UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL
MESTRADO EM ENGENHARIA CIVIL

JÉSSICA ANDRADE PRATA

DISSERTAÇÃO DE MESTRADO

PREDIÇÃO DE FALHA EM MOTOR DIESEL DE LOCOMOTIVA BASEADA NA ANÁLISE DO ÓLEO LUBRIFICANTE POR MEIO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA

VITÓRIA
2020

JÉSSICA ANDRADE PRATA

DISSERTAÇÃO DE MESTRADO

PREDIÇÃO DE FALHA EM MOTOR DIESEL DE LOCOMOTIVA BASEADA NA ANÁLISE DO ÓLEO LUBRIFICANTE POR MEIO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Engenharia Civil, na área de concentração Transportes.

Orientador: Prof. Dr. Rodrigo de Alvarenga Rosa.

Coorientador: Dr. Jodelson Aguilar Sabino

VITÓRIA
2020

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

---

Prata, Jéssica Andrade, 1992-
P912p Predição de falha em motor diesel de locomotiva baseada na análise do óleo lubrificante por meio de técnicas de aprendizado de máquina / Jéssica Andrade Prata. - 2020.
104 f. : il.

Orientador: Rodrigo de Alvarenga Rosa.
Coorientador: Jodelson Aguilar Sabino.
Dissertação (Mestrado em Engenharia Civil) - Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Motores a diesel. 2. Locomotivas. 3. Aprendizado do computador. 4. Inteligência artificial. I. Rosa, Rodrigo de Alvarenga. II. Sabino, Jodelson Aguilar. III. Universidade Federal do Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 624

PREDIÇÃO DE FALHA EM MOTOR DIESEL DE LOCOMOTIVA BASEADA NA ANÁLISE DO ÓLEO LUBRIFICANTE POR MEIO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA

Dissertação apresentada ao Curso de Mestrado em Engenharia Civil do Programa de Pós-Graduação em Engenharia Civil da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Engenharia Civil, área de Transportes.

Aprovada em _____ de _____ de 2020 por:

_____
Prof. Dr. Rodrigo de Alvarenga Rosa
Doutor em Engenharia Elétrica
Orientador

_____
Dr. Jodelson Aguilar Sabino
Vale S/A
Coorientador

_____
Prof. Dr. Patrício José Moreira Pires
Universidade Federal do Espírito Santo
Examinador Interno

_____
Prof. Dr. Daniel Cruz Cavalieri
Instituto Federal do Espírito Santo
Examinador Externo

Vitória – ES, março de 2020

AGRADECIMENTOS

A Deus e as boas energias do universo que me guiam. Aos meus pais Maria Edna e José Luiz pelos valores repassados, amor e por me permitirem voar. Ao meu irmão João, por entender minha ausência. A minha tia Camu por ser a melhor amiga e impulsionadora durante esses dois anos de mestrado e também nos demais anos da minha vida.

Ao orientador e amigo professor Dr. Rodrigo Rosa, por acreditar nesse desafio e sonhar junto. Pelo lado humano e pelo conhecimento que transmite com excelência.

Ao coorientador Jodelson Sabino pela oportunidade e por todo conhecimento, tempo e treinamentos compartilhados.

A Vale S/A por me permitir desenvolver este trabalho, me fornecendo as informações necessárias, em especial as equipes da Oficina de Locomotivas de Tubarão, TI e AI Center pela participação e atenção.

Aos amigos, familiares e a todos que direta ou indiretamente contribuíram com a realização deste trabalho.

“Na verdade, a compreensão de que a natureza de todas as coisas muda constantemente é a chave para a verdadeira felicidade, pois significa que por pior que seja a situação, ela mudará.”

Daisaku Ideka

RESUMO

As locomotivas diesel-elétricas são as mais utilizadas nas ferrovias de carga brasileiras, sendo o motor diesel parte indispensável de seu funcionamento. Caso o motor diesel apresente falhas, o trem tracionado pelas locomotivas não circulará, podendo levar a um não cumprimento dos planos de transporte e custos para as ferrovias. Esta dissertação teve como objetivo investigar dois algoritmos de aprendizado de máquina: a Rede Neural Artificial (RNA) e a Árvore de Decisão XGBoost, na predição de falhas em motores diesel de locomotivas por meio de dados reais adquiridos na Estrada de Ferro Vitória-Minas, localizada na Região Sudeste do Brasil. Os dados de entrada contemplaram os resultados de análises de óleo lubrificante, como os dos ensaios de espectrometria e físico-químico. O histórico de falhas que já ocorreram nos motores, o consumo de combustível das locomotivas e a movimentação de motores entre locomotivas também foram utilizados como entrada. Foram realizadas análises de correlação, normalização, e balanceamento dos dados. Além disso, para ambos os métodos de predição, foi aplicada a validação cruzada k-fold. Os classificadores foram avaliados pela acurácia, *recall* e pelas indicações de classificação na matriz de confusão. Ambos os algoritmos obtiveram bom desempenho para predizer falhas nos motores diesel, podendo auxiliar o planejador das manutenções dos motores nas oficinas de locomotivas.

Palavras Chaves: Motor diesel de locomotiva. Rede Neural Artificial. Árvore de Decisão XGBoost. Óleo lubrificante de motor diesel.

ABSTRACT

Diesel-electric locomotives are the most used in Brazilian freight railroads, with the diesel engine being an indispensable part of its operation. If the diesel engine fails, the train pulled by the locomotives will not circulate, which may lead to non-compliance with transport plans and costs for the railways. This dissertation aimed to investigate two machine learning algorithms: the Artificial Neural Network (RNA) and the XGBoost Decision Tree, in the prediction of locomotive diesel engine failures through real data acquired on the Vitória-Minas Railway, located in the Southeast Region of Brazil. The input data included the results of lubricating oil analyzes, such as those of spectrometry and physical-chemical tests. The history of engine failures that have already occurred, the fuel consumption of the locomotives and the movement of engines between locomotives were also used as input. Correlation, normalization, and data balancing analyzes were performed. In addition, for both prediction methods, k-fold cross-validation was applied. The classifiers were evaluated for accuracy, recall and classification indications in the confusion matrix. Both algorithms achieved good performance to predict failures in diesel engines, being able to assist the engine maintenance planner in locomotive workshops.

Keywords: Locomotive diesel engine. Artificial Neural Network. XGBoost Decision Tree. Diesel engine lubricanting oil.

LISTA DE FIGURAS

Figura 1 – Pseudocódigo da RNA em Backpropagation 24
Figura 2 – Divisão dos dados por validação cruzada k-*fold* 26
Figura 3 – Matriz de confusão 28
Figura 4 - Estrutura da rede e seu processo de modelagem 41
Figura 5 – Montagem básica de uma locomotiva diesel-elétrica 52
Figura 6 – Motor diesel GE 7FDL16 aplicado nas locomotivas DASH9 W.....53
Figura 7 – Quantidade de OM criadas relacionadas a motor diesel de 2009 a 2018 67
Figura 8 – RNA investigadas para predição de falhas em motores diesel de locomotivas 80
Figura 9 – Matriz de confusão da etapa de validação e teste da configuração 2 com a predição da falha em até 30 dias 86
Figura 10 – Matriz de confusão da etapa de validação e teste da configuração 4 com a predição da falha em até 30 dias 87
Figura 11 – Matriz de confusão da etapa de validação e teste da configuração 6 com a predição da falha em até 30 dias 87
Figura 12 – Matriz de confusão da etapa de validação e teste da configuração 13 com a predição da falha em até 30 dias 91
Figura 13– Matriz de confusão da etapa de validação e teste da configuração 14 com a predição da falha em até 30 dias 92
Figura 14 – Comparação entre as configurações que obtiveram melhores desempenhos da RNA e da Árvore XGBoost.....95

LISTA DE QUADROS

Quadro 1 – Sintomas, defeitos e sistemas do motor diesel relacionados às análises de óleo lubrificante 35

Quadro 1 – Sintomas, defeitos e sistema do motor diesel relacionados às análises de óleo lubrificante 36

Quadro 2 – Códigos selecionados para compor os dados de falhas (histórico de OM) do motor diesel 62

Quadro 3 – Períodos e sistemas informatizados utilizados para armazenar os históricos de dados 65

Quadro 4 – Exemplo da tabela de base de dados limitados em 6 instâncias 75

Quadro 5 – Parâmetros definidos para o desenvolvimento da Árvore XGBoost 81

LISTA DE TABELAS

Tabela 1 – Comparativo entre às aplicações e métodos utilizados nas referências bibliográficas encontradas e está dissertação 44

Tabela 2 – Elementos químicos com teores elevados (espectrometria) e os possíveis componentes do motor diesel de locomotivas GE que podem estar sofrendo desgaste 47

Tabela 3 – Valores encontrados nos ensaios físico-químicos e os possíveis componentes do motor diesel de locomotivas GE que podem estar sofrendo desgaste 47

Tabela 4 – Faixas com os limites normais e críticos dos valores encontrados no ensaio de espectrometria 48

Tabela 5 – Faixas com os limites normais e críticos dos valores encontrados nos ensaios físico-químicos 48

Tabela 6 – Dados que compõem as planilhas dos ensaios físico-químicos e espectrométrico 58

Tabela 7 – Dados que compõem as planilhas de falhas no motor diesel 61

Tabela 8 – Dados que compõem a planilha de identificação das locomotivas 63

Tabela 9 – Dados que compõem a planilha de consumo de combustível das locomotivas 64

Tabela 10 – Dados que compõem a planilha de movimentação de motores entre locomotivas 65

Tabela 11 – Correlações das variáveis com a predição de falha no motor diesel 69

Tabela 12 – Variáveis que apresentaram altas correlações entre si 69

Tabela 13 – Configurações da RNA para falha em até 30 dias 78

Tabela 14– Configurações da RNA para falha em até 45 dias 78

Tabela 15 – Configurações da Árvore XGBoost para falha em até 30 dias e 45 dias 82

Tabela 16 – Resultados de validação das RNA para predição de falha em até 30 dias 83

Tabela 17 – Resultados da validação das RNA para predição de falha em até 45 dias 84

Tabela 18 – Resultados das configurações 2, 4 e 6 na etapa de validação e teste 85

Tabela 19 – Resultados das matrizes de confusão das configurações 2, 4 e 6 na etapa de teste 88

Tabela 20 – Resultados das configurações 13, 14, 15 e 16 na etapa de validação 90

Tabela 21 – Resultados das configurações da Árvore XGBoost nas etapas de validação e teste 91

Tabela 22 – Resultados das matrizes de confusão das configurações 13 e 14 na etapa de teste 93

Tabela 23 – Acurácia e *recall* das configurações 2, 4, 6, 13 e 14 da fase de testes . 94

Tabela 24 – Comparação entre as configurações que obtiveram melhores desempenhos da RNA e da Árvore XGBoost..... 94

LISTA DE SIGLAS

Acc – Acurácia

AES – *Atomic Emission Spectrometry*

BPNN – *Backpropagation Neural Network*

EFVM – Estrada de Ferro Vitória – Minas

ER – Erros relativos

FAPES – Fundação de Amparo à Pesquisa e Inovação do Espírito Santo

FN – Falso negativo

FP – Falso positivo

FRF – Função de resposta em frequência

FTIR – *Fourier Transformation Infra Red*

GBM – Gradient Boosting Machines

GE – *General Electric*

IA – Inteligência Artificial

k-NN – *K-Nearest Neighbor*

LM – *Levenberg-Marquardt*

LNF – *Laser Net Finder*

Mg KOH/g - *milligrams of potassium hydroxide per gram*

MS – Média de serviços

NN – *Nearest Neighbor*

NORMAX – *Non-linear Autoregressive Moving Average with Exogenous Inputs*

OM – Ordem de Manutenção

ORC - *Organic Rankine Cycle*

PHM – *Proportional Hazards Modeling*

PNN – *Probabilistic Neural Network*

ppm – partes por milhão

R – Coeficiente de correlação

RML – Regressão Múltipla Linear

RNA – Rede Neural Artificial

SMOTE – *Synthetic Minority Over-sampling Technique*

TBN – *Total Base Number*

TU – toneladas úteis

UFES – Universidade Federal do Espírito Santo

VP – Verdadeiro positivo

VN – Verdadeiro negativo

XGBoost – Extreme Gradient Boosting

SUMÁRIO

1 INTRODUÇÃO 14
1.1 Objetivos 17
1.1.1 Objetivo Geral 17
1.1.2 Objetivos Específicos 17
1.2 Justificativa 18
1.3 Estrutura da proposta de dissertação 19
2 REFERÊNCIAL TEÓRICO 20
2.1 Técnicas de aprendizado de máquina 20
2.1.1 Rede neural artificial 20
2.1.2 Árvore XGBoost 30
2.2 Revisão bibliográfica envolvendo predição de falhas em motores diesel ... 32
3 DESCRIÇÃO E CONCEITUAÇÃO DO PROBLEMA 45
3.1 O problema analisado 45
3.2 Motor diesel de locomotivas 51
3.3 Óleo lubrificante de motores diesel 53
4 LEVANTAMENTO E TRATAMENTO DOS DADOS 57
5 REDE NEURAL ARTIFICIAL DE CLASSIFICAÇÃO PARA PREDIÇÃO DE FALHAS EM MOTORES DIESEL DE LOCOMOTIVAS COM BASE NA ANÁLISE DO ÓLEO LUBRIFICANTE 76
6 ÁRVORE XGBOOST DE CLASSIFICAÇÃO PARA PREDIÇÃO DE FALHAS EM MOTORES DIESEL DE LOCOMOTIVAS COM BASE NA ANÁLISE DO ÓLEO LUBRIFICANTE 81
7 RESULTADOS EXPERIMENTAIS 83
7.1 Resultados e análises da Rede Neural Artificial (RNA) 83
7.2 Teste da Rede Neural Artificial (RNA) 85
7.3 Resultados e análises da Árvore XGBoost 89
7.4 Teste da Árvore XGBoost 90
7.5 Comparação entre a Rede Neural Artificial (RNA) e a Árvore XGBoost 93
8 CONCLUSÕES 96
8.1 Trabalhos Futuros 98
REFERÊNCIAS 99

1 INTRODUÇÃO

No Brasil existem atualmente 12 concessionárias que realizam transportes de cargas por meio da ferrovia. No ano de 2017 todas estas concessionárias juntas transportaram 538.480 milhões de toneladas úteis (TU) de materiais como: minério de ferro, grãos (milho, soja, farelo de milho, etc.), carvão, fertilizantes, entre outros (ANTT, 2018). Mesmo diante de um crescimento anual de cargas transportadas por ferrovia no Brasil, existe ainda a necessidade de desenvolvimento do setor no país, não somente na construção de novas ferrovias, mas na melhoria contínua das existentes, principalmente no que envolve elevar o nível de segurança e eficácia das operações e das áreas envolvidas neste processo (MARCHETTI e WANKE, 2017).

A movimentação de cargas na ferrovia se dá por meio de trens, compostos por locomotivas e vagões, que circulam pela via férrea. No Brasil as ferrovias de carga utilizam, em sua grande maioria, locomotivas diesel-elétricas, as quais geram sua própria energia elétrica a partir de um motor diesel que aciona um gerador de energia elétrica. Cada locomotiva diesel-elétrica possui um motor diesel e vários motores elétricos, conhecidos como motores de tração, podendo percorrer grandes percursos sem interrupção para manutenção ou reabastecimento (ROSA e RIBEIRO, 2017).

O motor diesel é parte essencial da locomotiva pois gera a energia elétrica que alimenta os motores de tração. Sendo assim, monitorar seus componentes buscando um sistema confiável para operação da locomotiva é fundamental para se evitar perdas expressivas e planejar as manutenções. Analisar as falhas que já ocorreram no motor diesel, suas características e investigá-las é uma forma eficaz de identificar a causa do problema e tratar o sistema para que não ocorram próximas falhas. Para identificar essas falhas, por exemplo, podem-se utilizar análises laboratoriais do óleo lubrificante do motor (FONTE et al., 2017).

O óleo lubrificante é um dos materiais que está em contato com o motor diesel de forma obrigatória, ou seja, necessário ao seu funcionamento, e sua análise laboratorial é uma das maneiras de monitorar o motor diesel, identificando os primeiros sintomas de desgaste de seus componentes. A análise laboratorial envolve, por exemplo, a identificação da quantidade de partículas, composição, forma e tamanho, análise de viscosidade e Número Total de Basicidade (*Total Base Number* - TBN). Essas

informações indicam as condições dos componentes do motor diesel sem a necessidade de parar a locomotiva para que seja realizada uma avaliação (KIMURA, 2010).

Essa dissertação propõe uma Rede Neural Artificial (RNA) e um algoritmo Árvore XGBoost (*Extreme Gradient Boosting*), para predição de falhas em motores diesel de locomotivas diesel-elétricas por meio dos resultados das análises do óleo lubrificante usado nesses motores. A predição consiste em indicar se haverá falha, ou não, no motor diesel no período de 30 e 45 dias, sendo esses períodos definidos em conjunto com as equipes que realizam o planejamento e execução das manutenções.

As variáveis utilizadas nas RNA desenvolvidas e testadas nesta dissertação, são contínuas e categóricas, sendo que na camada de entrada da RNA tem-se 20 variáveis, ou seja, 20 neurônios: 1)Destilação, 2)Diluição, 3)Viscosidade, 4)TBN, 5)Ag, 6)Al, 7)Cr, 8)Cu, 9)Fe, 10)Na, 11)Ni, 12)Pb, 13)Si, 14)Sn, 15)Zn, 16)Mg, 17)Mo, 18)Mn, 19)Ca e 20)Litros abastecidos.

Para definição da quantidade de neurônios na camada intermediária, foram variadas as configurações em 1 e 2 camadas intermediárias, e definiu-se que cada camada teria a mesma quantidade de neurônios. Foram feitos testes utilizando 5, 10, 13, 15, 17, 20, 25, 30, 35, 35, 40, 50, 70, 100 e 200 neurônios em cada uma dessas duas camadas intermediárias. Das 15 opções de valores para as duas camadas intermediárias foram selecionadas as que apresentaram os melhores resultados para apresentação, sendo elas as camadas intermediárias com 13, 35 e 40 neurônios.

Ao total foram apresentadas 12 configurações de RNA, sendo 6 para a predição da falha em até 30 dias e as outras 6 para a predição da falha em até 45 dias. Em ambos os períodos (30 e 45 dias) foram apresentadas 3 configurações com uma camada intermediária variando entre 13, 35 e 40 neurônios, e 3 configurações com 2 camadas intermediárias também variando em 13, 35 e 40 neurônios em cada camada.

A camada de saída indica a predição de falha no motor diesel, objetivo principal dessa dissertação. Como a predição é binária, na qual “0” indica que não haverá falha e “1” que haverá falha, então a camada de saída apresenta um único neurônio.

Os mesmos 20 neurônios de entrada utilizados nas 12 configurações das RNA para predição das falhas em motores diesel, foram utilizados nos testes com o algoritmo Árvore XGBoost.

Os dados utilizados na RNA e na Árvore XGBoost foram os históricos de análises de óleo lubrificante utilizado na lubrificação dos motores diesel; os históricos de falhas que já ocorreram nesses mesmos motores; a movimentação dos motores diesel entre as locomotivas, o que inclui os motores novos e os motores que passaram por algum reparo e retornaram para a locomotiva de origem ou outra; e o consumo de diesel de cada locomotiva, o que indica o nível de utilização de cada uma das locomotivas.

As análises de óleo lubrificante que compõem o histórico disponibilizado pela ferrovia incluem dois tipos de ensaios laboratoriais: ensaios físico-químicos e ensaio espectrométrico. Para realização dos ensaios físico-químicos são necessários equipamentos específicos para cada tipo de análise, como por exemplo análise de creptação, destilação, diluição, TBN, viscosidade e insolúveis. Para o ensaio espectrométrico é necessário um único equipamento, o qual é capaz de identificar a concentração de elementos químicos presentes no óleo lubrificante em partes por milhão (ppm), como, por exemplo a concentração de chumbo, ferro e alumínio.

Os dados utilizados nessa dissertação, disponibilizados de 2009 a 2018 pela Estrada de Ferro Vitória-Minas (EFVM), são informações históricas registradas das 216 locomotivas da fabricante General Eletric (GE) e seus motores diesel, bem como das análises de óleo realizadas no Laboratório de Análises Químicas da Vale S/A, localizado em Vitória (ES – Brasil). Com 905km de extensão, a EFVM interliga os pontos de carregamento localizados no estado de Minas Gerais ao porto de Tubarão, localizado também em Vitória.

É necessário ressaltar que os históricos foram obtidos por intermédio da empresa Vale S/A, por meio do convênio firmado junto a Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES), número 75528452-2016, no qual a Vale S/A autoriza a utilização de alguns de seus dados operacionais em projetos acadêmicos. Porém para manter em sigilo os dados confidencias da empresa, os valores reais utilizados nessa dissertação foram substituídos, fazendo uso de uma função não divulgada.

Os quatro históricos já apresentados, mais um quinto documento contendo o registro das 216 locomotivas, no qual é possível diferenciá-las das demais locomotivas que trafegam na EFVM, formam as cinco bases de dados utilizadas nessa dissertação. As cinco bases de dados correspondem a 898.553 registros tratados. Porém com a criação de instâncias limitadas a 6 análises de óleo entre falhas, como será explicado no decorrer dessa dissertação, o *dataset* final (conjunto de dados contendo todas as 5 bases de dados) caiu para 45.469 registros, sendo que 40.220 desses registros foram usados para treinamento e validação e os 5.249 registros restantes foram usados para o teste das configurações da RNA e da Árvore XGBoost.

Para implementação dos dados foi utilizado o IBM SPSS Modeler, versão 18.1 (IBM, 2019). Para o modelo da RNA foi utilizada a estrutura tipo *Perceptron* de múltiplas camadas (PMC), com o neurônio artificial *Perceptron* e o algoritmo *Backpropagation*. Para divisão dos conjuntos de registros foi utilizada, em ambos os modelos, RNA e Árvore XGboost, uma validação cruzada denominada k-*fold,* sendo que nessa dissertação o *f* foi igual a 10.

Para analisar os resultados dos dois modelos, RNA e Árvore XGBoost, a fim de verificar principalmente o desempenho da RNA, foram selecionados os melhores resultados de ambos os modelos por meio das métricas acurácia e *recall.* Após identificar os melhores resultados nos dois modelos em separado, foi feita uma comparação entre a RNA e a Árvore XGBoost.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Analisar algoritmos de aprendizado de máquina para predição de falha em motor diesel de locomotiva com base na análise do óleo lubrificante utilizado nesses motores.

1.1.2 Objetivos Específicos

Os objetivos específicos são:

- Analisar diferentes configurações dos modelos de RNA e Árvore XGBoost a fim de encontrar a configuração com maior assertividade na predição das falhas nos motores diesel de locomotivas;
- Identificar qual dos dois modelos, RNA e Árvore XGBoost, apresentou maior competência em predizer falhas em motores diesel de locomotivas;
- Identificar possíveis melhorias nos processos que envolvem as análises de óleo lubrificante e os apontamentos de falhas dos motores diesel na ferrovia;
- Contribuir com o planejamento futuro das manutenções das locomotivas da ferrovia, no que envolve os motores diesel.

1.2 JUSTIFICATIVA

Para conhecimento prévio quanto à utilização de aprendizado de máquina para predição de falhas em motores, foi feita uma revisão na literatura. A maior parte das pesquisas encontradas envolviam motores de automóveis, navios e outros meios que utilizam motores diesel, nenhuma delas apresentou, com resultados, uma predição de falhas em motores diesel de locomotivas baseada em análise de óleo lubrificante.

Foram encontradas pesquisas envolvendo predição de janela de manutenção para motores diesel de equipamentos industriais por meio de RNA, outros trabalhos que almejavam prever o consumo de diesel de motores de automóveis e navios por meio de RNA e k-NN e um único estudo da revisão da literatura buscou predizer falhas em motores diesel de locomotivas por meio do óleo lubrificante, porém não obteve resultados, sendo algumas das justificativas a falta de quantidade e qualidade dos dados.

Além da contribuição para o meio científico, no que envolve, por exemplo, a identificação de particularidades no processo de análise de ocorrência de falhas no motor diesel e sua relação com o óleo lubrificante, esta dissertação também se justifica por apresentar resultados que colaboram com o planejamento das manutenções das locomotivas e das equipes que realizam as reposições dos componentes do motor; com a diminuição das paradas não programadas e interrupções nas vias férreas e com a simplificação e redução dos custos com as análises de óleo lubrificante.

1.3 ESTRUTURA DA PROPOSTA DE DISSERTAÇÃO

O trabalho será apresentado em 8 capítulos, sendo o primeiro deles a introdução contendo um explicativo breve do problema que será tratado ao longo da dissertação, e os dois modelos que serão desenvolvidos.

O Capítulo 2 será composto por um referencial teórico que aborda os principais conceitos de RNA e Árvore XGBoost e uma revisão bibliográfica sobre a utilização de modelos de predição em casos envolvendo motores diesel.

O Capítulo 3 apresentará o problema analisado, destacando o motor diesel e o óleo lubrificante e a descrição das locomotivas e da ferrovia estudada.

O Capítulo 4 apresentará o levantamento e o tratamento de dados históricos utilizados nesta dissertação, como as falhas que já ocorreram nos motores diesel de locomotivas e as análises do óleo lubrificante.

O Capítulo 5 irá explorar as configurações investigadas das RNA de classificação para predição de falhas nos motores diesel das locomotivas com base na análise do óleo lubrificante.

O Capítulo 6 apresentará as configurações investigadas da Árvore XGBoost de classificação para predição de falhas nos motores diesel das locomotivas com base na análise do óleo lubrificante.

O Capítulo 7 irá apresentar os resultados experimentais e análises das configurações investigadas para as RNA e para a Árvore XGBoost, e a comparação entre os dois modelos de predição.

