo setor de monitoramento de operações a partir de alertas de temperatura gerados pela telemetria. Esses alertas são produzidos automaticamente quando o valor do sensor de temperatura do líquido de arrefecimento do motor ultrapassa limites pré-definidos no sistema de monitoramento da frota, configurados de acordo com critérios operacionais e de segurança. Ao exceder esse limiar, o sistema registra um evento associado ao veículo e ao instante da ocorrência, que é então encaminhado ao centro de operações, onde pode motivar a abertura de uma ordem de serviço para inspeção e intervenção.

Observou-se que os registros baseados em relatos apresentaram maior incidência de diagnósticos inconclusivos ou de falsos positivos, enquanto as ordens originadas diretamente da telemetria mostraram-se mais consistentes e aderentes às condições reais dos veículos.

Com base nessas evidências, definiu-se que o desenvolvimento dos modelos preditivos utilizaria exclusivamente as ordens de serviço associadas a eventos de superaquecimento registrados pela telemetria. Essa decisão buscou reduzir a presença de registros espúrios e fortalecer a correspondência entre os eventos monitorados e os rótulos de falha, contribuindo para maior robustez e coerência metodológica dos modelos desenvolvidos.

A próxima seção descreve os procedimentos de consulta e extração dos dados de telemetria e das ordens de serviço, detalhando a estrutura das bases utilizadas e as etapas de preparação dos dados para aplicação nos modelos preditivos.

3.2 Consulta e Extração de Dados

Foram consideradas duas fontes principais de dados: (i) os registros de telemetria provenientes dos dispositivos embarcados nos veículos, e (ii) os dados de manutenção obtidos a partir das ordens de serviço geradas pela empresa operadora da frota. Como tais informações são armazenadas em plataformas distintas, a primeira gerida por uma empresa terceirizada e a segunda pelo sistema interno da frota, os processos de consulta e extração são descritos separadamente.

3.2.1 Dados de Telemetria

O processo de coleta das informações operacionais é realizado por dispositivos de telemetria embarcados, fornecidos por uma empresa multinacional especializada no monitoramento de veículos comerciais pesados e já instalados nos ônibus. Esses dispositivos integram-se aos sistemas eletrônicos dos veículos por meio de protocolos padronizados de comunicação e diagnóstico, como o SAE J1939 (baseado na rede CAN), o que garante interoperabilidade entre diferentes sistemas e viabiliza a consolidação das variáveis provenientes dos sensores instalados.

As informações coletadas são transmitidas para uma plataforma digital dessa mesma empresa de telemetria, utilizada pela operadora para o monitoramento e gestão da frota. Esse sistema disponibiliza relatórios gerenciais e acesso via *Application Programming Interface* (API), permitindo consultas estruturadas e integração com sistemas externos. Neste trabalho, os dados foram extraídos a partir do banco de dados vinculado a essa plataforma, o que dispensou o acesso direto ao barramento CAN ou o desenvolvimento de *hardware* próprio.

Essa estratégia foi adotada por três motivos principais:

- a) reduzir custos e tempo de implementação;
- b) evitar riscos técnicos associados à validação de novos dispositivos de coleta;

- c) garantir que a metodologia desenvolvida seja compatível com práticas atualmente aplicadas por empresas de transporte.

Embora essa escolha limite o nível de granularidade dos dados acessados, uma vez que somente os eventos registrados pela plataforma podem ser analisados, ela reforça o caráter aplicado do estudo, por reproduzir as condições reais de acesso à informação enfrentadas por gestores de frotas, que comumente dependem de provedores terceirizados de telemetria.

O fluxo de dados pode ser descrito da seguinte forma: (i) sensores embarcados coletam as variáveis operacionais do veículo; (ii) os módulos de telemetria, integrados ao barramento CAN, processam essas informações; (iii) os dados são transmitidos via rede móvel para os servidores da fornecedora; (iv) a plataforma em nuvem organiza e disponibiliza os registros em painéis interativos ou via API; e (v) os dados selecionados são extraídos para compor a base utilizada nesta pesquisa. A Figura 3.4 ilustra esse processo.

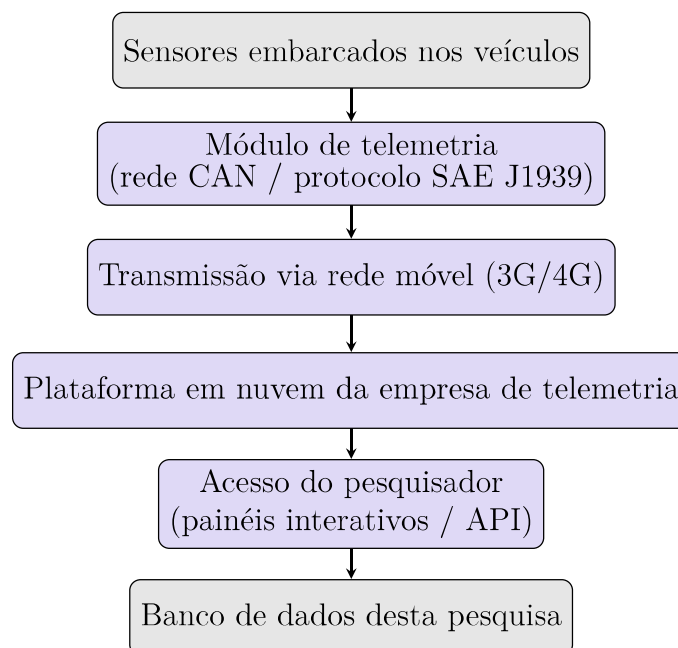


Figura 3.4 – Fluxograma do processo de coleta, transmissão e extração de dados de telemetria.

Os dados empregados neste estudo correspondem a registros orientados a eventos, caracterizando uma telemetria de natureza reativa. Isso significa que as informações são registradas somente quando determinados gatilhos são ativados pelo sistema de monitoramento. Esses gatilhos tem dois tipos principais: (i) eventos por exceção, acionados quando valores monitorados por sensores excedem limites estabelecidos de forma empírica por especialistas da empresa gestora da frota; e (ii) eventos por consolidação, processados ao final de cada viagem (*trip*), definida como o intervalo entre o acionamento e o desligamento da ignição do veículo, os quais agregam informações como temperatura média do motor,

pressão média do óleo, tempo em movimento, tempo parado, velocidade média, entre outras variáveis operacionais.

Embora essa estratégia reduza o volume de dados redundantes e esteja alinhada às práticas consolidadas de gestão da frota, ela impõe limitações à modelagem preditiva ao restringir a granularidade temporal das informações. Situações intermediárias ou tendências evolutivas podem não ser capturadas, especialmente nos casos em que a falha não ultrapassa imediatamente os limites configurados. Ainda assim, essa abordagem evidencia a necessidade de estudos futuros voltados à coleta contínua ou em maior resolução temporal diretamente da rede CAN, com potencial de aprimorar os modelos preditivos baseados em variáveis contínuas.

O período de análise considerado nesta pesquisa compreende o intervalo entre outubro de 2024 e setembro de 2025. Nesse período, o sistema de telemetria da empresa operadora registrou 34 tipos distintos de eventos, cada um armazenado em uma tabela específica no banco de dados da plataforma. Esses eventos são organizados em três categorias funcionais, *detail*, *summary* e *notify*, que determinam o conjunto de variáveis registradas em cada ocorrência. Dentre esses eventos, destacam-se:

- a) **Aceleração Brusca:** ocorre quando o veículo apresenta uma aceleração maior ou igual a 9 km/h por segundo;
- b) **Acelerando Parado:** a ignição está ligada, a velocidade é igual a 0 km/h e a rotação do motor é maior ou igual a 1100 rpm;
- c) **Banguela (nomenclatura adotada pelo sistema de telemetria):** velocidade superior a 30 km/h, com a rotação do motor entre 100 rpm e 800 rpm;
- d) **Fora da Faixa Verde:** ocorre quando a velocidade do veículo é superior a 25 km/h e a rotação do motor ultrapassa 2100 rpm;
- e) **Marcha Lenta:** caracteriza-se pela ignição ligada, velocidade inferior a 5 km/h e rotação do motor entre 0 rpm e 900 rpm;
- f) **Pressão Média do Óleo:** ao ligar a ignição, inicia-se a medição dos valores da pressão durante o trajeto. Ao desligar a ignição, calcula-se a média dessas leituras, registrando o valor resultante no banco de dados;
- g) **Rotação Excessiva:** ocorre quando a rotação do motor está entre 2560 rpm e 4000 rpm;
- h) **Temperatura Máxima Crítica do Motor:** a ignição está ligada, o motor opera a uma rotação superior a 400 rpm, e a temperatura do líquido de arrefecimento do motor ultrapassa 90 °C;

- i) **Temperatura Média do Motor:** ao ligar a ignição, inicia-se a medição das temperaturas do líquido de arrefecimento do motor durante o trajeto. Ao desligar a ignição, calcula-se a média dessas leituras, registrando o valor resultante no banco de dados;
- j) **Temperatura Média do Óleo:** ao ligar a ignição, inicia-se a medição das temperaturas do óleo do motor durante o trajeto. Ao desligar a ignição, calcula-se a média dessas leituras, registrando o valor resultante no banco de dados.

Os registros de telemetria são estruturados em formato tabular, contendo valores de sensores, identificadores do veículo, tipo de evento e marcações temporais. Dependendo da categoria do evento, cada tabela pode conter até 43 colunas, que descrevem os atributos associados àquela ocorrência específica. A Tabela 3.2 apresenta um subconjunto dessas colunas, selecionadas com o objetivo de ilustrar o formato dos registros individuais de eventos utilizados nas análises desenvolvidas neste estudo.

Tabela 3.2 – Exemplo de registros de eventos de telemetria disponibilizados pela plataforma utilizada

EventId	EventTypeId	AssetId	StartDateTime	Value
3359839294851540488	6454149451280645233	1344832993491865600	2024-10-15 02:12:25	11
3359854521658685321	6454149451280645233	1344832993491865600	2024-10-15 03:11:30	9
3359915837259624558	6454149451280645233	1344830822972653568	2024-10-15 07:09:27	9
3359933672675634524	6454149451280645233	1344832938072080384	2024-10-15 08:18:39	16
3359956300532526526	6454149451280645233	1344832993491865600	2024-10-15 09:46:28	12
3359963483082793608	6454149451280645233	1344832993491865600	2024-10-15 10:14:20	11

As principais colunas apresentadas na Tabela 3.2 são descritas a seguir:

- a) **EventId:** Identificador único de cada ocorrência de evento registrada no sistema;
- b) **EventTypeId:** Identificador do tipo de evento associado ao registro;
- c) **AssetId:** Identificador do veículo ao qual o evento está associado;
- d) **StartDateTime:** Data e hora de início da ocorrência do evento;
- e) **Value:** Valor numérico associado ao evento, resultante do processamento dos dados dos sensores.

Além disso, o evento denominado “temperatura do motor máxima crítica” está associado a limites de temperatura definidos de acordo com o projeto térmico de cada

modelo de veículo. Para os ônibus *Volkswagen*, o limite adotado é de 90 °C, para os *Iveco*, 96 °C, e para os *Mercedes-Benz*, 105 °C, todos referentes à temperatura do líquido de arrefecimento do motor registrada pelo sistema de telemetria.

Essas diferenças refletem as características de projeto dos motores e de seus sistemas de arrefecimento, bem como o padrão de emissões atendido por cada modelo, uma vez que motores mais modernos, como os Euro V, operam tipicamente em faixas térmicas mais elevadas do que motores de gerações anteriores, como os Euro III. Quando a temperatura medida ultrapassa o respectivo limite, o sistema registra automaticamente a ocorrência do evento no banco de dados e emite um alerta aos gestores da frota.

Outra base de dados utilizada neste trabalho foi o relatório de viagens (*Trips*), que consolida informações agregadas de cada viagem realizada pelos veículos da frota. Cada registro corresponde a uma viagem individual, definida como o intervalo compreendido entre o acionamento e o desligamento da ignição do veículo.

O conjunto de dados disponibilizado nesse relatório é composto por 48 colunas, contemplando informações temporais, operacionais e de identificação dos veículos. Para fins de ilustração do formato dos registros utilizados nesta pesquisa, a Tabela 3.3 apresenta um subconjunto dessas colunas.

Tabela 3.3 – Exemplo de registros de viagens (*trips*) disponibilizados pela plataforma de telemetria

TripId	AssetId	TripStart	TripEnd	Distance (km)
3470051509188821585	1589151677707223040	2025-08-08 02:12:06	2025-08-08 02:13:34	0.125
3470051906857080330	1593869423028727808	2025-08-08 02:13:38	2025-08-08 02:14:43	0.099
3470052179625560713	1589150083245199360	2025-08-08 02:14:42	2025-08-08 02:18:47	0.202
3470052396651373900	1593869423028727808	2025-08-08 02:15:33	2025-08-08 02:49:29	0.167
3470052476660025644	1593871154664407040	2025-08-08 02:15:51	2025-08-08 03:02:43	0.079

As principais colunas consideradas na Tabela 3.3 são descritas a seguir:

- a) **TripId**: Identificador único de cada viagem registrada no sistema;
- b) **AssetId**: Identificador do veículo associado à viagem;
- c) **TripStart**: Data e hora de início da viagem;
- d) **TripEnd**: Data e hora de término da viagem;
- e) **DistanceKilometers**: Distância total percorrida pelo veículo durante a viagem.

Com base nos horários de início e término de cada viagem, realizou-se o processamento dos relatórios de eventos, de modo a cruzar os dados e identificar quais eventos

ocorreram durante o intervalo correspondente a cada viagem e a cada veículo. Dessa forma, foi possível associar cada evento à respectiva viagem, enriquecendo o relatório de viagens com informações adicionais sobre a ocorrência de eventos críticos e parâmetros operacionais relevantes.

3.2.2 Dados de Manutenção (Ordens de Serviço)

Além dos dados de telemetria, esta pesquisa utilizou registros do sistema de gestão de manutenção da empresa responsável pela frota, referentes às Ordens de Serviço (OS). Esses registros documentam inspeções e intervenções realizadas nos veículos ao longo do período analisado.

As ordens de serviço foram utilizadas para identificar veículos associados a episódios de superaquecimento do motor e as respectivas datas de registro. A data de abertura da OS foi adotada como referência temporal para a rotulação das viagens analisadas, considerando horizontes de 15, 10 e 5 dias ou de 100, 50 e 20 viagens anteriores. Esse procedimento permitiu classificar viagens potencialmente problemáticas para o treinamento e a avaliação dos modelos de aprendizado de máquina.

Os dados de manutenção estão organizados em tabelas compostas por 35 colunas. A Tabela 3.4 apresenta um subconjunto dessas colunas, selecionadas para ilustrar o formato dos registros utilizados neste estudo.

Tabela 3.4 – Exemplo de registros de ordens de serviço utilizados no estudo

Numero da OS	Codigo do Veiculo	Data de Abertura da OS	Descricao do Servico	Complemento do Servico
1341940	50749	2024-01-06 13:49:57	Motor Esquentando	Trocada a válvula termostática. Sistema de arrefecimento verificado e veículo mantido em teste.
1340762	50732	2024-01-03 15:36:52	Motor Esquentando	Observação Totem:
1340753	50732	2024-01-03 14:27:55	Alerta de temperatura do motor – Telemetria	Substituição do radiador de água e da válvula termostática, com limpeza completa do sistema de arrefecimento.
1665086	50742	2025-09-05 18:24:52	Motor Esquentando	Identificado vazamento de água no terceiro selo do motor. Executado reparo paliativo e programada a substituição definitiva do componente.
1665064	50742	2025-09-05 17:02:14	Alerta de temperatura do motor – Telemetria	98 °C.

As principais colunas apresentadas na Tabela 3.4 são:

- a) **Numero da OS:** Identificador único da ordem de serviço;
- b) **Codigo do Veiculo:** Identificador do veículo;
- c) **Data de Abertura da OS:** Marco temporal adotado nas análises;
- d) **Descricao do Servico:** Descrição do sintoma registrado;
- e) **Complemento do Servico:** Descrição técnica do desfecho da inspeção.

3.3 Processamento dos Dados

Esta seção descreve o fluxo de processamento e pré-processamento dos dados adotado nesta pesquisa para a construção do conjunto de dados utilizado na modelagem

preditiva. O procedimento seguiu uma sequência estruturada de etapas, desde a integração das bases de dados até a classificação das viagens críticas, conforme ilustrado na Figura 3.5.

