



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES

ANTONIO LUCAS GABRIEL MARIANO

**USO DE APRENDIZADO DE MÁQUINA INTERPRETÁVEL PARA AVALIAÇÃO
DA DEFORMAÇÃO PERMANENTE EM MISTURAS ASFÁLTICAS**

FORTALEZA

2023

ANTONIO LUCAS GABRIEL MARIANO

USO DE APRENDIZADO DE MÁQUINA INTERPRETÁVEL PARA AVALIAÇÃO DA
DEFORMAÇÃO PERMANENTE EM MISTURAS ASFÁLTICAS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Transportes da Universidade Federal do Ceará, como requisito para obtenção do título de Mestre em Engenharia de Transportes. Área de concentração: Infraestrutura de Transportes.

Orientador: Prof. Jorge Barbosa Soares.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- M286u Mariano, Antonio Lucas Gabriel.
Uso de aprendizado de máquina interpretável para avaliação da deformação permanente em misturas asfálticas / Antonio Lucas Gabriel Mariano. – 2023.
103 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Transportes, Fortaleza, 2023.
Orientação: Prof. Dr. Jorge Barbosa Soares.
1. Aprendizado de máquina. 2. Deformação permanente. 3. Redes neurais . 4. XGBoost. 5. Shapley additive explanations (SHAP). I. Título.

CDD 388

USO DE APRENDIZADO DE MÁQUINA INTERPRETÁVEL PARA AVALIAÇÃO DA
DEFORMAÇÃO PERMANENTE EM MISTURAS ASFÁLTICAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Transportes da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Engenharia de Transportes. Área de concentração: Infraestrutura de Transportes.

Orientador: Prof. Jorge Barbosa Soares.

BANCA EXAMINADORA

Prof. Dr. Jorge Barbosa Soares (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Suelly Helena de Araújo Barroso
Universidade Federal do Ceará (UFC)

Prof. Dr. Iuri Sidney Bessa
Universidade Federal do Ceará (UFC)

Prof. Dr. César Lincoln Cavalcante Mattos
Universidade Federal do Ceará (UFC)

Prof. Dr. Francisco Thiago Sacramento Aragão
Universidade Federal do Rio de Janeiro (UFRJ)

AGRADECIMENTOS

Aos meus pais, Antonio e Liduina, por sempre terem me apoiado em minhas decisões e por me ensinar valores que me fizeram ser quem sou hoje.

Obrigado à professora Suelly Barroso por fazer parte da banca e por sempre ter se mostrado uma pessoa disponível para ajudar e aconselhar seus alunos.

Agradeço ao professor César Lincoln por ter me ajudado a usar a disciplina de aprendizado de máquina em prol da minha pesquisa. Sou grato pela sua enorme contribuição na minha pesquisa e por estar sempre disponível para tirar dúvidas e propor sugestões. Obrigado por comparecer à banca.

Agradeço aos professores Iuri Bessa e Thiago Aragão por prontamente terem aceitado a participar da banca.

Agradeço ao meu professor Jorge Soares pela orientação e pelo direcionamento para uma área de estudo que eu até então não cogitava, mas que depois me despertou interesse e curiosidade pelas possibilidades que ela pode oferecer. Sou grato pela confiança depositada em mim e pelas palavras de incentivo durante o mestrado. Durante essa caminhada, eu sou muito grato pela sua paciência, compreensão e pelos conselhos para além da vida acadêmica.

Gostaria de agradecer também ao Carlos David, cuja participação foi fundamental para o desenvolvimento do meu trabalho. Agradeço as sugestões e a disponibilidade.

Agradeço a todos os professores que fazem parte do PETRAN, em particular ao professor Uchoa que participou da minha qualificação e agregou ao trabalho com sugestões.

Sou grato também a todos, sem exceção, do Laboratório de Mecânica dos Pavimentos.