Por fim, no Capítulo 8 tem-se as conclusões desta dissertação e sugestões de trabalhos futuros a serem realizados.

2 REFERÊNCIAL TEÓRICO

Neste capítulo serão apresentados conceitos referentes à duas técnicas de aprendizado de máquina: Rede Neural Artificial (RNA) e Árvore XGBoost. Por fim será a apresentada a revisão bibliográfica encontrada envolvendo modelos de predição de falhas em motores diesel.

2.1 TÉCNICAS DE APRENDIZADO DE MÁQUINA

2.1.1 Rede neural artificial

Para realizar a predição de falhas em um motor diesel baseada na análise de óleo lubrificante, foi necessário criar uma RNA que mescle os dados de entrada e aprenda com essas informações, relacionando os valores dos parâmetros de análise de óleo, consumo de diesel e movimentação de motores entre locomotivas com a ocorrência da falha, predizendo assim se o motor diesel irá falhar ou não. O termo RNA está intimamente ligado ao termo Inteligência Artificial (IA) e está presente em inúmeros estudos científicos e projetos que envolvem predição.

A utilização da IA para resolução de problemas reais começou a ter uma maior expressividade na década de 1970. Esses problemas eram então tratados em computadores com o uso do conhecimento de especialistas em um determinado assunto, os quais criavam as regras de decisão utilizadas nesses computadores, resultando em uma solução caracterizada pela subjetividade e desconfiança de sua eficácia (FACELI et al., 2011). Este tipo de sistema computacional, implementado pela observação e intuição de especialistas, é ainda muito utilizado em diferentes empresas e atividades, como exemplo tem-se a predição atual das falhas nos motores diesel da EFVM.

Com o aumento da complexidade dos problemas a serem resolvidos e o volume de dados gerados em diversas atividades, tornou-se necessário o uso de ferramentas computacionais mais independentes dos especialistas, capazes de criar uma função ou hipótese, que consiga resolver determinado problema, apenas com as experiências passadas a elas. Para esse processo, no qual por meio de experiências passadas,

uma hipótese é induzida, ou uma função é aproximada, dá-se o nome de Aprendizado de Máquina (FACELI et al., 2011).

O aprendizado de máquina apresenta uma hierarquia de aprendizado para modelos preditivos, como o tratado nessa dissertação, definida por meio dos dados recebidos, os quais representam as experiências passadas à máquina. O aprendizado nesse caso é denominado supervisionado, ou seja, o algoritmo aprende uma função que mapeia uma entrada para uma saída com base em exemplos que relacionam essas entradas e saídas. No aprendizado supervisionado os algoritmos podem ser de classificação e regressão.

Nos aprendizados de classificação os exemplos são de reconhecimentos de padrões, ou seja, o algoritmo prediz qual conjunto de categorias uma dada observação pertence com base em dados de treinamento contendo essas observações e suas categorias. Nos aprendizados de regressão o modelo prediz com base em um conjunto de processos estatísticos que procuram estimar a relação entre uma variável dependente e variáveis independentes, também chamadas de variáveis preditoras. Nesta dissertação o aprendizado será de classificação, ou seja, o algoritmo classificará as locomotivas em categorias nas quais houve falhas nos motores (sim), ou não houve falhas nos motores (não).

Um dos principais motivos da escolha da RNA, como algoritmo para predição de falha em motor diesel foi a sua capacidade de alcançar bons resultados de predição em casos de aprendizagem supervisionada, em especial de classificação.

Como exemplo de utilização da RNA, com obtenção de resultados aplicáveis, tem-se: predição de desempenho e emissão de particulados de um motor diesel marítimo (NIU et al., 2017); comparação de técnicas de modelagem para predição de emissão de monóxido de carbono de motores diesel, com resultado superior da RNA em comparação com a regressão linear e não-linear (TOSUN et al., 2017); predição de desempenho e otimização de um *Organic Rankine Cycle* (ORC) para recuperação do calor residual de motores diesel considerando o erro quadrático médio e o coeficiente de correlação entre as variáveis (YANG et al., 2018).

Quanto a definição da RNA, entende-se que são modelos computacionais inspirados no sistema nervoso de seres vivos. A compreensão dessa analogia se dá pela forma

que a RNA aprende com os dados a ela apresentados, utilizando sinapses artificias (interconexões) entre unidades de processamento denominadas neurônios artificiais. Essas sinapses são representadas por vetores ou matrizes de pesos sinápticos. Esses neurônios realizam funções como captar sinais em suas entradas (dados/experiências passadas à rede de neurônios), agregá-los de acordo com sua função operacional e produzir uma resposta (saída) (SILVA, SPATTI e FLAUZINO, 2010).

A primeira coisa a se fazer para desenvolver um algoritmo para predição é separar o conjunto de dados em três partes, denominadas: treinamento, validação e teste. A parte chamada treinamento é usada para treinar a rede, na qual uma função de erro que geralmente é o erro quadrático médio ou o erro relativo absoluto médio é minimizada com relação às constantes de peso em iterações sucessivas. Como é definida a parcela dos dados destinada ao treinamento é um dos pontos mais importantes para obter a confiabilidade de um modelo de RNA e determinar o número de camadas e neurônios nessas camadas (ALIBAKSHI, 2018).

Um conjunto de dados de validação é utilizado para ajustar os parâmetros do classificador. O conjunto de validação, no caso de uma RNA por exemplo, seria usado para encontrar o número “ótimo” de unidades ocultas. Já o conjunto usado na etapa de teste é utilizado apenas para avaliar o desempenho de um classificador totalmente treinado. No caso de uma RNA serviria para estimar a taxa de erro após escolher o modelo final (ALIBAKSHI, 2018).

Outra atividade primordial no desenvolvimento de uma RNA, além de separar os dados nas três partes (treinamento, validação e teste) é definir sua arquitetura, no que envolve selecionar o número de camadas, número de neurônios em cada camada, funções de transferência, algoritmo de treinamento, valores iniciais dos pesos e viés (estado de inicialização) e dar continuidade com o processo de treinamento da RNA. As camadas que formam a RNA são denominadas de: camada de entrada, camada oculta e camada de saída (SILVA, SPATTI e FLAUZINO, 2010).

A arquitetura da RNA depende da disposição dos neurônios nas camadas, as formas de interligação entre eles e como as camadas são utilizadas. Um dos tipos de arquitetura é a *feedforward* de camadas múltiplas, representada pela utilização de

uma ou mais camadas ocultas e muito empregada em casos de classificação de padrões (SILVA, SPATTI e FLAUZINO, 2010).

Quanto a quantidade de entradas, não é necessário restringi-las, elas podem ser em qualquer quantidade, bem como as saídas por elas geradas (NOOR et al., 2015). Quanto ao número de neurônios nas camadas ocultas, é usual que seja igual a 2/3 do número de neurônios da camada de entrada, entre o número de neurônios nas camadas de entrada e saída ou menor que o dobro do número de neurônios na camada de entrada. Porém não se pode negligenciar fatores individuais e as configurações da RNA (ALIBAKSHI, 2018).

As funções de transferência são consideradas parâmetros da RNA que podem impactar de forma significativa o desempenho da mesma. As funções de transferência comumente utilizadas em modelos de RNA são: *hyperbolic tangent sigmoid*, *log-sigmoid*, *Hard-limit*, *Positive linear*, e *Radial basis* utilizadas nas camadas ocultas, e linear transfer function para a camada de saída (ALIBAKSHI, 2018).

O *algoritmo Backpropagation*, conhecido também como Regra Delta Generalizada, em sua fase de treinamento, é o algoritmo utilizado entre as arquiteturas *feedforward,* nas quais destacam-se os casos de predição supervisionada com redes *Perceptron* de múltiplas camadas (do inglês *Multilayer Perceptron* – MLP). A grande utilização deste tipo de rede pode ser explicada pela diminuição dos erros produzidos pelas respostas comparado a outros tipos de redes, devido principalmente aos vários ajustes realizados pelos pesos sinápticos na busca por resultados mais precisos (SILVA, SPATTI e FLAUZINO, 2010). O pseudocódigo de treinamento *Backpropagation* de uma RNA pode ser observado na Figura 1.

Figura 1 – Pseudocódigo da RNA em Backpropagation

```
1   Leia e normalize a base de dados
2   Separe os dados de treino e teste
3   Determine o número de camadas de entrada, ocultas e de saída
4   Determine os pesos para os neurônios das camadas de entrada, ocultas e de saída
5   Calcule a saída (y) para cada neurônio da camada de saída
6   Atualize os parâmetros da MLP
7   Repita os passos 5-6 até completar o algoritmo
8   Salve os resultados
9   Fim do modelo híbrido
10  Apresentando resultados
11  Fim
```

Fonte: Sudha, Ramyachitra e Manikandan (2018, adaptado)

Para que a RNA possa predizer algo é necessário treiná-la com os dados pertinentes, para que então ela consiga compreender a relação entre esses dados e predizer a falha no motor diesel de locomotivas. Baseado na descida de gradiente, o treinamento da RNA é interrompido quando se atinge o critério de parada, geralmente estabelecido pelo erro mínimo de validação do treinamento (DUDA, HART e STORK, 2012).

A divisão dos conjuntos de registros tem por finalidade treinar o modelo computacional com uma maior quantidade de dados (etapa de treinamento), validar os resultados deste treinamento (etapa de validação), para que só então dados antes nunca apresentados ao modelo, possam servir de teste para avaliar a capacidade do modelo de prever dados reais (etapa de teste).

Quanto a separação dos dados utilizados na RNA para análise de seus resultados, tem-se, a parte de dados para treinamento, validação e teste da RNA. Na maioria das vezes são utilizados 70% dos dados para treinamento da RNA, e os 30% restantes são divididos meio a meio, 15% para a fase de validação e 15% para a fase de teste. Separando os dados nesse formato observou-se uma melhora nos resultados (ELKHOURY et al., 2018). Porém, existem também desvantagens na utilização dessa divisão, no sentido de que dados que foram colocados na fase de validação e teste, poderiam ser relevantes na fase de treinamento para uma melhor generalização da RNA. Desta forma, um outro formato de divisão dos dados passou a ser utilizado pelos desenvolvedores, denominado validação cruzada.

A validação cruzada é uma técnica popular para seleção de algoritmos. A estratégia consiste na divisão de dados, uma ou várias vezes, para estimar o risco de cada algoritmo. Uma parte dos dados é chamada de amostra de treinamento, e é usada para treinar cada algoritmo. A parte restante, chamada de amostra de validação, é usada para estimar o risco do algoritmo. O objetivo final é selecionar o algoritmo com o menor risco estimado. A popularidade da validação cruzada pode ser explicada pela "universalidade" das heurísticas de divisão de dados e comparada ao erro de re-substituição, a sua utilização evita o ajuste excessivo *(overfitting)* porque a amostra de treinamento é independente da amostra de validação (ARLOT, 2010).

Um estimador do erro de um classificador é uma variável aleatória e sua qualidade é geralmente medida por meio de seu viés e variância. Existem vários estimadores do erro de classificação e uma dessas técnicas, e provavelmente a mais popular, é a validação cruzada k-fold. A estimativa do erro é o valor médio dos erros cometidos em cada k-fold. Assim, o estimador de erro depende de dois fatores: o conjunto de treinamento e a partição em k-folds (RODRIGUEZ, PEREZ, e LOZANO, 2009).

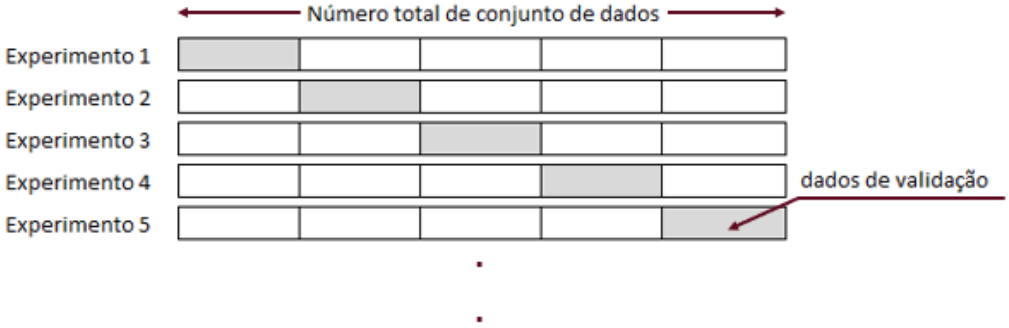
Para casos de algoritmos como a RNA o *overfitting* acontece quando os padrões de treinamento passam a ser memorizados pela RNA, o que pode ser explicado pela seleção de preditores que não executam nenhuma função útil ou até mesmo pela escolha de um algoritmo complexo demais para uma predição simples demais. É observado que modelos que incluem preditores desnecessários levam a decisões piores, o que pode ser explicado pelo fato de que os coeficientes ajustados aos preditores adicionam variação aleatória às previsões subsequentes (HAWKINS, 2004).

Quanto à utilização da validação cruzada como critério de parada para a RNA, tem-se que quando o erro da parte de validação aumenta ocorre a interrupção da fase de treinamento, sendo que o erro mais utilizado é o erro quadrático. Porém não existe garantia de que o treinamento será interrompido, por isso o ideal é estabelecer um valor limite para o erro, vetor gradiente ou um limite de épocas (DUDA, HART e STORK, 2012). No caso da RNA, uma época será toda a apresentação dos padrões do treinamento.

A validação cruzada *k-fold ou k-fold cross validation* é a forma básica de validação cruzada, sendo que outras validações cruzadas são casos especiais ou repetidas

rodadas da validação cruzada *k-fold*. Nesse tipo de validação cruzada um conjunto de dados (*k-fold*) é particionado em *k* segmentos de tamanhos iguais (ou quase iguais), de forma aleatória, em uma parte de treinamento e outra parte de validação (*f*). Na sequência iterações de treinamento e validação são executadas de modo que, em cada iteração, um *fold* diferente dos dados seja mantida para validação, enquanto os *k* – 1 *folds* restantes são usadas para aprendizado (REFAEILZADEH, TANG E LIU 2009). A Figura 2 apresenta um exemplo da divisão por validação cruzada k-*fold*.

Figura 2 – Divisão dos dados por validação cruzada k-*fold*



Fonte: O autor (2020)

Para que a precisão média resultante da validação cruzada *k-fold* em um determinando conjunto de dados não seja constante, existe na validação cruzada um mecanismo aleatório para *k-folds*. As duas medidas principais usadas para investigar o impacto desse mecanismo aleatório é o viés e a variância, sendo que o viés é a diferença esperada entre a estimativa de imprecisão e a precisão real, e a variação representa a variabilidade de uma estimativa de precisão. Quatro fatores podem afetar uma estimativa de precisão obtida pela validação cruzada k-fold, sendo elas: o número de *folds*, o nível da média, o número de instâncias em um *fold* e a repetição da validação cruzada k-fold (WONG, 2015).

A escolha do valor *f* é realizada levando em consideração o t*radeoff* entre a quantidade de conjuntos de dados particionados, com o tempo computacional do algoritmo de classificação. Para *f* igual a 2, por exemplo, tem-se partes muito pequenas para treinamento e teste, o que interfere na performance da RNA. Geralmente são utilizados *f* iguais a 5 ou 10, e se a utilização de *f* igual a 10 não tornar o tempo computacional muito superior em comparação com *f* igual a 5, é recomendado que *f* igual a 10 seja utilizado (YADAV e SHUKLA, 2016).

Quando se utiliza *k-fold*, o treinamento das RNA ocorre da seguinte forma: divide-se um conjunto de dados em *f* subconjuntos independentes de mesmo tamanho, os quais todos, exceto um desses subconjuntos, são usados para treinar um classificador, enquanto o restante é usado para avaliar o erro de generalização. Após o treinamento, é possível calcular, para cada um dos classificadores treinados, um limite superior do erro de generalização (ANGUITA et al., 2009). É importante observar que o número indicado para *f* interfere não apenas nos tempos de treinamento, mas também na troca entre a estabilidade da média do erro e o tamanho do termo de confiança. Os valores popularmente utilizados para *f* são 5 ou 10 devido aos melhores resultados encontrados utilizando esses valores em comparação com outros valores para *f* (HSU et al., 2003).

No caso onde o valor escolhido para *f* é 10, os dados utilizados para treinamento e validação são particionados de forma aleatória em 10 conjuntos de dados iguais. Destes 10 conjuntos, 9 são utilizados para treinamento e um único para validação. O procedimento é realizado alternando os conjuntos de dados de treinamento com o único conjunto de validação, até que se atinja 10 rodadas (YADAV e SHUKLA, 2016).

Para avaliar os resultados do algoritmo de predição, pode ser elaborada uma matriz de confusão. A matriz de confusão é uma matriz bidimensional que sintetiza a performance da classificação de um classificador com relação a alguns dados de teste. Essa matriz é configurada em uma dimensão pela classe verdadeira, de real ocorrência (casos *observados*), e na outra pela classe que o classificador atribui (casos *preditos*) (TING, 2017).

Como exemplo tem-se que o método de predição tem por objetivo predizer uma falha. Os casos em que o método de predição acerta são denominados verdadeiros. *Verdadeiro Positivo* (VP) é quando na realidade a falha ocorreu, e o modelo prediz essa ocorrência, já o *Verdadeiro Negativo* (VN) é quando na realidade a falha não ocorreu e o método, da mesma forma, não indica a ocorrência. Os casos falsos correspondem aos erros do método. *Falso negativo* (FN) é quando a falha ocorreu, mas o método não conseguiu predizer e *Falso positivo* (FP) é quando a falha não ocorreu, mas o método erroneamente indicou que ocorreu. A Figura 3 apresenta a distribuição da matriz de confusão em classes positivas e classes negativas.

Figura 3 – Matriz de confusão

		PREDITO	
		Não haverá falha	Haverá falha
OBSERVADO	Não houve falha	Verdadeiro Negativo (VN)	Falso Positivo (FP)
	Houve falha	Falso Negativo (FN)	Verdadeiro Positivo (VP)

Fonte: O autor (2020)

Com base na matriz de confusão é possível determinar a acurácia, uma métrica de avaliação de resultados muito utilizada por sua simplicidade, porém que deve ser acrescida de outras métricas para uma análise detalhada dos resultados. A acurácia é determinada pela porcentagem de acertos da RNA em relação ao total de casos que ela classificou. É possível obter um valor de acurácia para o treinamento, para validação e para o teste da RNA. A equação (1) representa a equação da acurácia (A_{cc}).

$$Acc = \frac{v_p + v_n}{total\ da\ amostra} \tag{1}$$

Além da acurácia, existem outras métricas que contribuem na análise dos resultados de predição em casos de classificação, como o F1-*score*, uma média harmônica de *precision* e *recall* calculada por meio do número de falsas indicações encontradas na predição realizada pela máquina e pelo avaliador humano, sendo 1 o melhor resultado possível de ser encontrado e 0 o pior. A métrica *precision* refere-se ao número de casos positivos corretos dividido pelo número de todos os casos positivos indicados pelo classificador (máquina), e *recall* é o número de casos positivos corretos dividido pelo número de todos os dados que deveriam ter sido classificados como positivos. Uma melhor compreensão das métricas pode ser adquirida na observação de suas equações.

As equações (1) e (2) correspondem às métricas *precision*, *recall* respectivamente, sendo que os parâmetros v_p, f_p, v_n e f_n são equivalentes aos parâmetros VP, FP, VN

e FN na mesma sequência, retratados na matriz de confusão da Figura 3 (HUANG et al., 2015).

$$Precision = \frac{v_p}{v_p + f_p} \tag{1}$$

$$Recall = \frac{v_p}{v_p + f_n} \tag{2}$$

As métricas de avaliação de resultados da matriz de confusão assim como as métricas *precision* e *recall* podem ser eficazes, no entanto, deve-se sempre avaliar o *tradeoff* entre elas, ou seja, deve haver uma troca empírica entre os valores, uma das métricas apresentará melhores resultados comparada às demais e vice-versa, o que deve ser avaliado é qual o objetivo final e o quais valores serão aceitos em prol desse objetivo (HUANG et al., 2015).

No caso dessa dissertação, foi escolhida a métrica de avaliação *recall* para apresentação dos resultados da RNA, pois as equipes do negócio indicaram que para atender as necessidades do projeto seria preciso minimizar ao máximo a quantidade de casos de falsos negativos (FN), visto que como as locomotivas passam para abastecimento de forma recorrente na Oficina, não seria um grande transtorno inspecioná-las quando o modelo computacional indicasse uma possível falha no motor. Já os casos nos quais a locomotiva falha em trânsito, por razão da falha no motor diesel não ter sido indicada pelo modelo computacional, denominados de casos FN, o transtorno e perdas é maior. Como o *recall* calcula exatamente essa relação entre o FN e os casos de falhas reais, ele foi o escolhido para ser uma das duas métricas de avaliação da RNA e da Árvore XGBoost.

É importante ressaltar que cada k-*fold* da RNA terá sua acurácia e *recall,* e que o k-*fold* escolhido como o de melhor potencial, ou seja, cujos pesos de generalização serão utilizados com outros conjuntos de dados para obtenção de novos resultados

para a RNA, será o k-*fold* que apresentar uma acurácia média e um *recall* médio dentre as demais partes de treinamento e validação.

A literatura apresenta uma quantidade razoável de trabalhos envolvendo RNA para resolver problemas das mais diversas áreas de conhecimento, porém não é grande o número de artigos que envolvem a RNA com o motor diesel e o óleo lubrificante, e ainda mais restritiva é a pesquisa quando são adicionados termos relacionados à ferrovia, como por exemplo, as locomotivas.

A capacidade preditiva de uma RNA quando envolve parâmetros de um motor diesel, está relacionada ao treinamento de seus dados experimentais e posterior validação por meio de dados independentes. Alguns testes laboratoriais envolvendo motores diesel são dispendiosos e consomem tempo, porém a recompensa na utilização de uma RNA para identificação das relações e ganhos nos procedimentos torna a análise periódica dos componentes do motor especialmente válida (NOOR et al., 2015).

Após entendimento da utilização da RNA, pode-se considerar que ela é capaz de predizer falhas em motores diesel de locomotivas, a partir de um conjunto de dados com variáveis preditoras que apresentem correlação com a falha do motor, qualidade e quantidade suficientes para o desenvolvimento da RNA.

2.1.2 Árvore XGBoost

Modelos baseados em árvores de decisão receberam como uma das últimas melhorias incorporadas a modelos anteriores, o chamado método Árvore XGBoost ou somente XGBoost (*Extreme Gradient Boosting*) em 2016, o qual, rapidamente tornou-se um dos métodos preferidos pelos profissionais da área de aprendizado de máquina. Este método é baseado em máquinas *Gradient Boosting* (do inglês *Gradient Boosting Machines* - GBM). O GBM usa um algoritmo que otimiza uma função de perda diferenciável para criar um conjunto de modelos de previsão, porém o XGBoost tem se mostrado mais eficiente e escalável que os demais modelos de árvore de decisão baseados no GBM (DE PISÓN et al., 2017).

O XGBoost melhora o desempenho do Gradient Boosting, incorporando mais estratégias de regularização para reduzir a complexidade do modelo e o excesso de ajustes. Estratégias essas que variam nas fases de construção e crescimento da

árvore de decisão. Na construção tem-se os parâmetros: *método da árvore* e *rodada de número boost*. O método da árvore pode ser selecionado de acordo com o tamanho da base de dados que será usada para treinamento do modelo, sendo *exact* a função usada para bases pequenas para médias, e *approx* para as grandes bases de dados. No SPSS Modeler 18.1 existe a opção *auto*, que escolhe para o usuário a melhor opção entre *exact* e *approx*. Para a *rodada de número boost* existe um padrão fixado, geralmente o valor é 10 (IBM, 2018).

Entre os parâmetros da fase de crescimento da árvore XGBoost, existe a delimitação do peso mínimo e a delimitação da profundidade máxima (DE PISÓN et al., 2017). O peso mínimo, também chamado peso filho mínimo, entre outros, é o peso mínimo ou número de instâncias estabelecidas para iniciar um processo dentro da árvore de decisão. Quanto maior o peso, mais conservador será o algoritmo. O valor comumente utilizado para o parâmetro é 1 (IBM, 2018). A profundidade máxima é o número máximo de divisões que a árvore irá estabelecer, sendo 6 o valor padrão e 3 e 4 outros valores também usados afim de reduzir a complexidade do modelo (TORLAY et al, 2017).

Além das estratégias nas fases de construção e crescimento, também foram implementadas melhorias para evitar ajustes excessivos (*overfitting*), como a incorporação dos termos de regularização L1 (Lasso) também chamado de *alfa*, e L2 (Ridge) também chamado de *lambda*. Quanto maiores forem os valores de *alfa* (valor padrão = 0) e *lambda* (valor padrão = 1), mais conservador será o modelo. Outro parâmetro também utilizado na Árvore XGBoost é o *eta*. O *eta* controla os ajustes excessivos por meio de taxas de aprendizagem (DE PISÓN et al., 2017). A taxa de aprendizado é normalmente definida entre 0,01 e 0,3, sendo 0,3 o valor padrão de utilização, pois valores muito menores podem levar a uma melhor generalização, porém podem também acabar não cumprindo a sua função de evitar super ajustes (TORLAY et al, 2017).

Outro parâmetro incorporado a Árvore XGBoost afim de evitar super ajustes foi a sub-amostragem. Com o intuito de selecionar uma parcela (amostra) do conjunto de dados de treinamento, o valor do parâmetro sub-amostragem varia entre 0 e 1, no qual o valor 1 é o padrão e valores como 0,7 e 0,8 também são utilizados (TORLAY et al, 2017).

Vale ressaltar que durante o processo de ajustes, o alto número de parâmetros do modelo pode sobrecarregar os esforços computacionais. Além disso, ainda que os métodos baseados em árvores de decisão tenham bom desempenho, a inclusão de recursos irrelevantes ou ruidosos pode degradar a precisão desses modelos (DE PISÓN et al., 2017).