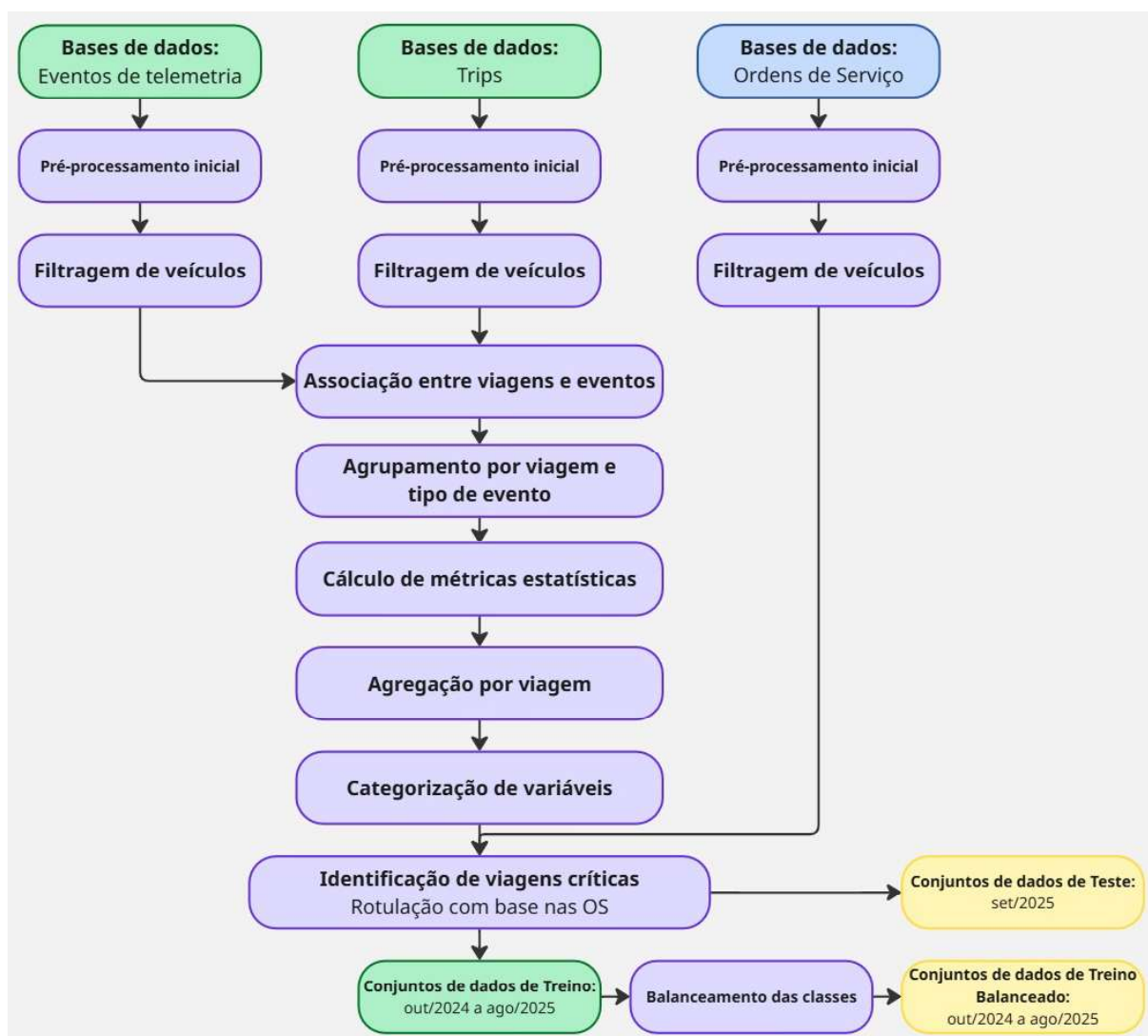


Figura 3.5 – Fluxo de processamento e pré-processamento dos dados adotado na pesquisa

3.3.1 Pré-processamento inicial e Filtragem de veículos

As etapas iniciais de processamento envolveram a integração das bases de dados de telemetria, viagens e ordens de serviço, bem como a remoção de registros duplicados, a padronização dos identificadores dos veículos e a verificação da consistência dos valores associados aos eventos de telemetria. Adicionalmente, foi aplicado um filtro temporal para restringir a análise ao período compreendido entre outubro de 2024 e setembro de 2025, conforme definido para este estudo.

Conforme discutido na Seção 3.1, a análise foi restrita aos veículos do modelo *Iveco 170S28*. Dessa forma, os conjuntos de dados foram previamente filtrados para incluir exclusivamente registros associados a esse modelo veicular.

Na etapa de pré-processamento, as tabelas de eventos, viagens e ordens de serviço foram tratadas para padronização do formato de data e hora e conversão dos valores para tipos compatíveis, de modo a assegurar a integração consistente das diferentes fontes de dados utilizadas nesta pesquisa.

3.3.1.1 Filtragem de Viagens Válidas

Os registros de viagem são gerados a partir do acionamento e desligamento da ignição do veículo e podem incluir deslocamentos que não representam a operação regular da frota, como movimentações internas em garagens, terminais, testes realizados em oficina ou períodos prolongados de ignição indevidamente agregados pelo sistema.

Com o objetivo de considerar predominantemente ciclos operacionais representativos da circulação dos ônibus em linhas regulares do sistema de transporte público, foram definidos critérios de filtragem baseados na análise estatística exploratória da duração e da distância das viagens registradas. Observou-se forte assimetria e elevada dispersão em ambas as variáveis, com valores extremos incompatíveis com a operação urbana. Por exemplo, embora a mediana da distância por viagem seja de aproximadamente 10,8 km, há registros superiores a 700 km, que não correspondem a deslocamentos reais de ônibus em serviço. De forma análoga, a duração apresentou valores máximos de vários dias, decorrentes de falhas de agregação de ignição ou de permanência do veículo ligado em pátios e oficinas.

Com base nesse diagnóstico, foram mantidas apenas as viagens com duração igual ou superior a 15 minutos e distância percorrida entre 5 e 200 km. O limite inferior de 5 km e 15 minutos permite excluir manobras internas, deslocamentos de curta duração e testes operacionais, enquanto o limite superior de 200 km preserva praticamente toda a faixa interquartil das viagens operacionais, removendo apenas registros extremos e estatisticamente atípicos. Viagens fora desses limites foram descartadas por não caracterizarem trajetos operacionais típicos ou por constituírem valores incompatíveis com a dinâmica do transporte coletivo urbano.

3.3.2 Associação entre Viagens e Eventos

Após as etapas de extração e pré-processamento, realizou-se a associação entre os registros de eventos de telemetria e as respectivas viagens. A plataforma de telemetria disponibiliza 34 tipos distintos de eventos, cada um armazenado em uma tabela específica, os quais descrevem diferentes aspectos do comportamento do veículo e de seus sistemas, como aceleração, temperatura, pressão, rotação, entre outros.

Cada registro de evento possui um identificador próprio (*EventId*) e informações temporais, mas não inclui explicitamente o identificador da viagem (*TripId*). Por essa

razão, tornou-se necessária uma etapa de vinculação entre os bancos de eventos e o banco de viagens com base na correspondência temporal.

Cada viagem é definida pelo intervalo de tempo compreendido entre o acionamento e o desligamento da ignição do veículo. Assim, para cada ônibus, os instantes de ocorrência dos eventos foram comparados com os horários de início e término de cada viagem, de modo que todo evento cujo carimbo temporal estivesse contido dentro de um intervalo de viagem fosse atribuído ao respectivo *TripId*. Esse procedimento está ilustrado na Figura 3.6, na qual eventos de diferentes tipos são associados às viagens a partir de sua sobreposição temporal.

Figura 3.6 – Associação entre eventos de telemetria e viagens por meio de correspondência temporal.

Hora	07h	08h	09h	10h	11h	12h	13h	14h	15h	16h
VIAGENS	Tripld: 3443625521909237592				Tripld: 3443518514437189179					
Aceleração Brusca										
Marcha Lenta										
Pressão Média do Óleo										
Temperatura Média do Motor										
Temperatura Média do Óleo										

(a) Eventos e viagens registrados de forma independente, sem vínculo explícito entre si.

Hora	07h	08h	09h	10h	11h	12h	13h	14h	15h	16h
VIAGENS	Tripld: 3443625521909237592				Tripld: 3443518514437189179					
Aceleração Brusca										
Marcha Lenta										
Pressão Média do Óleo										
Temperatura Média do Motor										
Temperatura Média do Óleo										

(b) Eventos associados às viagens a partir da sobreposição temporal dos registros.

Esse processo permite converter uma base de eventos independentes no tempo em um conjunto estruturado de informações por viagem, garantindo que todas as ocorrências de aceleração brusca, marcha lenta, temperatura e pressão registradas durante um percurso específico sejam corretamente vinculadas a ele.

3.3.3 Agrupamento, cálculo de métricas e agregação dos eventos por viagem

Uma vez associados aos respectivos *TripId*, os eventos foram agrupados por identificador de viagem e por tipo de evento (*EventTypeId*). Esse agrupamento permite distinguir as diferentes grandezas monitoradas, como temperatura média do motor, pressão do óleo, ocorrência de acelerações bruscas e períodos em marcha lenta.

Para cada par (*TripId*, *EventTypeId*), foram calculadas estatísticas descritivas que sintetizam o comportamento do veículo ao longo da viagem. As métricas extraídas incluem média, desvio padrão, número de ocorrências, valores mínimo e máximo dos registros de cada evento. A Figura 3.7 ilustra esse processo, no qual múltiplos eventos ao longo de uma viagem são convertidos em um conjunto compacto de atributos representativos.

Figura 3.7 – Agregação dos eventos por viagem e construção das variáveis analíticas.

Hora	11h	12h	13h	14h	15h	16h
VIAGENS	Tripld: 3443518514437189179					
Aceleração Brusca	Tripld, EventId e Value				Tripld, EventId e Value	
Marcha Lenta			Tripld, EventId e Value			
Pressão Média do Óleo	Tripld, EventId e Value					
Temperatura Média do Motor	Tripld, EventId e Value			Tripld, EventId e Value		
Temperatura Média do Óleo	Tripld, EventId e Value					

(a) Eventos associados à viagem, ainda representados como registros individuais.

Hora	11h	12h	13h	14h	15h	16h
VIAGENS	Tripld: 3443518514437189179					
Aceleração Brusca	Tripld, Média, Desvio Padrão, Ocorrência, Máximo e Mínimo					
Marcha Lenta	Tripld, Média, Desvio Padrão, Ocorrência, Máximo e Mínimo					
Pressão Média do Óleo	Tripld, Média, Desvio Padrão, Ocorrência, Máximo e Mínimo					
Temperatura Média do Motor	Tripld, Média, Desvio Padrão, Ocorrência, Máximo e Mínimo					
Temperatura Média do Óleo	Tripld, Média, Desvio Padrão, Ocorrência, Máximo e Mínimo					

(b) Eventos agregados por tipo, gerando estatísticas resumidas por viagem.

Essas estatísticas agregadas foram então integradas ao banco de viagens, resultando em um conjunto de dados ampliado no qual cada linha representa uma viagem individual e cada coluna corresponde a uma variável operacional ou a um atributo derivado dos eventos de telemetria. Esse procedimento é esquematizado na Figura 3.8, que mostra a fusão entre o banco de viagens e o banco de eventos, originando o banco de viagens ampliado utilizado como base para o treinamento e avaliação dos modelos preditivos.

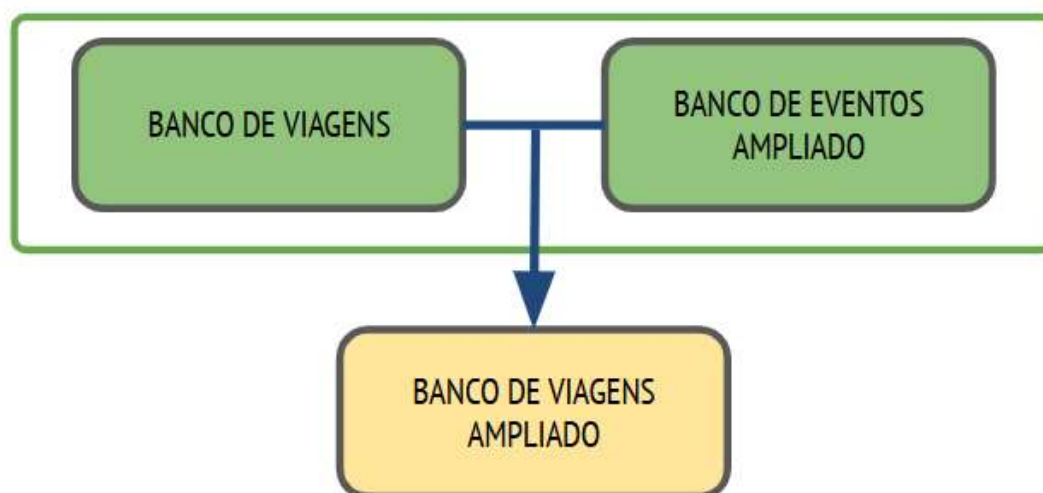


Figura 3.8 – Fluxo de integração entre o banco de viagens e o banco de eventos, resultando no banco de viagens ampliado.

3.3.4 Categorização de Variáveis

Com o objetivo de reduzir a complexidade dos dados de entrada, preservar padrões operacionais relevantes e facilitar a comparação entre viagens realizadas sob diferentes condições de operação, foram criadas variáveis categóricas derivadas a partir de atributos contínuos e temporais do relatório de viagens.

O conjunto de dados agregado por viagem é composto por métricas estatísticas dos eventos de telemetria, além de informações temporais e operacionais. A partir desse conjunto, três novas colunas categóricas foram adicionadas à tabela de viagens, representando regimes operacionais sob os quais cada percurso foi realizado. Essas variáveis categóricas não substituem as variáveis originais, mas atuam como atributos derivados, complementando a descrição de cada viagem e permitindo que os modelos considerem diferenças estruturais de operação.

As variáveis categorizadas foram definidas da seguinte forma:

- a) **Tipo de dia:** classificado como *dia útil*, *fim de semana* ou *feriado*. Essa classificação foi adotada conforme a prática operacional da empresa gestora da frota, que distingue formalmente esses três regimes devido às diferenças sistemáticas de demanda, carga de passageiros, intervalos entre viagens e padrões de tráfego. Embora sábados e domingos apresentem características próprias, ambos são tratados operacionalmente como regime não útil e, por isso, foram agrupados em uma única categoria.
- b) **Horário da viagem:** categorizado como *horário de pico* ou *fora do horário de pico*, de acordo com os intervalos oficialmente definidos pela empresa gestora da frota.

Esses períodos correspondem aos horários de maior demanda e congestionamento, nos quais os veículos operam sob maior carga térmica e mecânica.

- c) **Velocidade média da viagem:** classificada em níveis *baixo*, *médio* ou *alto*, com base na distribuição empírica das velocidades médias observadas no conjunto de dados. Os limiares foram definidos de modo a separar viagens predominantemente congestionadas, viagens em regime urbano regular e viagens com maior fluidez, buscando representar diferentes regimes de esforço do motor ao longo do percurso.

3.3.5 Definição de viagens críticas a partir das ordens de serviço

As viagens foram rotuladas como *críticas* quando se encontravam temporalmente associadas à ocorrência de uma ordem de serviço de superaquecimento do motor, aberta a partir de alertas de temperatura do líquido de arrefecimento registrados pelo sistema de telemetria. Essas ordens de serviço correspondem aos casos em que a temperatura do motor ultrapassou os limites específicos definidos para cada modelo de veículo, conforme descrito na Seção 3.1, caracterizando uma condição de risco operacional.

Todas as demais viagens, que não se enquadraram em nenhuma janela preditiva associada a essas ordens de serviço, foram classificadas como *normais*. Essa rotulação foi registrada no conjunto de dados por meio da variável `Pre_OS`, que assume o valor “sim” para viagens críticas e “não” para viagens normais.

A construção das classes foi realizada a partir de diferentes horizontes preditivos, com o objetivo de avaliar a capacidade dos modelos em antecipar a ocorrência da falha em distintos intervalos de observação. Para cada ordem de serviço de superaquecimento, foram definidos dois tipos de janelas retrospectivas:

- **Horizontes temporais**, correspondentes a intervalos fixos de 5, 10 e 15 dias anteriores à abertura da ordem de serviço;
- **Horizontes baseados em número de viagens**, correspondentes às últimas 20, 50 e 100 viagens realizadas pelo veículo antes da ocorrência.

A Figura 3.9 apresenta a sequência de viagens de um veículo representada em uma escala temporal relativa à abertura da ordem de serviço associada ao superaquecimento do motor. Nessa representação, o instante da abertura da ordem de serviço é definido como tempo zero, permitindo identificar de forma clara as viagens realizadas antes e depois do evento. Os valores negativos correspondem às viagens anteriores à falha, enquanto os valores positivos indicam viagens posteriores, facilitando a definição dos horizontes preditivos adotados neste estudo.

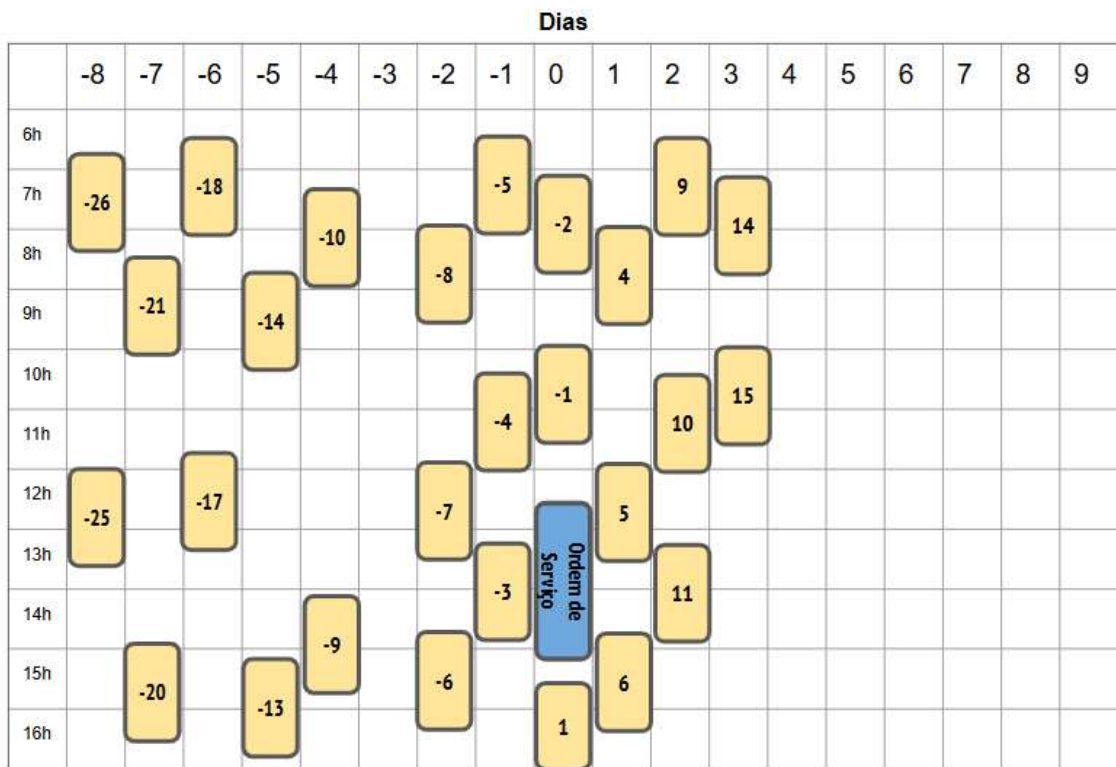


Figura 3.9 – Representação das viagens em escala temporal relativa à ordem de serviço, considerada como tempo zero. Os valores negativos indicam viagens anteriores à falha, enquanto os positivos indicam viagens posteriores.