Agradeço em especial aos amigos de mestrado que deixaram essa caminhada mais leve, Andressa, Pedro Luiz, Thiago, Otávio, Sasaki e Assis.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES pela bolsa concedida para a realização desta pesquisa

RESUMO

A deformação permanente (DP) é um dos principais defeitos observados na estrutura de pavimentos asfálticos. O novo método de dimensionamento de pavimentos nacional, o MeDiNa, trata a deformação permanente da camada asfáltica por meio do *flow number* (FN), cujos valores são indicados para classificar as misturas em conformidade com as condições de solicitação de tráfego. Esse método trata a deformação permanente como um defeito a ser mitigado ainda na etapa de dosagem. Portanto, a existência de um modelo de previsão para o FN poderia ser de grande utilidade no sentido de nortear a metodologia de dosagem. O uso do aprendizado de máquina tem sido empregado em diversas áreas, incluindo a infraestrutura viária e se apresenta como uma alternativa para realizar previsões acerca do comportamento das misturas asfálticas em relação à deformação permanente. No entanto, os algoritmos de aprendizado de máquina são vistos como “caixa-preta”, devido a dificuldade em entender como eles realizam suas previsões. Em diversas aplicações, compreender a previsão do modelo pode ser tão relevante quanto à precisão dessa previsão. O presente trabalho propõe o desenvolvimento de uma modelagem para classificação de misturas asfálticas em relação à deformação permanente de acordo com o tipo de tráfego, utilizando a ferramenta de aprendizado de máquina explicável, *SHAP*ley Additive exPlanations (*SHAP*). Foram utilizadas as redes neurais artificiais (RNA) e também o *eXtreme Gradient Boosting* (*XGBoost*). O banco de dados construído possui informações de 251 misturas asfálticas. Foram propostas 3 configurações em relação as variáveis de entrada (C1, C2 e C3). O modelo apresentou um acurácia de 84% medida com o uso da validação cruzada na configuração C3. A análise com os valores *SHAP* mostrou como as variáveis consideradas influenciam o comportamento da mistura em relação à deformação permanente e, por conseguinte, da categoria de tráfego correspondente, possibilitando um melhor entendimento e consideração desse importante fenômeno.

Palavras-chave: aprendizado de máquina; deformação permanente; redes neurais; *XGBoost*; *shapley additive explanations* (*SHAP*).

ABSTRACT

Permanent deformation or rutting is one of the main distresses observed in asphalt pavements. The new Brazilian pavement design method, called MeDiNa, considers permanent deformation of the surface course through the flow number (FN), indicated to classify asphalt mixtures in accordance with the solicited traffic. This method treats permanent deformation as a defect to be mitigated during the mixture design stage. Therefore, the existence of a prediction model for FN could be of great use in guiding the design methodology. Several works have used machine learning (ML) in different areas, including road infrastructure, and the technique presents itself as an alternative to make predictions about the behavior of asphalt mixtures in relation to permanent deformation. Many machine learning systems are essentially considered “black boxes”, because of the difficulty in understanding how the code works. In several applications, understanding the model’s prediction can be as relevant as the accuracy of that prediction. The work presented herein proposed the development of a modeling for classification of asphalt mixtures in relation to permanent deformation according to the corresponding traffic, using the explainable machine learning tool, *SHAP*ley Additive exPlanations (*SHAP*). Artificial neural networks (ANN) and eXtreme Gradient Boosting (XGBoost) were used. The constructed database has information from 251 asphalt mixtures. Three configurations were proposed in relation to input variables (C1, C2 and C3). The model using XGBoost presented an accuracy of 84% for the predictions of the traffic class corresponding to the mixture using cross validation in configuration C3. Moreover, an analysis with the *SHAP* values has shown how the variables considered in the research affect the mixture behavior with respect to permanent deformation and consequently to the traffic to which the mixture is fit, therefore increasing the understanding and the way to treat this relevant distress for the asphalt pavement industry.

Keywords: machine learning, permanent deformation, neural networks, XGBoost, shapley additive explanations (SHAP).