2.2 REVISÃO BIBLIOGRÁFICA ENVOLVENDO PREDIÇÃO DE FALHAS EM MOTORES DIESEL

Para realização do levantamento bibliográfico foram cumpridas etapas, como a escolha das palavras-chaves, definição das bases de pesquisa, avaliação dos resumos dos estudos e escolha e análise dos estudos relevantes.

Como palavras chaves para a realização da revisão bibliográfica foram utilizadas combinações em português e inglês das seguintes palavras: predição de falhas, rede neural artificial, motor diesel, óleo lubrificante e locomotivas. As bases de pesquisa utilizadas foram: Periódicos CAPES, *Science Direct*, *Scorpus* e *Google Scholar*.

Por meio da pesquisa bibliográfica, buscando referências que envolvessem RNA e motores diesel em um mesmo trabalho, foram selecionados os nove estudos encontrados para apresentação em resumo. Diante da dificuldade de encontrar estudos científicos envolvendo exatamente o tema proposto nesta dissertação, foram selecionados estudos de 1989 a 2018.

Jardine et al. (1989) analisaram se o método *Proportional Hazards Modeling* (PHM) poderia contribuir ou não com a tomada de decisão dos analistas ou especialistas que determinam, por meio da análise de óleo lubrificante utilizado nos motores diesel de locomotivas, quando esse motor irá falhar. Na ferrovia onde estas locomotivas transitavam no Canadá, a política de manutenção determinava que após 12 anos de uso todo o motor diesel deveria ser substituído por outro novo, que as manutenções de rotina, como limpeza, troca de filtro/óleo, e outras, deveriam ser feitas a cada 90 dias, e a coleta do óleo lubrificante destinado à 5 laboratórios de análise, deveria ser realizada a cada 7 dias. Nos laboratórios eram feitos dois tipos de testes nos óleos lubrificantes, um teste físico para examinar contaminações por água por exemplo, e um teste espectrométrico, que identifica substâncias químicas presentes no óleo,

como chumbo e ferro. O resultado da análise não era observado individualmente, ou seja, era realizada uma comparação do resultado anterior com o atual para criar tendências, altos níveis de determinado elemento poderiam sugerir desgaste anormal dos componentes do motor. Os dados de análise do óleo eram inseridos em um sistema computacional da ferrovia que por meio das tendências sinalizava aos analistas que o motor poderia estar com alguma falha ou não. Os alertas possuíam uma assertividade de 70%. Desta forma, os autores propuseram melhorar os acertos (índices acima de 70%) utilizando o método PHM.

A função criada pelos autores, denominada função de risco (risco de falha) tinha como objetivo identificar o número de dias desde o início do período de observação até a falha e foi representada por um gráfico de *Weibull*. Os autores definiram que as falhas simples, que não exigiam reboque da locomotiva até a oficina não seriam incluídas no estudo, justamente por não serem tão custosas a ferrovia, e as falhas graves, envolvendo por exemplo virabrequim e rolamentos, seriam utilizadas na função. Desta forma, as variáveis explicativas utilizadas na função foram: resultados dos testes espectrométricos do óleo lubrificante (não havia quantidade suficiente de resultados dos testes físicos), registros de histórico de manutenção do motor diesel e idade do motor. Para a coleta destes dados foram utilizadas informações de 118 locomotivas SD40 da fabricante General Motors. Na fase de tratamento dos dados, houve uma eliminação daqueles dados de análise de óleo que não variavam com o tempo, principalmente os que eram próximos de zero. Com isso, os únicos dados de análise de óleo que foram utilizados na função foram: cobre, alumínio, ferro, sódio, zinco e níquel. Quanto ao tratamento dos dados do histórico de falhas, aquelas falhas que fossem similares ou que ocorressem concomitantemente em peças diferentes do motor, seriam colocadas em um mesmo grupo, e o grupo passaria a ser usado na função, no lugar da falha de forma individual (JARDINE et al., 1989).

Além dos dados já apresentados pelos autores, outras combinações de falhas e análises de óleo foram utilizadas na função, resultando em 55 observações (23 falhas e 32 não falhas) em 15 covariáveis (criadas pelas relações entre os elementos químicos presentes no óleo com eles mesmos). Além do PHM também foi utilizado o MINITAB para relacionar os elementos presentes no óleo e observar suas variações quando ocorre, ou não, falha no motor. Os testes mostraram que os elementos Cobre, Prata e Níquel apresentavam níveis menores na falha do que na “não falha” e que

para o zinco e o cálcio, a resposta era o inverso. Para os outros elementos não houve diferença significativa. Como conclusão do estudo, não foi identificada nos resultados da função de risco uma relação entre os dados de análise do óleo e as falhas, semelhantes às que ocorriam na realidade, diante disso foi levantada a possível necessidade de aumentar a amostra de dados, pois o problema poderia estar na quantidade pequena deles, visto que a quantidade de não falhas foi superior à de falhas (105 locomotivas das 118 não tiveram falhas entre uma revisão e outra) (JARDINE et al., 1989).

O artigo de Cunha et al. (1994) propôs a criação de uma metodologia para elaboração de um plano de manutenção que flexibilize as manutenções programas de motores diesel de locomotivas, afim de utilizar mais os equipamentos em momentos de grandes demandas e/ou retê-los na Oficina de forma preventiva antes que falhas catastróficas ou falhas menores, mas que levam até uma falha catastrófica, aconteçam. Para isto foi desenvolvida uma metodologia para criação de um algoritmo baseado em Lógica Difusa (Lógica Fuzzy) por meio de um Sistema Especialista. Os dados utilizados foram obtidos na EFVM, em um momento de grande aumento da demanda do transporte, ou seja, a retenção de locomotivas nas oficinas deveria ser mínima, em contrapartida elas estavam se esforçando mais devido a necessidade de atendimento a demanda. Diante desta situação, viu-se necessário haver critérios técnicos que decidissem pelo adiantamento, ou não, de uma revisão nas locomotivas, retendo-as de uma forma que a taxa de falhas não aumentasse. Os autores analisaram 12 locomotivas que apresentavam mais falhas em seus motores diesel, e as relações das falhas com as análises de óleo lubrificante, identificando assim, que seria possível prever uma falha por meio destas análises de óleo, sem que fosse preciso parar as locomotivas (as coletas de óleo eram realizadas por meio de engate rápido), ato que seria necessário se fossem analisar as falhas através da vibração do motor diesel, por exemplo. Os autores, então, criaram um quadro (Quadro 1) com a relação entre os valores correspondentes das análises de óleo lubrificante e os componentes do motor diesel que poderiam estar falhando ou na eminencia da falha.

Quadro 1 – Sintomas, defeitos e sistemas do motor diesel relacionados às análises de óleo lubrificante (continua)

SINTOMAS	DEFEITOS	SISTEMA
Viscosidade do óleo lubrificante alta	Quebra do gorne do pistão	Conjunto de força
	Quebra do pistão	Conjunto de força
	Quebra do anel	Conjunto de força
	Vazamento pelas válvulas dos cabeçotes	Conjunto de força
	Ovalização de camisas	Conjunto de força
Viscosidade do óleo lubrificante baixa	Diluição de combustível no lubrificante	Sistema combustível
Ponto de fulgor do óleo lubrificante baixo	Diluição de combustível no lubrificante	Sistema combustível
Taxa de variação do TBN negativa e alta	Diluição de combustível no lubrificante	Sistema combustível
	Vazamento pelas válvulas dos cabeçotes	Conjunto de força
Taxa de variação do TBN positiva Concentração de insolúveis no óleo lubrificante alta	Vazamento de óleo	Sistema de lubrificação
	Vazamento pelas válvulas dos cabeçotes	Conjunto de força
	Ovalização de camisas	Conjunto de força
Taxa de variação da concentração de insolúveis no óleo lubrificante alta e positiva	Quebra de anel	Conjunto de força
	Vazamento pelas válvulas dos cabeçotes	Conjunto de força
	Deficiência de filtragem do lubrificante	Sistema de lubrificação
Alta concentração de Si no óleo lubrificante (taxa de variação da concentração positiva e alta)	Deficiência de filtragem do ar de admissão	Sistema central de ar
Alta concentração de Fe no óleo lubrificante (taxa de variação da concentração positiva e alta)	Deficiência de filtragem do ar de admissão	Sistema central de ar
	Quebra de anel	Conjunto de força
	Desgaste acelerado de anéis	Conjunto de força
Alta concentração de Pb no óleo lubrificante (taxa de variação da concentração positiva e alta)	Desgaste acelerado de casquilhos das bielas	Conjunto de força
	Desgaste acelerado de casquilhos do eixo comando ou virabrequim	Motor diesel
Alta concentração de Cu no óleo lubrificante (taxa de variação da concentração positiva e alta)	Desgaste em arruela de carregador	Conjunto de força
Alta concentração de Cr no óleo lubrificante (taxa de variação da concentração positiva e alta)	Desgaste acelerado de camisas	Conjunto de força
	Desgaste acelerado de eixos cromados	Motor diesel

Fonte: Cunha et al., (1994)
Nota: Dados adaptados pelo autor

Quadro 1 – Sintomas, defeitos e sistema do motor diesel relacionados às análises de óleo lubrificante (conclusão)

Alta concentração de Sn no óleo lubrificante (taxa de variação da concentração positiva e alta)	Desgaste acelerado de casquilhos das bielas	Conjunto de força
	Desgaste acelerado de casquilhos do eixo comando ou virabrequim	Motor diesel
	Desgaste acelerado de saia de pistão (os que são recobertos com Sn)	Conjunto de força
Alto consumo de óleo lubrificante (taxa de variação do consumo positiva e alta)	Vazamento de lubrificantes	Sistema de lubrificação
	Desgaste acelerado de anéis	Conjunto de força
	Camisas desgastadas	Conjunto de força
	Camisas ovalizadas	Conjunto de força
	Anel quebrado	Conjunto de força

Fonte: Cunha et al., (1994)
Nota: Dados adaptados pelo autor

Após criação do Quadro 1 com as relações entre óleo e falha, os autores definiram na metodologia quais seriam os itens essências para compor o algoritmo de avaliação de desempenho que forneceria como saída um único valor para cada locomotiva, de forma que fosse possível identificar quais delas, se adiadas suas manutenções programadas, representariam maior risco. Para construção do algoritmo os autores indicaram a utilização não somente das falhas, mas também de seus sintomas, sendo que a saída deve ter por objetivo definir a qualidade da locomotiva, numa classificação entre 0 e 10, sendo essa qualidade estabelecida pela Lógica Difusa por meio de informações adquiridas por estudos matemáticos e com os dados das análises de óleo, como por exemplo a viscosidade do óleo, e os históricos de falhas que já ocorreram nos motores diesel. Para avaliação da eficácia dos resultados do algoritmo, os autores criaram dois índices. O primeiro deles, denominado IC, busca comparar as locomotivas em dia com a manutenção programada com locomotivas em atraso com a manutenção programada, em uma observação antes e após a utilização dos resultados do algoritmo e o segundo índice, denominado Média de Serviços (MS), relacionava não conformidades executadas durante a manutenção programada, em um determinando sistema, em locomotivas para as quais aquele sistema é considerado crítico, com a mesma MS dos casos onde o sistema não é considerado crítico. Como resultados das comparações entre estes dois índices por meio dos históricos da EFVM e demonstradas no artigo pelos autores, foi possível concluir que há espaço para aplicação desta metodologia em próximo estudo (CUNHA et al., 1994).

O estudo de Mancuzo (1999) implementou uma ferramenta de detecção automática de falhas em motores de combustão interna de automóveis a partir da análise de vibrações nos mesmos. Com o objetivo de identificar as bandas de frequência de interesse na caracterização das falhas dos motores, fez-se necessário o conhecimento da função de resposta em frequência (FRF) entre cada elemento gerador de falha e o ponto de medição do sinal de vibração. O método utilizado para a análise foi a *probabilistic neural network* (PNN). A arquitetura da PNN foi treinada com 2 entradas, 3 classes e 2 casos de treinamento para cada classe. A camada de entrada é fictícia, pois seus neurônios não realizam nenhuma operação. A camada de classificação é constituída de um neurônio para cada caso do conjunto de treinamento. O processo de classificação inicia com a apresentação da amostra desconhecida para todos os neurônios da camada de classificação. Cada neurônio de classificação calcula a distância entre a amostra de entrada e o caso do conjunto de treinamento ao qual ele está associado. Esta distância será o argumento da função de ativação do neurônio, que é essencialmente uma janela *Parzen*. Quanto as entradas, foram envolvidas as falhas de batida de válvula e defeitos nos tensores, enquanto as saídas da rede foram: (P1) - motor com tensor da correia sincronizadora rumoroso, (P2) - motor com tensor da correia do alternador rumoroso, (P3) - motor sem defeito e (P4) - motor com ambos os tensores rumorosos (MANCUZO, 1999).

O autor indica que na etapa de validação da rede, foi calculada a matriz de erro classificatório, que é um dos indicadores de como a rede aprendeu os padrões a ela apresentados. Tanto o estudo cinemático dos mecanismos geradores de defeito no motor, quanto os resultados da análise modal parcial executada, auxiliaram significativamente na compreensão do comportamento vibratório do motor. Comprovou-se a grande utilidade da técnica do envelope no diagnóstico de falhas, cujo potencial é aumentado com a utilização de um modelo de geração de vibrações em rolamentos. Conseguiu-se, para o restrito banco de dados adquirido, o sucesso no diagnóstico automático de rumorosidade nos tensores e de batida de válvula. Apesar do êxito no diagnóstico do problema de batida de válvula, não se conseguiu identificar qual das válvulas do motor era a causa do defeito, o que, se alcançado, melhoraria a produtividade da linha de montagem dos motores. Com base nos resultados, confirmou-se o bom desempenho da rede PNN, utilizada como classificadora (MANCUZO, 1999).

O artigo de Caneca et al. (2006) teve como objetivo monitorar a condição de óleos lubrificantes usados em motores diesel de automóveis por meio de duas estratégias. A primeira estratégia foi uma abordagem qualitativa, categorizando as amostras em três classes de acordo com seu estágio de uso (curto, médio e longo prazo) e classificando-as por análise discriminante. Para comparação, um classificador convencional K-NN também foi empregado. A segunda estratégia foi uma abordagem quantitativa que emprega espectroscopia de infravermelho e técnicas de calibração multivariada para prever a viscosidade, que é o principal parâmetro de controle para lubrificantes em serviço. Neste caso, técnicas de regressão múltipla linear (RML) foram empregadas na calibração. O efeito de diferentes procedimentos de pré-processamento, bem como a utilidade da seleção de variáveis, foi avaliado por meio de um estudo de planejamento fatorial.

O autor destaca que no primeiro momento foi realizada uma análise qualitativa, pois foi necessário estabelecer uma função de densidade de probabilidade, utilizando um método de classificação que pressupõe que os objetos seguem uma distribuição gaussiana dentro de cada classe. O algoritmo de seleção proposto neste trabalho foi o *stepwise* que leva em conta tanto o poder discriminante de cada variável quanto a colinearidade entre as variáveis. As variáveis altamente correlacionadas com aquelas já selecionadas foram descartadas para evitar problemas de colinearidade. O algoritmo cessou quando não havia mais variáveis disponíveis. O conjunto de variáveis que foi escolhido para seguir, foi aquele que resultou no menor número de erros de validação cruzada. Para a realização da análise foi utilizado um conjunto de 114 amostras de óleo lubrificante para motores diesel em diferentes estágios. As 112 amostras que não foram consideradas *outliers* (apresentaram valores anormais de viscosidade) foram divididas em conjuntos de calibração, validação e previsão com 64, 25 e 23 amostras, respectivamente. A melhor solução resultante das 20 realizações do algoritmo foi adotada e o método utilizado foi considerado eficaz (CANECA et al., 2006).

O estudo de Junior, Machado e Barreto (2006) tem como objetivo predizer quando os equipamentos de uma planta industrial vão parar, ou seja, vão falhar. Para isso os autores criaram uma RNA formada por meio de uma série temporal diária, com base em um levantamento de cinco anos de intervenções de paradas para manutenções em diversos sistemas produtivos da planta industrial. Optou-se por escolher apenas 1

dos 23 motivos de paradas para verificar se o comportamento diário do motivo poderia ser previsto por redes neurais. Os dados foram organizados em uma série de 1826 dias corridos com suas respectivas horas paradas. Após a normalização, os dados relativos a cada código foram divididos em três blocos: treinamento (01 a 608 dos dados), validação (609 a 1216 dos dados) e teste (1217 a 1824 dos dados). O *software* utilizado foi o *Matlab*® por meio da Neural Network Toolbox. Este programa utiliza para redes *feedforward*, o algoritmo *backpropagation* usando um gradiente decrescente nomeado por *trainlm*.

Para acelerar a fase de treinamento, os autores normalizaram os dados, proporcionando dados de entrada entre 0 e 1. A função de ativação escolhida foi a *log-sigmoid*, obtendo-se saídas também compreendidas entre 0 e 1 facilitando as medidas de acurácia. Foi pré-definido que para todas as redes testadas seriam fixados dois valores para o erro médio, 0,01 e 0,005 e um limite de 30.000 interações. Para a taxa de aprendizado, foram utilizados dois importantes parâmetros que auxiliam na aceleração do treinamento, o *default* do *Matlab*®, e o *traingd*. Para a arquitetura inicial da rede, foi utilizada apenas uma camada oculta, e os menores valores de erros encontrados determinaram as melhores redes. Para cada configuração foram repetidas no mínimo dez vezes cada arquitetura. Como resultado às análises verificou-se a necessidade de, para a mesma arquitetura, realizar várias simulações com o intuito de encontrar o melhor grupo de resultados. Três neurônios na camada oculta, rede para 30.000 interações e limite de Erro Médio Quadrático igual a 0,005 foi a topologia de melhor resultado. Conclui-se que os resultados confirmaram a aplicabilidade das redes neurais nas previsões de paradas para a manutenção na área industrial da Petroflex. Confirmou-se, ainda, a habilidade das RNA em efetuar predições em um conjunto de dados com forte não-linearidade (JUNIOR, MACHADO e BARRETO, 2006).

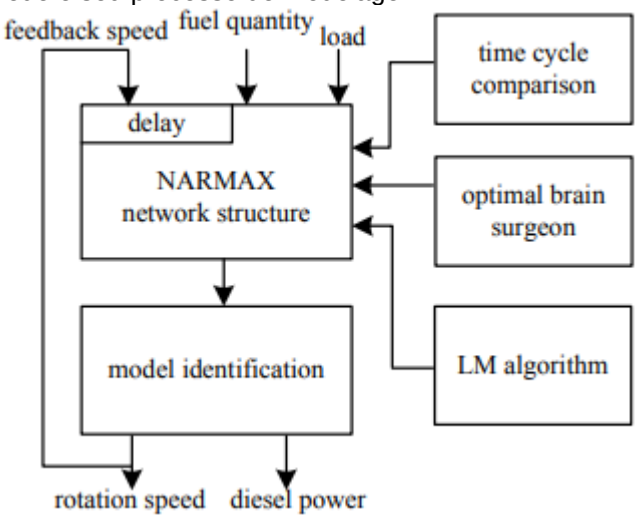
O objetivo do artigo de Parlak et al. (2006) é investigar a aplicabilidade de uma RNA para predizer parâmetros, como consumo específico de combustível e temperatura de exaustão em motores diesel de automóveis. O algoritmo escolhido para a análise foi o *Backpropagation,* compondo uma rede com uma camada de entrada, uma camada oculta e uma camada de saída. Para desenvolvimento da RNA, 70% dos 80 dados experimentais utilizados foram usados para treinamento da rede, e 30% para verificar a capacidade de generalização da rede. As entradas foram: velocidade do motor,

pressão efetiva média de ruptura e tempo de injeção. As saídas foram: consumo de combustível específico do freio e temperatura de exaustão. O Neural Networks Toolbox do MATLAB 5.2 foi usado para implementar a RNA. A função de transferência *log-sigmoid* foi usada na camada oculta e na camada de saída. Sete neurônios foram usados na camada oculta.

Os autores normalizaram todos os valores de entrada e saída na faixa de 0,1 a 0,9. Os erros relativos (ER) no período de teste entre os valores previstos e medidos foram encontrados entre 1 e 9%. O erro relativo médio durante o teste foi de 1,93% na previsão do consumo específico de combustível e de 2,36% na previsão da temperatura de exaustão. A capacidade preditiva da rede para o consumo específico de combustível e temperatura de exaustão foi considerada satisfatória (PARLAK el al., 2006).

O artigo de Biao et al. (2009) envolveu o estudo de uma RNA para identificação de um sistema de injeção de combustível em um motor diesel de locomotiva. O primeiro passo foi determinar a estrutura do modelo a ser utilizada, sendo escolhido como principal estrutura o NARMAX (*Non-linear AutoRegressive Moving Average with eXogenous inputs*). O segundo passo seria confirmar e otimizar a estrutura. Para confirmar foi determinado três entradas: quantidade de injeção de combustível de cada cilindro, carga do gerador principal e velocidade de rotação de feedback. E dois sinais de saída: velocidade de rotação e potência do diesel. A rede foi composta por duas camadas ocultas, cada uma com 50 neurônios. Todos os neurônios ocultos usaram a não-linearidade *Sigmiod*, e os neurônios de saída usaram a função linear pura. Os pesos iniciais foram aleatórios de 0 a 1, e para finalmente confirmar a classificação do modelo foi usada comparação de Ciclo de Tempo. Para otimizar a estrutura de rede foi utilizada a estratégia *Optimal Brain Surgeon*. A Figura 4 mostra a estrutura da rede e seu processo de modelagem.

Figura 4 - Estrutura da rede e seu processo de modelagem



Fonte: Biao, L. et al. (2009, p. 3418)

No terceiro passo, os autores escolheram o algoritmo *Levenberg-Marquardt* (LM) para treinar a rede, algoritmo que retrata a forma melhorada do método de Gauss-Newton. Após treinar o algoritmo, chegou-se na quarta etapa, que tem por objetivo identificar o sistema dinâmico. Um grande número de dados dinâmicos formou curvas de desempenho contínuo que foram obtidas na bancada de testes do motor diesel de locomotivas de 16 cilindros. A fim de melhorar a precisão e simplificar o cálculo, foi usada a função PREMNMX no MATLAB para padronizar os dados de entrada e saída. Com a comparação do ciclo de tempo, obtiveram-se as classificações da seguinte forma: a classificação da quantidade de injeção de combustível é 0; a classificação de carga é 1 e a velocidade de rotação de feedback é 2. O número total de nós de entrada é 6. Originalmente, existiam 100 nós ocultos e 2900 pesos nas redes. Quando o algoritmo terminou, o número de nós ocultos foi reduzido para 52 e os pesos foram reduzidos para 613. Conclui-se que o algoritmo foi bom para reduzir a redundância das redes. Usaram-se as amostras de teste para testar as redes antigas e as redes de otimização separadamente, e houve uma diferença muito pequena entre as duas redes. Como última etapa obtiveram-se os resultados, mostrando que o uso do algoritmo LM, para reduzir o cálculo complexo foi válido, bem como os resultados da simulação comprovaram a racionalidade e aplicabilidade do modelo (BIAO, et al., 2009).

O objetivo do estudo de Noor et al. (2015) foi o de predizer o desempenho de motores diesel marítimos utilizando RNA. Predizer o desempenho, significa predizer a potência

de frenagem, o torque de saída, o consumo de combustível específico do freio, a eficiência térmica do freio e a eficiência volumétrica. Para desenvolver e treinar a rede foi utilizando o *software* Neural Network Toolbox do MATLAB. A modelagem da RNA compreendeu duas fases: a primeira foi treinar o modelo de rede, enquanto a segunda fase foi valida-lo com novos dados que não foram usados para treinamento. Os dados de entrada para o treinamento da rede foram coletados a partir de testes laboratoriais de motores que rodam a várias rotações.

Os autores explicam que em muitas aplicações de RNA, a arquitetura *backpropagation neural network* (BPNN) com uma camada oculta é suficiente. A fim de encontrar uma arquitetura ideal, diferentes números de neurônios na camada oculta foram considerados e o erro de predição para cada rede foi calculado. O BPNN foi baseado na regra de aprendizado e correção de erros. A operação do modelo de RNA foi dividida em duas etapas principais: computação avançada e aprendizado retroativo. Na computação, os padrões de entrada aplicados aos neurônios da primeira camada foram apenas um estímulo para a rede (NOOR et al., 2015).

Como conclusão das análises, os autores observaram que o resultado da previsão da RNA, com 22 neurônios na camada oculta, foi encontrado em boa concordância com os dados experimentais. A distribuição de dados para a RNA foi quase similar e aos dados experimentais reais com coeficiente de correlação (R) na faixa de alcance de 0,9 a 1,0, indicando que a RNA desenvolvida é capaz de fazer a predição, do desempenho dos motores diesel marítimos, com razoável precisão (NOOR et al., 2015).

O objetivo do estudo de Valis, Zak e Pokora (2015) foi o de prever limiares de falha em motores diesel de automóveis baseado em dados reais de óleo, descrito em um modelo de degradação no tempo (processo estocástico) com os metais, presentes nos elementos do motor, Fe (ferro) e Pb (chumbo). Para isso foram usados os resultados da análise de regressão linear como uma entrada básica, ajudando determinar um curso linear de geração de partículas de Fe e Pb, curso este apresentado com intervalos de confiança de 95% para um valor de um veículo individual. Esses valores são aplicados posteriormente para a modelagem de processo *Wiener* com um desvio positivo. Os dados arquivados foram estipulados por meio de métodos como o *Atomic Emission Spectrometry* (AES) e pelas observações

morfológicas *Fourier Transformation Infra Red* (FTIR) e *Laser Net Finder* (LNF). A degradação do óleo no sistema foi expressa por uma função, em um processo contínuo no tempo, que adquire valores aleatórios de uma medida de degradação. O importante neste estudo foi definir o tempo derivado da função em um histograma e a densidade de probabilidade com a maior precisão possível e de maneira não distorcida. O motivo pelo qual foram selecionados os elementos Fe e Pb foi que a taxa de correlação matemática testada do tempo de operação foi a mais alta. Também foram implementados métodos de *Clustering*, como hierárquico e não hierárquico, para suportar o cálculo da taxa de correlação.