A Figura 3.10 apresenta os diferentes horizontes utilizados neste estudo. A partir deles, as viagens localizadas dentro das janelas preditivas foram rotuladas como críticas, enquanto aquelas fora dos limites dos horizontes foram rotuladas como normais. Esse processo é exemplificado nas Figuras 3.11 e 3.12, que ilustram, respectivamente, a aplicação de um horizonte de 5 dias e de um horizonte de 20 viagens sobre a mesma sequência de percursos.



Figura 3.10 – Horizontes preditivos considerados na pesquisa, definidos em termos de janelas temporais (5, 10 e 15 dias) e número de viagens (20, 50 e 100 viagens anteriores à ordem de serviço).

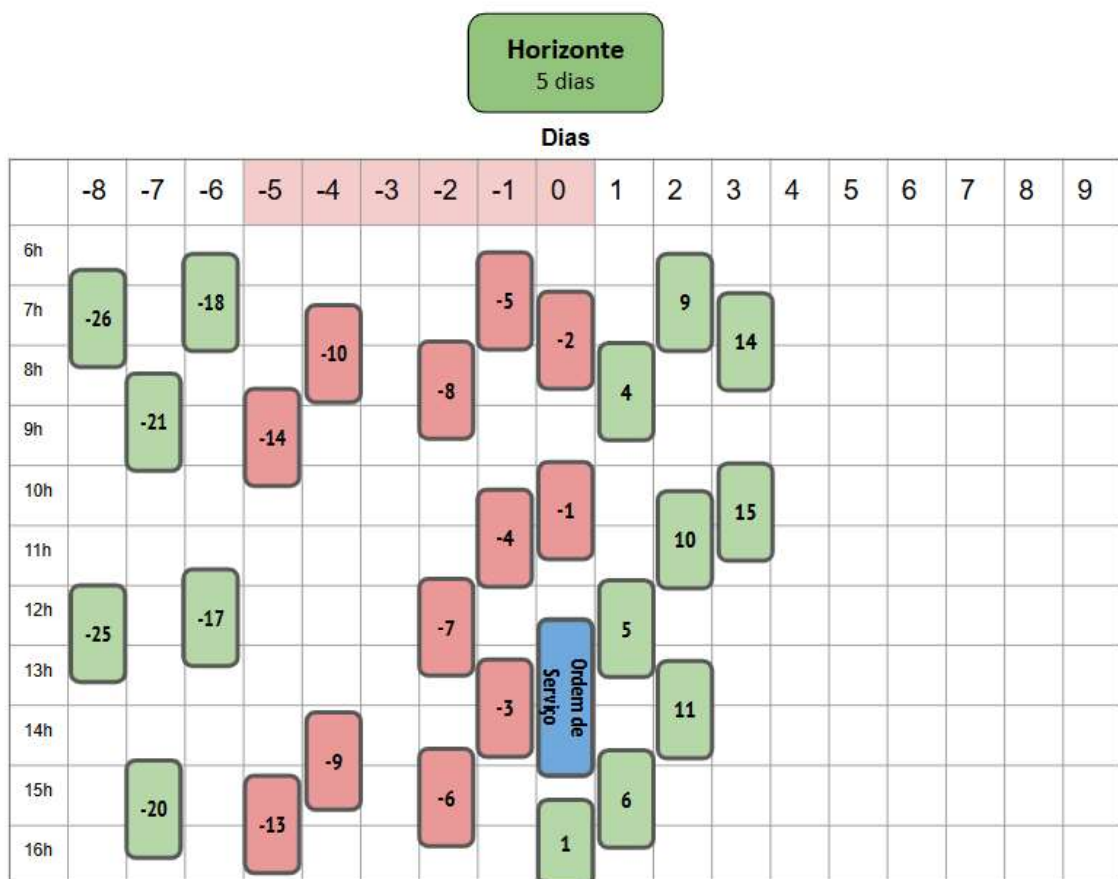


Figura 3.11 – Exemplo da aplicação de um horizonte temporal de 5 dias. As viagens destacadas em vermelho correspondem às que foram rotuladas como críticas por estarem dentro da janela anterior à ordem de serviço, enquanto as demais permanecem como normais.

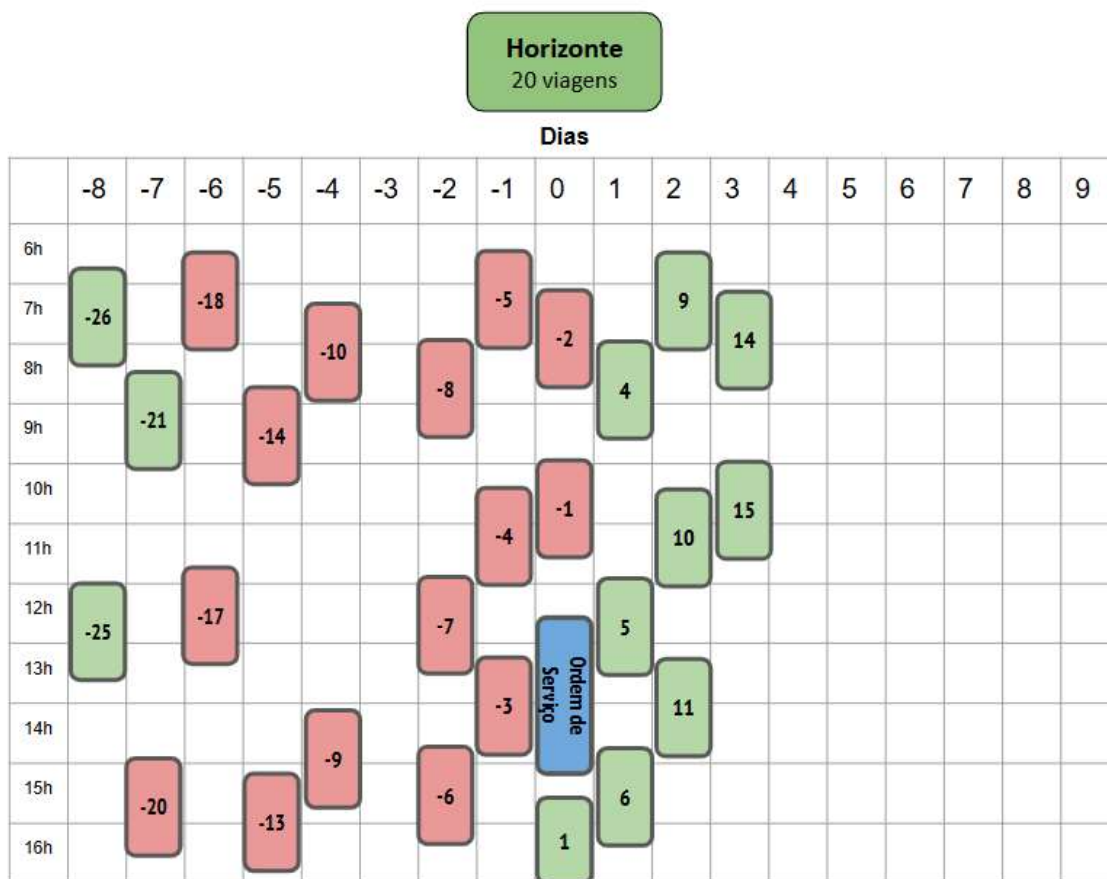


Figura 3.12 – Exemplo da aplicação de um horizonte baseado nas 20 viagens anteriores à ordem de serviço. As viagens destacadas em vermelho foram rotuladas como críticas, por estarem mais próximas da ocorrência do superaquecimento.

Esse procedimento resulta em múltiplos conjuntos de dados derivados, um para cada horizonte considerado, permitindo analisar de forma sistemática como a distância temporal ou operacional em relação à falha afeta o desempenho dos modelos preditivos.

3.3.6 Conjunto de Dados Agregado por Viagem e Particionamento

Após a execução das etapas de processamento descritas anteriormente, obteve-se um conjunto de dados agregado por viagem, no qual cada linha representa uma viagem individual identificada por um *TripId*. Esse conjunto integra métricas operacionais do relatório de viagens, estatísticas derivadas dos eventos de telemetria e informações temporais associadas ao início e término de cada percurso.

A Tabela 3.5 apresenta uma amostra reduzida desse conjunto de dados, com o objetivo de ilustrar a estrutura final adotada para a modelagem preditiva.

Tabela 3.5 – Exemplo de conjunto de dados agregado por viagem

TripId	TripStart	Distância (km)	Temp. Média (°C)	Alertas Críticos	Pre_OS
3363989834053116900	26/10/2024 06:38	24.92	86.4	0	não
3364020867388984300	26/10/2024 08:39	18.62	87.1	0	não
3364044037608589300	26/10/2024 10:08	37.51	88.3	0	não
3364118415046915100	26/10/2024 14:57	53.49	91.6	1	sim
3364162982267287600	26/10/2024 17:50	160.28	92.8	2	sim
3364267629835227100	27/10/2024 00:36	82.15	93.4	1	sim

Para a etapa de modelagem, o conjunto de dados foi particionado com base em critério temporal. O conjunto de treinamento compreendeu as viagens realizadas entre outubro de 2024 e agosto de 2025. O conjunto de teste foi composto exclusivamente pelas viagens registradas em setembro de 2025, mantido integralmente sem aplicação de técnicas de balanceamento, de modo a representar as condições reais de operação da frota.

3.3.7 Balanceamento das Classes no Conjunto de Treinamento

A análise inicial dos dados evidenciou um desbalanceamento significativo entre as classes, com predominância de viagens normais em relação às viagens classificadas como críticas. Esse cenário é recorrente em problemas de manutenção preditiva, nos quais eventos de falha representam uma parcela reduzida do histórico operacional.

Conjuntos de dados fortemente desbalanceados tendem a induzir viés nos modelos de classificação, uma vez que muitos algoritmos buscam minimizar o erro global, favorecendo a predição da classe majoritária. Como consequência, modelos treinados nessas condições podem apresentar elevada exatidão aparente, mas desempenho insatisfatório na identificação da classe minoritária, que neste estudo corresponde às viagens associadas ao padrão de superaquecimento do motor.

Diante desse cenário, adotou-se uma estratégia de balanceamento aplicada exclusivamente ao conjunto de treinamento, com base em pareamento temporal das viagens. Para cada viagem classificada como crítica, foi mantida uma única viagem normal realizada no mesmo dia e em faixa horária equivalente. As demais viagens normais, que não possuíam correspondência temporal direta com eventos críticos, foram descartadas do conjunto de treinamento.

Essa estratégia teve como objetivo reduzir o viés associado ao desbalanceamento das classes, ao mesmo tempo em que preservou condições operacionais comparáveis entre as observações utilizadas no treinamento dos modelos, como período do dia, demanda típica e contexto operacional da frota. A Figura 3.13 ilustra a distribuição original das viagens normais e críticas ao longo do dia para diferentes veículos, evidenciando a predominância

da classe normal. Já a Figura 3.14 apresenta o resultado do balanceamento adotado, destacando a seleção de viagens normais pareadas temporalmente com as viagens críticas.

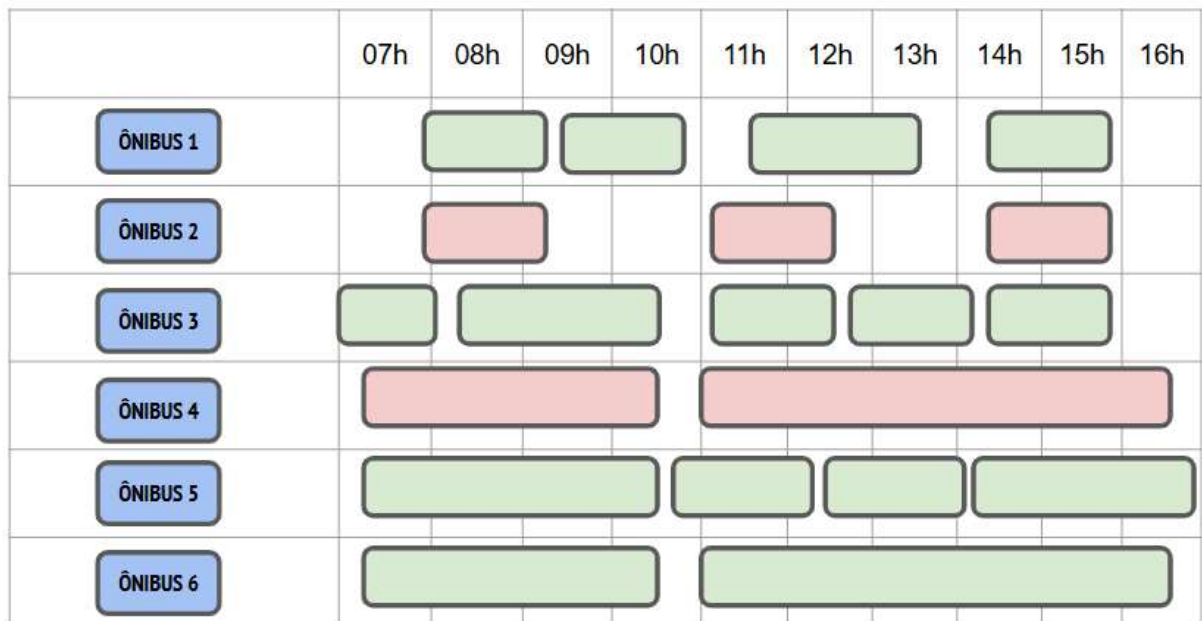


Figura 3.13 – Distribuição original das viagens normais e críticas ao longo do tempo e entre diferentes veículos, evidenciando o desbalanceamento entre as classes no conjunto de dados.

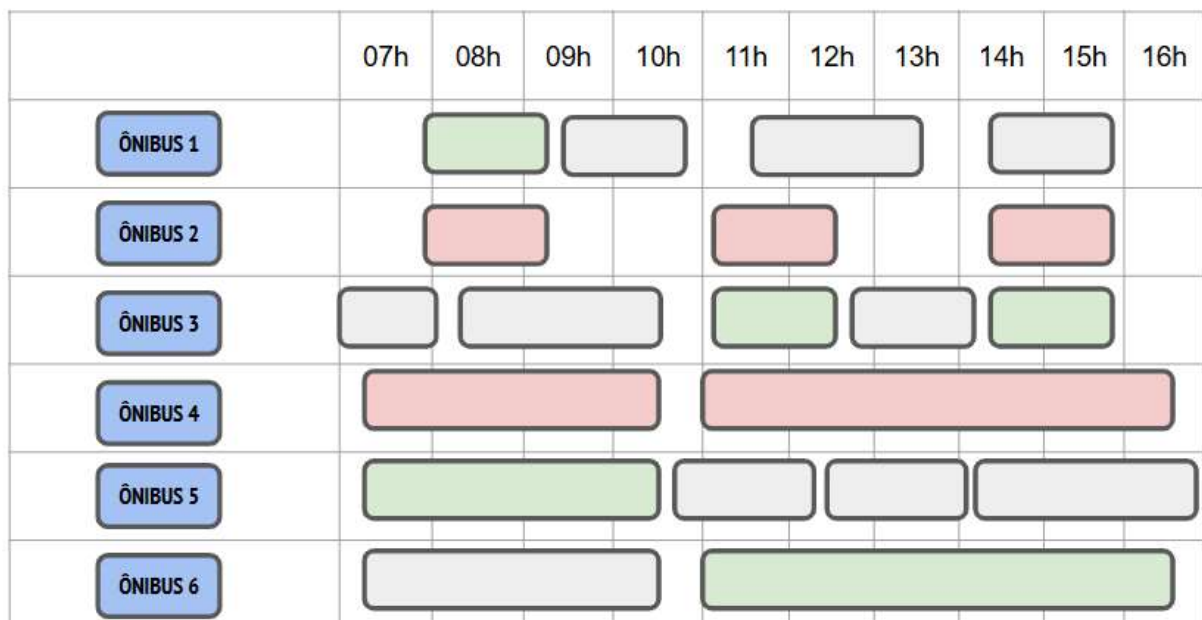


Figura 3.14 – Resultado do balanceamento das classes no conjunto de treinamento por meio de pareamento temporal, com a seleção de viagens normais realizadas no mesmo dia e em faixas horárias equivalentes às viagens críticas.

Ressalta-se que esse procedimento não foi aplicado ao conjunto de teste, o qual manteve a distribuição original das classes. Dessa forma, assegura-se que a avaliação do

desempenho dos modelos reflita um cenário operacional realista, caracterizado por eventos críticos raros em relação ao volume total de viagens monitoradas.

Após o particionamento dos dados, os conjuntos de treinamento e teste foram exportados em formato `.csv` e utilizados no ambiente *Orange Data Mining*, ferramenta adotada para a implementação, treinamento e avaliação dos modelos de aprendizado de máquina neste estudo.

3.4 Identificação e Seleção de Variáveis

Após as etapas de processamento e agregação dos eventos de telemetria por viagem, o conjunto de dados resultante passou a conter aproximadamente 198 variáveis. Essas variáveis incluem tanto medições diretas provenientes dos sensores quanto métricas estatísticas derivadas dos eventos registrados ao longo de cada percurso.

Antes da aplicação de métodos estatísticos de seleção de variáveis, realizou-se uma etapa preliminar de limpeza estrutural das variáveis de entrada. Nessa etapa, foram removidos identificadores de viagem, veículo e motorista, bem como campos de natureza administrativa e temporal que não representam fenômenos físicos ou operacionais do veículo. A exclusão dessas variáveis teve como objetivo evitar que os modelos de aprendizado capturassem padrões específicos do sistema de registro ou da operação, em vez de relações associadas a um padrão de superaquecimento do motor.

Essa decisão metodológica visa reduzir vieses indesejados e prevenir vazamento de informação, preservando a capacidade de generalização dos modelos preditivos para cenários distintos daqueles observados no conjunto de treinamento.