LISTA DE FIGURAS

Figura 1 – Afundamento em trilha de roda.....	17
Figura 2 – Trilha de roda em pavimentos.....	20
Figura 3 – Correlação entre viscosidade rotacional e FN.....	22
Figura 4 – Contato entre grãos em um conjunto de partículas.....	23
Figura 5 – Correlação $\eta_{FAD} \times FN$	25
Figura 6 – Ensaio uniaxial de carga repetida.....	25
Figura 7 – Representação dos ciclos durante o ensaio.....	26
Figura 8 – Zonas da curva de deformação plástica e <i>flow number</i>	26
Figura 9 – Modelo neural McCulloch.....	29
Figura 10 – Modelo não linear de um neurônio.....	30
Figura 11 – Função linear.....	31
Figura 12 – Função sigmoide ou logística.....	32
Figura 13 – Função tangente hiperbólica.....	33
Figura 14 – Função ReLU.....	34
Figura 15 – Esquematização de uma rede perceptron multicamadas.....	34
Figura 16 – Ilustração de um problema não-linearmente separável.....	35
Figura 17 – Processo de aprendizado Classificador <i>XGBoost</i>	37
Figura 18 – Distribuição das misturas do banco de dados em relação ao tipo de ligante.....	45
Figura 19 – Fluxograma das etapas metodológicas deste trabalho.....	46
Figura 20 – Conversão da variável categórica em variáveis numéricas.....	47
Figura 21 – Esquematização da validação cruzada <i>k-fold</i>	54
Figura 22 – Diagramas de caixa dos valores de FN separados por tipo de ligante.....	56
Figura 23 – Histogramas das variáveis no conjunto de dados.....	58
Figura 24 – Matriz de correlação Pearson das variáveis do banco de dados.....	59
Figura 25 – Matriz de correlação Pearson das variáveis do ligante CAP.....	61
Figura 26 – Matriz de correlação Pearson das variáveis do ligante AMP.....	62
Figura 27 – Matriz de correlação Pearson das variáveis do ligante AB.....	63

Figura 28 – Matriz de correlação Pearson das variáveis do ligante AM	64
Figura 29 – Matriz de confusão para o modelo de RNA na configuração C1	65
Figura 30 – Matriz de confusão para o modelo de RNA na configuração C2	67
Figura 31 – Matriz de confusão para o modelo de RNA na configuração C3	68
Figura 32 – Matriz de confusão para o modelo <i>XGBoost</i> na configuração C1	70
Figura 33 – Relevância das variáveis obtido no modelo <i>XGBoost</i> usado na configuração C1	70
Figura 34 – Matriz de confusão para o modelo <i>XGBoost</i> na configuração C2	71
Figura 35 – Relevância das variáveis obtido no modelo <i>XGBoost</i> usado na configuração C2	72
Figura 36 – Matriz de confusão para o modelo <i>XGBoost</i> na configuração C3	73
Figura 37 – Relevância das variáveis obtido no modelo <i>XGBoost</i> usado na configuração C3	74
Figura 38 – Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (EP), definida como alvo	77
Figura 39 – Gráfico cachoeira relacionado à mistura descrita na Tabela 15	79
Figura 40 – Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (P), definida como alvo	80
Figura 41– Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (M), definida como alvo	80
Figura 42 – Interpretação global do modelo referente as misturas da classe (M)	81
Figura 43 – Gráfico da média absoluta dos valores <i>SHAP</i> para classe (M)	82
Figura 44 – Interpretação global do modelo referente as misturas da classe (P)	83
Figura 45 – Gráfico da média absoluta dos valores <i>SHAP</i> para classe (P)	84
Figura 46 – Interpretação global do modelo referente as misturas da classe (EP)	85
Figura 47 – Gráfico da média absoluta dos valores <i>SHAP</i> para classe (EP)	86
Figura 48 – Gráfico da soma das médias absolutas dos valores <i>SHAP</i> de todas as classes	86
Figura 49 – Comparação entre os resultados obtidos pelo <i>XGBoost</i> e pela biblioteca <i>SHAP</i> , referente a importância das variáveis	88