Os três autores indicaram algumas entradas para uso além das análises de óleo lubrificante, como os dados vibro-acústicos, dados de emissão térmica, de consumo de combustível; de custo do ciclo de vida e dados de tribo-diagnóstico. Acredita-se que tais dados tenham o potencial para prever ocorrência de falha branda ou suave, formando bons insumos para otimização de manutenção preventiva e aumentando as chances de se prever a vida útil residual do motor. Pode-se concluir que por meio deste estudo é possível determinar a previsão de tempo de uma ocorrência de falha, o intervalo ótimo do plano de manutenção, o tempo recomendado/permitido para a conclusão da manutenção e a otimização na análise de custos do ciclo de vida do motor (VALIS, ZAK e POKORA, 2015).

Após buscas nas bases de pesquisa já mencionadas, nenhum estudo (artigo, dissertação ou tese) foi encontrado envolvendo exatamente a predição de falhas baseada na análise do óleo lubrificante em motor diesel de locomotivas com a obtenção de resultados. Para uma melhor comparação entre os estudos encontrados, apresentados anteriormente, foi criada a Tabela 1, na qual será possível identificar as diferenças quanto às aplicações e os métodos utilizados entre os estudos e também compará-los a está dissertação.

Tabela 1 – Comparativo entre às aplicações e métodos utilizados nas referências bibliográficas encontradas e está dissertação

Ano	Autores	Aplicação	Objetivo	Método	Meio	O objetivo foi alcançado?
1989	Jardine et al.	Motor diesel de locomotivas	Predizer se haverá falhas nos motores diesel	PHM	Análise do óleo lubrificante	Não
1994	Cunha et al.	Motor diesel de locomotivas	Metodologia para criação de um algoritmo (Lógica Difusa) que avalie a qualidade das locomotivas	Relações motor e óleo e índices de desempenho	Análise do óleo lubrificante	Sim
1999	Mancuzo	Motor de combustão interna de automóveis	Predizer se haverá falhas nos motores diesel	PNN	Análise de vibração dos motores	Sim
2006	Caneca et al.	Motor diesel de automóveis	Avaliar as condições do óleo lubrificante	K-NN e RML	Análise do óleo lubrificante	Sim
2006	Junior, Machado e Barreto	Equipamentos industriais	Predizer se haverá falhas nos equipamentos	RNA	Intervenções de paradas para manutenção	Sim
2006	Parlak et al.	Motor diesel de automóveis	Determinar o consumo específico de combustível e a temperatura de exaustão dos motores	RNA	Velocidade do motor, pressão de ruptura e tempo de injeção do motor	Sim
2009	Biao et al.	Motor diesel de locomotivas	Determinar a velocidade de rotação e a potência do sistema de injeção de combustível	RNA	Injeção de combustível, carga do gerador principal e velocidade de rotação do motor	Sim
2015	Noor et al.	Motor diesel marítimo	Determinar a potência e eficiência de frenagem, torque de saída e o consumo de combustível	RNA	Testes de rotação dos motores	Sim
2015	Valis, Zak e Pokora	Motor diesel de automóveis	Predizer se haverá falhas nos motores diesel	Regressão Linear	Análise do óleo lubrificante, dados vibro-acústicos, de emissão térmica, de consumo de combustível, custo do ciclo de vida e tribo-diagnóstico	Sim
2019	Esta dissertação	Motor diesel de locomotivas	Predizer se haverá falhas nos motores diesel	RNA	Análise do óleo lubrificante	Sim

Legenda: Proportional Hazards Modelling (PHM); Probabilistic Neural Network (pnn); k-Nearest Neighbor (K-NN); Regressão Múltipla Linear (RML); Rede Neural Artificial (RNA).
Fonte: O autor (2019)

3 DESCRIÇÃO E CONCEITUAÇÃO DO PROBLEMA

3.1 O PROBLEMA ANALISADO

O problema estudado envolveu a necessidade de predizer falhas em motores diesel de locomotivas. Atualmente o planejamento de manutenção destes motores na EFVM é realizada por meio de um plano, denominado Estratégia de Manutenção, realizado de forma manual com base na experiência dos engenheiros e técnicos responsáveis.

Da fabricante GE existem atualmente na EFVM alguns modelos diferentes de locomotivas, e neste estudo utilizamos os dados de 216 locomotivas de dois modelos diferentes, denominados de forma habitual como DASH-8 e DASH-9. A escolha dos modelos foi feita em conjunto com as equipes de manutenção, pois são as locomotivas mais utilizadas e em maior número da ferrovia.

Dentre as formas de se analisar a existência de desgastes nos motores diesel das locomotivas e que alguma falha tende a ocorrer nos mesmos, são feitas análises no óleo lubrificante utilizados nos motores diesel. O óleo é coletado no Carter, compartimento localizado na parte de baixo do motor diesel. Nas locomotivas GE essa coleta acontece de 15 em 15 dias e é realizada nos momentos em que as locomotivas passam pelo local de abastecimento. Existe uma pequena torneira saindo do Carter por onde se retira a amostra de óleo, conforme um procedimento específico da ferrovia, que busca eliminar as possibilidades de contaminação do óleo por fatores externos e internos durante a coleta. O recipiente que armazena a amostra de óleo é etiquetado com a identificação da locomotiva, o local no qual a locomotiva estava, o dia em que a coleta foi feita e quem realizou a mesma.

Após realizada a coleta, as amostras de óleo lubrificante seguem para o Laboratório da Oficina de Tubarão onde serão feitos o ensaio espectrométrico e os ensaios físico-químicos. Um dos motivos na utilização do ensaio espectrométrico, é que por meio da verificação da concentração dos elementos químicos presentes no óleo pode-se observar desgastes das partes móveis do motor, falhas na filtragem do ar de admissão, entre outras complicações. Já com os ensaios físico-químicos é possível verificar a presença de água e particulados no óleo lubrificante, mostrando à equipe

de manutenção que existem problemas no motor diesel, como vazamentos e contaminações.

Quanto aos desgastes que podem ser identificados pelos elementos químicos presentes no óleo no ensaio de espectrometria, entende-se que de forma direta, alguns destes elementos já podem indicar um possível desgaste, como é o caso do Ferro (Fe), Cromo (Cr), Cobre (Cu), Chumbo (Pb), Estanho (Sn), Prata (Ag), Níquel (Ni), Alumínio (Al) e Molibdênio (Mo), por estarem presentes na composição das peças do motor diesel. Outros elementos indicam o desgaste no motor por meio da detecção de contaminações internas e externas no óleo, como é o caso do Sódio (Na), Silício (Si) e novamente o Al. Podendo o Al indicar desgaste e também contaminação, foi criada uma relação entre elementos para uma correta interpretação, ou seja, quando os resultados de Si forem maiores que os do Al, é provável que tenha uma contaminação externa no óleo. Quando os resultados de Si forem menores os que do Al, é possível que esteja havendo um desgaste prematuro dos componentes do motor diesel.

Os três elementos restantes, Cálcio (Ca), Magnésio (Mg) e Zinco (Zn), são elementos que indicam condições da aditivação do óleo lubrificante, não somente a quantidade dos aditivos presentes no óleo, mas também possíveis falhas operacionais envolvendo o abastecimento do óleo lubrificante e misturas feitas durante sua utilização. É possível observar com as indicações acima, que as correlações entre os elementos químicos também podem contribuir com a detecção de um possível desgaste no motor diesel. Outras análises que, nas devidas proporções dos elementos, podem auxiliar nesta detecção, são as análises combinadas entre: Na e Si; Pb, Sn e Cu; Fe e Si; Cr e Si. Algumas correlações entre ensaios físico-químicos também são válidas, como por exemplo a relação entre o aumento da viscosidade com presença de água no óleo e o aumento de insolúveis com a diminuição dos valores de TBN.

As Tabelas 2 e 3 apresentarão alguns exemplos nos quais a interpretação dos resultados do ensaio espectrométrico e dos ensaios físico-químicos podem prever possíveis desgastes nos componentes dos motores diesel. As relações apresentadas foram identificadas pelos mecânicos e analistas em mais de 10 anos de observações

entre os resultados das análises de óleo e as condições das peças (composição) dos motores diesel.

Tabela 2 – Elementos químicos com teores elevados (espectrometria) e os possíveis componentes do motor diesel de locomotivas GE que podem estar sofrendo desgaste

Elementos químicos com teores elevados	Possíveis locais de desgaste no motor diesel
Ferro	Cilindros do motor diesel, camisas, anéis de segmento, engrenagens da extremidade livre, eixo comando de válvulas, cruzetas
Zinco	Cárter e interior do motor diesel (soltura pintura interna)
Cromo	Camisas, anéis de segmento dos cilindros, eixo comando de válvulas, eixos de manivelas e roletes das cruzetas
Cobre	Buchas dos balancins, virabrequim do motor diesel, buchas do eixo de comando de válvulas
Cobre, chumbo e estanho	Mancais fixos inferiores e superiores, anel de encosto do eixo de manivelas, mancais móveis inferiores e superiores, buchas das bielas articuladas e buchas dos pistões
Alumínio	Saias dos pistões, guias das cruzetas, mancais do eixo comando de válvulas, mancal de tampa de extremidade livre e conjunto retentor de óleo da extremidade do alternador principal de tração
Sódio	Sistema de resfriamento para o motor diesel
Silício	Filtro sacola e caixa de ar do motor diesel

Fonte: Adaptado acervo Vale (2019)

Tabela 3 – Valores encontrados nos ensaios físico-químicos e os possíveis componentes do motor diesel de locomotivas GE que podem estar sofrendo desgaste

Ensaios físico-químicos	Possíveis indicações de desgaste no motor diesel
Valores abaixo do normal de TBN	Corrosão nos componentes metálicos do motor
Valor acima do normal de insolúveis em pentano	Anormalidades nos compressores e guias de válvulas
Creptação e destilação (presença de água no óleo)	Anormalidades no bico injetor, tampa e junta da jaqueta e do cárter e vazamentos no sistema de resfriamento para o motor diesel.
Valores acima do normal de diluição	Anormalidades no bico injetor

Fonte: Adaptado acervo Vale (2019)

Todas as análises realizadas nas amostras de óleo foram feitas em um mesmo laboratório sob uma mesma metodologia. Todos os itens analisados no ensaio

espectrométrico e nos ensaios físico-químicos, bem como o nome do equipamento utilizado em cada ensaio, as unidades de medida e as faixas com os valores normais e críticos de cada parâmetro, serão apresentados nas Tabela 4 e 5 respectivamente. No caso do ensaio espectrométrico, um único aparelho denominado espectrômetro de emissão ótica a plasma, é utilizado.

Tabela 4 – Faixas com os limites normais e críticos dos valores encontrados no ensaio de espectrometria

Elementos Químicos	Unidade	Valor normal	Valor Alerta	Valor Critico
Cu	ppm	0 a 20	20 a 30	>30
Pb	ppm	0 a 8,25	8,25 a 11,97	>11,97
Sn	ppm	0 a 4,5	4,5 a 5	>5
Na	ppm	0 a 126	126 a 223	>223
Cr	ppm	0 a 15	15 a 20	>20
Si	ppm	0 a 18	18 a 21	>21
Fe	ppm	0 a 28	28 a 34	>34
Al	ppm	0 a 5	5 a 10	>10
Ag	ppm	0 a 1	1 a 2	>2
Zn	ppm	0 a 10	10 a 20	>20
Ni	ppm	0 a 3	3 a 5	>5
Ca	ppm	4000 a 6000	2000 a 4000	<2000
Mg	ppm	0 a 5	5 a 10	>10
Mn	ppm	-	-	-
Mo	ppm	-	-	-

Fonte: Adaptado acervo Vale (2019)

Tabela 5 – Faixas com os limites normais e críticos dos valores encontrados nos ensaios físico-químicos

Ensaio	Equipamento	Unidade	Valor crítico inferior	Valor alerta inferior	Valor normal	Valor alerta superior	Valor crítico superior
Destilação	Conjunto de vidrarias e termômetro	%	NA	NA	0	0,1	0,2
Viscosidade	Viscosímetro	cSt	0 a 12,4	12,5 a 13,5	Entre 13,6 e 16,9	17 a 19	>19
TBN	Titulador potenciométrico	mgKOH/g	3	5	>5	NA	NA
Insolúveis	Filtração a vácuo	%	NA	NA	<2	2	4
Creptação	Chapa de aquecimento	sim/não	sim	sim	não	sim	sim
Diluição	Viscosímetro	%	NA	NA	<2	2	5

Fonte: Adaptado acervo Vale (2019)

O ensaio espectrométrico, envolvendo os elementos Mn e Mo, são realizados pelo Laboratório Químico da ferrovia, mas não são utilizados nas interpretações de possíveis falhas no motor diesel. Desta forma, as faixas de seus limites não são conhecidas pelas equipes da ferrovia, e não são comumente fornecidos pelos manuais disponíveis por fornecedores de óleo e equipamentos específicos para ferrovia. Porém, uma análise de correlação foi feita com os dados destes elementos, e foi verificado que eles possuem boas correlações com desgastes nos componentes dos motores diesel. Desta forma, eles continuaram fazendo parte do escopo desta dissertação, ainda que antes não fossem conhecidos seus valores críticos e normais em presentes no óleo lubrificante.

Após realizada as análises, a equipe de laboratório e os analistas técnicos de manutenção realizam as interpretações de seus resultados, destacando os valores anormais que podem indicar possíveis desgastes nos motores, e principalmente se o nível desses desgastes constitui uma possível falha catastrófica no motor na sequência.

Após a realização das interpretações dos dados pelos analistas, caso seja identificado algum risco na relação da análise do óleo lubrificante com o motor diesel de determinada locomotiva, é solicitada uma inspeção no motor diesel, por meio da criação de uma ordem de manutenção (OM) no Sistema Informatizado da Manutenção.

As OM abertas no sistema de manutenção da ferrovia podem ser de três tipos: ordens preventivas (YPM), corretivas (YCM) e emergenciais (YEM). As YPM são abertas com periodicidade definida nos planos de manutenção para inspeção das locomotivas e avaliação de seus componentes. Caso durante uma inspeção seja identificado desgaste em algum componente é aberta uma YCM para detalhar o que será corrigido ou trocado. As YEM são ordens criadas quando o motor diesel falha durante a operação da locomotiva, ou seja, no tráfego da ferrovia. Para inserir uma YEM no sistema de manutenção o operador da locomotiva entra em contato com o centro de controle de trens da ferrovia, indicando detalhes da falha e solicitando a abertura da YEM no sistema. Os três tipos de ordens são utilizados nessa dissertação, porém a falha é relacionada apenas com os tipos YCM e YEM.

As ordens do tipo YEM são de muito valor para predição, já que indicam uma falha exatamente no momento em que ela ocorreu. Entende-se também que a maioria das falhas tendem a ocorrer quando a locomotiva está em funcionamento (em viagem), pois são nestes momentos que ela está realizando um maior esforço, principalmente nos elementos geradores de energia, como o motor diesel. Toda essa explicação não elimina o valor das ordens YCM, já que elas também indicam falhas, as quais foram observadas na Oficina de Locomotivas ou em outro momento de utilização da locomotiva.

Além da análise do óleo lubrificante, os dados de consumo de diesel e os dados de movimentação de motores diesel entre locomotivas, também foram utilizados como dados para predição das falhas nos motores. O consumo de diesel indica o grau de utilização da locomotiva, e consequentemente do motor diesel, ou seja, quanto maior a quantidade em litros de diesel abastecida em uma determinada locomotiva, mais utilizada ela foi, e desta forma, maior foi a utilização do motor diesel inserido nessa locomotiva.

A movimentação de motor entre locomotivas refere-se a um arquivo simples que indica qual a locomotiva que determinado motor diesel está inserido e a data na qual ele esteve nessa locomotiva. Esse documento é necessário pois um mesmo motor pode ser utilizado por várias locomotivas. A explicação desse fato vem da condição que uma locomotiva(1) não pode ficar muito tempo retida na Oficina, sendo assim o motor diesel retirado dela para manutenção fica na Oficina sendo retificado, e um outro motor, que possivelmente foi retificado há pouco e estava em outra locomotiva(2), é inserido na locomotiva(1), liberando a mesma para voltar para a operação.

Por meio do documento de movimentação de motores entre locomotivas é possível acompanhar não só as locomotivas, mas cada motor diesel até o final de sua vida útil, o que garante que todos os dados utilizados nas relações com as falhas em determinado motor foram relacionados no período correspondente a locomotiva que o motor estava inserido.

Diante do entendimento dos processos apresentados e da necessidade de uma equipe com competência técnica e experiência na relação existente com os motores diesel, bem como na dificuldade em se obter resultados de interpretação rápidos,

entende-se que a espera pela entrega dos resultados pelo laboratório mais as interpretações dos analistas torna a previsão das falhas, as vezes tardia, ou seja, a falha ocorre antes que seus sintomas sejam identificados. Face a toda essa complexidade, e na necessidade de se obter uma maior confiabilidade da ferrovia, a equipe de manutenção espera que seja criado um método baseado em IA para auxiliar o analista na predição de falhas em motores diesel de locomotivas. Eles ainda esperam que este método seja mais eficiente, rápido e independa do conhecimento específico dos analistas.

Sabendo-se que as duas últimas bases de dados apresentadas (consumo de diesel e movimentação de motores entre locomotivas) são simples, não requerendo maiores explicações, nos próximos subcapítulos, 3.2 e 3.3, serão abordadas revisões técnicas especificas sobre motor diesel de locomotivas e análise de óleo lubrificante de motor diesel.

3.2 MOTOR DIESEL DE LOCOMOTIVAS

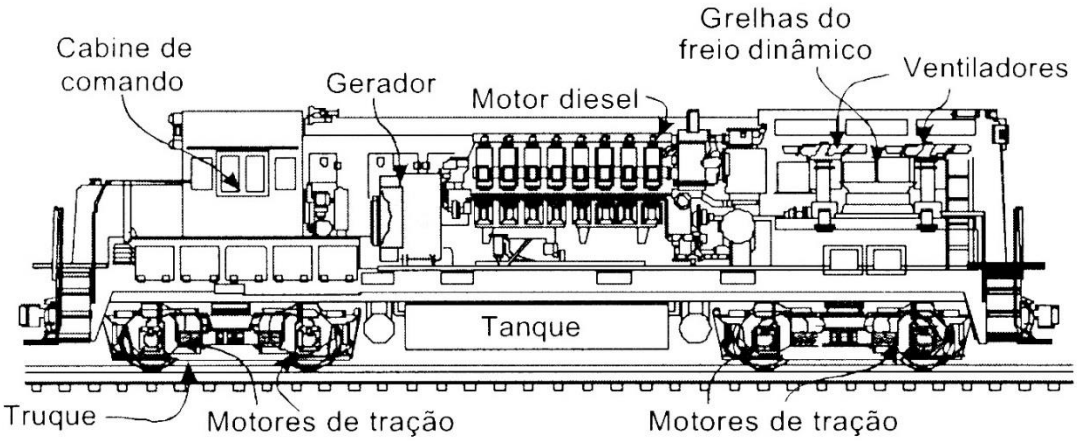
O motor diesel surgiu em 1893, foi patenteado por Rudolf Diesel, sendo identificado como um novo tipo de motor de combustão interna. A aplicação do motor diesel teve início no começo dos anos 90, com relatos de sua utilização em navios, seguido dos automóveis, e por volta de 1925 em locomotivas (MOLLENHAUER e TSCHÖKE, 2010).

O esforço trator e a potência são dois fatores importantes para que a locomotiva consiga mover o trem. O esforço trator ou tração aciona engrenagens que efetivamente giram as rodas do trem, e para que isso ocorra é necessário haver a combustão do motor diesel, que produz a energia necessária para que um gerador elétrico abasteça motores elétricos de tração (ROSA e RIBEIRO, 2017).

A Figura 5 apresenta uma melhor compreensão da montagem de uma locomotiva diesel-elétrica e onde o motor diesel fica localizado. Os motores de tração são acionados pela energia gerada pelo conjunto composto pelo motor diesel e pelo gerador, movimentando o trem a partir do momento que o maquinista, da cabine de comando, seleciona a potência desejada. Os ventiladores são acionados caso o motor

chegue a uma determinada temperatura, na qual o calor produzido precise ser vertido para a atmosfera.

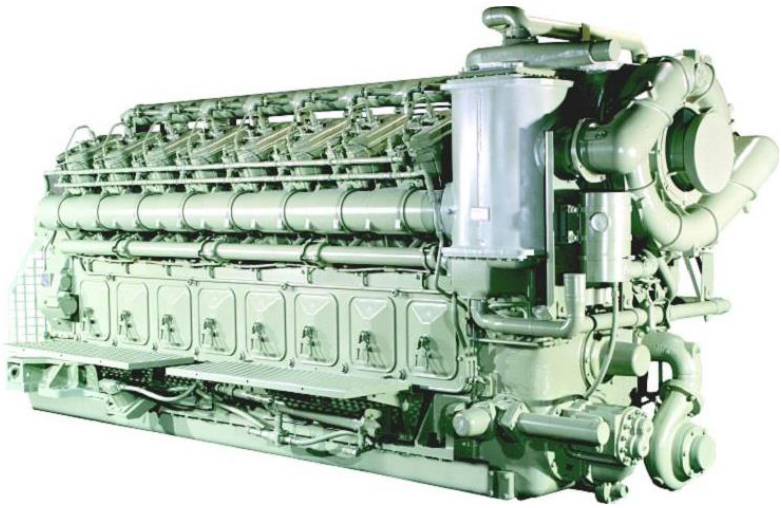
Figura 5 – Montagem básica de uma locomotiva diesel-elétrica



Fonte: Rosa e Ribeiro (2016, p. 162)

O motor diesel é o motor de combustão interna mais utilizado em áreas como: mecanismos de instalação fixa, veículos comerciais, locomotivas, construção civil, máquinas agrícolas e navios. Isso se deve, principalmente, pelo alto grau de eficiência e à economia de combustível resultante. Outro fator que estimula a utilização do motor diesel é que, ao contrário do motor a gasolina, ele suporta mais choques, possui grande potência e segurança operacional, principalmente por utilizar óleos menos inflamáveis (REIF, 2014). Toda essa aplicabilidade é positiva, visto que esta dissertação pode ser aplicada a outros equipamentos que utilizam o motor diesel, e de maneira necessária, também o óleo lubrificante. A Figura 6 apresenta um motor diesel utilizado em uma locomotiva DASH9 W, possibilitando uma maior compreensão das dimensões e aparência do motor diesel.

Figura 6 – Motor diesel GE 7FDL16 aplicado nas locomotivas DASH9 W



Fonte: Romanha Junior (2016, p.16)

O motor diesel é formado por subsistemas essenciais a geração de energia, sendo os principais: sistema de controle, de lubrificação, de injeção, de admissão de ar e sistema de arrefecimento. Para que um motor diesel funcione adequadamente todos os seus subsistemas devem estar em um bom estado de utilização, o que contribui com a redução no consumo de diesel e com um aumento da vida útil do equipamento (XIN, 2011).

Dentre os subsistemas principais do motor diesel, o de lubrificação tem um papel importante no sentido de manter os elementos móveis do motor separados, protegendo as peças dos impactos gerados no funcionamento do equipamento e por consequência protegendo o motor de corrosões e demais falhas (VAN BASSHUYSEN e SCHÄFER, 2004).

3.3 ÓLEO LUBRIFICANTE DE MOTORES DIESEL

Lubrificante é qualquer substância interposta entre superfícies que estão em contato e em movimento, com o objetivo de diminuir o desgaste e a fricção entre elas, além de protege-las contra corrosão e auxiliar na vedação (CUNHA, 2005). Nos motores diesel, o óleo lubrificante atua como controlador do atrito, inibindo assim um maior desgaste se seus componentes, também atua como controlador da temperatura, pois absorve o calor gerado pelo contato das superfícies, e como controlador da corrosão, pois evita que os ácidos presentes no ambiente em contato com o motor corroam os

metais de suas peças, entre outras funções que contribuem com o bom funcionamento do motor diesel e sua vida útil. Por consequência a falta de lubrificação causa problemas no motor diesel e em qualquer outro equipamento que necessite de óleo lubrificante, tais como: aquecimento, desgaste e dilatação, ruídos, desalinhamento e ruptura das peças (BORBA e ZUCOLOTTI, 2010).

Para que o óleo lubrificante atue com suas propriedades de preservação do motor, ele deve ser inserido em um sistema formado por vários elementos, ressaltando o reservatório, a bomba de óleo e o resfriador de óleo. Nos mancais de um motor diesel, e demais partes do motor por onde o óleo circula existe uma série de acontecimentos resultantes do processo de funcionamento do motor, que podem ser observados por meio da análise físico-química e espectrométrica do óleo lubrificante, como geração de partículas e outras substâncias resultantes do desgaste dos elementos do motor. Por meio da interpretação dos resultados das análises de óleo, é possível definir processos de manutenção e identificar a condição de vida do motor (CARMO, 2018).

Quando é feita uma análise no óleo lubrificante de um motor diesel de locomotiva e nela é observada uma quantidade de partículas que indicam variações de suas propriedades químicas e físicas, e um desgaste deste motor, automaticamente a locomotiva é levada para realizar o reparo, pois é provável que a qualquer momento vai haver uma falha catastrófica no equipamento, sendo necessário rebocar a locomotiva, ou seja, utilizar outra locomotiva para conduzi-la até a oficina. Realizando uma análise de custos em um período de um ano, conclui-se que os custos evitados com falhas catastróficas serão, em escala, expressivamente maiores que os custos para analisar o óleo lubrificante da mesma frota de locomotivas no mesmo período (SPECTRO INC, 2012).