Após a limpeza estrutural, aplicou-se uma etapa de seleção de variáveis baseada em métodos do tipo filtro, com o objetivo de reduzir a dimensionalidade do conjunto de dados e priorizar atributos mais informativos em relação à variável-alvo. Esse tipo de abordagem foi escolhido por sua independência em relação ao algoritmo de aprendizado utilizado, pela transparência do processo de avaliação e pelo baixo custo computacional, entendido como menor complexidade algorítmica e maior eficiência na realização de múltiplos experimentos comparativos.

A avaliação da relevância das variáveis foi realizada por meio do *widget Rank*, disponível no ambiente *Orange Data Mining*. Esse recurso permite ranquear os atributos de acordo com diferentes critérios estatísticos de associação com a variável-alvo, definida neste estudo como a classificação das viagens em críticas ou normais quanto a antecedência de superaquecimento do motor.

Foram consideradas as seguintes métricas de seleção:

1. **Índice de Gini**, que avalia a capacidade de uma variável em reduzir a impureza das partições dos dados;
2. **Qui-quadrado** (χ^2), que mede o grau de dependência estatística entre variáveis categóricas;
3. **Ganho de Informação**, que quantifica a redução da incerteza da variável-alvo ao se conhecer o valor de um atributo;
4. **Razão de Ganho**, que normaliza o ganho de informação, penalizando variáveis com grande número de categorias;
5. **ReliefF**, que estima a relevância dos atributos com base na capacidade de distinguir instâncias próximas pertencentes a classes diferentes.

A utilização de múltiplos critérios de ranqueamento permitiu avaliar a consistência da relevância das variáveis, reduzindo a dependência de um único método de seleção. A partir dos resultados obtidos, foram definidos subconjuntos de variáveis que serviram de base para os experimentos de modelagem preditiva apresentados no capítulo seguinte.

A lista completa das variáveis consideradas nessa etapa, bem como seus respectivos escores segundo cada métrica de seleção, é apresentada no Apêndice X.

3.5 Escolha e Configuração dos Algoritmos de Aprendizado de Máquina

Nesta etapa, definiram-se os algoritmos de aprendizado de máquina utilizados para a construção e avaliação dos modelos preditivos. Foram selecionados três métodos amplamente empregados em problemas de classificação supervisionada com dados tabulares: Floresta Aleatória, Máquinas de Vetores de Suporte e Redes Neurais Artificiais, conforme discutido no Capítulo 2.

A escolha desses algoritmos fundamenta-se em suas características complementares. A Floresta Aleatória representa uma abordagem baseada em conjuntos de árvores de decisão, robusta a ruídos e capaz de lidar com variáveis heterogêneas e relações não lineares. As SVM são reconhecidas por sua eficácia em cenários de alta dimensionalidade e pela flexibilidade proporcionada pelo uso de diferentes funções de *kernel*. Já as Redes Neurais Artificiais permitem modelar relações complexas entre variáveis por meio de transformações não lineares sucessivas, sendo amplamente utilizadas em tarefas de classificação com múltiplos atributos.

O objetivo da modelagem adotada neste estudo não foi a otimização exaustiva de hiperparâmetros, mas a avaliação do comportamento desses algoritmos quando aplicados

a dados reais de telemetria comercial, considerando diferentes subconjuntos de variáveis e horizontes preditivos. Essa escolha metodológica está alinhada ao caráter exploratório da pesquisa e ao interesse em compreender as limitações e potencialidades dos modelos em condições próximas à prática operacional.

Por esse motivo, adotou-se como ponto de partida a configuração padrão disponibilizada pelo ambiente *Orange Data Mining*, a qual representa um uso típico dessas técnicas em contextos aplicados e operacionais. Essa configuração foi utilizada como linha de base para comparação entre os algoritmos.

Adicionalmente, para cada algoritmo foi considerada uma segunda configuração, envolvendo pequenas variações em hiperparâmetros considerados relevantes. O objetivo dessas variações foi verificar a sensibilidade dos resultados a mudanças pontuais na parametrização, permitindo avaliar se o desempenho observado estava associado a uma configuração específica ou se refletia limitações mais estruturais do conjunto de dados e da representação adotada.

No caso das Máquinas de Vetores de Suporte, foram exploradas variações no tipo de *kernel* e no parâmetro de regularização C . O *kernel* define a forma como os dados são projetados em um espaço de maior dimensionalidade, possibilitando a separação de padrões não linearmente separáveis no espaço original. O parâmetro C controla o compromisso entre maximizar a margem de separação e penalizar erros de classificação, influenciando a complexidade do modelo e sua capacidade de generalização.

Para as Redes Neurais Artificiais, investigaram-se variações na função de ativação, no método de otimização, no número máximo de iterações e no valor do parâmetro de regularização. Optou-se por uma arquitetura de camada oculta única com 50 neurônios, escolha compatível com a natureza tabular dos dados e com o objetivo de evitar modelos excessivamente complexos, que poderiam levar ao sobreajuste. Essa configuração buscou um equilíbrio entre capacidade de representação e estabilidade do treinamento, considerando o tamanho do conjunto de dados disponível.

No algoritmo de Floresta Aleatória, foram consideradas variações na profundidade máxima das árvores e nos critérios mínimos de divisão dos nós. Esses parâmetros influenciam diretamente a complexidade dos modelos individuais e o grau de generalização do conjunto, permitindo avaliar o impacto do controle de complexidade no desempenho preditivo.

A adoção conjunta de configurações padrão e ajustadas permitiu avaliar a robustez dos algoritmos frente a alterações controladas de hiperparâmetros, bem como identificar tendências de desempenho consistentes ao longo dos diferentes experimentos realizados. As configurações utilizadas em cada experimento são apresentadas nas Tabelas 3.6, 3.7 e 3.8.

Tabela 3.6 – Configurações de parâmetros para *Random Forest*.

Parâmetro	Config 1	Config 2
Nº Árvores	100	100
Máx. Profundidade	Sem limite	5
Mín. Amostras p/ Divisão	Sem restrição	10
Balanceamento de Classes	Não	Não
Treino Replicável	Sim	Sim

Tabela 3.7 – Configurações de parâmetros para Redes Neurais.

Parâmetro	Config 1	Config 2
Neurônios	50	50
Ativação	ReLU	Tanh
Solver	Adam	SGD
Regularização (α)	0.01	0.05
Iterações	250	500
Replicável	Sim	Sim

Tabela 3.8 – Configurações de parâmetros para SVM.

Parâmetro	Config 1	Config 2
Kernel	Sigmoid	Linear
Custo (C)	0.10	1.00
Kernel g	auto	—
Kernel c	1.00	—
Tolerância Numérica	0.0010	0.0010
Limite Iterações	1000	1000

A avaliação do desempenho dos modelos foi realizada por meio de métricas de classificação adequadas a cenários com desbalanceamento entre classes, incluindo AUC,

F1-score, precisão, *recall*, exatidão e o coeficiente de correlação de Matthews (MCC), conforme detalhado no capítulo de resultados.

3.6 Avaliação e Validação do Modelo Preditivo

A avaliação e validação dos modelos preditivos teve como objetivo analisar a capacidade dos algoritmos em identificar viagens associadas à falha de superaquecimento do motor, considerando diferentes horizontes preditivos, estratégias de seleção de variáveis e configurações de modelos. Para isso, o conjunto de treinamento foi organizado em seis conjuntos de dados distintos, definidos a partir de janelas temporais anteriores à abertura da ordem de serviço (15, 10 e 5 dias) e do número de viagens anteriores consideradas (100, 50 e 20 viagens).

A Figura 3.15 apresenta o *workflow* implementado no ambiente *Orange Data Mining*, ilustrando o encadeamento entre os conjuntos de dados, as estratégias de seleção de variáveis, os algoritmos de classificação e os módulos de avaliação de desempenho.

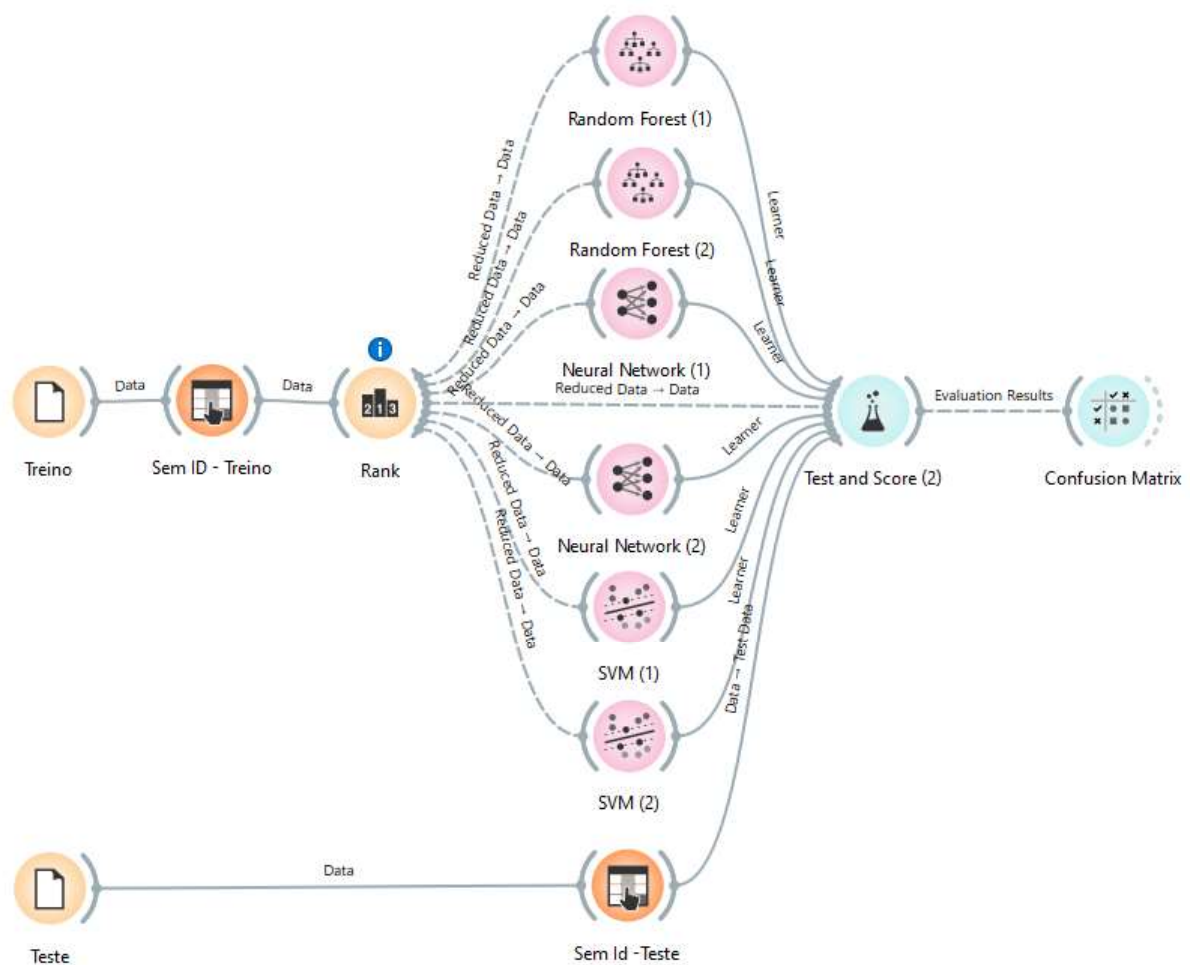


Figura 3.15 – *Workflow* utilizado para treinamento e avaliação dos modelos no *Orange Data Mining*.

Para cada conjunto de dados, aplicou-se o procedimento de seleção de variáveis descrito na Seção 3.4. Os critérios estatísticos *Gini Index*, *Chi-Square*, *Information Gain*, *Gain Ratio* e *ReliefF* foram utilizados como filtros independentes, com o objetivo de gerar diferentes subconjuntos de variáveis de entrada. Esses filtros não constituem resultados em si, mas representam condições experimentais distintas, permitindo avaliar a sensibilidade dos modelos à escolha do método de seleção de atributos.

Para cada critério de seleção, foram escolhidos os trinta atributos mais bem ranqueados, originando subconjuntos alternativos de variáveis que alimentaram os modelos preditivos. Adicionalmente, considerou-se um cenário de referência no qual todas as variáveis disponíveis foram utilizadas, sem aplicação de filtro estatístico.

Concluída a etapa de seleção de variáveis, os algoritmos de aprendizado de máquina — Floresta Aleatória, Máquinas de Vetores de Suporte e Redes Neurais Artificiais — foram treinados utilizando duas configurações distintas para cada técnica, conforme descrito na Seção 3.5. Cada configuração foi tratada como um modelo independente, totalizando seis modelos avaliados para cada combinação de horizonte preditivo e estratégia de seleção de variáveis.

A avaliação dos modelos foi conduzida por meio do *widget Test & Score*, que confronta os modelos treinados com o conjunto de teste previamente definido, sem aplicação de validação cruzada ou reamostragem, preservando assim a separação temporal entre dados de treinamento e teste. O desempenho foi avaliado utilizando as métricas AUC, exatidão, *F1-score*, precisão, *recall* e coeficiente de correlação de Matthews (MCC), considerando como classe positiva as viagens críticas.

Adicionalmente, o módulo *Confusion Matrix* foi empregado para analisar a distribuição de acertos e erros de classificação, permitindo uma avaliação detalhada do comportamento dos modelos na identificação de viagens críticas.

O conjunto de experimentos conduzidos nesta etapa, envolvendo diferentes horizontes preditivos, critérios de seleção de variáveis e configurações de modelos, é sintetizado no diagrama apresentado na Figura 3.16. Esse diagrama representa o desenho experimental adotado e explicita as combinações avaliadas ao longo do estudo.

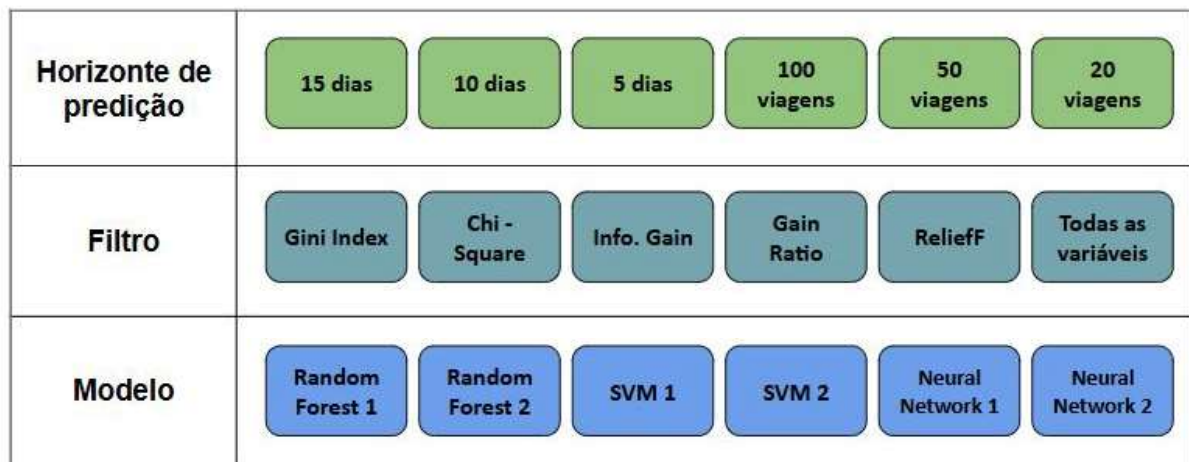


Figura 3.16 – Diagrama da etapa de avaliação dos modelos de aprendizado de máquina.

Considerando a combinação entre os diferentes elementos do desenho experimental adotado, o conjunto total de experimentos realizados neste estudo resulta da interação entre: (i) seis horizontes preditivos distintos (15, 10 e 5 dias, bem como 100, 50 e 20 viagens anteriores), (ii) seis estratégias de seleção de variáveis (cinco filtros estatísticos e um cenário sem aplicação de filtro) e (iii) seis modelos preditivos, correspondentes a duas configurações para cada um dos três algoritmos avaliados. Dessa forma, foram conduzidos ao todo $6 \times 6 \times 6 = 216$ experimentos independentes, cujos resultados são analisados de forma sistemática no capítulo seguinte.

4 Resultados e Discussões

Este capítulo apresenta e discute os resultados obtidos a partir da aplicação dos modelos de aprendizado de máquina aos diferentes cenários experimentais definidos na metodologia. O objetivo central é avaliar a capacidade dos modelos em identificar viagens associadas ao risco de superaquecimento do motor, analisando como o desempenho preditivo é influenciado por três elementos fundamentais: o horizonte preditivo adotado, o método de seleção de variáveis e as configurações dos algoritmos de classificação.

Os experimentos foram conduzidos no ambiente *Orange Data Mining*, utilizando seis conjuntos distintos de dados derivados de diferentes horizontes preditivos. Esses horizontes foram definidos tanto em termos temporais, considerando janelas de observação de 15, 10 e 5 dias anteriores à abertura de uma ordem de serviço associada ao superaquecimento do motor, quanto em termos operacionais, com base em históricos compostos pelas 100, 50 e 20 viagens imediatamente anteriores a esse evento. Cada conjunto de dados foi submetido a cinco critérios de seleção de atributos, a saber, *Gini Index*, *Chi-Square*, *Information Gain*, *Gain Ratio* e *ReliefF*. Em seguida, foram treinados três algoritmos de aprendizado supervisionado, Floresta Aleatória, Máquinas de Vetores de Suporte e Redes Neurais Artificiais, cada um avaliado sob duas configurações distintas de parâmetros. Esse arranjo experimental resultou em um total de 216 combinações, permitindo uma análise abrangente do comportamento dos modelos frente a variações no horizonte preditivo, na seleção de variáveis e nas estratégias de modelagem.

As análises apresentadas nas seções subsequentes foram organizadas de forma progressiva, iniciando pela caracterização das variáveis e dos atributos selecionados, passando pela avaliação do efeito do horizonte preditivo, pela comparação entre algoritmos e pelo impacto da seleção de variáveis no desempenho dos modelos, e culminando em análises globais e na discussão integrada dos resultados. Essa organização busca responder a questões diretamente relacionadas aos objetivos desta pesquisa, conforme descrito a seguir:

- a) **Desempenho preditivo:** em que medida os algoritmos de aprendizado de máquina conseguem identificar viagens críticas associadas ao risco de superaquecimento do motor?
- b) **Padrões e variáveis relevantes:** quais atributos derivados da telemetria contribuem de forma consistente para a discriminação de comportamentos associados à falha, considerando diferentes horizontes preditivos?
- c) **Efeito do horizonte preditivo:** de que forma janelas temporais de 15, 10 e 5 dias,

bem como históricos de 100, 50 e 20 viagens anteriores, influenciam a capacidade dos modelos em detectar sinais antecipados de superaquecimento?