LISTA DE TABELAS

Tabela 1 – Critérios de FN para diferentes níveis de tráfego	27
Tabela 2 – Variáveis numéricas e categóricas de entrada do banco de dados desenvolvido ...	43
Tabela 3 – Configurações consideradas para as variáveis de entrada.....	48
Tabela 4 – Exemplo de normalização dos dados	49
Tabela 5 – Hiperparâmetros utilizados no modelo <i>XGBoost</i>	52
Tabela 6 – Hiperparâmetros utilizados na modelagem da RNA	53
Tabela 7 – Estatística descritiva do banco de dados	57
Tabela 8 – Tabela de desempenho do modelo de RNA para a configuração C1	67
Tabela 9 – Tabela de desempenho do modelo de RNA para a configuração C2	68
Tabela 10 – Tabela de desempenho do modelo de RNA para a configuração C3	69
Tabela 11 – Tabela de desempenho do modelo <i>XGBoost</i> para a configuração C1.....	70
Tabela 12 – Tabela de desempenho do modelo <i>XGBoost</i> para a configuração C2.....	72
Tabela 13 – Tabela de desempenho do modelo <i>XGBoost</i> para a configuração C2.....	73
Tabela 14 – Comparação entre os desempenhos dos modelos para as 3 configurações.....	75
Tabela 15 – Valores das variáveis da mistura utilizada no gráfico de força	77

LISTA DE SIGLAS E ABREVIACOES

AB	Asfalto Borracha
ABNT	Associao Brasileira de Normas Tcnicas
AMP	Asfalto Modificado por Polmero
AM	Asfalto de alto Mdulo
ATR	Afundamento de Trilha de Roda
CAP	Cimento Asfltico de Petrleo
CNT	Confederao Nacional do Transporte
CP	Corpo de Prova
CI	Componentes Intersticiais
DP	Deformao Permanente
DNIT	Departamento Nacional de Infraestrutura de Transportes
E*	Mdulo dinmico das misturas asflticas
FAD	Faixa de Agregados Dominantes
FN	<i>Flow number</i>
FNe	Falso negativo
FP	Falso positivo
G*	Mdulo de Cisalhamento dos Ligantes
MLP	<i>Multilayer Perceptron</i>
NBR	Norma Brasileira
P34	Percentual passante na peneira 3/4”
P38	Percentual passante na peneira 3/8”
P4	Percentual passante na peneira #4
P200	Percentual passante na peneira #200
PA	Ponto de Amolecimento
PEN	Penetrao
PG	<i>Performance grade</i>
RNA	Redes Neurais Artificiais
RNC	Rede Neural Convolutiva
<i>ReLU</i>	<i>Rectified Linear Unit</i>
TL	Teor de ligante

TMN	Tamanho Máximo Nominal
VIS 135	Viscosidade <i>Brookfield 135°C</i>
VN	Verdadeiro negativo
VP	Verdadeiro positivo
Vv%	Volume de vazios
<i>XGBoost</i>	<i>eXtreme Gradient Boosting</i>

SUMÁRIO

1 INTRODUÇÃO	16
1.1 Considerações iniciais.....	16
1.2 Contextualização	16
1.3 Problema de pesquisa	18
1.4 Objetivos.....	19
1.4.1 <i>Objetivo Geral</i>	19
1.4.2 <i>Objetivos Específicos</i>	19
2 REVISÃO BIBLIOGRÁFICA	20
2.1 Deformação permanente em pavimentos asfálticos	20
2.2 Impacto do ligante na deformação permanente	21
2.3 Influência do agregado na deformação permanente	23
2.4 Ensaio uniaxial de carga repetida.....	25
2.5 Modelo de Francken	28
2.6 Aprendizado de máquina	28
2.7 Redes Neurais Artificiais (RNAs)	29
2.7.1 <i>Funções de ativação</i>	31
2.7.1.1 <i>Função linear</i>	31
2.7.1.2 <i>Função sigmoide ou logística</i>	31
2.7.1.3 <i>Função tangente hiperbólica</i>	32
2.7.1.4 <i>Função ReLU</i>	33
2.7.2 <i>Perceptron Multicamadas</i>	34
2.7.3 <i>Algoritmo de retropropagação do erro (Backpropagation)</i>	35
2.8 XGBoost (eXtreme Gradient Boosting)	35
2.9 Relevância do atributo	37
2.10 Interpretabilidade em aprendizado de máquina	37
2.11 Uso de IA em infraestrutura de transportes	39
3 MATERIAIS E MÉTODOS	42
3.1 Materiais.....	42
3.2 Método empregado	46
3.2.1 <i>Coleta de Dados</i>	46