Para avaliar, por meio do óleo lubrificante, o que contaminou o motor, deve-se identificar de que forma e nível ocorreu essa contaminação. Algumas das formas mais comuns de contaminação e as consequências destas para o motor diesel, estão: contaminação por água, diluições encontradas nos combustíveis e poeira. A função lubrificante do óleo é severamente afetada pela presença de água, pois esta pode gerar corrosões no motor. Da mesma forma quando o contato ocorre entre o óleo e os combustíveis, a viscosidade do óleo cai de forma potencial, acelerando o desgaste. Assim como a água, a poeira também acelera o desgaste dos componentes do motor,

entrando em contato com o óleo por meio de uma filtração de ar ineficiente e/ou falta de cuidado com o estoque, durante o manuseio do recipiente contendo o óleo lubrificante (NEWELL, 1999).

As contaminações por água, combustível e poeira ocorrem quando estas substâncias entram em contato com o próprio óleo lubrificante utilizado no motor diesel. Vazamento de água para o compartimento que contém o óleo, oxida o mesmo e ocasiona a corrosão nas partes móveis do motor; vazamento de combustível (diluição), que em contato com o óleo e o enxofre ali presente, também contribui com a corrosão dos mancais do motor diesel; passagem de compressão, efeito que aumenta a quantidade de partículas de resíduos de carbono e a viscosidade do óleo, fazendo com que a filtragem do mesmo tenha que ser mais frequente. Para identificar estes efeitos sem a necessidade de parar a locomotiva e inspecionar o motor diesel, é feito a análise periódica do óleo lubrificante coletado no motor diesel (BORBA e ZUCOLOTTI, 2010).

As análises que devem ser realizadas no óleo lubrificante utilizado em motor diesel, de forma preferencial, de acordo com os laboratórios de análise e fabricantes de óleo e equipamentos, para que se consiga identificar algum possível desgaste no motor diesel, são de dois tipos, ensaio espectrométrico e ensaios físico-químicos. O ensaio espectrométrico utiliza um único equipamento e identifica a concentração de elementos químicos presentes no óleo, como níquel, alumínio, ferro, chumbo, entre outros. Os ensaios físico-químicos, utilizam diferentes equipamentos para avaliar a creptação, destilação, diluição, TBN, insolúveis e viscosidade do óleo lubrificante. A creptação determina qualitativamente se existe água no óleo ou não, caso exista, a quantidade de água é medida no ensaio de destilação. A diluição é um ensaio que determina se existe diesel no óleo lubrificante, enquanto o ensaio de TBN determina a quantidade e o caráter básico destes óleos, o que geralmente contribui com a identificação do consumo de aditivos presentes no óleo. A análise de insolúveis identifica substâncias que não são solúveis no óleo (vernizes, borras, gomas, resinas, etc.) resultantes da degradação do motor diesel ou contaminações provenientes do ambiente interno ou externo da locomotiva. A viscosidade do óleo está relacionada a resistência do fluido ao escoamento, e dentre todas as análises pode ser considerada a mais importante, pois identifica as reais taxas de desgaste dos elementos do motor e a qualidade lubrificante do óleo.

A necessidade de fazer uso das análises de óleo lubrificante para priorizar manutenção em vez de realizar os gastos com marcos em cronogramas pré-estabelecidos, em equipamentos que não necessariamente precisam daquela manutenção naquele momento, se torna cada dia mais evidente para as empresas de transportes que realizam os cálculos dos custos e de eficiência de seus planos de manutenção (SPECTRO INC, 2012). Em manutenções preditivas envolvendo o óleo lubrificante são necessários instrumentos específicos de monitoração e uma interpretação técnica avançada, o que torna este processo oneroso e menos vantajoso. Frente a essa realidade, iniciar os processos de planejamento da manutenção fazendo uso de métodos de predição torna-se ainda mais necessário e válido (KIMURA, 2010).

4 LEVANTAMENTO E TRATAMENTO DOS DADOS

O levantamento dos dados utilizados nessa dissertação foi feito na Oficina de Locomotivas de Tubarão e no Laboratório Químico de Tubarão. Todos os dados foram obtidos de sistemas informatizados da EFVM.

Após entendimento do assunto que seria tratado nesta dissertação, ficou claro que para compor o método de predição seriam necessárias três bases de dados essenciais e mais duas para complementar e tentar melhorar a performance do modelo de predição. A primeira seria a base de dados das análises de óleo lubrificante propriamente ditas, realizadas no laboratório da ferrovia. A segunda seria composta pelas falhas que já ocorreram nos motores diesel das locomotivas da EFVM, as OM, ressaltando que essas falhas devem possuir uma relação entre o motor diesel e o óleo lubrificante, ou seja, deve ser possível prever a falha com informações da análise do óleo lubrificante. Devido aos diferentes modelos de locomotivas utilizados na EFVM, uma terceira base de dados foi necessária para identificar as 216 locomotivas selecionadas para o estudo em meio as demais. Quando essas três bases são unidas é possível identificar para cada locomotiva seu histórico de falhas e os resultados das análises de óleo lubrificante realizadas em seus motores diesel.

As outras duas bases de dados utilizadas para complementar o estudo referem-se ao consumo de diesel das locomotivas e a planilha com as informações da movimentação dos motores diesel entre locomotivas. Ambas as bases foram utilizadas na tentativa de melhorar os resultados encontrados com as três bases essenciais.

A primeira base de dados, de análise do óleo lubrificante, foi adquirida em duas partes, ensaios físico-químicos e ensaio espectrométrico. Estes dados atualmente são armazenados em um sistema informatizado, porém até março de 2018 eles eram armazenados em um outro sistema, desta forma, veio deste último a maior parte dos dados. A Tabela 6 apresenta os títulos das colunas das planilhas contendo os resultados dos ensaios físico-químicos e do ensaio espectrométrico. Além do título das colunas, é apresentado o seu significado, e se a coluna irá compor ou não, o grupo de dados dessa dissertação.

Tabela 6 – Dados que compõem as planilhas dos ensaios físico-químicos e espectrométrico

Nº da coluna	Título da coluna	Significado do título da coluna	Será utilizado?
1	Id. da coleta	Nº de identificação da coleta	Não
2	Local de coleta	Nº de identificação da locomotiva	Sim
3	Data Coleta	Data em que foi feita a coleta do óleo	Sim
4	Data de recebimento	Data que o laboratório recebeu a amostra de óleo	Não
5	Tipo de amostra	Tipo de material que está sendo analisado	Não
6	Nome da classe	Equipamentos que compõem a locomotiva	Não
7	Método	O método utilizado para realizar o ensaio	Não
8	Unidade	Unidade de medida dos valores dos ensaios	Não
9	Creptação	Indica se existe água no óleo	Não
10	Destilação	Indica a quantidade de água no óleo	Sim
11	Diluição	Indica se existe combustível no óleo	Sim
12	Insolúveis	Indica a existência de particulados no óleo	Não
13	TBN	Indica o teor de basicidade do óleo	Sim
14	Viscosidade	Indica a viscosidade do óleo (usado)	Sim
15	Ag	Quantidade de prata presente no óleo	Sim
16	Al	Quantidade de alumínio presente no óleo	Sim
17	Ca	Quantidade de cálcio presente no óleo	Sim
18	Cr	Quantidade de cromo presente no óleo	Sim
19	Cu	Quantidade de cobre presente no óleo	Sim
20	Fe	Quantidade de ferro presente no óleo	Sim
21	Mg	Quantidade de magnésio presente no óleo	Sim
22	Mn	Quantidade de manganês presente no óleo	Sim
23	Mo	Quantidade de molibdênio presente no óleo	Sim
24	Na	Quantidade de sódio presente no óleo	Sim
25	Ni	Quantidade de níquel presente no óleo	Sim
26	Pb	Quantidade de chumbo presente no óleo	Sim
27	Si	Quantidade de silício presente no óleo	Sim
28	Sn	Quantidade de estanho presente no óleo	Sim
29	Zn	Quantidade de zinco presente no óleo	Sim

Fonte: Adaptado acervo Vale (2019).

A decisão de quais variáveis seriam utilizadas no primeiro momento da análise foi das equipes de manutenção dos motores em conjunto com os desenvolvedores do modelo. Apenas as variáveis que de forma óbvia não seriam necessárias foram descartas. Com as variáveis escolhidas nessa fase foi feita uma segunda “triagem”, que contou não só com as indicações das equipes de manutenção dos motores, como também de soluções provenientes de um software especifico, como será explicado nos próximos parágrafos deste capítulo. Para todas as bases de dados utilizadas foi feita essa análise prévia e depois a análise mais aprofundada das variáveis.

Para correlacionar todas as cinco bases de dados é necessário que exista um dado em comum entre elas, neste caso será a identificação da locomotiva, semelhante ao que seria a placa de um carro, ou seja, o óleo lubrificante analisado em uma determinada locomotiva deve ser vinculado à falha que pode ter ocorrido nesta mesma locomotiva, tudo sendo ligado por meio do código da locomotiva em comum. Desta forma, outros dados de identificação, como o número da coleta, não serão necessários.

A coluna 3 indica a data da coleta, informação necessária para que seja possível correlacionar as análises de óleo lubrificante com a falha do motor diesel em uma mesma locomotiva. A correlação consiste na interpretação dos resultados de análise do óleo realizadas em uma amostra coletada em uma data imediatamente antes da data de uma falha em uma mesma locomotiva. Com essa lógica será possível entender se valores anormais presentes nestes resultados de análise de óleo podem explicar a falha ocorrida no motor diesel.

A coluna 4, data de recebimento, indica quando o laboratório recebeu a amostra de óleo coletada, geralmente é uma data igual ou bem próxima da data de coleta, sendo este dado desnecessário para este estudo. As colunas 5 e 6, tipo de amostra e nome da classe indicam que a amostra é de óleo lubrificante novo ou usado e a classe do equipamento é um motor diesel. Estas colunas existem, pois, o Laboratório Químico da Oficina de Tubarão realiza experimentos em diversos materiais e para diferentes equipamentos, porém neste caso as planilhas já foram tratadas, e os dados de materiais que não sejam o óleo lubrificante usado no motor diesel, como por exemplo o óleo hidráulico, água, areia entre outros materiais usados na locomotiva, foram descartados de imediato. O mesmo ocorre para as classes de equipamentos que não correspondam ao motor diesel, como o compressor, sistema de refrigeração, caixa multiplicadora, gerador principal, entre outros equipamentos que, muitas vezes, também utilizam óleo lubrificante em seus sistemas.

As colunas 7 e 8, método do ensaio e unidade de medida de cada ensaio, não serão utilizadas pois não são necessárias para que ocorra a correlação dos resultados de análise de óleo lubrificante com o histórico de falhas no motor diesel. As colunas 9 a 29 são os ensaios de laboratório compostas pelos resultados dos mesmos, dados

fundamentais para predizer as falhas no motor diesel. Porém a coluna 9 e 12 não serão utilizadas.

A coluna 9 refere-se ao ensaio físico-químico de creptação, que determina se existe água no óleo lubrificante, assim como o ensaio de destilação, porém a destilação indica a quantidade exata de água, e a creptação apenas se existe a presença ou não. Assim como será explicado mais a frente, dois ensaios que identificam o mesmo resultado não são necessários para a predição de falha no motor diesel. A coluna 12, que indica o ensaio de insolúveis, não será utilizada pois durante a análise dos resultados desse ensaio foi identificado muitos *outliers* (valores anormais aos valores possíveis para este ensaio), e em reunião com a equipe do Laboratório Químico foi reconhecido um erro no equipamento que realiza o ensaio, fazendo com que seus resultados não sejam válidos para utilização na predição.

A segunda base de dados é formada pelas falhas que ocorreram no motor diesel, armazenadas no formato de OM, que são inseridas em um sistema computacional pelos responsáveis. As OM são necessárias para que a equipe de manutenção identifique que existe uma falha no motor diesel prejudicando o funcionamento dele e da locomotiva em que ele está inserido, e que o procedimento de manutenção deve ser iniciado. De 2006 a 2009 todas as OM eram armazenadas em um sistema informatizado, incluindo todos os tipos de OM: YPM, YCM e YEM. Em 2009 iniciou-se a abertura de OM do tipo YEM pelo centro de controle de trens, dessa forma de 2009 até o presente momento as ordens YEM ficam em um sistema especifico. Já as ordens YPM e YCM migraram para um sistema novo de manipulação de ordens de 2015 até o presente momento.

A Tabela 7 apresentará as colunas das planilhas de falhas, o significado de seus títulos e se as colunas serão utilizadas ou não para compor o grupo de dados desta dissertação.

Tabela 7 – Dados que compõem as planilhas de falhas no motor diesel

Nº da coluna	Título da coluna	Significado do título da coluna	Será utilizado?
1	OM	Nº da ordem de manutenção	Não
2	Tipo ordem	Ordem preventiva, corretiva ou emergencial	Sim
3	Locomotiva	Nº de identificação da locomotiva	Sim
4	Data entrada	Data da criação da OM	Sim
5	Texto breve	Descrição da manutenção	Não
6	Criado por	Quem criou a OM no sistema	Não
7	Centro de custo	Área responsável pela manutenção	Não
8	Status	Se a OM foi realizada ou não	Não
9	Data início	Data que a manutenção foi iniciada	Não
10	Data fim	Data que a manutenção foi finalizada	Não
11	Hora início	Hora que a manutenção foi iniciada	Não
12	Hora fim	Hora que a manutenção foi finalizada	Não
13	Plano	Sigla do componente tratado na manutenção	Não

Fonte: Adaptado acervo Vale (2019)

A coluna 1 da planilha de falhas consiste nos números das OM, códigos irrelevantes para esta dissertação, já as colunas 2, 3 e 4, respectivamente o tipo da ordem, a identificação da locomotiva e a data de criação da OM, são relevantes e necessárias para criação da correlação entre a análise de óleo e a falha. A identificação da locomotiva, coluna 3, é a mesma identificação utilizada nas planilhas de análise de óleo, ou seja, o dado em comum entre todas as planilhas utilizadas nesta dissertação.

A coluna 4, data entrada, é a data que a OM foi criada no sistema. Essa data geralmente é a mesma que a falha foi observada, ou durante uma viagem da locomotiva ou quando ela estava nos locais de inspeção. Desta forma, a data de criação da OM será considerada a data da falha nos casos das manutenções corretivas.

As colunas de 5 a 12 são compostas por dados não relevantes para esta dissertação e que por isso foram eliminadas na fase de tratamento dos dados. Essas colunas descrevem informações para realização da manutenção pelas equipes da ferrovia, como o texto breve com a descrição da manutenção, geralmente explicando quais foram os sintomas constatados e a peça que será corrigida ou trocada, a identificação do colaborador da ferrovia que criou a OM no sistema, a data e a hora de início e fim

das operações de manutenção, bem como se elas estão totalmente finalizadas ou não.

Como as OM são criadas para realização da manutenção de qualquer peça da locomotiva, e não existe um filtro rápido que elimine de forma especifica as OM que não possuem relação com o motor diesel e as outras bases de dados utilizadas, a coluna 13 foi criada para realizar esta seleção. Nesta coluna foram criados códigos relacionados à peça que está sendo recuperada ou trocada, e por este código foi possível selecionar apenas as OM que possuem relação com o motor diesel e o óleo lubrificante. Porém essa coluna 13 não será utilizada para compor o método de predição, pois essa seleção foi feita de forma manual, antes de imputador os dados para criação do modelo. Desta forma, todas as OM imputadas para criação do modelo já foram selecionadas antes de serem inseridas no mesmo. O Quadro 2 apresenta os códigos dos componentes do motor diesel que foram selecionados para uso no método de predição e seus significados.

Quadro 2 – Códigos selecionados para compor os dados de falhas (histórico de OM) do motor diesel

Código	Significado
BA	Bomba d'água
BBINJ	Bomba injetora (conjunto)
BL	Bomba de limpeza
BP	Bomba de óleo lubrificante (pressão)
CF	Conjunto de força
MD	Motor diesel
ROLEO	Resfriador de óleo
SPMD	Soprador do motor diesel (Blower)
TB	Turbina
JAQ	Jaqueta
FMI	Filtro do óleo lubrificante
FLT	Filtro da turbina

Fonte: Adaptado acervo Vale (2019)

Com as bases de dados de falhas, representadas pelas OM, e os dados de análise de óleo lubrificante, era necessário estabelecer uma forma de filtrar apenas as locomotivas que fazem parte do escopo dessa dissertação, nesse caso dos modelos de locomotivas conhecidos como Dash-8 e Dash-9. Para isso uma planilha com os dados de todas as locomotivas que operam na EFVM foi unida às bases de falhas e análise de óleo por meio dos códigos das locomotivas Dash-8 e Dash-9, excluindo todos os dados pertinentes a outros modelos de locomotivas. A Tabela 8 apresenta

os títulos das colunas da planilha com as informações das locomotivas, seus significados e quais dados serão usados nesta dissertação.

Tabela 8 – Dados que compõem a planilha de identificação das locomotivas

Nº da coluna	Título da coluna	Significado do título da coluna	Será utilizado?
1	Código loco	Nº de identificação da locomotiva	Sim
2	Código loco SAP	Qual ferrovia a locomotiva transita e seu modelo	Sim
3	Código modelo loco	Nome do modelo da locomotiva	Não
4	Peso total	Peso total da locomotiva	Não
5	Quantidade de eixos	Quantidade de eixos da locomotiva	Não
6	Empresa proprietária	Código da empresa proprietária das locomotivas	Não
7	Data fabricação	Data de fabricação da locomotiva	Não

Fonte: Adaptado acervo Vale (2019)

A Tabela 8 apresenta além das informações sobre todas as locomotivas que transitam na EFVM, como também a identificação numérica das mesmas, a frota e o modelo a qual pertencem, entre outras características. Nessa dissertação é preciso apenas identificar o número de identificação da locomotiva e qual o seu modelo, para que fosse selecionado para o desenvolvimento do modelo apenas as locomotivas escolhidas na fase de estudo do problema. Para compor essa dissertação será necessário utilizar apenas os dados que compõem às colunas 1 e 2 da Tabela 8. As demais colunas não serão necessárias.

A data de fabricação da locomotiva parece ser um dado relevante, visto que quanto mais a locomotiva operou mais chances ela tem de falhar, ainda que o motor diesel instalado nela seja novo. Porém, como a idade da locomotiva não pode ser vinculada ao seu grau de utilização, já que na ferrovia as locomotivas desempenham diferentes atividades, como exemplo, uma transita mais em pátios, outras mais no transporte de cargas, ou seja, nem sempre uma locomotiva mais velha terá sido mais utilizada em comparação com uma mais nova, e dessa forma entendeu-se que não seria adequado utilizar essa variável.

Após a obtenção das duas bases de dados principais, falhas e análise de óleo, identificar apenas as locomotivas relevantes, realizar o tratamento dos dados de forma manual e no software onde será desenvolvido o modelo, e realizar inúmeros testes com o fluxo criado para predição, foi constatado que mais duas bases de dados poderiam ser utilizadas para melhorar os resultados. A quarta base de dados é

composta pelos dados de consumo de combustível da locomotiva, de 2009 a 2018, sendo que de 2009 a 2015 os dados estiveram armazenados em um sistema informatizado e de 2015 até o momento em outro sistema informatizado.

A quarta base de dados indica, assim, o grau de utilização de cada locomotiva, pois entende-se que uma locomotiva mais operante, exige mais de seu motor diesel. Desta forma, a relação da falha e até mesmo dos resultados das análises de óleo com o motor diesel, é diferente em uma locomotiva que está sempre em trânsito com uma locomotiva que, na maioria das vezes, está parada ou na oficina. A Tabela 9 apresenta as colunas da planilha de consumo de combustível, o significado de seus títulos e se as colunas serão utilizadas ou não para compor o grupo de dados dessa dissertação.

Tabela 9 – Dados que compõem a planilha de consumo de combustível das locomotivas

Nº da coluna	Título da coluna	Significado do título da coluna	Será utilizado?
1	Empresa	Nome da ferrovia	Não
2	Posto	Nome do posto de abastecimento	Não
3	Locomotiva	Nº de identificação da locomotiva	Sim
4	Volume abastecido	Volume de combustível abastecido	Sim
5	Nível combustível	Nível atual de combustível no tanque	Não
6	Capacidade tanque	Capacidade do tanque de combustível	Não

Fonte: Adaptado acervo Vale (2019)

Das informações adquiridas na planilha apresentada na Tabela 9, apenas as colunas 3 e 4 serão utilizadas, contendo respectivamente: a identificação da locomotiva (dado em comum com as demais bases de dados) e o volume de combustível abastecido para cada locomotiva. Entende-se que quanto maior o volume de combustível abastecido na locomotiva, maior a sua utilização, e maiores são as chances de ocorrer uma falha. As demais colunas não acrescentam informações relevantes à dissertação.

A outra base de dados selecionada para testes, na tentativa de alcançar melhores resultados de predição, foi a planilha com as movimentações de motores diesel entre locomotivas, com dados de 2000 até 2018. A base tornou-se essencial no sentido de que identificar os períodos que cada motor esteve em determinada locomotiva e poder relacionar os demais dados dentro das faixas desses períodos é primordial para que o algoritmo de predição interprete corretamente os dados. A Tabela 10 apresenta as colunas da planilha de movimentação de motores entre locomotivas, o significado de

seus títulos e se as colunas serão utilizadas ou não para compor o grupo de dados dessa dissertação.

Tabela 10 – Dados que compõem a planilha de movimentação de motores entre locomotivas

Nº da coluna	Título da coluna	Significado do título da coluna	Será utilizado?
1	Empresa	Nome da ferrovia	Não
2	Motor	Nº de identificação do motor	Sim
3	Locomotiva	Nº de identificação da locomotiva	Sim
4	Data entrada	Data que o motor entrou na locomotiva	Sim
5	Data saída	Data que o motor saiu da locomotiva	Sim

Fonte: Adaptado acervo Vale (2019)

Dos dados apresentados na Tabela 10, somente as colunas 2,3,4 e 5 serão utilizadas. O número de identificação da locomotiva é a chave que liga essa base de dados às demais bases utilizadas, é a informação que elas possuem em comum. O número de identificação do motor é usado para acompanhar qual motor está instalado em determinada locomotiva, e a data de entrada e saída corresponde a data que o motor foi instalado em determinada locomotiva e a data que ele foi retirado da locomotiva respectivamente.

Observa-se que as cinco bases de dados, apresentadas até aqui, possuem históricos em diferentes períodos, e muitas vezes foram armazenadas em diferentes sistemas informatizados com o tempo. Para que fique mais claro os períodos e sistemas que cada base de dados foi armazenada, o Quadro 3 apresentará o resumo dessas informações.

Quadro 3 – Períodos e sistemas informatizados utilizados para armazenar os históricos de dados

Histórico de dados	Histórico de dados - detalhado	Sistema informatizado	Período disponível de armazenamento dos dados	
			Data de início	Data fim
Análises de óleo lub.	Análises de óleo lub.	Sistema A	2006	março 2018
Análises de óleo lub.	Análises de óleo lub.	Sistema B	março2018	até o momento
Falhas motor - OM	OM tipos: YPM, YCM e YEM	Sistema C	2006	2009
Falhas motor - OM	OM tipos: YPM e YCM	Sistema C	2009	2015
Falhas motor - OM	OM tipo: YEM	Sistema D	2009	até o momento
Falhas motor - OM	OM tipos: YPM e YCM	Sistema E	2015	até o momento
Ativos da EFVM	Cadastro das locomotivas	Sistema F	2000	até o momento
Consumo de combustível	Consumo de combustível	Sistema G	2009	2015
Consumo de combustível	Consumo de combustível	Sistema H	2015	até o momento
Movimentação motores	Movimentação motores	Sistema I	2000	até o momento

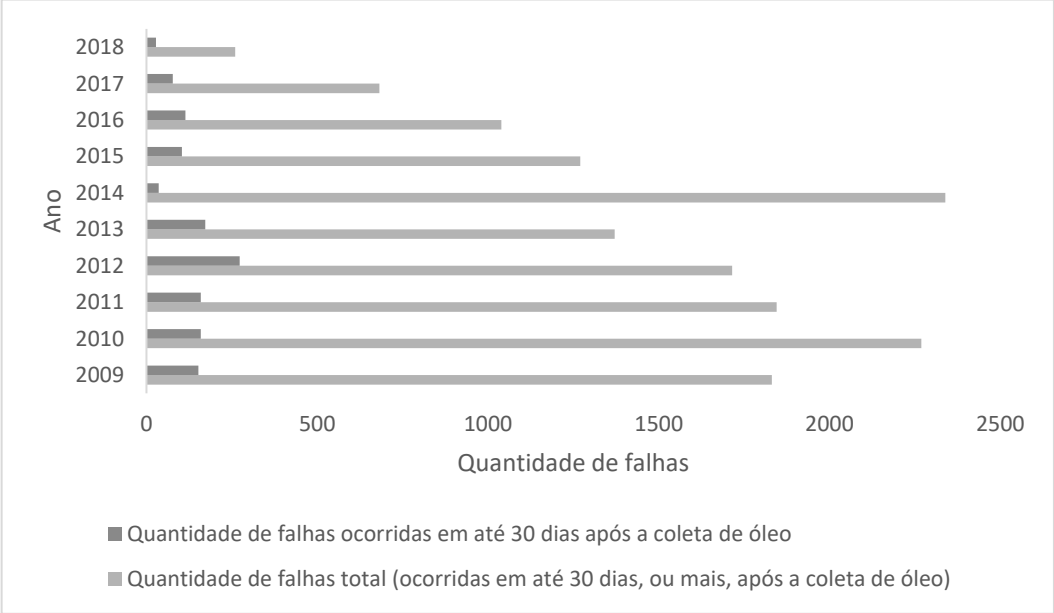
Fonte: Adaptado acervo Vale (2019)

Como pode ser observado no quadro 3, as bases de dados não coincidem em disponibilidade inicial, ou seja, algumas delas estão disponíveis para uso desde 2000, já outras só a partir de 2009. Essa diferença nos períodos limita a utilização dos dados mais antigos, ou seja, todas as bases de dados serão utilizadas a partir da data de início da base mais atual, neste caso a partir de 2009. Desta forma, todos os dados utilizados nessa dissertação são de 2009 a dezembro de 2018.