- d) **Comparação entre algoritmos:** quais técnicas apresentam comportamento mais estável ou robusto diante das diferentes condições operacionais representadas nos cenários avaliados?
- e) **Aplicabilidade prática:** os resultados obtidos indicam potencial de uso operacional como ferramenta de apoio à manutenção da frota, considerando as limitações inerentes aos dados de telemetria disponíveis?

A abordagem adotada neste capítulo vai além da simples apresentação numérica dos resultados. O foco recai sobre a interpretação dos padrões observados, a análise crítica de suas possíveis causas e a discussão de suas implicações operacionais e científicas. Ao final, são apresentadas considerações sobre a viabilidade da abordagem proposta e indicam direções para o aprimoramento de estudos futuros.

4.1 Caracterização das Variáveis e Seleção de Atributos

A análise das variáveis disponíveis é fundamental para compreender o comportamento dos modelos de aprendizado de máquina e sua capacidade de discriminação entre viagens normais e críticas. Em aplicações de manutenção preditiva baseadas em dados reais de operação, a composição do espaço de atributos influencia diretamente a qualidade das predições, sobretudo em cenários caracterizados por desbalanceamento de classes e elevada heterogeneidade operacional.

O conjunto de dados utilizado nesta pesquisa é composto por 154 variáveis que descrevem cada viagem a partir de informações operacionais do veículo, registros de eventos de telemetria e indicadores estatísticos derivados na etapa de processamento dos dados. Essas variáveis apresentam naturezas distintas, incluindo atributos contínuos, categóricos e binários, bem como diferentes origens funcionais, conforme descrito no Capítulo 3. A descrição detalhada dessas variáveis é apresentada no Apêndice X.

Antes da aplicação dos métodos de seleção de atributos, foi realizada uma filtragem inicial, definida na metodologia, com o objetivo de remover variáveis de identificação, atributos temporais absolutos, informações espaciais e campos diretamente associados ao processo de rotulação ou indisponíveis no momento da predição. Após essa etapa, o conjunto final considerado para análise passou a conter 154 atributos.

O elevado número de atributos amplia o espaço de busca dos algoritmos de aprendizado de máquina e pode introduzir redundância e correlação entre variáveis, especialmente devido à presença de múltiplas estatísticas derivadas de um mesmo tipo de evento de

telemetria. Esse aspecto reforça a necessidade de uma análise sistemática da relevância dos atributos antes da etapa de modelagem.

Nesta seção são apresentados os resultados obtidos com a aplicação dos critérios de ranqueamento de variáveis definidos no Capítulo 3, utilizando o *widget Rank* do *software Orange Data Mining*. As subseções a seguir discutem os padrões observados nos rankings de atributos e suas implicações para a interpretação do comportamento dos modelos.

4.1.1 Critérios de seleção de variáveis

A aplicação de filtros de seleção de variáveis corresponde à primeira análise apresentada neste capítulo com o objetivo de investigar o papel dos atributos no processo de modelagem preditiva. Essa análise busca examinar como diferentes critérios de relevância organizam o conjunto de variáveis disponíveis, bem como identificar padrões recorrentes e possíveis sobreposições entre os atributos mais bem posicionados nos rankings.

Foram considerados os filtros disponíveis no *software Orange*, a saber, *Gini Index*, *Chi-Square*, *Information Gain*, *Gain Ratio* e *ReliefF*, além do cenário sem aplicação de filtro. Cada critério foi aplicado aos seis horizontes de amostragem utilizados no estudo, resultando em 36 listas distintas de atributos empregadas no treinamento e na avaliação dos modelos.

A análise dessas listas foi conduzida de forma comparativa, considerando a recorrência dos atributos entre diferentes critérios de seleção, a sensibilidade dos rankings aos horizontes de amostragem e a presença de padrões estruturais comuns entre os atributos mais frequentemente selecionados. Os resultados dessa análise são apresentados e discutidos nas subseções seguintes.

4.1.2 Atributos Discriminativos

A análise dos rankings produzidos pelos filtros de seleção de variáveis permite identificar os atributos mais recorrentes entre as primeiras posições nas diferentes estratégias de amostragem e horizontes preditivos avaliados. Para fins de interpretação dos resultados, a Tabela 4.1 sintetiza os vinte atributos que aparecem com maior frequência entre os mais bem ranqueados, considerando as combinações entre os cinco critérios de seleção de variáveis e os seis horizontes de amostragem adotados neste estudo.

Tabela 4.1 – Top 20 atributos mais frequentes entre todos os filtros e estratégias de amostragem.

Atributo	Frequência
ra_curva_brusca_max	15
x_mix_excesso_de_velocidade_os_count	15
ra_curva_brusca_mean	15
Tipo_dia	15
ra_temperatura_do_motor_em_viagem_valor_maximo_count	14
ra_uso_indevido_do_pedal_de_acelerador_85_max	14
ra_uso_indevido_do_pedal_de_acelerador_85_std	14
ra_temperatura_do_motor_em_viagem_valor_maximo_max	14
ra_temperatura_do_oleo_media_min	14
ra_temperatura_do_oleo_media_max	14
ra_temperatura_do_oleo_media_mean	14
ra_temperatura_do_motor_media_mean	13
ra_marcha_lenta_count	13
ra_pressao_do_oleo_media_iveco_min	13
ra_falha_de_shutdown_count	13
ra_acelerando_parado_mean	13
ra_temperatura_do_motor_media_min	13
StartEngineSeconds	13
EndEngineSeconds	13

4.1.3 Atributos Discriminativos

A análise dos rankings produzidos pelos filtros de seleção de variáveis permite identificar os atributos que aparecem com maior frequência entre as primeiras posições nas diferentes estratégias de amostragem e horizontes preditivos avaliados. Para fins de interpretação dos resultados, a Tabela 4.1 sintetiza os vinte atributos mais recorrentes,

considerando as combinações entre os cinco critérios de seleção de variáveis e os seis horizontes de amostragem adotados neste estudo.

Com o objetivo de facilitar a análise qualitativa desses resultados, os atributos recorrentes foram organizados em grupos conceituais de acordo com sua natureza funcional e com o tipo de informação operacional que representam. Ressalta-se que esse agrupamento não decorre de um procedimento automático de clusterização, mas de uma categorização interpretativa adotada exclusivamente para apoiar a análise dos resultados.

O primeiro grupo reúne atributos associados a padrões de condução mais exigentes, relacionados ao comportamento do condutor durante a operação do veículo. Nesse grupo estão incluídas variáveis como eventos de curva brusca, excesso de velocidade e uso inadequado do pedal do acelerador, que figuram de forma consistente entre os atributos mais bem posicionados nos rankings. A presença recorrente dessas variáveis sugere que padrões de condução mais agressivos ou irregulares estão associados a diferenças relevantes entre viagens normais e críticas.

O segundo grupo é composto por atributos diretamente relacionados ao comportamento térmico e ao funcionamento do conjunto motriz, incluindo variáveis de temperatura do motor, temperatura do óleo e indicadores de falha de desligamento. A recorrência desses atributos entre os mais bem ranqueados indica que, mesmo após a agregação estatística por viagem, permanecem sinais térmicos discriminativos relevantes para a antecipação de eventos de superaquecimento.

O terceiro grupo engloba atributos operacionais contextuais, que caracterizam condições gerais da viagem, independentemente de eventos específicos de condução ou falha. Variáveis como *Tipo_dia* e indicadores de tempo de operação figuram entre os atributos recorrentes, sugerindo que fatores contextuais e temporais também contribuem para a diferenciação entre viagens normais e críticas.

De forma geral, os rankings obtidos a partir dos diferentes critérios de seleção revelam um conjunto heterogêneo de atributos relevantes, abrangendo variáveis comportamentais, térmicas e contextuais. Esses resultados reforçam a complexidade da base de dados e fornecem subsídios para a análise do impacto da seleção de variáveis sobre o comportamento dos modelos de predição.

4.1.4 Síntese e Implicações da Seleção de Variáveis

Os resultados apresentados nas subseções anteriores indicam que os diferentes critérios de seleção de variáveis priorizam conjuntos heterogêneos de atributos, refletindo a diversidade informacional presente na base de dados analisada. Embora alguns filtros tenham destacado atributos diretamente associados ao comportamento térmico do motor, enquanto outros enfatizaram variáveis operacionais e contextuais, não se observa a formação

de um subconjunto reduzido de atributos claramente dominante sob todos os critérios e horizontes avaliados.

Essa heterogeneidade sugere que o fenômeno de superaquecimento do motor, no contexto dos dados de telemetria disponíveis, é influenciado por múltiplas dimensões operacionais, de condução e de funcionamento do veículo. Como consequência, a recorrência de determinados atributos nos rankings de seleção deve ser interpretada mais como um indicativo da estrutura dos dados do que como evidência direta de separabilidade robusta entre viagens normais e críticas.

Dessa forma, a seleção de variáveis desempenha neste estudo um papel predominantemente exploratório e interpretativo, contribuindo para a compreensão dos tipos de informações mais frequentemente associados ao risco de falha, mas sem, por si só, estabelecer um conjunto de atributos universalmente discriminativo. A análise dos efeitos dessas escolhas sobre o desempenho preditivo dos modelos é apresentada em seções posteriores do capítulo.

Na sequência, o capítulo avança para a avaliação do efeito das diferentes estratégias de amostragem adotadas, examinando como a definição do horizonte preditivo influencia a capacidade dos modelos em capturar sinais antecipados de superaquecimento.

4.2 Desempenho dos Modelos por Horizonte Preditivo

A influência do horizonte preditivo sobre o desempenho dos modelos de aprendizado de máquina é particularmente relevante, como o observado neste estudo. A Tabela 4.2 apresenta a distribuição de viagens normais e críticas nos conjuntos de treino e teste para cada estratégia de amostragem, evidenciando variações substanciais na proporção de instâncias críticas entre os diferentes cenários. Essas variações afetam diretamente a composição informacional dos conjuntos de treinamento e a estabilidade das métricas de desempenho.

Tabela 4.2 – Distribuição de viagens normais e críticas nos conjuntos de treino e teste para cada estratégia de amostragem.

Amostragem	Treino			Teste		
	Normal	Crítica	Total	Normal	Crítica	Total
5 dias	892	988	1880	6915	261	7176
10 dias	1827	2024	3851	6750	426	7176
15 dias	2862	3151	6013	6678	498	7176
20 viagens	770	846	1616	6946	230	7176
50 viagens	1889	2094	3983	6739	437	7176
100 viagens	3890	4294	8184	6640	536	7176

Esse comportamento está associado tanto à forma de definição da classe crítica a partir da abertura da ordem de serviço quanto à heterogeneidade das viagens incluídas sob esse rótulo. Em estratégias baseadas em janelas mais amplas, como aquelas definidas por 20, 50 ou 100 viagens anteriores à falha, uma parcela significativa das instâncias rotuladas como críticas apresenta perfis operacionais semelhantes aos da classe normal. Esse fenômeno caracteriza uma diluição da classe crítica, na qual apenas uma fração das instâncias contém sinais potencialmente discriminativos relacionados ao risco de superaquecimento.

A diluição da classe crítica reduz a separabilidade entre as classes, uma vez que as distribuições das variáveis tornam-se progressivamente mais semelhantes. Nesse contexto, métricas como AUC e *recall* tendem a assumir valores médios mais elevados em janelas amplas, não necessariamente em função de maior capacidade preditiva, mas porque os rótulos críticos passam a representar majoritariamente viagens com comportamento típico da classe normal.

A Tabela 4.3 evidencia esse efeito ao apresentar a variabilidade das métricas entre as diferentes estratégias de amostragem. Observa-se que o *F1-score* apresenta médias mais elevadas no cenário de 100 viagens, embora esse conjunto inclua uma proporção reduzida de instâncias efetivamente associadas a sinais de pré-falha. De forma semelhante, a AUC exhibe valores médios superiores acompanhados por elevados desvios padrão, indicando instabilidade no desempenho dos modelos. Métricas mais sensíveis ao desbalanceamento, como o coeficiente de correlação de Matthews, também apresentam flutuações relevantes, reforçando os efeitos da diluição sobre a qualidade preditiva.

Tabela 4.3 – Desempenho médio dos modelos para diferentes estratégias de amostragem.

Métrica	Estatística	15 dias	10 dias	5 dias	100 viagens	50 viagens	20 viagens
AUC	Média	0,588	0,591	0,582	0,622	0,565	0,622
AUC	Desvio Padrão	0,129	0,120	0,105	0,137	0,111	0,095
AUC	Máximo	0,829	0,794	0,742	0,869	0,836	0,753
AUC	Mínimo	0,343	0,329	0,321	0,329	0,487	0,494
CA	Média	0,397	0,482	0,354	0,433	0,297	0,396
CA	Desvio Padrão	0,298	0,302	0,302	0,283	0,308	0,317
CA	Máximo	0,000	0,000	0,000	0,000	0,000	0,000
CA	Mínimo	0,071	0,060	0,036	0,077	0,064	0,033
F1	Média	0,161	0,135	0,091	0,173	0,136	0,086
F1	Desvio Padrão	0,055	0,051	0,031	0,070	0,056	0,030
F1	Máximo	0,291	0,258	0,154	0,354	0,312	0,146
F1	Mínimo	0,087	0,011	0,057	0,031	0,024	0,052
Precisão	Média	0,095	0,081	0,050	0,105	0,078	0,047
Precisão	Desvio Padrão	0,039	0,032	0,020	0,044	0,036	0,019
Precisão	Máximo	0,190	0,168	0,095	0,236	0,203	0,086
Precisão	Mínimo	0,064	0,024	0,030	0,051	0,050	0,028
Recall	Média	0,749	0,637	0,763	0,744	0,823	0,746
Recall	Desvio Padrão	0,253	0,289	0,216	0,274	0,252	0,234
Recall	Máximo	1,000	0,995	1,000	1,000	1,000	1,000
Recall	Mínimo	0,110	0,007	0,410	0,019	0,016	0,287
MCC	Média	0,069	0,059	0,047	0,088	0,049	0,056
MCC	Desvio Padrão	0,089	0,075	0,056	0,106	0,087	0,055
MCC	Máximo	0,257	0,228	0,145	0,331	0,299	0,145
MCC	Mínimo	-0,057	-0,061	-0,055	-0,091	-0,049	-0,039

Em contrapartida, nas janelas temporais mais curtas, como aquelas definidas em 5 ou 10 dias anteriores à abertura da ordem de serviço, a quantidade de viagens rotuladas como críticas é significativamente menor. Essa redução tende a aumentar a coerência interna da classe crítica, porém à custa de um número absoluto reduzido de exemplos, tornando o treinamento altamente sensível à seleção específica das amostras e elevando a variância das métricas. Assim, horizontes curtos enfrentam limitações associadas à insuficiência amostral, enquanto horizontes longos sofrem com a diluição da classe crítica.

4.3 Comparação entre Algoritmos

Embora o desempenho preditivo tenha sido limitado na maioria dos cenários avaliados, a comparação entre os três algoritmos de aprendizado de máquina permite identificar padrões consistentes sobre como cada abordagem responde às características da base de dados. A Tabela 4.4 apresenta estatísticas descritivas agregadas dos 36 experimentos realizados com cada algoritmo, sintetizando seu comportamento geral frente às diferentes combinações de horizontes preditivos, filtros de seleção de variáveis e configurações de parâmetros.

Tabela 4.4 – Estatísticas descritivas das métricas de desempenho obtidas para cada algoritmo e configuração

Parâmetro	Estatística	RF (1)	RF (2)	RNA (1)	RNA (2)	SVM (1)	SVM (2)
AUC	Média	0.692	0.664	0.589	0.573	0.528	0.523
AUC	Desvio Padrão	0.123	0.093	0.104	0.088	0.094	0.095
AUC	Máximo	0.875	0.788	0.803	0.759	0.689	0.719
AUC	Mínimo	0.490	0.489	0.473	0.488	0.321	0.329
CA	Média	0.614	0.604	0.462	0.265	0.115	0.306
CA	Desvio Padrão	0.240	0.271	0.306	0.219	0.099	0.279
CA	Máximo	0.819	0.836	0.924	0.638	0.470	0.887
CA	Mínimo	0.069	0.055	0.054	0.046	0.033	0.041
F1	Média	0.180	0.163	0.119	0.120	0.104	0.097
F1	Desvio Padrão	0.081	0.053	0.058	0.040	0.030	0.030
F1	Máximo	0.362	0.278	0.243	0.226	0.156	0.154
F1	Mínimo	0.063	0.063	0.011	0.062	0.054	0.052
Prec	Média	0.109	0.099	0.072	0.065	0.056	0.055
Prec	Desvio Padrão	0.056	0.039	0.031	0.024	0.017	0.017
Prec	Máximo	0.243	0.186	0.143	0.131	0.088	0.084
Prec	Mínimo	0.032	0.033	0.024	0.032	0.028	0.029
Recall	Média	0.654	0.603	0.680	0.865	0.933	0.729
Recall	Desvio Padrão	0.168	0.219	0.310	0.145	0.118	0.336
Recall	Máximo	0.986	0.998	0.995	0.995	1.000	1.000
Recall	Mínimo	0.347	0.302	0.007	0.557	0.582	0.055
MCC	Média	0.140	0.113	0.063	0.046	0.000	0.012
MCC	Desvio Padrão	0.100	0.063	0.070	0.053	0.027	0.027
MCC	Máximo	0.339	0.233	0.223	0.200	0.058	0.085
MCC	Mínimo	0.012	0.011	-0.020	-0.023	-0.057	-0.061

A análise dessas métricas evidencia diferenças sistemáticas entre os algoritmos. No entanto, os resultados indicam que nenhuma das abordagens avaliadas alcança, de maneira

simultânea e estável, níveis de AUC, *F1-score* e MCC que, aliados a um controle adequado de falsos positivos, sustentem sua aplicação operacional no contexto analisado.