3.2.2 Pré-processamento e seleção das variáveis.....	47
3.2.3 Separação banco de dados, Treino e Teste.....	48
3.2.4 Normalização das variáveis.....	49
3.2.5 SMOTE (Técnica de Super Amostragem Sintética da Minoria).....	50
3.2.6 Modelagem Via RNA e XGBOOST.....	50
3.2.6.1 Ajuste de hiperparâmetros do XGBOOST.....	50
3.2.6.2 Ajuste de hiperparâmetros da RNA.....	52
3.2.7 Validação cruzada k-fold.....	53
3.2.8 Análise dos resultados.....	54
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS	56
4.1 Análise Descritiva dos Dados	56
4.1.1 Relações Estatísticas entre as Variáveis.....	59
4.1.1.1 Relações estatísticas para todo o conjunto de dados.....	59
4.1.1.2 Relações estatísticas para ligante puro (CAP).....	60
4.1.1.3 Relações estatísticas para asfalto modificado por polímero.....	61
4.1.1.4 Relações estatísticas para o asfalto borracha.....	62
4.1.1.5 Relações estatísticas para o ligante de alto módulo.....	63
4.2 Desempenho da Rede Neural Artificial	64
4.2.1 Modelo de RNA para a configuração C1.....	65
4.2.2 Modelo de RNA para a configuração C2.....	67
4.2.3 Modelo de RNA para o modelo C3.....	68
4.3 Desempenho do XGBoost	69
4.3.1 Modelo utilizando o XGBoost para a configuração C1.....	69
4.3.2 Modelo utilizando o XGBoost para a configuração C2.....	71
4.3.3 Modelo utilizando o XGBoost para a configuração C3.....	72
4.4 Comparação entre os modelos	74
4.5 Gráficos SHAP.....	76
4.5.1 Interpretação local – gráficos de força e gráfico cachoeira.....	76
4.5.2 Interpretabilidade global.....	80
4.6 Comparação entre a relevância dos atributos e os valores SHAP.....	87
5 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	89
5.1 Principal Contribuição da Pesquisa	89
5.2 Principais Limitações da Pesquisa	89

5.3 Principais Constatações da Pesquisa.....	90
5.4 Sugestões para pesquisas futuras.....	92
REFERÊNCIAS	93
APÊNDICE A – Banco de dados das 251 misturas utilizadas na pesquisa	99

1 INTRODUÇÃO

1.1 Considerações iniciais

O primeiro capítulo faz uma breve contextualização referente ao atual panorama das rodovias brasileiras. Trata ainda sobre conceitos introdutórios relacionados à deformação permanente (DP), as consequências desse defeito e o ensaio utilizado para a avaliação da mistura asfáltica em relação a esse fenômeno. Em seguida, é apresentada uma sugestão do uso do aprendizado de máquina interpretável como forma de realizar previsões acerca do comportamento da mistura asfáltica e analisar o impacto das variáveis dessa mistura. Por fim, são apresentados o objetivo geral e os objetivos específicos.

1.2 Contextualização

No Brasil, as rodovias são as principais responsáveis pelos deslocamentos de pessoas e de cargas. Segundo a pesquisa da Confederação Nacional do Transporte (CNT) de 2022, 95% do tráfego de pessoas e mais de 61% do transporte de cargas são realizados através das rodovias. Há de se destacar o intenso crescimento na frota de veículos, 57,1% no período de 2011 a 2021, segundo o Mapa da Motorização Individual no Brasil de 2021.