Segundo as equipes de manutenção da Oficina de Locomotivas e do Laboratório Químico da Ferrovia, as características de manutenção dos motores diesel e os procedimentos de análise de óleo lubrificante não tiveram mudanças significativas de 2009 a 2018, e por isso esse período é satisfatório para ser utilizado. Foi determinado que para atender à necessidade dos usuários do modelo, neste caso a Oficina de Locomotivas no planejamento das manutenções dos motores diesel, a predição da falha no motor deveria ocorrer em um prazo de até 30 dias, ou seja, ao imputar as análises de óleo no modelo, o último deveria fornecer a indicação de falha do motor nos próximos 30 dias após a análise de óleo. Desta forma as equipes poderiam planejar a inspeção da locomotiva que iria falhar e realizar as devidas manutenções de seus componentes antes que os mesmos falhassem durante a operação da locomotiva.

A Figura 7 apresenta a quantidade de falhas ocorridas nos motores diesel anualmente, de 2009 até 2018. Percebe-se na Figura a quantidade de falhas ocorridas no prazo de até 30 dias após a coleta e análise de óleo e também a quantidade total de falhas ocorridas independentemente do prazo, ou seja, as que ocorreram em até 30, 45, 60 dias e assim por diante.

Figura 7 – Quantidade de OM criadas relacionadas a motor diesel de 2009 a 2018



Fonte: O autor (2019)

Para avaliar uma segunda opção de prazo na predição da falha no motor, foram testados modelos com predição em até 45 dias. Segundo as equipes da Oficina de Locomotivas, caso estes modelos apresentassem melhores resultados que os modelos de predição em até 30 dias, eles poderiam ser aceitos e utilizados pelas equipes no planejamento das manutenções. É importante ressaltar que não houveram bons resultados nas tentativas de predição em períodos superiores a 45 dias, e por isso esse foi o período máximo escolhido para os testes finais.

Quanto a Figura 7, é notável que a quantidade de falhas relacionadas a motor diesel decresceu expressivamente de 2014 para 2015 em diante. A explicação para esse fato envolve inúmeros fatores, entre eles as melhorias realizadas nas equipes de manutenção e nos equipamentos utilizados nas Oficinas. Porém a relação da análise de óleo com a falha ainda persiste, e desde 2009 essas relações trouxeram ganhos no planejamento das manutenções dos motores, no que envolve identificar desgastes no motor por meio das interpretações das análises de óleo lubrificante.

Todos os dados apresentados neste capítulo foram selecionados em conjunto com a experiência das equipes de manutenção dos motores diesel, e a escolha de cada um deles foi feita com base na relação com o problema e com a finalidade desta dissertação, sendo estes os dados que irão compor o método de predição. Para identificar os dados que serão utilizados foi preciso observar a qualidade, a confiabilidade e a correlação dos mesmos com o problema, sendo que nem sempre

um dado com uma menor qualidade foi descartado, pois existe a possibilidade de tratar *outliers* identificados.

Nessa dissertação a fase de tratamento dos dados foi em parte manual, como por exemplo a exclusão de dados não necessários das planilhas recebidas, e em outra parte foi necessário o uso do *software* de predição IBM SPSS Modeler 18.1, para auxiliar no tratamento e na tentativa de melhorar a qualidade dos dados. Ainda que a maioria dos dados estejam armazenados em sistemas informatizados criados dentro da própria EFVM, e sejam computados nestes sistemas manualmente pelas equipes da ferrovia, foi identificada uma boa qualidade dos mesmos, visto que o número de registros (linhas das planilhas de dados) nulos é baixo, assim como o número de espúrios (valores muito diferentes dos habituais daquele dado). Desta forma, entende-se que a confiabilidade dos dados também é considerada aceitável para compor essa dissertação.

Na fase de tratamento dos dados, além da exclusão de dados não necessários foi preciso padronizar as nomenclaturas, pois diversas colunas possuíam seus dados em formatos diferentes, por exemplo, a identificação da locomotiva envolvia letras e números em alguns casos, e em outros apenas números. Adequar esses dados é fundamental para que o modelo reconheça os dados, e por isso foi utilizado o IBM SPSS Modeler 18.1 e suas funções para que essa adequação fosse realizada.

É importante ressaltar que uma análise de correlação foi feita entre todas as variáveis com a falha do motor diesel, e entre as próprias variáveis. A correlação das variáveis com a falha no motor diesel deve ser alta para que o algoritmo de predição compreenda a relação e realize a predição. O contrário ocorre com a correlação entre as próprias variáveis. Se duas variáveis apresentarem alta correlação o indicado é que se utilize apenas uma delas na predição. A análise de correlação contribui com a fase de tratamento dos dados, podendo diminuir o tempo de processamento e melhorar os resultados obtidos. A Tabela 11 apresenta o nível de correlação encontrado entre as variáveis das bases de dados utilizadas com a predição de falha no motor diesel.

Tabela 11 – Correlações das variáveis com a predição de falha no motor diesel

Variável	Nível de correlação	Variável	Nível de correlação
Destilação	-0.003	Ni	-0.010
Viscosidade	0.036	Pb	-0.002
TBN	-0.012	Si	-0.007
Diluição	0.004	Sn	0.017
Ag	-0.021	Zn	-0.003
Al	-0.009	Mg	-0.020
Cr	-0.016	Mo	0.019
Cu	-0.003	Mn	-0.034
Fe	-0.004	Ca	0.043
Na	0.030	Litros	0.027

Fonte: O autor (2019)

A Tabela 12 apresenta o nível de correlação entre as próprias variáveis obtidas. Foi selecionado para apresentação apenas as variáveis que apresentaram fortes correlações, as quais foram levadas às equipes da Oficina de Locomotivas para uma análise detalhada de sua utilização na predição.

Tabela 12 – Variáveis que apresentaram altas correlações entre si

Variável	Variável correlacionada	Nível correlação
TBN	Viscosidade	-0.758
Fe	Viscosidade	0.557
Fe	TBN	-0.663
Cr	Ag	0.546
Mn	Ag	0.590
Fe	Cr	0.597
Pb	Cr	0.558
Mn	Cr	0.725
Pb	Fe	0.535

Fonte: O autor (2019)

A correlação criada e apresentada nas Tabelas 11 e 12 foi desenvolvida no IBM SPSS *Modeler* e é denominada correlação de Pearson. Na estatística descritiva ela mede o grau de correlação entre duas variáveis, assumindo valores entre 1 e -1. Quanto mais próximo de 1 mais a variável possui uma correlação positiva com a outra variável. Valores negativos tendendo a -1, indicam que quando a variável aumenta a outra sempre diminui, e vice-versa. E valores próximos de 0 indicam que duas variáveis não dependem de forma linear uma da outra, porém pode existir uma dependência não linear entre elas e o resultado deve ser analisado (IBM, 2019).

Foi determinado que variáveis muito correlacionadas devem ser analisadas junto aos mecânicos e laboratoristas da Oficina de Locomotivas, para que fosse escolhida apenas uma das variáveis para predição. A análise deve ser feita em conjunto com as equipes, pois a variável pode ser muito importante para detecção de falhas nos motores, e mesmo muito correlacionada a outra variável não deve ser descartada. O mesmo acontece com as variáveis que não possuem muita correlação com a falha no motor diesel.

Nesse caso, as variáveis que apresentam elementos químicos, como já mencionado, são obtidas de um único aparelho de análise, o espectrômetro. Dessa forma, excluir uma ou outra variável não faria diferença no processo de análise, e não contribui de forma expressiva com a diminuição do tempo de processamento. Já os ensaios físico-químicos, feitos individualmente em aparelhos distintos podem ser analisados, e dependendo do que as equipes avaliarem, podem ser descartados como variáveis de predição.

No caso das correlações entre as variáveis e a falha no motor diesel, a destilação, diluição, Cu, Fe, Pb e Zn apresentaram baixa correlação com a falha, porém as equipes da Oficina indicaram que as mesmas são importantes na identificação de falhas, e solicitaram que fossem feitos testes, com e sem elas. Nos casos das correlações entre as variáveis foi selecionado apenas os valores acima de 0.5 como fortes. A viscosidade e o TBN, por exemplo, apresentaram alta correlação entre si, e uma delas poderia ser descartada do algoritmo de predição, porém as equipes da Oficina também solicitaram testes com e sem uma das duas variáveis, para que fosse escolhida a configuração que apresentasse o melhor resultado de predição.

Após serem feitos os testes, assim como será explicado no Capítulo 5, foi observado que apenas a creptação, quando retirada do grupo das variáveis de predição indicava melhores resultados. A creptação indica qualitativamente a presença de água no óleo, assim como a destilação, porém a última apresenta a quantidade de água presente no óleo, em um formato numérico. Assim como já foi explicado no levantamento dos dados, como elas representam a mesma coisa, só que de formas diferentes, elas são altamente correlacionadas e uma delas poderia ser desprezada, como foi comprovado nos testes.

Desta forma, após realização dos testes, observou-se que os resultados não variaram expressivamente com a eliminação de alguns dos ensaios, com e sem os que apresentaram correlações especiais. Ao final das análises apenas os dados dos ensaios de insolúveis e creptação não foram utilizados para compor o modelo de predição. No caso dos dados do ensaio de insolúveis, como já foi mencionado, o índice de *outliers* ultrapassou o limite estabelecido como aceitável, e no caso da creptação a sua alta correlação com o ensaio de destilação fez com que ele não fosse necessário.

Realizando a união das cinco bases de dados e efetuando todo o tratamento necessário dos dados chegou-se a um total de 898.553 registros para compor o modelo de predição. Porém, registros livres de valores nulos e espúrios não é garantia de que os dados sejam representados por uma distribuição normal padrão (gaussiana), ou seja, com média nula e desvio-padrão unitário. Quando os dados se encontram normalizados, eles podem ser facilmente comparados entre si, sem a influência das diferentes escalas dos valores de cada dado. Desta forma, foi realizada uma normalização, utilizando a técnica *z-score,* em todos os dados utilizados no método de predição, por meio do IBM SPSS Modeler.

A técnica *z-score* é uma variação de escala que representa o número de desvios-padrão (a partir) da média. Desta forma, é garantido que as distribuições terão média igual a zero e desvio-padrão igual a 1, após aplicação da técnica. A escolha da técnica z-core geralmente acontece quando existem valores discrepantes, mas não tão extremos que seja necessário eliminá-los. Como a preparação dos dados foi feita antes de aplicar a normalização, essa técnica se mostrou pertinente. Outras técnicas de normalização poderiam ter sido utilizadas, como a transformação logarítmica, que calcula o log dos valores em casos nos quais algumas variáveis possuem muitas classificações e outras menos, padronizando o tamanho das faixas de classificação. No caso dessa dissertação o próprio IBM SPSS Modeler seleciona o método, e não deixa opções de escolha para o desenvolvedor.

Outro aspecto que deve ser observado nos dados de predição é a existência, ou não, de equilíbrio entre as quantidades de casos nos quais houve falhas com a quantidade de casos nos quais não houveram falhas nos motores diesel. Entende-se que se as

quantidades estiverem desbalanceadas, ou seja, o total de casos onde não houveram falhas for muito maior que o total de casos onde houveram falhas, ou vice-versa, o método de predição poderá ser influenciado pelo caso em maior quantidade. Nesta dissertação, o histórico de dados final apresentou 37,9 vezes menos casos nos quais houveram falhas em comparação com os casos nos quais não houveram falhas. Desta forma, foi necessário realizar um balanceamento dos dados, sendo necessário, primeiro, identificar qual a melhor técnica de balanceamento para esta situação.

Entre as técnicas de balanceamento mais utilizadas, tem-se balanceamento por *over-sampling,* como a técnica *Synthetic Minority Over-sampling Technique* (SMOTE) e balanceamento por *under-sampling.* A técnica *over-sampling* duplica resultados de dados pertencentes a uma classe minoritária, sub-representada na amostra, para uma quantidade mais equilibrada em relação a classe majoritária (classe com maior quantidade de dados). Essa técnica é indicada para casos onde a quantidade de registros no total, não é muito expressiva, evitando, assim, a perda de registros importantes que poderiam ser eliminados no uso da segunda técnica, *under-sampling*, que remove amostras da classe majoritária. Em contrapartida o método *over-sampling* consume mais tempo no processamento, pois aumenta a quantidade de dados (IBM, 2019). Independente das indicações, é interessante utilizar as duas técnicas e avaliar quais obtiveram melhores resultados de predição. Neste caso, utilizando a técnica simples do *over-sampling* abteve-se melhores resultados, e por isso foi a técnica escolhida e utilizada.

Além da normalização e balanceamento, após união das bases de dados em uma tabela única, estes dados foram correlacionados, na tentativa de indicar ao método de predição possíveis relações da análise de óleo com as falhas nos motores diesel. Desta forma, foram criadas novas variáveis a partir das relações entre os valores das análises de óleo e das OM, utilizando: 1) estatística descritiva; 2) faixas limites dos ensaios de óleo lubrificante; 3) variações diárias dos valores de análises de óleo lubrificante; e 4) cálculos envolvendo as datas das falhas nos motores e as datas das coletas de óleo lubrificante.

Com a estatística descritiva (1) foi possível criar novas colunas com valores como a média, mínimo, máximo e desvio padrão dos dados das colunas referentes à análise

de óleo lubrificante. As faixas limites dos ensaios de óleo lubrificante (2) classificaram seus valores em valores aceitáveis, alertas e críticos. Se o valor do ensaio estiver dentro dos limites estabelecidos nos manuais de análise de óleo, ele é considerado aceitável e o valor que aparecerá na nova coluna será igual a 1, se ele for próximo do valor aceitável será 2, caso contrário será 3. Um exemplo é o ensaio que detecta a concentração do elemento cobre no óleo, sendo o valor aceitável dessa concentração uma faixa em torno de 0 a 20 ppm. Caso os valores estejam na faixa entre 21 e 30 ppm, são considerados alertas. Valores que não estão nestas duas faixas, são considerados críticos.

As variações diárias dos valores das análises de óleo lubrificante (3) para uma mesma locomotiva, dizem respeito a diferença entre os valores de cada análise de óleo da tabela (valores de uma coleta de óleo atual menos os valores de uma coleta anterior). Essa nova coluna criada possibilita identificar tendências, como por exemplo, se ocorreram falhas nos motores nas épocas em que as variações entre as coletas foram maiores. Por fim, os cálculos envolvendo as datas das falhas e as datas de coleta do óleo lubrificante (4) são realizados para determinar o tempo entre estas datas, ou seja, a relação direta entre a data da falha no motor e os resultados das análises de óleo imediatamente antes dessa falha, o que pode contribuir com o entendimento da razão da ocorrência da falha.

Após realizar testes de predição com essas novas variáveis, em específico com as três primeiras apresentadas acima (estatística descritiva, faixas limites de análise de óleo lubrificante e variação diária) não foi identificado melhores resultados, em comparação com a utilização das variáveis iniciais das bases de dados apresentadas. Desta forma, as variáveis escolhidas para utilização no modelo de predição não contaram com as novas colunas (variáveis) criadas. Porém a quarta variável criada (cálculo entre a data da falha e a data da coleta) foi necessária para criação do resultado do modelo, e por isso foi utilizada.

É importante entender como as variáveis utilizadas foram relacionadas, para que o modelo de predição compreendesse as relações das análises de óleo, do consumo de diesel e das trocas de motores entre as locomotivas com a falha no motor diesel. Como a análise é a falha no motor, então todas os períodos de coleta de óleo,

consumo de diesel e o próprio histórico de falhas dos motores, foram vinculados em intervalos coincidentes com os períodos em que os motores estavam em determinada locomotiva. Ou seja, entre falhas ocorridas em um determinado motor houveram análises de óleo, a cada 15 dias aproximadamente.

É possível que entre duas falhas de um determinando motor tenha sido feita uma única análise do óleo lubrificante utilizado nele, ou duas, ou três e assim por diante, dependendo da “distância” em dias entre as falhas. O mesmo é pensado para o consumo de diesel da locomotiva, que varia de acordo com a sua utilização e foi atribuído nos períodos entre falhas. Com essa lógica é possível indicar para o modelo de predição em seu treinamento, quais foram os valores das análises de óleo e o quanto aquela locomotiva foi utilizada antes de haver a próxima falha no motor diesel instalado nela.

Como existe uma grande variação na quantidade de análises de óleo entre falhas entre locomotivas, por exemplo: uma determinada locomotiva possui 20 análises de óleo entre as primeiras duas falhas ocorridas no motor diesel. Porém, uma segunda locomotiva possui apenas 7 análises de óleo entre as duas primeiras falhas ocorridas no motor diesel. Além da variação entre locomotivas, existe também variações em uma mesma locomotiva e seus intervalos de falhas.

Analisando os dados e ouvindo aqueles que de forma habitual manipulam esses dados na Oficina de Locomotivas, foi estabelecido que instâncias de até 6 análises de óleo entre falhas seria uma quantidade razoável de análises de óleo para que o modelo de predição compreendesse as relações com as falhas e as tendências. Desta forma, foram mantidas as 6 últimas análises de óleo mais próximas da falha mais atual, e por meio da criação das instâncias, e eliminação das análises de óleo das instâncias 7, em diante, a quantidade de registros caiu para 45.469, compondo a base final de dados a serem destinados ao modelo de predição. O Quadro 4 mostra a lógica das instâncias e dos períodos em que o motor diesel esteve em determinada locomotiva.

Quadro 4 – Exemplo da tabela de base de dados limitados em 6 instâncias

Dias até a proxima falha	43	32	157	.	.	.
Litros abastecidos (acumulados)	344.202.600	344.202.600	1.766.148.000	.	.	.
Data de abastecimento	01/03/2009	01/03/2009	01/04/2009	.	.	.
Instâncias	6	5	4	3	2	1
Data de saída do motor da locomotiva	15/03/2009	15/03/2009	28/08/2013	.	.	.
Data de entrada do motor na locomotiva	20/01/2005	20/01/2005	15/03/2009	.	.	.
Código do motor	288.503	288.503	296.843	.	.	.
(demais análises de óleo)	...	...	...	...	...	...
Viscosidade	15.032	16.012	14.674	.	.	.
Data da coleta	03/01/2009	07/02/2009	20/03/2009	.	.	.
Data da falha anterior	15/12/2008	15/12/2008	11/03/2009	.	.	.
Tipo da ordem	YCM	YCM	YPM	.	.	.
Data da falha	11/03/2009	11/03/2009	24/08/2009	.	.	.
Modelo da locomotiva	DASH-8	DASH-8	DASH-8	.	.	.
Código da locomotiva	10	10	10	.	.	.

Fonte: O autor (2019)

5 REDE NEURAL ARTIFICIAL DE CLASSIFICAÇÃO PARA PREDIÇÃO DE FALHAS EM MOTORES DIESEL DE LOCOMOTIVAS COM BASE NA ANÁLISE DO ÓLEO LUBRIFICANTE

Com ênfase nos estudos encontrados envolvendo o uso de RNA para predizer problemáticas nas mais diversas áreas, que utilizam, nos casos de predição supervisionada, as redes *Perceptron* de múltiplas camadas (PMC) e o algoritmo de treinamento *Backpropagation*, e o grande êxito que obtiveram em seus resultados, esta dissertação utilizará a mesma estrutura em sua RNA. É importante destacar que a semente de aleatoriedade foi fixada em ambos os algoritmos apresentados nessa dissertação.

O software IBM SPSS *Modeler* 1.8 foi o software escolhido para simulação da RNA, ele oferece uma análise estatística avançada e uma vasta biblioteca de algoritmos que podem fornecer resultados de predição com apresentações em diferentes formatos (IBM, 2019).

Como foi indicado no Capítulo 4, instâncias de até 6 análises de óleo entre falhas foram criadas e testes foram feitos utilizando os dados sem a criação das instâncias e com a criação das instâncias de acordo com as configurações que serão apresentadas a seguir. Por meio dos resultados encontrados foi observado que a utilização das instâncias contribuiu com melhores resultados além de diminuir o tempo de processamento. Desta forma, todos os testes apresentados nesse capítulo e nos próximos serão feitos com a base de dados dividida em instâncias de até 6 análises de óleo entre falhas.

Como o objetivo da dissertação é predizer se haverá falha ou não no motor diesel de locomotiva em um período determinado, as configurações das RNA propostas foram analisadas e testadas em dois diferentes períodos: 1) se a falha ocorrerá ou não, em até 30 dias após a última coleta de óleo; 2) se a falha ocorrerá ou não, em até 45 dias após a última coleta de óleo.

A arquitetura escolhida para a RNA contou com 3 camadas de neurônios. A camada de entrada, camada intermediária e a camada de saída. A primeira camada, denominada camada de entrada, foi composta por 20 neurônios, sendo eles:

1)Destilação, 2)Diluição, 3)Viscosidade, 4)TBN, 5)Ag, 6)Al, 7)Cr, 8)Cu, 9)Fe, 10)Na, 11)Ni, 12)Pb, 13)Si, 14)Sn, 15)Zn, 16)Mg, 17)Mo, 18)Mn, 19)Ca, 20)Litros abastecidos.

Para a camada intermediária, é válido que sejam testadas diferentes configurações. Desta forma, foi variada a quantidade de camadas intermediárias e a quantidade de neurônios nestas camadas. O objetivo da diversificação dos números de camadas e neurônios intermediários é obter um resultado que descreva uma maior aproximação com as indicações reais de falha, ou seja, uma melhor predição. O software utilizado para teste das RNA é limitado em no máximo duas camadas intermediárias, por isso este foi o número utilizado. Foram testados um número maior que 200 neurônios nas duas camadas intermediárias, porém os resultados encontrados não apresentaram grandes melhorias em comparação com as demais configurações indicadas, além de aumentar o tempo de processamento da RNA.

Como foi explicado no Capítulo 2.1, a quantidade de neurônios nas camadas intermediárias pode ser definida com base no número de neurônios na camada de entrada e saída, e dessa forma foram selecionadas as 2 camadas intermediárias alternando o número de neurônios em: 5, 10, 13, 15, 17, 20, 25, 30, 35, 40, 50, 70, 100 e 200 neurônios. No total tem-se 24 configurações, ou seja, para cada uma das duas camadas intermediárias tem-se 14 variações de neurônios. Foi realizado o treinamento e validação dessas configurações para o período de 30 dias e foram selecionadas as de melhor resultado com base no k-*fold* médio de cada uma delas. Como foi observado na literatura, é mais provável que melhores resultados estejam entre 10 e 40 neurônios em ambas as camadas intermediárias.

Os melhores resultados encontrados resultantes da fase de validação foram das configurações usando 13, 35 e 40 neurônios nas camadas intermediárias. Desta forma, 12 configurações de RNA serão apresentadas, sendo 2 variações de camadas intermediárias, 3 variações de quantidade de neurônios nessas duas camadas intermediárias, e 2 períodos de análise, 30 e 45 dias.

Quanto à camada de saída, independente dos períodos escolhidos, possui apenas um neurônio, *sim* e *não*, ou seja, se indicar *sim* significa que haverá falha no motor diesel e se indicar *não*, significa que não haverá falha no motor diesel. Esse tipo de

análise é característica de casos de classificação binária, ou seja, na qual existe apenas duas possibilidades de respostas.

Para que fique clara as configurações de camadas e neurônios geradas para utilização na RNA, foram criadas as Tabelas 13 e 14 nas quais é possível identificar as configurações criadas para cada período de predição de falha no motor diesel (30 e 45 dias). Observa-se que as configurações são idênticas para os dois períodos, porém é necessário nomear cada configuração com um número para facilitar na análise dos resultados no Capítulo 6.