- a) **Random Forest** apresentou o comportamento mais consistente entre os algoritmos avaliados. As maiores médias de AUC foram observadas nesse método, com valores de 0,692 na configuração (1) e 0,664 na configuração (2). O algoritmo também obteve os maiores valores médios de *F1-score*, precisão e coeficiente MCC, ainda que modestos em termos absolutos. Os desvios padrão relativamente menores indicam maior estabilidade frente às diferentes combinações de filtros e horizontes preditivos, sugerindo maior robustez à heterogeneidade e ao ruído presentes nos dados.
- b) **Redes Neurais Artificiais** apresentaram desempenho intermediário, acompanhado de elevada sensibilidade às variações experimentais. As médias de AUC situam-se entre aquelas obtidas pelo *Random Forest* e pela SVM. Observa-se ampla variação do *recall*, com valores mínimos próximos de zero e máximos próximos de um, indicando forte dependência das janelas de observação e dos limiares de decisão. As métricas de *F1-score*, precisão e MCC permanecem reduzidas, refletindo dificuldades em equilibrar sensibilidade e controle de falsos positivos.
- c) **SVM** apresentou, de forma sistemática, os menores valores médios de AUC, *F1-score*, precisão e MCC. Em alguns cenários, os valores de MCC assumiram valores negativos, caracterizando desempenho inferior ao de uma classificação aleatória. Embora o *recall* tenha atingido valores elevados em determinados casos, esse comportamento decorreu, em geral, da classificação da maioria das instâncias como positivas, o que compromete a utilidade prática do modelo.

Esses resultados indicam que as limitações observadas estão associadas, em grande medida, a características estruturais da base de dados e do contexto operacional no qual os dados foram gerados, mais do que à escolha do algoritmo em si. Entre esses fatores destacam-se o uso de registros de telemetria baseados em eventos, em vez de sinais contínuos ou séries temporais de alta granularidade, a ausência de medições diretas provenientes da rede CAN dos veículos e a definição das viagens como ciclos de ignição, o que agrega condições operacionais distintas em uma mesma instância.

Adicionalmente, incertezas associadas ao processo de abertura e fechamento das ordens de serviço, bem como à persistência de falhas após intervenções de manutenção, afetam a consistência do rótulo de viagem crítica, reduzindo a separabilidade efetiva entre as classes. Essas características estruturais limitam o contraste informacional disponível para os modelos e impõem restrições comuns a todas as técnicas avaliadas, independentemente de sua complexidade ou capacidade de modelagem.

4.3.1 Caso representativo: melhor configuração do Random Forest

O melhor desempenho observado para o algoritmo *Random Forest* ocorreu na estratégia de amostragem baseada em 100 viagens, combinada com o filtro *Chi-Square*. As métricas obtidas nessa execução foram $AUC = 0,869$, exatidão = 0,806, $F1-score = 0,354$, precisão = 0,236, $recall = 0,715$ e coeficiente de correlação de Matthews igual a 0,331. Esses valores caracterizam essa configuração como o melhor cenário entre os 216 experimentos avaliados.

A matriz de confusão correspondente, apresentada na Tabela 4.5, indica que, dentre 7.176 observações, 5.396 foram corretamente classificadas como não críticas e 384 como críticas. Foram registrados 1.241 falsos positivos e 155 falsos negativos, refletindo a baixa prevalência do evento crítico no conjunto de teste, aproximadamente 7,5%. Os percentuais apresentados na matriz correspondem a proporções por coluna, evidenciando a distribuição das classes em cada grupo de predição.

A Tabela 4.5 apresenta a matriz de confusão correspondente.

Tabela 4.5 – Matriz de confusão *Random Forest* (100 viagens, *Chi-Square*).

Real	Predito		Σ
	Não	Sim	
Não (viagem normal)	5.396 (97,2%)	1.241 (76,4%)	6.637
Sim (viagem crítica)	155 (2,8%)	384 (23,6%)	539
Σ	5.551	1.625	7.176

Esse resultado evidencia elevada sensibilidade do modelo, com identificação de cerca de 71,5% das viagens críticas. Em contrapartida, a precisão permaneceu reduzida, indicando que apenas 23,6% das instâncias classificadas como críticas correspondem efetivamente a eventos reais. Na prática, esse padrão implica a geração de um número elevado de alarmes incorretos, o que limita a utilidade operacional do modelo. O valor de MCC igual a 0,331 confirma a existência de correlação entre as predições e os rótulos verdadeiros, embora em magnitude ainda insuficiente para aplicação automatizada em ambiente real.

Esse cenário representa o limite superior de desempenho observado com a base de dados utilizada, ilustrando que, mesmo na melhor configuração, a redução da taxa de falsos positivos permanece um desafio central para a viabilidade prática da abordagem.

Na sequência, são apresentados casos representativos dos demais algoritmos, com o objetivo de comparar os padrões qualitativos de erro observados em diferentes técnicas.

4.3.2 Padrões extremos de predição nos modelos não lineares

Os modelos não lineares avaliados neste estudo, especificamente as Redes Neurais Artificiais e as Máquinas de Vetores de Suporte, apresentaram comportamentos instáveis, caracterizados por padrões extremos de predição. Esses padrões manifestam-se em cenários nos quais os modelos tendem a concentrar a maior parte das predições em uma única classe, classificando quase todas as instâncias como críticas ou como não críticas.

Esse comportamento foi observado em diferentes combinações de filtros de seleção de variáveis e horizontes de amostragem, refletindo a baixa separabilidade entre as classes e o forte desbalanceamento do problema, marcado pela reduzida prevalência de viagens críticas. Para ilustrar esses efeitos, são analisadas a seguir quatro matrizes de confusão representativas, que evidenciam os padrões extremos identificados nos modelos não lineares.

4.3.2.1 RNA prevendo quase tudo como positivo.

O cenário do algoritmo RNA (2), com horizonte de 10 dias e filtro *ReliefF*, apresentou *recall* igual a 0,995 e precisão de apenas 0,059. Esse resultado indica que o modelo identificou praticamente todos os casos críticos, porém à custa de um elevado número de falsos positivos, conforme evidenciado pela matriz de confusão apresentada na Tabela 4.6. O coeficiente de correlação de Matthews, MCC igual a 0,005, confirma que, apesar do *recall* elevado, não há correlação significativa entre as predições e os rótulos verdadeiros.

Tabela 4.6 – Matriz de confusão da RNA (2), horizonte de 10 dias, filtro *ReliefF*.

Real	Predito		Σ
	Não	Sim	
Não (normal)	43 (95,6%)	6.710 (94,1%)	6.753
Sim (crítica)	2 (4,4%)	421 (5,9%)	423
Σ	45	7.131	7.176

Esse padrão caracteriza um colapso da predição para a classe positiva. O modelo passa a classificar a maior parte das observações como críticas, o que inflaciona o *recall*, mas resulta em precisão extremamente baixa. Do ponto de vista operacional, esse comportamento implicaria a geração de um volume excessivo de falsos positivos, comprometendo a utilidade prática do modelo.

4.3.2.2 RNA prevendo quase tudo como negativo.

O cenário da configuração RNA (1), amostragem de 10 dias e filtro *ReliefF*, representa o extremo oposto ao caso anterior. Nesse experimento, o modelo apresentou exatidão igual a 0,924, porém *recall* de apenas 0,007, indicando que praticamente nenhum caso

crítico foi corretamente identificado. A matriz de confusão correspondente é apresentada na Tabela 4.7.

Tabela 4.7 – Matriz de confusão da RNA (1), horizonte de 10 dias, filtro *ReliefF*.

Real	Predito		Σ
	Não	Sim	
Não (normal)	6.626 (94,0%)	124 (97,6%)	6.750
Sim (crítica)	423 (6,0%)	3 (2,4%)	426
Σ	7.049	127	7.176

Nesse caso, o modelo colapsou para a classe negativa, classificando quase todas as observações como normais. Esse comportamento leva a um número elevado de falsos negativos e inviabiliza qualquer aplicação preditiva, pois os casos críticos, justamente os mais relevantes, deixam de ser detectados. O MCC negativo (-0,021) confirma que o modelo apresenta desempenho inferior ao de uma estratégia aleatória.

4.3.2.3 SVM com predição massivamente positiva

O comportamento extremo da SVM também foi observado em cenários nos quais o modelo passou a classificar a maior parte das observações como pertencentes à classe crítica. Um exemplo representativo ocorre na configuração SVM (2), horizonte de 20 viagens e filtro *Information Gain*, na qual foram obtidos *recall* igual a 1,0 e precisão de 0,032. A matriz de confusão correspondente, apresentada na Tabela 4.8, indica que o modelo gerou 7.111 predições positivas e apenas 65 predições negativas. Todas as predições negativas corresponderam a viagens não críticas, enquanto, entre as predições positivas, apenas 230 representaram casos reais e 6.881 foram falsos positivos.

Tabela 4.8 – Matriz de confusão da SVM (2), 20 viagens, *Information Gain*.

Real	Predito		Σ
	Não	Sim	
Não (normal)	65 (100,0%)	6.881 (96,8%)	6.946
Sim (crítica)	0 (0,0%)	230 (3,2%)	230
Σ	65	7.111	7.176

Esse padrão caracteriza um colapso da predição para a classe positiva. O modelo identifica todos os casos críticos, o que explica o *recall* máximo, porém à custa de um volume excessivo de predições incorretas. A precisão extremamente baixa reflete a dificuldade da SVM em separar as classes no espaço de atributos disponível. O coeficiente de correlação de Matthews, MCC igual a 0,017, confirma que, apesar da elevada sensibilidade, a correlação global entre predições e rótulos verdadeiros é praticamente inexistente. Esse padrão de

erro compromete a viabilidade operacional do modelo nesse cenário, em função do volume excessivo de alarmes falsos gerados.

4.3.2.4 SVM prevendo quase tudo como negativo

Em outro conjunto de experimentos, a SVM apresentou comportamento oposto ao observado no caso anterior, passando a classificar a quase totalidade das observações como não críticas. Um exemplo representativo é a configuração SVM (2), horizonte de 50 viagens e filtro *ReliefF*, que resultou em *recall* igual a 0,055, precisão de 0,056 e exatidão de 0,887. A matriz de confusão correspondente é apresentada na Tabela 4.9.

Tabela 4.9 – Matriz de confusão da SVM (2), 50 viagens, *ReliefF*.

Real	Predito		Σ
	Não	Sim	
Não (normal)	6.339 (93,9%)	401 (94,4%)	6.740
Sim (crítica)	412 (6,1%)	24 (5,6%)	436
Σ	6.751	425	7.176

Esse padrão caracteriza um colapso da predição para a classe negativa. O modelo reproduz majoritariamente a distribuição da classe normal, o que resulta em elevado número de falsos negativos e em valores de *recall* extremamente baixos. A precisão também permanece reduzida, indicando que mesmo as poucas predições positivas raramente correspondem a casos críticos reais. O coeficiente de correlação de Matthews ligeiramente negativo confirma a ausência de correlação útil entre as predições e os rótulos verdadeiros. Na prática, esse comportamento inviabiliza o uso do modelo para detecção de risco, uma vez que a maioria dos eventos críticos deixa de ser identificada.

4.4 Impacto da Seleção de Variáveis no Desempenho dos Algoritmos de Aprendizado de Máquina

A avaliação do impacto dos filtros de seleção de variáveis sobre o desempenho dos algoritmos permite analisar se a priorização de subconjuntos de atributos considerados mais relevantes resulta em ganhos mensuráveis de desempenho preditivo. A Tabela 4.10 apresenta métricas consolidadas de desempenho, incluindo AUC, exatidão, *F1-score*, precisão, *recall* e coeficiente de correlação de Matthews, para cada critério de seleção avaliado, considerando conjuntamente as seis estratégias de amostragem, os três algoritmos de aprendizado supervisionado e suas diferentes configurações.

Tabela 4.10 – Desempenho dos modelos por filtros de seleção de variáveis.

Métrica	Estatística	ReliefF	Chi-Square	Sem Filtro	Gini Index	Gain Ratio	Information Gain
AUC	Média	0,553	0,678	0,600	0,597	0,522	0,620
AUC	Desvio Padrão	0,065	0,120	0,129	0,127	0,088	0,103
AUC	Máximo	0,717	0,869	0,863	0,864	0,679	0,714
AUC	Mínimo	0,414	0,391	0,386	0,393	0,329	0,321
CA	Média	0,385	0,570	0,428	0,410	0,172	0,393
CA	Desvio Padrão	0,253	0,217	0,346	0,345	0,200	0,200
CA	Máximo	0,924	0,819	0,921	0,922	0,816	0,836
CA	Mínimo	0,060	0,170	0,033	0,033	0,036	0,041
F1	Média	0,111	0,165	0,129	0,130	0,114	0,133
F1	Desvio Padrão	0,035	0,077	0,071	0,068	0,039	0,040
F1	Máximo	0,168	0,354	0,328	0,328	0,202	0,214
F1	Mínimo	0,011	0,054	0,031	0,024	0,062	0,063
Precisão	Média	0,062	0,099	0,078	0,077	0,062	0,077
Precisão	Desvio Padrão	0,019	0,054	0,043	0,043	0,025	0,030
Precisão	Máximo	0,093	0,236	0,205	0,206	0,125	0,165
Precisão	Mínimo	0,024	0,028	0,029	0,030	0,032	0,032
<i>Recall</i>	Média	0,706	0,652	0,706	0,720	0,923	0,758
<i>Recall</i>	Desvio Padrão	0,257	0,141	0,315	0,305	0,150	0,246
<i>Recall</i>	Máximo	1,000	0,952	1,000	1,000	1,000	1,000
<i>Recall</i>	Mínimo	0,007	0,289	0,019	0,016	0,350	0,302
MCC	Média	0,032	0,113	0,060	0,058	0,037	0,070
MCC	Desvio Padrão	0,041	0,108	0,096	0,094	0,037	0,049
MCC	Máximo	0,142	0,331	0,324	0,321	0,136	0,163
MCC	Mínimo	-0,057	-0,091	-0,025	-0,025	0,000	0,009

Ressalta-se que as métricas apresentadas na Tabela 4.10 assumem valores normalizados, nos quais valores mais próximos de 1 indicam melhor capacidade de discriminação entre as classes, enquanto valores próximos de 0 refletem desempenho equivalente ao de uma classificação aleatória. No caso do coeficiente de correlação de Matthews, valores negativos indicam desempenho inferior ao aleatório.

De forma geral, os resultados indicam que a aplicação dos filtros de seleção de variáveis não produz ganhos expressivos de desempenho quando comparada ao cenário sem seleção de atributos. Entre os métodos avaliados, o filtro *Chi-square* apresenta desempenho superior aos demais em diferentes métricas. A AUC média obtida com esse filtro, de 0,678, supera a observada no cenário sem filtro, de 0,600, assim como os valores médios de

F1-score e do coeficiente de correlação de Matthews, que também figuram entre os mais elevados.

Os resultados associados ao *Chi-square* são consistentes com os padrões observados na análise dos rankings de atributos, nos quais esse critério apresentou maior recorrência de variáveis relacionadas ao comportamento térmico do motor. Ainda assim, mesmo nesse caso, observa-se elevada variabilidade das métricas, refletida nos desvios padrão e nos valores extremos, indicando que o desempenho permanece fortemente dependente das estratégias de amostragem e das configurações dos modelos.

Em contraste, filtros como *Gain Ratio* e *Gini Index* apresentam desempenho inferior ao cenário sem seleção de variáveis. O *Gain Ratio*, em particular, registra a menor AUC média e valores reduzidos de MCC, indicando baixa capacidade de discriminação entre viagens normais e críticas quando considerado o conjunto total de experimentos.

O cenário sem aplicação de filtro, embora não apresente os melhores valores individuais em nenhuma métrica, mantém desempenho comparável ao observado com a maioria dos métodos de seleção. Esse resultado sugere que a utilização do conjunto completo de atributos não compromete significativamente a capacidade preditiva dos modelos e que os filtros avaliados oferecem benefícios limitados quando analisados de forma global.

De modo geral, a análise consolidada indica que o impacto da seleção de variáveis sobre o desempenho dos modelos é modesto neste estudo, com melhoria perceptível apenas para o filtro *Chi-square*. Ressalta-se, contudo, que a agregação das métricas nesta seção tem caráter exploratório e visa identificar tendências gerais associadas aos critérios de seleção, não substituindo as análises específicas por horizonte preditivo e por algoritmo apresentadas nas seções anteriores.

4.5 Panorama Geral do Desempenho

O panorama geral do desempenho dos modelos foi construído a partir da consolidação dos 216 experimentos realizados ao longo do estudo. Dada a diversidade de combinações avaliadas, esta subseção adota uma análise baseada em estatísticas agregadas, com o objetivo de fornecer uma visão global do comportamento das métricas de avaliação e situar os limites gerais de desempenho observados no problema analisado.

A Tabela 4.11 apresenta estatísticas descritivas das métricas AUC, exatidão, *F1-score*, precisão, *recall* e coeficiente de correlação de Matthews, incluindo valores médios, desvios padrão e extremos. Essas estatísticas sintetizam o conjunto completo de experimentos e devem ser interpretadas como indicadores de tendência geral e variabilidade, não como desempenho esperado de configurações específicas.

Tabela 4.11 – Estatísticas descritivas das métricas dos 216 experimentos de classificação.

Métrica	AUC	CA	F1	Precisão	Recall	MCC
Média	0.595	0.394	0.131	0.076	0.744	0.062
Desvio Padrão	0.118	0.303	0.060	0.039	0.257	0.080
Máximo	0.875	0.924	0.362	0.243	1.000	0.339
Mínimo	0.321	0.033	0.011	0.024	0.007	-0.061
Ocorrências	216	216	216	216	216	216

Do ponto de vista da capacidade discriminativa, a AUC média de 0,595 indica separabilidade apenas marginalmente superior ao acaso. A ampla variação observada, com valores entre 0,321 e 0,875, evidencia elevada instabilidade entre os experimentos, sugerindo que apenas um número reduzido de configurações atinge níveis moderados de discriminação, enquanto a maioria opera próxima ao limiar não informativo.