Tabela 13 – Configurações da RNA para falha em até 30 dias

Configuração	Quantidade de neurônios na camada de entrada	Quantidade de camadas intermediárias	Quantidade de neurônios nas camadas intermediárias	Quantidade de neurônios na camada de saída
1	20	1	13	1
2		2	13	
3	20	1	35	1
4		2	35	
5	20	1	40	1
6		2	40	

Fonte: O autor (2019)

Tabela 14– Configurações da RNA para falha em até 45 dias

Configuração	Quantidade de neurônios na camada de entrada	Quantidade de camadas intermediárias	Quantidade de neurônios nas camadas intermediárias	Quantidade de neurônios na camada de saída
7	20	1	13	1
8		2	13	
9	20	1	35	1
10		2	35	
11	20	1	40	1
12		2	40	

Fonte: O autor (2019)

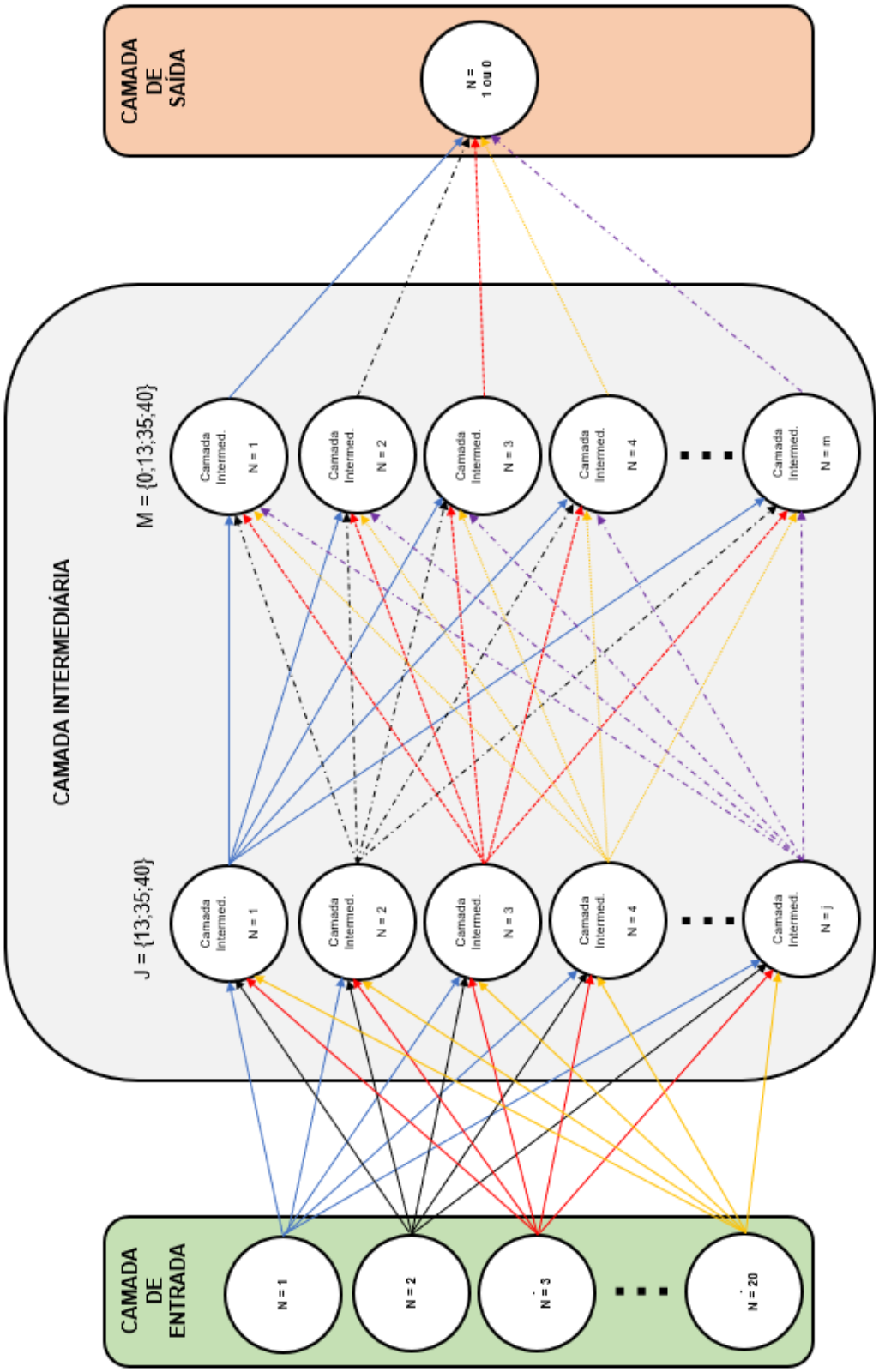
Para utilização dos dados nas RNA, foram separados 20% dos dados para testes, e os outros 80% dos dados foram separados entre dados de treinamento e validação por meio da validação cruzada k-*fold*, sendo f igual a 10. Dessa forma, os dados foram divididos em 10 partes iguais, onde cada uma das 10 partes foi usada para realizar treinamentos e validações. Como métricas de avaliação dos resultados foi utilizada a acurácia e o *recall*, portanto para cada parcela destinada a treinamento e validação foi obtido um valor de acurácia e de *recall*.

Como na validação cruzada k-*fold* foi usado *f*=10, então o total de registros usados no treinamento e validação da RNA deveria ser múltiplo de *k*, para que a divisão por 10 retratasse nos resultados uma melhor performance. Desta forma, do total de 45.469 registros, 40.220 registros foram usados para treinamento e validação das configurações apresentadas, e o restante de 5.249 registros foi usado na realização dos testes.

A escolha das melhores configurações da RNA para 30 e 45 dias de predição de falha, e posterior utilização na fase de testes, se deu inicialmente pela análise da acurácia e do *recall*. Porém é válido lembrar que não são os melhores resultados encontrados que foram selecionados e suas configurações usadas na fase de testes. As configurações escolhidas foram as que apresentaram valores médios de acurácia e *recall* entre todas as divisões do k-*fold*, pois são essas as configurações que apresentam os melhores pesos para generalização da RNA.

Definidas as camadas de entrada, intermediária e de saída da RNA, e a quantidade de seus respectivos neurônios, foi possível realizar as simulações. A Figura 8 apresenta as RNA que serão simuladas para predizer a falha no motor diesel das locomotivas. É importante ressaltar que o valor zero no conjunto M significa que quando a RNA possuir apenas uma camada intermediária, as demais camadas não existirão.

Figura 8 – RNA investigadas para predição de falhas em motores diesel de locomotivas



Fonte: O autor (2019)

6 ÁRVORE XGBOOST DE CLASSIFICAÇÃO PARA PREDIÇÃO DE FALHAS EM MOTORES DIESEL DE LOCOMOTIVAS COM BASE NA ANÁLISE DO ÓLEO LUBRIFICANTE

Para comparar os resultados encontrados com a RNA, foi escolhido o método de classificação Árvore XGBoost. O mesmo conjunto de dados usados no treinamento, validação e teste da RNA, foi utilizado na Árvore XGBoost para que os resultados pudessem ser comparados e os métodos avaliados.

O software IBM SPSS *Modeler* 1.8 foi o software utilizado para simulação da Árvore XGBoost, assim como foi usado nas simulações das RNA, e com o mesmo objetivo das RNA, o de predizer a falha no motor diesel de locomotivas em até 30 e 45 dias.

Árvore XGBoost é um modelo de árvore de decisão que melhora o desempenho do Gradient Boosting, incorporando mais estratégias de regularização para reduzir a complexidade do modelo e o excesso de ajustes. Como foi explicado na revisão da literatura, existem alguns parâmetros que devem ser estipulados para o desenvolvimento da Árvore XGBoost. O Quadro 5 apresenta os parâmetros para os três processos conhecidos da Árvore XGBoost: construção do modelo, crescimento da árvore e controles de super ajustes.

Quadro 5 – Parâmetros definidos para o desenvolvimento da Árvore XGBoost

Construção do modelo		**Crescimento da árvore**		**Controle de super ajustes**			
Método da árvore:	auto	Profundidade máxima:	3 e 6	Sub-amostragem:	1	Alfa:	0
Rodada de número boost:	10	Peso mínimo:	1	Eta:	0,3	Lambda:	1

Fonte: O autor (2020)

Todos os valores utilizados nestes parâmetros foram definidos com base na revisão da literatura apresentada no Subcapítulo 2.1.2, demais estudos envolvendo Árvore XGBoost e com base no objetivo dessa dissertação. O método da árvore é *auto*, pois o próprio SPSS *Modeler* seleciona o melhor meio de se construir a árvore de acordo com o tamanho da base de dados utilizada no modelo, neste caso entre pequena e média. A *rodada de número boost,* como mencionado na revisão da literatura tem como valor padrão 10, e é esse o valor utilizado na maioria dos estudos encontrados. A profundidade máxima, a qual determina a quantidade de divisões que a árvore

poderá estabelecer, pode variar de 0 a 6, sendo o valor padrão igual a 6, porém alguns estudos testam valores menores, por isso nesta dissertação será testado também o valor 3. O valor padrão do peso mínimo é 1, e esse será o valor adotado para essa Árvore. Como a base de dados utilizada é entre pequena e média, a *sub-amostragem* foi 1, ou seja, foi utilizada toda a base de dados. Os valores de *eta*, *alfa* e *lambda* foram conservadores, mantidos semelhantes às literaturas encontradas.

Todos os tratamentos realizados nas bases de dados, incluindo a normalização e balanceamento *over-sampling* e a validação cruzada k-*fold,* feita no conjunto de dados para utilização na RNA, também foram utilizados na simulação da Árvore XGBoost. Sendo que na validação cruzada k-*fold*, o k foi igual a 10 e toda a separação da base de dados e o procedimento dessa separação foi idêntica àquela feita para a simulação da RNA. Ao final tem-se para cada configuração 10 resultados de acurácia e *recall* de acordo com as duas variações de profundidade máxima, 6 e 3. Sendo que as duas configurações serão simuladas para 30 e 45 dias de predição de falhas, como pode ser observado na Tabela 15.

Tabela 15 – Configurações da Árvore XGBoost para falha em até 30 dias e 45 dias

Configuração	Período de análise da falha	Profundidade máxima
13	Em até 30 dias	3
14		6
15	Em até 45 dias	3
16		6

Fonte: O autor (2019)

Assim como para a RNA, para as duas configurações da Árvore XGBoost foram usados 40.220 registros para treinamento e validação e 5.249 registros para testes.

7 RESULTADOS EXPERIMENTAIS

O capítulo 7 tem por objetivo apresentar os resultados encontrados e os testes realizados na predição de falhas dos motores diesel de locomotivas utilizando RNA e Árvore XGBoost, bem como uma comparação entre os dois métodos de classificação.

7.1 RESULTADOS E ANÁLISES DA REDE NEURAL ARTIFICIAL (RNA)

Com as variáveis definidas e tratadas e os parâmetros da RNA e da Ávore XGBoost selecionados, foi iniciado o processamento dos modelos de predição em um computador com dois processadores Intel® Xeon® CPU E5-2640 v3 2.60GHz e 64GB de memória RAM.

Para as RNA e suas configurações, foram utilizados 40.220 registros para treinamento e validação. Para cada uma das configurações foi obtido um valor de acurácia e de *recall*. Todos os resultados encontrados serão apresentados nas Tabelas 16 e 17 de acordo com as configurações de 1 a 12 já demonstradas, sendo que de 1 a 6 referem-se aos resultados para predição de falha em até 30 dias e de 7 a 12 para 45 dias.

Essas tabelas incluem além dos números das configurações, a quantidade de camadas intermediárias, a quantidade de neurônios nessas camadas e os resultados de menor, maior e média acurácia, além de menor, maior e médio *recall* na etapa de validação. O tempo de processamento das configurações que envolveram as acurácias médias na validação também foi apresentado nas Tabelas. Observando as Tabelas conclui-se que as configurações 2, 4 e 6 apresentaram os melhores desempenhos para uma predição em até 30 dias, analisando as três variações de neurônios nas camadas intermediárias.

Tabela 16 – Resultados de validação das RNA para predição de falha em até 30 dias

Configuração	Quantidade de neurônios na camada de entrada	Quantidade de camadas intermediárias	Quantidade de neurônios nas camadas intermediárias	Maior acurácia	Menor acurácia	Média das acurácias	Maior *recall*	Menor *recall*	Média *recall*	Tempo de processamento das acurácias médias (min)
1	20	1	13	79,89%	72,97%	77,12%	0,88	0,82	0,87	0,23
2		2	13	82,21%	77,32%	79,20%	0,95	0,87	0,90	0,43
3	20	1	35	86,52%	76,75%	81,31%	0,95	0,87	0,90	1,28
4		2	35	92,54%	78,20%	87,73%	0,98	0,87	0,97	1,97
5	20	1	40	85,13%	79,09%	81,39%	0,94	0,88	0,91	2,03
6		2	40	92,07%	82,39%	87,27%	0,99	0,90	0,96	2,13

Fonte: O autor (2019)

Tabela 17 – Resultados da validação das RNA para predição de falha em até 45 dias

Configuração	Quantidade de neurônios na camada de entrada	Quantidade de camadas intermediárias	Quantidade de neurônios nas camadas intermediárias	Maior acurácia	Menor acurácia	Média das acurácias	Maior recall	Menor recall	Média recall	Tempo de processamento das acurácias médias (min)
7	20	1	13	74,79%	70,12%	72,85%	0,83	0,76	0,81	0,20
8		2	13	75,80%	73,57%	74,76%	0,86	0,81	0,83	0,51
9	20	1	35	81,58%	72,21%	76,12%	0,87	0,79	0,81	1,36
10		2	35	79,20%	74,68%	77,44%	0,87	0,80	0,84	1,75
11	20	1	40	77,19%	72,07%	75,13%	0,84	0,79	0,82	2,53
12		2	40	78,97%	74,14%	76,92%	0,84	0,79	0,84	3,10

Fonte: O autor (2019)

Os resultados com a predição em até 45 dias não foram satisfatórios como os da predição em até 30 dias, e por isso não serão utilizados na fase de testes.

Analisando os resultados da validação das configurações 2, 4 e 6, observa-se que o maior valor de acurácia, 92,54% pertence a configuração 4, seguido da acurácia de 92,07% da configuração 6 e por fim a de 82,21% da configuração 2. Porém quando observado o maior valor de *recall*, percebe-se que a configuração 6 apresentou um valor maior, de 0,99, que a configuração 4, a qual apresentou o valor de 0,98 e que a configuração 2, de 0,95. Mesmo diante destes resultados é preciso lembrar que as configurações que devem ser identificadas para utilização na fase de testes são as que apresentaram os maiores valores médios de acurácia e *recall*. Dessa forma, foi observado que as mesmas configurações, 4, 6 e 2 apresentaram os melhores valores médios de acurácia e *recall* respectivamente para 35, 40 e 13 neurônios nas camadas intermediárias.

Os maiores valores médios de acurácia são os das configurações 4 e 6, respectivamente 87,73% e 87,27%, seguido por último pela configuração 2 com 79,20%. O *recall* segue a mesma indicação, 0,97 para configuração 4, 0,96 para configuração 6 e 0,90 para configuração 2. Quanto ao tempo de processamento da fase de validação destas três configurações, observa-se que a configuração 2 levou menos tempo para gerar seus resultados, em torno de 0,40 minutos, seguida da configuração 4 que levou 1,97 minutos, e da configuração 6 que levou 3,33 minutos para gerar os seus resultados.

Comparando as três configurações, 2, 4 e 6, é notório que a 4 apresentou os melhores resultados médios de acurácia e *recall*. Porém as três configurações serão utilizadas na fase de testes para obtenção de maiores oportunidades de avaliação dos resultados. Desta forma, nos testes foram utilizados os pesos salvos em arquivos

referentes aos neurônios das camadas intermediárias das configurações 2, 4 e 6. Os resultados encontrados na fase de testes utilizando os pesos referentes aos valores médios de acurácia e *recall* das configurações 2, 4 e 6 poderão ser observados no Subcapítulo 7.2.

7.2 TESTE DA REDE NEURAL ARTIFICIAL (RNA)

Para teste da RNA foram utilizadas as configurações 2, 4 e 6, respectivamente para predição de falha no motor diesel em até 30 dias, e seus pesos obtidos nas etapas de treinamento e validação do k-*fold* que apresentou valores médios de acurácia e *recall*, contemplando um total de 5.249 registros para a predição. É importante ressaltar que os dados utilizados nessa etapa de teste ainda não foram apresentados para a RNA nas etapas anteriores e os resultados serão representados pela acurácia e pelo *recall.* A Tabela 18 apresentará as configurações usadas na etapa de teste e seus resultados, bem como os resultados encontrados na etapa de validação dessas configurações para que seja possível comparar os desempenhos.

Tabela 18 – Resultados das configurações 2, 4 e 6 na etapa de validação e teste

Configuração	Acurácia (validação)	Acurácia (teste)	*Recall* (validação)	*Recall* (teste)
2	79,20%	66,83%	0,90	0,80
4	87,73%	77,82%	0,97	0,57
6	87,27%	76,59%	0,96	0,66

Fonte: O autor (2019)

É possível observar que os resultados encontrados na fase de testes não foram melhores que os resultados encontrados nas etapas de validação, o que pode ser considerado normal nas RNA, visto que a cada etapa realizada, treinamento, validação e teste os resultados tendem a decrescer em desempenho. Porém, ainda que a acurácia não tenha apresentado valores acima de 80% em todos os casos de teste, o *recall* da configuração 2 apresentou valor de 0,80, o que é considerado satisfatório e poderá ser melhor compreendido na análise das matrizes de confusão ao final deste Subcapítulo.

Para a predição da configuração 2 os resultados de teste foram de 66,83% para acurácia e 0,8 para *recall,* da configuração 4 foram de 77,82% de acurácia e 0,57 de *recall*, e por fim os resultados da fase de testes para a configuração 6 foi de 76,59% de acurácia e 0,66 de *recall*. Por meio dos resultados encontrados é possível considerar a capacidade da RNA de se generalizar para predizer falhas em motores

diesel de locomotivas com base na análise do óleo lubrificante, embora exista a oportunidade de se alcançar resultados melhores, principalmente pelo fato dos resultados de teste terem diminuído consideravelmente de performance em relação aos resultados de validação. É necessário um estudo aprofundado para entender as causas que levaram a essa redução de performance.

Para que se compreenda melhor os resultados apresentados, serão mostradas as matrizes de confusão da etapa de validação e de teste nas Figuras 9, 10 e 11 para as três configurações 2, 4 e 6 respectivamente, que buscam predizer falha em motor diesel de locomotiva em até 30 dias. Junto às porcentagens de cada quadrante da matriz de confusão, são apresentadas também as quantidades em número dos casos: verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN). Analisando os resultados na horizontal, tem-se os casos *observados*, ou seja, os casos reais (históricos de falhas), já analisando na vertical tem-se os casos *preditos*, ou seja, as indicações do modelo de predição, neste caso as RNA.

Figura 9 – Matriz de confusão da etapa de validação e teste da configuração 2 com a predição da falha em até 30 dias

		PREDITO (VALIDAÇÃO)	
		Não haverá falha	Haverá falha
OBSERVADO (VALIDAÇÃO)	Não houve falha	1.372 68,72%	624 31,28%
	Houve falha	212 10,44%	1.814 89,56%

		PREDITO (TESTE)	
		Não haverá falha	Haverá falha
OBSERVADO (TESTE)	Não houve falha	3.308 66,16%	1.692 33,84%
	Houve falha	49 19,68%	200 80,32%

Fonte: O autor (2020)

Figura 10 – Matriz de confusão da etapa de validação e teste da configuração 4 com a predição da falha em até 30 dias

		PREDITO (VALIDAÇÃO)	
		Não haverá falha	Haverá falha
OBSERVADO (VALIDAÇÃO)	Não houve falha	1.573 78,79%	423 21,21%
	Houve falha	69 3,39%	1.957 96,61%

		PREDITO (TESTE)	
		Não haverá falha	Haverá falha
OBSERVADO (TESTE)	Não houve falha	3.942 78,84%	1.058 21,16%
	Houve falha	106 42,57%	143 57,43%

Fonte: O autor (2020)

Figura 11 – Matriz de confusão da etapa de validação e teste da configuração 6 com a predição da falha em até 30 dias

		PREDITO (VALIDAÇÃO)	
		Não haverá falha	Haverá falha
OBSERVADO (VALIDAÇÃO)	Não houve falha	1.572 78,77%	424 21,23%
	Houve falha	83 4,11%	1943 95,89%

		PREDITO (TESTE)	
		Não haverá falha	Haverá falha
OBSERVADO (TESTE)	Não houve falha	3.856 77,12%	1.144 22,88%
	Houve falha	85 34,14%	164 65,86%

Fonte: O autor (2020)

Assim como foi explicado no Subcapítulo 2.1, a matriz de confusão possibilita analisar os acertos e erros do modelo, afim de se obter um entendimento mais preciso em direção ao objetivo do estudo. Os acertos do modelo estão apresentados nos quadrantes mais claros, e os erros nos quadrantes mais escuros das Figuras 11, 12 e 13.

É notório que os acertos do modelo nas fases de validação são superiores aos das fases de testes, principalmente pelo fato de que na fase de testes os dados que até então não haviam sido apresentados para o modelo, são apresentados sem balanceamento, diferente dos dados usados nas fases de treinamento e validação, balanceados por meio de um *over-sampling* simples, ou seja, criação de cópias da

classe minoritária pelo próprio SPSS *Modeler*, dessa forma os dados que são usados para teste são exatamente como os dados que serão imputados periodicamente no modelo, sem utilização de métodos de balanceamento.

É possível observar que a matriz de confusão que obteve menores diferenças comparando as fases de validação e testes foi a da 2ª configuração, apresentada na Figura 12, na qual a RNA é composta pela camada de entrada, de saída, e por duas camadas intermediárias, cada uma delas com 13 neurônios. É também dessa configuração o melhor resultado esperado de *recall* (0,8), na fase de testes, já que ela apresenta o menor valor para o quadrante de falso negativo (FN). Ainda que a acurácia da 2ª configuração seja a de menor valor entre as 3 configurações analisadas, sua matriz de confusão mostra resultados mais próximos do objetivo da dissertação que as demais configurações.

A Tabela 19 apresenta de forma comparativa os valores já observados nas matrizes de confusão da etapa de teste das três configurações, 2, 4 e 6.

Tabela 19 – Resultados das matrizes de confusão das configurações 2, 4 e 6 na etapa de teste

Configuração	Verdadeiro Positivo (teste)	Falso Positivo (teste)	Falso Negativo (teste)	Verdadeiro Positivo (teste) (%)	Falso Positivo (teste) (%)	Falso Negativo (teste) (%)
2	200	1.692	49	80,32%	33,84%	19,68%
4	143	1.058	106	57,43%	21,16%	42,57%
6	164	1.144	85	65,86%	22,88%	34,14%

Fonte: O autor (2020)

Os casos de verdadeiro positivo (VP) são apresentados para mostrar os acertos do modelo, ou seja, quanto maiores os valores, mais assertivo o modelo estará sendo. Significa que se houve uma falha, o modelo indicou que ela ocorreria. Quanto às colunas de falso negativo e falso positivo, representam os erros do modelo, ou seja, quanto menores os valores, mais assertivo estará sendo o modelo. Quando o motor diesel não apresenta falhas, mas ainda assim, o modelo indica uma falha nesse motor, é um erro do modelo denominado falso positivo (FP), nesse caso valores muito altos podem comprometer a utilização do modelo, já que o usuário do modelo terá que reter locomotivas sem falhas para inspeção do motor.

No caso dessa dissertação, as locomotivas utilizadas no estudo passam periodicamente pela Oficina para serem inspecionadas e abastecidas. Dessa forma, as equipes de manutenção deixaram claro que valores de falso positivo abaixo de

30% não seriam problema, caso o modelo conseguisse minimizar o pior dos casos para as equipes, os falsos negativos (FN). Os FN representam os casos nos quais os motores falham, muitas vezes durante o tráfego, mas o modelo não indica previamente as falhas. Esses casos geram transtornos envolvendo o não cumprimento dos contratos de transporte ferroviário, atrasos nos prazos de entrega das mercadorias, congestionamento de trens, entre outros problemas.

O que pode ser percebido em todas as 300 matrizes de confusão geradas, para cada k-*fold* de treinamento e validação, e também nas matrizes geradas nas fases de testes é que sempre vai haver um *tradeoff* entre os valores de FP e FN. Geralmente quando um dos valores aumenta o outro diminui e vice-versa. Dessa forma, buscando-se atender o usuário do modelo, é preciso minimizar ambos os erros, FP e FN, com preferência pelo menor FN. Observando as configurações 2, 4 e 6, a 2 é a que apresenta o menor valor de FN, e por isso maior valor de *recall* na fase de testes. Desta forma, a configuração 2 foi selecionada para ser comparada à configuração de melhor desempenho da Árvore XGBoost no subcapítulo 7.5.

7.3 RESULTADOS E ANÁLISES DA ÁRVORE XGBOOST

Para treinamento e validação da Árvore XGBoost foram utilizados os parâmetros e configurações apresentados no Capítulo 6. As configurações 13 e 14 destinadas a apresentar as predições em até 30 dias, e as configurações 15 e 16 para as predições em até 45 dias. O conjunto de dados utilizado foi o mesmo usado na criação das RNA, totalizando 40.220 registros para as fases de treinamento e validação para ambos os casos de predição, 30 e 45 dias. As métricas de análise dos resultados, para cada configuração, continuam sendo a acurácia e o *recall*.

Para uso na etapa de testes foram salvos os pesos estabelecidos nas configurações de 13 a 16. Porém os pesos utilizados de fato na fase de testes, foram selecionados entre as 10 partes do k-*fold* da fase de treinamento e validação que apresentaram valores médios de acurácia e *recall*. Na Tabela 20 serão apresentados os resultados das configurações de 13 a 16, na qual é possível observar as colunas que descrevem as configurações, os períodos de análise da falha (30 e 45 dias), os valores selecionados para a profundidade máxima da Árvore XGBoost, os resultados das

acurácias e *recall* e o tempo de processamento dos casos das acurácias e *recall* médios, de cada configuração na etapa de validação.

Tabela 20 – Resultados das configurações 13, 14, 15 e 16 na etapa de validação

Configuração	Períodos de análise da falha	Profundidade Máxima	Maior acurácia	Menor acurácia	Média das acurácias	Maior *recall*	Menor *recall*	Média *recall*	Tempo de processamento para acurácias médias (min)
13	30 dias	3	76,66%	74,37%	75,61%	0,96	0,93	0,95	0,12
14		6	85,48%	83,31%	84,43%	0,97	0,95	0,97	0,37
15	45 dias	3	74,79%	72,92%	73,49%	0,91	0,86	0,88	0,17
16		6	82,47%	80,83%	81,66%	0,92	0,90	0,91	0,42

Fonte: O autor (2020)

Observado a Tabela 20 é possível identificar qual das configurações apresentou os melhores resultados para acurácia e *recall*. Nota-se que independente do período de análise da falha, 30 ou 45 dias, os melhores resultados de acurácia foram das configurações nas quais a profundidade máxima testada foi igual a 6. No caso do *recall* houveram variações, mas ainda assim os melhores resultados, principalmente quanto a média do *recall* foi utilizando a profundidade máxima igual a 6 na configuração 14.