As métricas relacionadas à exatidão global e ao balanceamento de classes exigem interpretação cautelosa. Embora a exatidão apresente valores máximos elevados, a média global de 0,394 indica ausência de consistência ao longo dos experimentos. Esse comportamento está associado à diferença intencional entre a distribuição das classes no conjunto de treinamento, ajustada para viabilizar o aprendizado dos modelos, e o desbalanceamento preservado no conjunto de teste, que reflete as condições reais de operação. Como consequência, observa-se um *trade-off* recorrente entre sensibilidade e precisão, no qual alguns modelos tendem a priorizar a identificação da classe crítica, elevando o *recall* e reduzindo a precisão, enquanto outros privilegiam a classe normal, resultando em valores de exatidão aparentemente elevados, porém pouco informativos no contexto analisado.

Nesse contexto, métricas mais adequadas a cenários desbalanceados reforçam as limitações observadas. O *F1-score* apresenta média de 0,131, com valores mínimos próximos de zero, evidenciando dificuldade em conciliar sensibilidade e controle de falsos positivos. A precisão média de 0,076 indica elevada taxa de classificações incorretas da classe crítica, enquanto o *recall* médio de 0,744, com máximos iguais a 1, reflete cenários nos quais os modelos classificaram a maioria das instâncias como críticas.

O coeficiente de correlação de Matthews sintetiza esses efeitos de forma mais robusta. Com média de 0,062 e valores variando entre -0,061 e 0,339, observa-se baixa correlação entre as predições e os rótulos reais, incluindo casos de desempenho inferior ao acaso. Esses resultados indicam que, de maneira geral, os modelos apresentam limitações significativas em capturar padrões discriminativos consistentes no conjunto de dados analisado.

4.6 Melhores Desempenhos (Top 20 por AUC)

Após a análise estatística global, torna-se pertinente examinar os cenários que apresentaram os melhores desempenhos relativos. A Tabela 4.12 reúne os vinte maiores valores de AUC obtidos nos 216 experimentos, permitindo identificar padrões recorrentes entre os casos de melhor desempenho.

Tabela 4.12 – Top 20 valores de AUC entre os 216 experimentos.

Amostragem	Filtro	Modelo	AUC	CA	F1	Precisão	Recall	MCC
100 viagens	Chi-Square	RF (1)	0.875	0.813	0.362	0.243	0.711	0.339
100 viagens	Gini Index	RF (1)	0.864	0.753	0.328	0.206	0.808	0.321
100 viagens	Sem Filtro	RF (1)	0.863	0.749	0.328	0.205	0.821	0.324
50 viagens	Chi-Square	RF (1)	0.836	0.819	0.312	0.203	0.675	0.299
15 dias	Chi-Square	RF (1)	0.829	0.789	0.291	0.190	0.622	0.257
15 dias	Gini Index	RF (1)	0.822	0.752	0.277	0.173	0.685	0.251
15 dias	Sem Filtro	RF (1)	0.818	0.746	0.274	0.171	0.691	0.249
50 viagens	Sem Filtro	RF (1)	0.818	0.760	0.261	0.161	0.696	0.250
50 viagens	Gini Index	RF (1)	0.804	0.745	0.249	0.152	0.696	0.237
100 viagens	Chi-Square	NN (1)	0.803	0.624	0.243	0.143	0.810	0.223
10 dias	Sem Filtro	RF (1)	0.794	0.761	0.226	0.140	0.587	0.196
15 dias	Chi-Square	RF (2)	0.788	0.801	0.278	0.186	0.552	0.233
10 dias	Chi-Square	RF (1)	0.786	0.809	0.258	0.168	0.559	0.228
10 dias	Gini Index	RF (1)	0.783	0.759	0.217	0.134	0.563	0.183
100 viagens	Sem Filtro	RF (2)	0.779	0.588	0.221	0.129	0.784	0.188
100 viagens	Chi-Square	RF (2)	0.774	0.707	0.252	0.156	0.660	0.210
100 viagens	Gini Index	RF (2)	0.773	0.631	0.234	0.139	0.756	0.202
100 viagens	Chi-Square	NN (2)	0.759	0.580	0.226	0.131	0.819	0.200
15 dias	Sem Filtro	RF (2)	0.755	0.718	0.234	0.144	0.620	0.192
50 viagens	Chi-Square	NN (1)	0.753	0.650	0.203	0.118	0.730	0.185

A análise da Tabela 4.12 indica predominância do algoritmo *Random Forest* entre os melhores resultados, especialmente em sua primeira configuração, sugerindo maior robustez frente à heterogeneidade e ao ruído presentes nos dados. A SVM não figura entre os cenários de melhor desempenho, enquanto as Redes Neurais aparecem apenas em configurações específicas.

Observa-se também maior frequência de horizontes de amostragem mais longos, em especial aqueles baseados em 100 viagens e, em menor medida, 50 viagens e 15 dias. Embora esses cenários apresentem valores de AUC mais elevados, tal comportamento está

associado à forma de definição dos rótulos, que passa a incluir um conjunto maior de viagens anteriores à ordem de serviço. Esse efeito tende a aumentar artificialmente a AUC, sem necessariamente refletir maior capacidade de identificação de padrões de pré-falha.

Em relação à seleção de variáveis, os filtros *Chi-square* e *Gini Index* aparecem com maior frequência entre os melhores desempenhos, embora configurações sem aplicação de filtro também estejam presentes. Esse resultado reforça que os ganhos proporcionados pela seleção de atributos são limitados frente às restrições impostas pela estrutura dos dados.

Apesar de alguns cenários apresentarem valores de AUC superiores a 0,80, as demais métricas permanecem baixas. Mesmo o melhor desempenho observado, com AUC igual a 0,875, apresenta MCC de 0,339, valor modesto e insuficiente para aplicações operacionais. A combinação de baixa precisão, *F1-score* reduzido e elevada taxa de falsos positivos indica que, mesmo nos melhores casos, a separabilidade alcançada pelos modelos é limitada. Assim, a análise dos desempenhos mais elevados não altera o diagnóstico geral. Nenhum modelo apresentou desempenho robusto e consistente para a identificação confiável de viagens críticas em condições reais de operação.

4.7 Discussão dos Resultados

Os resultados apresentados neste capítulo evidenciam a complexidade do problema de predição de eventos de superaquecimento em ônibus urbanos a partir de dados reais de telemetria. A análise integrada do desempenho dos modelos, do efeito da seleção de variáveis e dos diferentes horizontes preditivos permite compreender os limites observados e situar adequadamente o potencial e as restrições da abordagem adotada.

De forma geral, observou-se que, embora algumas configurações tenham apresentado valores elevados em métricas específicas, como AUC ou *recall*, o desempenho global permaneceu modesto e caracterizado por elevada variabilidade. Resultados pontuais mais favoráveis, especialmente associados ao algoritmo de Floresta Aleatória, não se mantiveram de forma consistente entre diferentes estratégias de amostragem. A análise de casos representativos evidenciou ainda padrões extremos de predição, em particular nos modelos não lineares, nos quais se observaram colapsos para uma única classe. Esses comportamentos indicam dificuldades estruturais na separação entre viagens normais e críticas, mais relacionadas às características dos dados e à definição do problema do que à escolha do algoritmo.

A análise da seleção de variáveis reforça essa interpretação. Embora alguns critérios de ranqueamento tenham destacado atributos potencialmente associados ao comportamento térmico do motor, não se observou convergência consistente entre os subconjuntos de variáveis selecionados nem ganhos robustos de desempenho preditivo. De modo geral, os filtros priorizaram atributos operacionais e comportamentais, refletindo a natureza indireta

das informações disponíveis na base de telemetria. Esse resultado sugere que a limitação não reside na ausência de técnicas de seleção, mas na própria estrutura informacional dos dados utilizados.

O horizonte preditivo mostrou-se um fator crítico para a interpretação dos resultados. Em janelas de amostragem mais amplas, observou-se um processo de diluição da classe crítica, no qual muitas das viagens rotuladas como pré-falha apresentam comportamento operacional semelhante ao da classe normal. Esse efeito reduz a separabilidade entre as classes e pode levar a valores elevados de métricas como AUC ou *recall* sem correspondência direta com ganho preditivo efetivo. Por outro lado, janelas mais curtas concentram instâncias mais coerentes, porém em quantidade limitada, resultando em elevada variabilidade das métricas. Nenhum dos cenários avaliados conciliou simultaneamente volume amostral adequado e consistência informacional da classe crítica.

Cabe destacar que essa limitação não decorre da escolha específica da frota analisada, mas reflete restrições estruturais da base de telemetria disponível. Conforme discutido na metodologia, a seleção dos veículos visou maximizar a consistência relativa dos dados dentro das condições operacionais existentes, ainda que tais restrições não pudessem ser completamente eliminadas.

Adicionalmente, aspectos do processo de manutenção contribuem para o ruído observado nos dados. A utilização das ordens de serviço como referência temporal pressupõe a resolução da falha no momento do encerramento da OS, enquanto, na prática, eventos de superaquecimento podem ocorrer de forma intermitente ou sob condições específicas de operação. Esse desalinhamento potencial entre o estado real do veículo e o rótulo atribuído introduz incertezas adicionais na separação entre classes.

De forma mais ampla, os resultados indicam que as principais limitações observadas estão associadas à natureza dos dados de telemetria operacional, predominantemente baseados em eventos e variáveis agregadas, e à baixa granularidade temporal das medições térmicas disponíveis. Nesse contexto, os modelos avaliados operam com informações parciais sobre a dinâmica do fenômeno de superaquecimento, o que restringe a capacidade de aprendizado de padrões discriminativos estáveis.

Em síntese, o Capítulo 4 demonstra que, nas condições analisadas, a aplicação de modelos clássicos de aprendizado de máquina a dados reais de telemetria permite identificar tendências e limitações relevantes, mas não resulta em desempenho preditivo robusto para antecipação confiável de eventos de superaquecimento. Esses achados não apenas delimitam os limites da abordagem proposta, como também fornecem subsídios objetivos para orientar avanços futuros, a serem discutidos no capítulo seguinte, no que se refere à aquisição de dados, à modelagem temporal do fenômeno e à estratificação mais refinada das falhas analisadas.

5 Conclusão

Este trabalho investigou a viabilidade de antecipar eventos de superaquecimento do motor em ônibus urbanos a partir de dados reais de telemetria agregados por viagem. O estudo foi conduzido com base em um conjunto de dados construído especificamente para esse fim, integrando informações provenientes de sistemas de telemetria, registros de eventos operacionais e ordens de serviço de manutenção, originalmente não concebidos de forma integrada. A partir dessa base, foi proposta e aplicada uma estratégia de rotulagem retrospectiva de viagens, fundamentada em janelas temporais e operacionais definidas em relação à abertura de ordens de serviço associadas a alertas térmicos, permitindo a análise sistemática de diferentes horizontes preditivos.

A partir desse escopo, foram avaliadas múltiplas combinações de algoritmos de aprendizado de máquina, critérios de seleção de variáveis e estratégias de amostragem, sob condições de validação temporal compatíveis com o uso operacional. Os resultados apresentados no Capítulo 4 indicam que, embora algumas configurações específicas tenham alcançado valores elevados em métricas como AUC ou *recall*, o desempenho global dos modelos foi caracterizado por baixa consistência e elevada variabilidade entre os cenários avaliados. As análises realizadas sugerem que essa instabilidade está menos relacionada à escolha dos algoritmos em si e mais associada às características estruturais da base de dados e do contexto operacional no qual essas informações são geradas.

Em particular, a granularidade temporal dos dados disponíveis e a natureza predominantemente baseada em eventos da telemetria impõem restrições à observação direta da evolução térmica que antecede o superaquecimento. Nesse contexto, a estratégia de rotulagem adotada, embora operacionalmente viável e coerente com as informações disponíveis, herda a heterogeneidade e as incertezas inerentes a esses dados. Como consequência, apenas uma parcela das viagens rotuladas como críticas apresenta padrões operacionais ou térmicos distintivos de forma consistente, efeito que se intensifica com a ampliação das janelas de amostragem e resulta em uma redução progressiva da separabilidade entre as classes, conforme evidenciado ao longo do Capítulo 4.

A análise da seleção de variáveis reforçou esse diagnóstico. Embora alguns filtros tenham destacado atributos coerentes com o fenômeno investigado, como temperaturas máximas do motor e do óleo, nenhum subconjunto de variáveis apresentou comportamento discriminativo estável sob diferentes estratégias de amostragem. A recorrência de atributos operacionais e comportamentais nos *rankings* de relevância reflete a natureza da base analisada, predominantemente construída a partir de eventos discretos associados à ultrapassagem de limiares operacionais, e não de séries térmicas contínuas. Embora a

ocorrência de temperaturas acima de determinados patamares seja, em si, um indicativo direto do problema de superaquecimento, a ausência de informações mais detalhadas sobre a dinâmica térmica anterior a esses eventos limita a identificação de padrões progressivos que poderiam sinalizar o padrão de falha com maior antecedência.

Nesse cenário, os resultados do estudo indicam que, no estado atual da base de dados analisada, a predição antecipada de eventos de superaquecimento do motor apresenta limitações substanciais. Essas limitações estão associadas a fatores já identificados na etapa de análise exploratória e discutidos ao longo do trabalho, como a baixa granularidade das variáveis térmicas disponíveis, a escassez e a heterogeneidade dos eventos críticos e a redução da pureza da classe positiva em janelas temporais mais longas. Embora algoritmos como *Random Forest* e SVM tenham apresentado desempenho relativamente superior em alguns cenários, esses resultados não se mostraram estáveis ao longo das diferentes configurações e estratégias avaliadas, não atingindo níveis de confiabilidade compatíveis com a implantação de um sistema preditivo em ambiente operacional.

Além disso, a qualidade dos rótulos derivados das ordens de serviço, influenciada por práticas de diagnóstico e manutenção, mostrou-se determinante para a separabilidade entre as classes e para a estabilidade dos modelos. Em termos mais amplos, a base de telemetria analisada foi concebida prioritariamente para monitoramento reativo e acompanhamento operacional, e não para a antecipação de falhas, o que impõe limites estruturais à aplicação direta de abordagens preditivas baseadas exclusivamente nesses dados.

Nesse contexto, uma das principais contribuições deste trabalho reside na análise empírica e sistemática da viabilidade de aplicação de técnicas de aprendizado de máquina a dados reais de telemetria veicular, integrando de forma coerente as etapas de construção da base de dados, rotulagem, seleção de variáveis, modelagem e avaliação sob validação temporal realista. Ao explorar de maneira comparativa diferentes estratégias metodológicas, o estudo evidencia que os limites observados não se restringem a escolhas específicas de modelagem, mas decorrem principalmente das características estruturais dos dados disponíveis.

Por fim, os resultados obtidos fornecem subsídios técnicos objetivos para a tomada de decisão em empresas de transporte público urbano quanto à adoção de estratégias de manutenção preditiva. Em particular, o trabalho contribui ao explicitar as condições necessárias, as limitações esperadas e os principais gargalos associados à utilização de dados de telemetria operacional para a antecipação de falhas térmicas, orientando tanto o aprimoramento dos processos de coleta e instrumentação quanto o delineamento de abordagens analíticas mais adequadas para investigações futuras.

5.1 Trabalhos Futuros

Os resultados e limitações identificados neste estudo apontam direções claras para pesquisas futuras, tanto no aprimoramento da infraestrutura de dados quanto na evolução das abordagens analíticas aplicadas à manutenção preditiva em frotas de ônibus urbanos.

- a) **Aprimoramento da granularidade da telemetria:** a coleta de medições contínuas de temperatura do motor, pressão do óleo, rotação e outras variáveis relevantes ao longo da viagem permitiria capturar a dinâmica térmica do sistema de forma mais fiel. Essa maior resolução temporal é fundamental para identificar tendências progressivas de degradação, superando as limitações associadas a atributos agregados por viagem e a eventos disparados apenas por ultrapassagem de limites.
- b) **Ampliação do número e da diversidade de eventos críticos:** o aumento do histórico de falhas reais, associado à maior diversidade de causas de superaquecimento e condições operacionais, contribuiria para melhorar a representatividade da classe crítica. Uma base mais extensa permitiria reduzir a instabilidade dos modelos e possibilitar análises estratificadas por tipo de falha, mitigando a heterogeneidade interna observada neste estudo.
- c) **Estratégias de balanceamento da classe crítica:** técnicas de balanceamento, como *SMOTE*, *ADASYN* ou métodos híbridos de sub e superamostragem, podem ser exploradas em estudos futuros. No entanto, sua aplicação deve ser cuidadosamente avaliada, de modo a evitar a introdução de padrões artificiais que não correspondam ao comportamento físico real do sistema monitorado.
- d) **Integração de variáveis externas:** a incorporação de informações contextuais, como condições climáticas, topografia das rotas, carga transportada, perfil operacional das linhas e intensidade do tráfego, pode enriquecer a caracterização do esforço térmico imposto ao motor. Essas variáveis podem contribuir para reduzir a variabilidade não explicada e aumentar a capacidade discriminativa dos modelos.
- e) **Aprimoramento do sistema de ordens de serviço:** os resultados evidenciaram limitações na definição da classe crítica decorrentes da forma como as ordens de serviço são registradas. A ausência de informações padronizadas sobre diagnóstico técnico, causa raiz e ações corretivas dificulta a caracterização precisa dos eventos de superaquecimento. A adoção de procedimentos mais estruturados de registro, com validação técnica das falhas, permitiria construir rótulos mais confiáveis e alinhados ao fenômeno de interesse.
- f) **Reformulação do problema de modelagem:** melhorias na qualidade e no detalhamento dos registros de manutenção abririam novas possibilidades de enquadramento

do problema. Uma alternativa seria explorar comparações entre viagens anteriores e posteriores à intervenção de manutenção, desde que haja evidência objetiva de resolução da falha. Outra possibilidade é reformular o problema como a predição de tendências de degradação térmica, em vez de uma classificação binária, aproximando a modelagem de abordagens contínuas mais compatíveis com práticas de manutenção preditiva.