Para a fase de testes foram utilizados os pesos salvos em arquivos referentes aos das configurações 13 e 14, respectivamente as configurações que apresentaram melhores resultados na predição da falha em motor diesel em até 30 dias. Como as configurações 15 e 16, envolvendo a predição em até 45 dias obtiveram resultados inferiores aos das predições em até 30 dias, elas não seguiram para as fases de testes. Os resultados encontrados na fase de testes utilizando os pesos referentes aos valores médios de acurácia e *recall* das Configurações 13 e 14 poderão ser observados no subcapítulo 7.4.

7.4 TESTE DA ÁRVORE XGBOOST

Para a etapa de testes da Árvore XGBoost foi utilizado o mesmo conjunto de dados que foi utilizado nos testes da RNA, totalizando 5.249 registros para predição de falha no motor diesel em até 30 e 45 dias. Os pesos utilizados foram os das configurações 13 e 14 salvos das etapas de treinamento e validação da Árvore XGBoost. É importante ressaltar que os dados utilizados nessa etapa de teste ainda não foram apresentados para da Árvore XGBoost nas etapas anteriores e os resultados serão representados pela acurácia e pelo *recall.* A Tabela 21 apresentará as configurações

com as respectivas acurácias e *recall* nas fases de validação e teste das Árvores XGBoost.

Tabela 21 – Resultados das configurações da Árvore XGBoost nas etapas de validação e teste

Configuração	Acurácia (validação)	Acurácia (teste)	*Recall* (validação)	*Recall* (teste)
13	75,61%	55,95%	0,95	0,87
14	84,43%	71,33%	0,97	0,79

Fonte: O autor (2020)

É possível observar, como já era esperado, que os resultados encontrados na fase de testes não foram melhores que os resultados encontrados nas etapas de validação. Ao analisar as acurácias na fase de testes percebe-se que nenhuma delas obteve resultado superior a 80%, sendo a maior delas a da configuração 14, de 71,33%. Quanto ao *recall* percebe-se resultados mais satisfatórios, em torno da faixa de 79 a 87%.

Vale ressaltar que o melhor resultado de *recall,* da configuração 13, vem acompanhado do pior resultado de acurácia, o que poderá ser melhor compreendido por meio das matrizes de confusão das configurações 13 e 14 nas Figuras 12 e 13 abaixo. Os resultados da configuração 14 apresentaram valores mais equilibrados de acurácia e *recall*, 71,33% e 0,79 respectivamente, e ainda que não seja uma acurácia acima de 80%, pode-se considerar a capacidade da Árvore XGBoost de se generalizar para predizer falhas em motores diesel de locomotivas com base na análise do óleo lubrificante.

Figura 12 – Matriz de confusão da etapa de validação e teste da configuração 13 com a predição da falha em até 30 dias

		PREDITO (VALIDAÇÃO)	
		Não haverá falha	Haverá falha
OBSERVADO (VALIDAÇÃO)	Não houve falha	1.090 54,62%	906 45,38%
	Houve falha	98 4,83%	1.928 95,17%

		PREDITO (TESTE)	
		Não haverá falha	Haverá falha
OBSERVADO (TESTE)	Não houve falha	2.727 54,54%	2.273 45,46%
	Houve falha	33 13,42%	216 86,58%

Fonte: O autor (2020)

Figura 13– Matriz de confusão da etapa de validação e teste da configuração 14 com a predição da falha em até 30 dias

		PREDITO (VALIDAÇÃO)	
		Não haverá falha	Haverá falha
OBSERVADO (VALIDAÇÃO)	Não houve falha	1.447 72,50%	549 27,5%
	Houve falha	70 3,46%	1.956 96,54%

		PREDITO (TESTE)	
		Não haverá falha	Haverá falha
OBSERVADO (TESTE)	Não houve falha	3.550 71,00%	1.450 29,00%
	Houve falha	53 21,29%	196 78,71%

Fonte: O autor (2020)

Sabendo-se que os acertos do modelo estão apresentados nos quadrantes mais claros, e os erros nos quadrantes mais escuros das Figuras 12 e 13, e que analisando horizontalmente a matriz, tem-se na linha superior os casos *observados* (reais) onde não houve falha, e na linha inferior os casos *observados* onde houve falha, percebe-se que o modelo desempenhou menos entre os valores da linha superior, ou seja, dos casos onde não houve falha na realidade, mas de forma incorreta o modelo indicou que houveram as falhas, o que é denominado de falso positivo (FP). Na linha inferior, percebe-se um melhor desempenho do modelo, buscando minimizar o falso negativo (FN).

A escolha do melhor resultado entre as configurações partiu do entendimento de negócio obtido nas 54 reuniões realizadas com as equipes que realizam as manutenções e planos de manutenções das locomotivas e seus motores diesel. O modelo deve atender o usuário de forma que na prática ele seja útil. Neste caso, a configuração 14 apresentou uma matriz de confusão mais equilibrada, onde os erros do modelo FP e FN, respectivamente 29% e 21,29% seriam aceitos pelos usuários.

É válido observar que na configuração 13 seria possível predizer 20 casos de falhas a mais que na configuração 14, o que pode ser considerado muito vantajoso ás equipes de manutenção, porém o número de locomotivas que seriam inspecionadas devido aos casos de falsos positivos (FP) faz com que a opção de melhor configuração continue com a configuração 14. A Tabela 22 apresenta de forma comparativa os

valores já observados nas matrizes de confusão da etapa de teste das duas configurações 13 e 14.

Tabela 22 – Resultados das matrizes de confusão das configurações 13 e 14 na etapa de teste

Configuração	Verdadeiro Positivo (teste)	Falso Positivo (teste)	Falso Negativo (teste)	Verdadeiro Positivo (teste) (%)	Falso Positivo (teste) (%)	Falso Negativo (teste) (%)
13	216	2.273	33	86,58%	45,46%	13,42%
14	196	1.450	53	78,71%	29,00%	21,29%

Fonte: O autor (2020)

Sabe-se que quanto maior o valor de verdadeiro positivo (VP) melhor a assertividade do modelo, neste caso a configuração 13 é superior a configuração 14. O contrário ocorre com os valores de falso positivo (FP) e negativo (FN), que precisam ser minimizados para que o modelo tenha um bom desempenho. Percebe-se valores menores de falso positivo (FP) para a configuração 14 em relação a configuração 13, porém o mesmo não ocorre com os valores de falso negativo (FN). Essa dificuldade de comparar as configurações devido as perdas e ganhos entre os quadrantes das matrizes de confusão é o que torna o entendimento do negócio e suas variáveis tão importante no desenvolvimento da dissertação.

Após analisar as matrizes de confusão e os valores de acurácia e *recall* de ambas as configurações, 13 e 14, resultantes da Árvore XGBoost, selecionou-se como configuração de melhor desempenho a configuração 14. No Subcapítulo 7.5 será possível comparar a configuração 14 da Árvore XGBoost com a configuração escolhida como a de melhor desempenho pelo método da RNA.

7.5 COMPARAÇÃO ENTRE A REDE NEURAL ARTIFICIAL (RNA) E A ÁRVORE XGBOOST

Com base nos resultados apresentados das RNA e das Árvores XGBoost foi possível observar que, de maneira geral, não existiram grandes discrepâncias entre os métodos quando observados os valores encontrados de acurácia e *recall* da etapa de testes dos modelos, como pode ser identificado na Tabela 23. Desta forma é necessário analisar as matrizes de confusão de ambos os métodos de predição em conjunto com os resultados de acurácia e *recall* para identificação do método que obteve melhor desempenho na predição de falhas, lembrando que o *recall* continua sendo a métrica de maior peso na decisão. É importante lembrar que todos os

resultados apresentados nesse capítulo são referentes a predição em até 30 dias, já que as predições em até 45 dias não performaram para serem utilizadas na fases de testes.

Tabela 23 – Acurácia e *recall* das configurações 2, 4, 6, 13 e 14 da fase de testes

Método de predição	Configuração	Acurácia	*Recall*
RNA	2	66,83%	0,80
	4	77,82%	0,57
	6	76,59%	0,66
XGBoost	13	55,95%	0,87
	14	71,33%	0,79

Fonte: O autor (2020)

Após análise comparativa de todas as configurações que foram testadas nas RNA e nas Árvores XGBoost, será possível observar o comparativo apenas das configurações de melhor desempenho de ambos os métodos de predição, neste caso a configuração 2 e a configuração 14. A diferença da acurácia entre a Árvore XGBoost e a RNA foi de 4,5%, na qual a Árvore XGBoost obteve melhor desempenho. Já a diferença do *recall* entre a RNA e a Árvore XGBoost foi de 0,01, na qual a RNA obteve melhor desempenho. A Tabela 24 apresenta os resultados das métricas de avaliação para as configurações dos dois modelos de predição, bem como os valores das matrizes de confusão e o ganho, representado pela divisão entre os resultados da RNA pelos resultados da Árvore XGBoost.

Tabela 24 – Comparação entre as configurações que obtiveram melhores desempenhos da RNA e da Árvore XGBoost

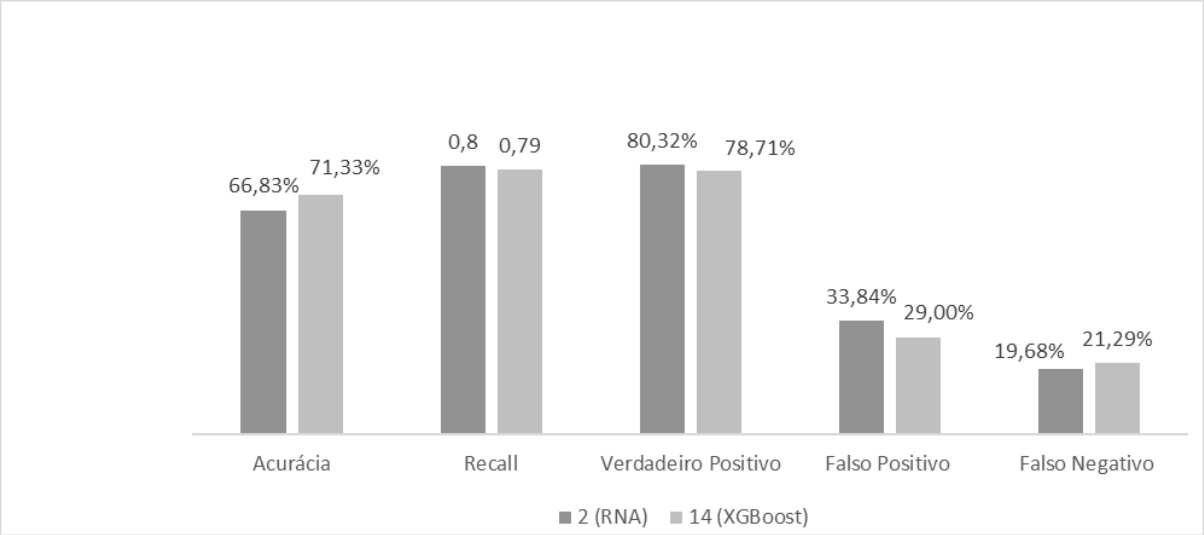
Configuração	Acurácia	*Recall*	Verdadeiro Positivo	Verdadeiro Positivo (%)	Falso Positivo	Falso Positivo (%)	Falso Negativo	Falso Negativo (%)
2 (RNA)	66,83%	0,80	200	80,32%	1.692	33,84%	49	19,68%
14 (XGBoost)	71,33%	0,79	196	78,71%	1.450	29,00%	53	21,29%
Ganho(%) (RNA/XGBoost)	-6,31%	1,26%	2,04%	2,04%	-16,7%	-16,7%	7,6%	7,6%

Fonte: O autor (2020)

Percebe-se que os valores positivos relacionados ao ganho não significam uma divisão positiva dos fatores, e sim que de fato a RNA foi superior à Árvore XGBoost em seus resultados. Da mesma forma, valores negativos de ganho significam que os resultados da Árvore XGBoost foram melhores que os resultados da RNA. Lembrando que no caso dos verdadeiros positivos quanto maior o valor, mais assertivo o modelo de predição, e nos casos dos falsos positivos e negativos, quanto menor o valor melhor o modelo desempenhou.

A Figura 14 apresenta um gráfico no qual é possível observar de forma mais clara as cinco variáveis usadas para comparar as configurações 2 e 14. Sabendo-se que o *recall* é a métrica que possui maior peso na decisão da escolha do melhor método, justamente por representar os valores de falso negativo, ou seja, os valores principais de minimização desta dissertação. Entende-se pela fórmula do *recall*, que quanto maior seu valor, melhor o modelo desempenhou, desta forma como a RNA possui um valor maior de *recall*, ela seria o método de melhor desempenho. Porém a diferença entre o *recall* da RNA e da Árvore XGBoost é de apenas 0,01 ou 1,26%, e por isso a acurácia e os valores de falso positivo (FP) também precisam ser levados em conta na análise.

Figura 14 – Comparação entre as configurações que obtiveram melhores desempenhos da RNA e da Árvore XGBoost



Fonte: O autor (2020)

A acurácia da configuração 14 (XGBoost) foi superior a acurácia da configuração 2 (RNA). Os valores de falso positivo (FP) da configuração 14 foram inferiores aos da configuração 2, ou seja, melhores em 16,7%. Entende-se dessa forma que a RNA apresentou melhores resultados de *recall* e falsos negativos (FN), e a Árvore XGBoost melhores resultados de acurácia e falso positivos.

Por meio dos resultados e análises realizadas nas configurações das RNA e Árvores XGboost, com o objetivo de se predizer falhas em motores diesel de locomotivas com base na análise do óleo lubrificante, pode-se afirmar que ambos os algoritmos podem ser utilizados como uma ferramenta de auxílio na realização dos planos de manutenção de motores diesel de locomotivas.

8 CONCLUSÕES

Esta dissertação teve como objetivo investigar dois algoritmos de aprendizado de máquina, a RNA e a Árvore XGBoost, na predição de falhas em motores diesel de locomotivas baseado na análise do óleo lubrificante usado nos motores, consumo de combustível e troca de motores entre locomotivas. Os dados foram adquiridos na Estrada de Ferro Vitória-Minas e passaram por tratamento como normalização, análise de correlação e balanceamento *over-sampling.*

Para a implementação das RNA foi utilizada a rede *Perceptron* de Múltiplas Camadas (PMC) e o algoritmo de treinamento *Backpropagation*. Diferentes arquiteturas de RNA foram simuladas, utilizando 20 neurônios na camada de entrada, variando as quantidades de neurônios em 1 e 2 camadas intermediárias, no objetivo de identificar se haveria falha, ou não no motor diesel, no único neurônio da camada de saída. Na Árvore XGBoost foi variada a profundidade máxima em 3 e 6 e utilizados outros 7 parâmetros fixados, 4 deles para controle de super ajustes. Em ambos os métodos de predição foi utilizada a validação cruzada k-fold, sendo *f* igual a 10.

Os resultados finais apresentaram como a RNA de melhor desempenho aquela com 2 camadas intermediárias e 13 neurônios em ambas as camadas intermediárias, a qual obteve uma acurácia de 66,83% e *recall* de 0,80 na fase de testes. Para a Árvore XGBoost o melhor desempenho foi na utilização da profundidade máxima igual a 6, obtendo uma acurácia de 71,33% e *recall* de 0,79 na fase de testes.

Comparando a RNA e a Árvore XGBoost que apresentaram os melhores desempenhos na fase de testes, conclui-se que a acurácia da Árvore XGBoost foi superior a acurácia da RNA em 6,31%. O *recall* da RNA foi superior ao da Árvore XGBoost em 1,26%. Levando em consideração a utilização do método de predição na prática dos planos de manutenção das locomotivas e seus motores, entende-se que a Árvore XGBoost poderá auxiliar melhor as decisões de manutenção, embora a diferença tenha sido muito pequena.

Por meio dos resultados e análises realizadas nas configurações das RNA e Árvores XGboost, com o objetivo de se predizer falhas em motores diesel de locomotivas com base na análise do óleo lubrificante, pode-se afirmar que ambos os algoritmos podem

ser utilizados como uma ferramenta de auxílio na realização dos planos de manutenção de motores diesel de locomotivas.

8.1 TRABALHOS FUTUROS

Como trabalhos futuros propõem-se predizer não só se vai haver a falha no motor diesel, como também qual componente do motor está relacionado à falha predita. Propõem-se também comparar os resultados obtidos da RNA e da Árvore XGBoost com outros métodos como: One-Class SVM, k-NN, e com um método de regressão. Como métrica de avaliação dos resultados, sugere-se utilizar a Curva ROC e demais métricas que contribuam com um melhor desempenho e análise do resultado, assim como a utilização de algum algoritmo para ajuste dos hiperparâmetros.

REFERÊNCIAS

ALIBAKSHI, Amin. Strategies to develop robust neural network models: Prediction of flash point as a case study. **Analytica chimica acta**, v. 1026, p. 69-76, 2018.

ADENIYI, David Adedayo; WEI, Zhaoqiang; YONGQUAN, Y. Automated web usage data mining and recommendation system using K-Nearest Neighbor (K-NN) classification method. **Applied Computing and Informatics**, v. 12, n. 1, p. 90-108, 2016.

AGÊNCIA NACIONAL DE TRANSPORTE TERRESTRE. **Anuário Estatístico 2018**. 2018. Disponível em: <http://www.antt.gov.br/ferrovias/arquivos/Anuario_Estatisticohtml.>. Acesso em: 14 janeiro 2019.

ANGUITA, Davide et al. K-Fold Cross Validation for Error Rate Estimate in Support Vector Machines. In: DMIN. 2009. p. 291-297.

ARLOT, Sylvain et al. A survey of cross-validation procedures for model selection. Statistics surveys, v. 4, p. 40-79, 2010.

BIAO, Liu et al. System identification of locomotive diesel engines with autoregressive neural network. In: **Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on**. IEEE, 2009. p. 3417-3421.

BORBA, J. L.; ZUCOLOTTI, C. **Lubrificantes aplicados a locomotivas**. Trilha Técnica VALER/GADFT-CEL/VALE S/A. Vitória, 2010.

CANECA, Arnobio Roberto et al. **Assessment of infrared spectroscopy and multivariate techniques for monitoring the service condition of diesel-engine lubricating oils**. Talanta, v. 70, n. 2, p. 344-352, 2006.

CARMO, Daniel Augusto do. **Detecção e diagnóstico de falhas em sistema de arrefecimento de motor Diesel ferroviário.** 2018. Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Espírito Santo, Vitória, 2018.

CUNHA, C. M. P. et al. Flexibilização da manutenção programada baseada na análise de desempenho: (1ª parte) – Projeto. In: VI SEMINÁRIO ESPÍRITO-SANTENSE DE MANUTENÇÃO. Vitória, 1994.

CUNHA, Rodrigo Carvalho. **Redutor de velocidade através da técnica de partículas de desgaste no óleo lubrificante auxiliada pela análise de vibrações**. 2005. Dissertação (mestrado em Engenharia Mecânica) – Programa de Pós-Graduação em Engenharia Mecânica, Universidade Estadual Paulista, Ilha Solteira, 2005.

DE PISÓN, Francisco Javier Martínez et al. (Ed.). **Hybrid Artificial Intelligent Systems: 12th International Conference, HAIS 2017, La Rioja, Spain, June 21-23, 2017, Proceedings**. Springer, 2017.

DUDA, Richard O.; HART, Peter E.; STORK, David G. Pattern classification. John Wiley & Sons, 2012.

ELKHOURY, Najwa; HITIHAMILLAGE, Lalith; MORIDPOUR, Sara; ROBERT, Dilan Degradation Prediction of Rail Tracks: A Review of the Existing Literature, **The Open Transportation Journal**, v.12, n. 1, p.88-104, fev. 2018.

FACELI, Katti et al. **Inteligência Artificial: Uma abordagem de aprendizado de máquina**. 2011.

FONTE, M. et al. Failure mode analysis of a diesel motor crankshaft. **Engineering Failure Analysis**, v. 82, p. 681-686, 2017.

Google Earth. **Oficina de Locomotivas de Tubarão**. Disponível em: <https://earth.google.com/web/>. Acesso em: 18 fevereiro 2019.

GOU, Jianping et al. A new distance-weighted k-nearest neighbor classifier. **J. Inf. Comput. Sci**, v. 9, n. 6, p. 1429-1436, 2012.

HAWKINS, Douglas M. The problem of overfitting. **Journal of chemical information and computer sciences**, v. 44, n. 1, p. 1-12, 2004.

HSU, Chih-Wei et al. A practical guide to support vector classification. 2003.

HUANG, Hao et al. Maximum F1-score discriminative training criterion for automatic mispronunciation detection. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 23, n. 4, p. 787-797, 2015.

IBM. IBM SPSS Modeler 18.1. 2018. Disponível em: <https://www.ibm.com/br-pt/products/spss-modeler>. Acesso em: 20 Agosto 2018.

JARDINE, A. K. S. et al. Proportional hazards analysis of diesel engine failure data. **Quality and Reliability Engineering International**, v. 5, n. 3, p. 207-216, 1989.

JÚNIOR, Rubião Gomes; MACHADO, Maria Augusta Soares; BARRETO, Jorge Muniz. Previsão de falhas em manutenção industrial usando redes neurais. In: III SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA. Resende, 2006.

KIMURA, Rogério Katsuharu. **Uso da técnica de análise de óleo lubrificante em motores Diesel estacionários, utilizando-se misturas de biodiesel e diferentes níveis de contaminação do lubrificante**. 2010. 128 f. Dissertação (mestrado em Engenharia Mecânica) – Programa de Pós-Graduação em Engenharia Mecânica, Universidade Estadual Paulista, Ilha Solteira, 2010.

MANCUZO, Mechelangelo Viana. **Detecção automática de falhas em motores de combustão interna utilizando rede neural**. 1999. Dissertação (mestrado em Engenharia Mecânica) – Programa de Pós-Graduação em Engenharia Mecânica, Universidade Federal de Uberlândia, Uberlândia, 1999.

MARCHETTI, Dalmo; WANKE, Peter. Brazil's rail freight transport: Efficiency analysis using two-stage DEA and cluster-driven public policies. **Socio-Economic Planning Sciences**, v. 59, p. 26-42, 2017.

MOLLENHAUER, Klaus; TSCHÖKE, Helmut (Ed.). Handbook of diesel engines. Springer Science & Business Media, 2010.

NEWELL, Gary E. Oil analysis cost-effective machine condition monitoring technique. **Industrial Lubrication and Tribology**, v. 51, n. 3, p. 119-124, 1999.

NIU, Xiaoxiao et al. Investigation of ANN and SVM based on limited samples for performance and emissions prediction of a CRDI-assisted marine diesel engine. **Applied Thermal Engineering**, v. 111, p. 1353-1364, 2017.

NOOR, CW Mohd et al. Application of artificial neural network for prediction of marine diesel engine performance. In: 3RD INTERNATIONAL CONFERENCE OF MECHANICAL ENGINEERING RESEARCH, 2015, Malaysia**. IOP Conference Series: Materials science and engineering.** Malaysia, 2015.

PARLAK, Adnan et al. Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a diesel engine. **Applied Thermal Engineering**, v. 26, n. 8-9, p. 824-828, 2006.

REFAEILZADEH, Payam; TANG, Lei; LIU, Huan. Cross-validation. **Encyclopedia of database systems**, p. 532-538, 2009.

RODRIGUEZ, Juan D.; PEREZ, Aritz; LOZANO, Jose A. Sensitivity analysis of k-fold cross validation in prediction error estimation. **IEEE transactions on pattern analysis and machine intelligence**, v. 32, n. 3, p. 569-575, 2009.

ROSA, Rodrigo de Alvarenga; RIBEIRO, Rômulo Castello Henrique. **Estradas de Ferro:** Projeto, Especificação & Construção. Vitória: Editora EDUFES, 2017.

REIF, Konrad (Ed.). **Diesel engine management**. Springer Vieweg, 2014.

SILVA, IN da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. Redes neurais artificiais para engenharia e ciências aplicadas. **São Paulo: Artliber**, 2010, vol. 23, no 5, p. 33-111.

SPECTRO INC. Oil Analysis helps keep the trains to running. Chelmsford. England, 2012.

SUDHA, P.; RAMYACHITRA, D.; MANIKANDAN, P. Enhanced artificial neural network for protein fold recognition and structural class prediction. **Gene Reports**, v. 12, p. 261-275, 2018.

TANUPABRUNGSUN, Sikana; ACHALAKUL, Tiranee. Feature reduction for anomaly detection in manufacturing with mapreduce GA/K-NN. In: **2013 International Conference on Parallel and Distributed Systems**. IEEE, 2013. p. 639-644.

TING, Kai Ming. Confusion matrix. **Encyclopedia of Machine Learning and Data Mining**, p. 260-260, 2017.

TORLAY, L. et al. Machine learning–XGBoost analysis of language networks to classify patients with epilepsy. **Brain informatics**, v. 4, n. 3, p. 159-169, 2017.

TOSUN, Erdi et al. Comparative analysis of various modelling techniques for emission prediction of diesel engine fueled by diesel fuel with nanoparticle additives. **European Mechanical Science**, v. 1, n. 1, p. 15-23, 2017.

VALIŠ, David; ŽÁK, Libor; POKORA, Ondřej. Failure prediction of diesel engine based on occurrence of selected wear particles in oil. **Engineering Failure Analysis**, v. 56, p. 501-511, 2015.

VAN BASSHUYSEN, Richard; SCHÄFER, Fred. **Internal combustion engine handbook-basics, components, systems and perspectives**. 2004.

WONG, Tzu-Tsung. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. **Pattern Recognition**, v. 48, n. 9, p. 2839-2846, 2015.

XIN, Qianfan. Diesel engine system design. Elsevier, 2011.

YADAV, Sanjay; SHUKLA, Sanyam. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In: **2016 IEEE 6th International Conference on Advanced Computing (IACC)**. IEEE, 2016. p. 78-83.

YANG, Fubin et al. Artificial neural network (ANN) based prediction and optimization of an organic Rankine cycle (ORC) for diesel engine waste heat recovery. **Energy conversion and management**, v. 164, p. 15-26, 2018.