Em conjunto, essas direções delineiam caminhos consistentes para o avanço de sistemas de monitoramento e predição de falhas em veículos urbanos. O aprimoramento da qualidade e da estrutura dos dados, aliado ao uso de abordagens analíticas mais adequadas à dinâmica do fenômeno, tem potencial para viabilizar modelos mais robustos, interpretáveis e aplicáveis na prática operacional, contribuindo para maior confiabilidade, segurança e eficiência no transporte coletivo.

Referências

ABNT NBR 5462. *NBR 5462: Confiabilidade e manutenção*. Rio de Janeiro, 1994. Norma Brasileira. Citado na página 24.

AGGARWAL, C. C. *Neural Networks and Deep Learning*. Springer, 2018. Disponível em: <http://ndl.ethernet.edu.et/bitstream/123456789/88552/1/2018_Book_NeuralNetworksAndDeepLearning.pdf>. Citado na página 30.

ANTP, A. N. de T. P. *Sistema de Informações da Mobilidade Urbana: Relatório Geral 2018*. São Paulo, Brasil, 2020. Acesso em: 11 dez. 2024. Disponível em: <<https://www.antp.org.br/relatorios-a-partir-de-2014-nova-metodologia.html>>. Citado na página 18.

BEN-HUR, A.; WESTON, J. A user's guide to support vector machines. In: *Methods in Molecular Biology*. s.l.: Humana Press, 2010. v. 609, p. 223–239. Disponível em: <<https://www.cs.colostate.edu/~asa/pdfs/howto.pdf>>. Citado na página 30.

BREIMAN, L. Random forests. *Machine Learning*, Kluwer Academic Publishers, v. 45, p. 5–32, 2001. Citado na página 30.

CAI, B. *et al.* Fault detection and diagnostic method of diesel engine by combining rule-based algorithm and bns/bpms. *Journal of Manufacturing Systems*, Elsevier, v. 57, p. 148–157, 2020. Citado 3 vezes nas páginas 25, 26 e 27.

CAI, J. *et al.* Feature selection in machine learning: A new perspective. *Neurocomputing*, Elsevier, v. 300, p. 70–79, 2018. Citado na página 31.

CAO, L. *et al.* Study on diesel engine status analysis and fault diagnosis based on sae j1939 protocol. In: *Control Engineering Key Laboratory Chongqing Communication Institute*. [S.l.: s.n.], 2015. Citado na página 25.

CARVALHO, T. P. *et al.* A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, Elsevier, v. 137, p. 106024, 2019. ISSN 0360-8352. Available online 05 September 2019. Disponível em: <<https://doi.org/10.1016/j.cie.2019.106024>>. Citado 3 vezes nas páginas 20, 24 e 29.

CHEN, W. *A Rule-based Expert System for Predictive Maintenance of a Hybrid Bus*. Dissertação (Mestrado) — University of Ottawa, Ottawa, Canada, 2020. Citado 5 vezes nas páginas 19, 20, 26, 27 e 28.

CNPQ. *Engenharias - Plataforma Lattes - CNPq*. 2025. Disponível em: <<http://lattes.cnpq.br/web/dgp/engenharias>>. Citado na página 41.

CORAZZA, M. V. *et al.* Testing an innovative predictive management system for bus fleets: outcomes from the ravenna case study. *IET Intelligent Transport Systems*, IET, v. 12, n. 4, p. 286–293, 2018. Citado 3 vezes nas páginas 20, 25 e 26.

DAVARI, N. *et al.* A survey on data-driven predictive maintenance for the railway industry. *Sensors*, MDPI AG, v. 21, n. 17, p. 5739, aug 2021. Disponível em: <<https://doi.org/10.3390%2Fs21175739>>. Citado na página 30.

- Diário da Manhã. *Iveco Bus entrega 60 chassis para empresa de transporte coletivo de Goiânia*. 2021. Disponível em: <<https://www.dm.com.br/dmautos/iveco-bus-entrega-60-chassi-para-empresa-de-transporte-coletivo-de-goiania/>>. Citado 2 vezes nas páginas 9 e 44.
- ERSÖZ, O. Ö. *et al.* A systematic literature review of the predictive maintenance from transportation systems aspect. *Sustainability*, MDPI, v. 14, n. 21, p. 14536, 2022. Citado 5 vezes nas páginas 19, 20, 24, 25 e 29.
- FAN, Y.; NOWACZYK, S.; RÖGNVALDSSON, T. Using histograms to find compressor deviations in bus fleet data. In: *The SAIS Workshop 2014 Proceedings*. [S.l.]: Swedish Artificial Intelligence Society (SAIS), 2014. p. 123–132. Citado na página 27.
- FARSI, M.; Schiffmann, W. R. An overview of controller area network. *Computing & Control Engineering Journal*, IEE, p. 113–120, June 1999. Citado 2 vezes nas páginas 27 e 28.
- FLEMING, W. J. Overview of automotive sensors. *IEEE Sensors Journal*, IEEE, v. 1, n. 4, p. 296–308, 2001. Citado na página 27.
- FRADEN, J. *Handbook of Modern Sensors: Physics, Designs, and Applications*. 5th. ed. Cham, Switzerland: Springer, 2015. ISBN 978-3-319-19302-1. Disponível em: <<https://link.springer.com/book/10.1007/978-3-319-19303-8>>. Citado na página 27.
- GIANNOULIDIS, A.; GOUNARIS, A. A context-aware unsupervised predictive maintenance solution for fleet management. *Journal of Intelligent Information Systems*, Springer, v. 60, p. 521–547, 2022. Citado 4 vezes nas páginas 20, 26, 27 e 37.
- GUO, Y.; XU, B.; XIAO, Q. Maintenance reliability index system and its analysis method of public transit. In: *IOP Conference Series: Materials Science and Engineering*. [S.l.]: IOP Publishing, 2018. v. 466, n. 1, p. 012058. Citado 2 vezes nas páginas 18 e 26.
- GÜVEN, ; ŞAHİN, H. Predictive maintenance based on machine learning in public transportation vehicles. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, Bandırma Onyedü Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, v. 4, n. 1, p. 89–98, 2022. Citado 3 vezes nas páginas 19, 20 e 39.
- JEVINGER, *et al.* Artificial intelligence for improving public transport: a mapping study. *Public Transport*, Springer Science and Business Media LLC, v. 0123456789, 2023. Disponível em: <<https://doi.org/10.1007/s12469-023-00334-7>>. Citado na página 29.
- JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, v. 349, n. 6245, p. 255–260, 2015. Citado na página 29.
- KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. *Artificial Intelligence*, Elsevier Science B.V., v. 97, p. 273–324, 1997. Citado na página 30.
- LI, C.; GUO, C.; XUE, Y. Early warning model of engine fault of liquefied natural gas bus. In: IEEE. *2018 4th International Conference on Transportation Information and Safety (ICTIS)*. Hangzhou, China, 2018. p. 1054–1060. Citado 2 vezes nas páginas 28 e 39.
- LI, J. *et al.* Research on construction of crude set model of critical fault information for bus based on CAN-BUS data. *IEEE Access*, v. 7, p. 141489–141499, 2019. Citado na página 26.

- MASSARO, A.; SELICATO, S.; GALIANO, A. Predictive maintenance of bus fleet by intelligent smart electronic board implementing artificial intelligence. *IoT*, MDPI, v. 1, n. 2, p. 180–197, 2020. Citado 3 vezes nas páginas 26, 28 e 39.
- MCKINLEY, T. *et al.* Identifying nox sensor failure for predictive maintenance of diesel engines using explainable ai. *PHM Society European Conference*, v. 5, n. 1, p. 11, 2020. Disponível em: <<https://doi.org/10.36001/phme.2020.v5i1.1267>>. Citado 2 vezes nas páginas 28 e 39.
- MENESES, R. *Fotografia de ônibus urbano em Goiânia*. 2023. Disponível em: Ônibus Brasil. Disponível em: <<https://onibusbrasil.com/RafaelMeneses/10842074>>. Citado 2 vezes nas páginas 9 e 44.
- MOHANKUMAR, P. *et al.* A review of micromachined sensors for automotive applications. *Measurement*, Elsevier, v. 140, p. 305–322, 2019. Citado na página 27.
- MOHER, D. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Journal of Chinese Integrative Medicine*, v. 7, p. 889–896, September 2009. Citado na página 17. Disponível em: <<https://doi.org/10.3736/jcim20090918>>. Citado na página 34.
- NTU, A. N. das Empresas de T. U. *Anuário NTU: 2023-2024*. Brasília, Brasil: NTU, 2024. 72 p. Acesso em: 11 dez. 2024. ISBN 2317-8868. Disponível em: <<http://www.ntu.org.br>>. Citado na página 18.
- RMTC, R. M. de T. C. *Boletim Diário de Fluxo e Demanda*. 2025. Disponível em: <<https://www.rmtcgoiania.com.br/blog/category/boletim-diario-de-fluxo-e-demanda/>>. Acesso em: 18 jan. 2025. Disponível em: <<https://www.rmtcgoiania.com.br/blog/category/boletim-diario-de-fluxo-e-demanda/>>. Citado na página 42.
- RMTC, R. M. de Transportes Coletivos de G. *Ônibus da empresa Rápido Araguaia estão de volta às ruas*. 2015. Disponível em: <<https://www.rmtcgoiania.com.br/blog/2015/12/23/onibus-da-empresa-rapido-araguaia-estao-de-volta-as-ruas/>>. Citado 2 vezes nas páginas 9 e 44.
- RÖGNVALDSSON, T. *et al.* Self-monitoring for maintenance of vehicle fleets. *Data Mining and Knowledge Discovery*, Springer, v. 32, n. 2, p. 344–384, 2018. Citado 4 vezes nas páginas 20, 26, 27 e 37.
- ROSETO, F. *et al.* Real-world fuel efficiency and emissions from an urban diesel bus engine under transient operating conditions. *Applied Energy*, Elsevier, v. 261, p. 114442, 2020. Citado na página 26.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 4th. ed. [S.l.]: Pearson, 2021. Citado na página 29.
- SARKER, I. H. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, Springer Nature Singapore Pte Ltd, v. 2, n. 1, p. 1–21, 2021. Citado na página 29.
- SIEGEL, J. E.; ERB, D. C.; SARMA, S. E. A survey of the connected vehicle landscape—architectures, enabling technologies, applications, and development areas. *IEEE Transactions on Intelligent Transportation Systems*, IEEE, v. 19, n. 8, p. 2391–2402, 2018. Citado na página 27.

SUSTO, G. A. *et al.* Machine learning for predictive maintenance: A multiple classifiers approach. *IEEE Transactions on Industrial Informatics*, IEEE, v. 11, n. 3, p. 812–820, 2015. Citado na página 29.

TURNER, J. D.; AUSTIN, L. Sensors for automotive telematics. *Measurement Science and Technology*, v. 11, n. 2, p. R58–R79, 2000. Citado 2 vezes nas páginas 26 e 28.

TURRIONI, J.; MELLO, C. *Metodologia de Pesquisa em Engenharia de Produção*. 1. ed. [S.l.]: UNIFEI, 2012. Citado na página 41.

ZHAI, Y. The impact of intelligent transportation technology on the operational efficiency of public transportation enterprises. *Academic journal of science and technology*, Darcy Roy Press Co. Ltd., v. 12, n. 3, p. 134–138, 2024. Citado 2 vezes nas páginas 19 e 20.

APÊNDICE A – Descrição das Variáveis Utilizadas

Este apêndice apresenta a descrição das variáveis utilizadas na construção do conjunto de dados empregado neste estudo. As variáveis listadas correspondem ao conjunto final de atributos considerados na etapa de modelagem preditiva, após filtragem inicial e pré-processamento dos dados descritos no Capítulo 3.

O conjunto de dados foi construído a partir de um banco de dados operacional que integra registros de telemetria veicular, eventos operacionais e informações de manutenção. A formação do *dataset*, incluindo os critérios de integração, agregação e rotulagem das variáveis ao nível de viagem, é detalhada no Capítulo 3.

As variáveis finais utilizadas nos modelos não correspondem diretamente a colunas originais do banco de dados, mas a indicadores estatísticos derivados de eventos registrados ao longo da duração de cada viagem. Dessa forma, o conjunto de dados analisado representa uma descrição agregada do comportamento operacional dos veículos, construída especificamente para fins de análise preditiva.

As variáveis seguem um padrão de nomenclatura que indica sua origem e forma de agregação. O prefixo *ra_* identifica eventos definidos pela empresa gestora da frota, a partir de regras operacionais internas e limites estabelecidos conforme critérios próprios de monitoramento e manutenção. Esses eventos são armazenados em tabelas específicas do banco de dados, como *ra_temperatura_do_motor*, *ra_pressao_do_oleo* e *ra_falha_de_shutdown*.

O prefixo *x_mix_* refere-se a eventos definidos pela empresa fornecedora do sistema de telemetria, com regras e limiares próprios, independentes do processo de manutenção. Esses eventos representam ocorrências como aceleração brusca, frenagem brusca e excesso de velocidade, conforme documentado nas tabelas *x_mix_aceleracao_brusca_os*, *x_mix_freada_brusca_os* e *x_mix_excesso_de_velocidade_os*. Em ambos os casos, os eventos são posteriormente agregados por viagem, conforme a metodologia descrita neste trabalho.

Ao longo das tabelas deste apêndice, o símbolo *** é utilizado como notação abreviada para representar conjuntos de variáveis que compartilham o mesmo prefixo semântico, diferenciando-se apenas pelo sufixo estatístico. Esses sufixos indicam estatísticas agregadas por viagem, incluindo *mean* (média), *max* (valor máximo), *min* (valor mínimo), *std* (desvio padrão) e *count* (quantidade de ocorrências do evento durante a viagem).

Por exemplo, a notação `ra_temperatura_do_motor_maxima_critica_*` representa o conjunto de variáveis derivadas das ocorrências de temperatura do motor em faixa crítica, agregadas segundo as estatísticas mencionadas.

As Tabelas a seguir apresentam as variáveis utilizadas neste trabalho, organizadas por grupos funcionais, de acordo com o tipo de informação operacional representada. Para cada grupo, são listados os nomes das variáveis e uma descrição resumida de seu significado no contexto da telemetria veicular e da manutenção preditiva.

Tabela A.1 – Variáveis térmicas do motor e do óleo.

Variável	Descrição
<code>ra_temperatura_do_motor_mean,</code> <code>ra_temperatura_do_motor_max,</code> <code>ra_temperatura_do_motor_min,</code> <code>ra_temperatura_do_motor_std,</code> <code>ra_temperatura_do_motor_count</code>	Estatísticas agregadas da temperatura do motor registradas durante a viagem.
<code>ra_temperatura_do_motor_media_mean,</code> <code>ra_temperatura_do_motor_media_max,</code> <code>ra_temperatura_do_motor_media_min,</code> <code>ra_temperatura_do_motor_media_std,</code> <code>ra_temperatura_do_motor_media_count</code>	Estatísticas agregadas da temperatura média do motor ao longo da viagem.
<code>ra_temperatura_do_motor_maxima_critica_*</code>	Indicadores estatísticos associados à ocorrência de temperatura do motor em faixa crítica.
<code>ra_temperatura_do_oleo_media_*</code>	Estatísticas agregadas da temperatura média do óleo do motor durante a viagem.
<code>ra_temp_do_motor_em_viagem_maximo_*</code>	Valores máximos de temperatura do motor registrados em eventos de telemetria durante a viagem.

Tabela A.2 – Variáveis de condição mecânica e funcionamento do motor.

Variável	Descrição
ra_pressao_do_oleo_media_iveco_*, ra_pressao_do_oleo_minima_critica_iveco_*	Estatísticas da pressão do óleo do motor e ocorrência de pressão em nível crítico.
ra_rotacao_excessiva_*, MaxRpm	Indicadores de rotação excessiva do motor e rotação máxima registrada.
ra_falha_de_shutdown_*	Eventos associados a falhas no desligamento do motor.
EngineSeconds, StartEngineSeconds, EndEngineSeconds	Tempo de funcionamento do motor durante a viagem.

Tabela A.3 – Variáveis de condução e comportamento operacional.

Variável	Descrição
ra_freada_brusca_*, ra_aceleracao_brusca_*, ra_curva_brusca_*	Indicadores de eventos de condução agressiva durante a viagem.
ra_uso_indevido_do_pedal_de_acelerador_85_*	Eventos de uso inadequado do acelerador acima de limiar definido.
ra_marcha_lenta_*, ra_marcha_lenta_excessiva_*	Estatísticas relacionadas ao tempo em marcha lenta e marcha lenta excessiva.
ra_banguela_*	Indicadores de condução em ponto morto.
ra_acelerando_parado_*, ra_acelerando_parado_garagem_*	Eventos de aceleração com o veículo parado.

Tabela A.4 – Variáveis de velocidade, deslocamento e operação da viagem.

Variável	Descrição
AverageSpeed, Grupo_AverageSpeed, MaxSpeedKilometersPerHour	Indicadores de velocidade média e máxima da viagem.
DistanceKilometers, Duration, DrivingTime, StandingTime	Distância percorrida e tempos associados à viagem.
FuelUsedLitres	Consumo de combustível durante a viagem.
MaxAccelerationKilometersPerHourPerSecond, MaxDecelerationKilometersPerHourPerSecond	Aceleração e desaceleração máximas registradas.
StartOdometerKilometers, EndOdometerKilometers	Leitura do odômetro no início e no fim da viagem.

Tabela A.5 – Variáveis contextuais e temporais.

Variável	Descrição
Dia_semana, Tipo_dia, Horario_trip	Indicadores categóricos de contexto temporal da viagem.
x_mix__os__	Indicadores estatísticos agregados por viagem a partir de eventos definidos pela empresa fornecedora do sistema de telemetria, com regras e limiares próprios.

As variáveis descritas neste apêndice compõem o espaço de atributos utilizado nos experimentos de aprendizado de máquina apresentados ao longo deste trabalho. A organização por grupos funcionais visa facilitar a interpretação dos resultados e a análise do impacto das diferentes categorias de variáveis no desempenho dos modelos.