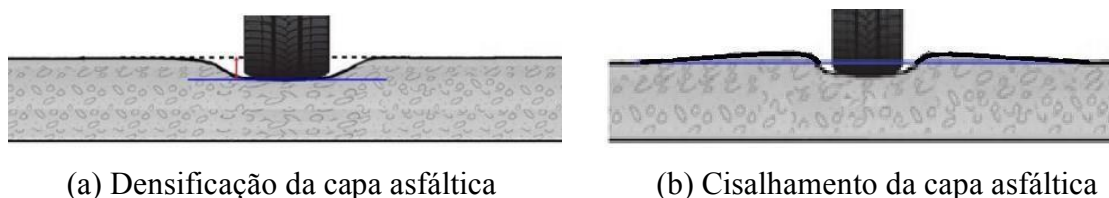
A infraestrutura brasileira não tem acompanhado o crescimento da frota veicular, sendo que o referido levantamento da CNT aponta um aumento de 0,5% da malha pavimentada. Outro dado a ser considerado é a densidade da malha viária nacional que é de 25km a cada 1000m², considerada baixa em comparação com países como: China (447km/1000m²) e Estados Unidos (437km/1000m²). Em relação a América Latina, esse número está abaixo de países como: Uruguai (43,9km/1000m²), Argentina (42,3km/1000m²) e Equador (31,4km/1000m²). No levantamento da CNT (2022), 52,2% dos trechos avaliados têm problemas no pavimento, tais como trincamentos, afundamentos e panelas.

A deformação permanente, conhecida também como afundamento de trilha de roda (ATR), é um fenômeno ocasionado pelas solicitações de cargas veiculares, podendo ocorrer na camada de revestimento e/ou nas subcamadas do pavimento. O trabalho aqui desenvolvido diz respeito apenas ao primeiro caso e alguns fatores podem ser apontados e explorados como possíveis causas desse defeito. Segundo Ferreira (2016), a granulometria selecionada influencia diretamente nesse fenômeno, pois os métodos tradicionais de escolha das faixas granulométricas dos agregados não levam em conta o impacto da distribuição dos agregados.

O ATR pode ser causado por densificação e/ou deformação por cisalhamento

(Bastos, 2016). Segundo Guimarães (2009), considera-se falha no revestimento as deformações acima de 12,5mm. Na Figura 1, estão ilustrados os dois tipos de afundamento.

Figura 1 – Afundamento em trilha de roda



Fonte: Ferreira (2015).

A avaliação do potencial de resistência ao acúmulo de deformação é tradicionalmente realizada a partir de ensaios laboratoriais. Um dos parâmetros mais utilizados é o *flow number* (FN), que é obtido como resultado do ensaio uniaxial de carga repetida (NBR 16505 e DNIT 184/2018 - ME) e auxilia no ranqueamento de misturas asfálticas. Algumas pesquisas (Witczak *et al.*, 2002; Bastos *et al.*, 2017) mostraram que há uma boa correlação entre o FN e a deformação permanente em campo. Portanto, o FN é compreendido como uma opção viável para avaliar a suscetibilidade à deformação permanente de materiais usados em revestimentos asfálticos.

O método de dimensionamento nacional (MeDiNa) de pavimentos recém implantado no Brasil trata a deformação permanente como um defeito a ser mitigado ainda na etapa de dosagem (Franco e Motta, 2018). Portanto, a existência de um modelo de previsão para o FN poderia ser de grande utilidade no sentido de nortear a metodologia de dosagem, bem como a escolha do ligante asfáltico e dos agregados, a fim de possibilitar uma mistura mais resistente ao ATR, além de permitir a otimização desses componentes para gerar uma solução de custo reduzido.

A análise da deformação permanente é essencial para investigar a resposta das estruturas de pavimentação. Um modelo capaz de prever o comportamento da mistura ao longo dos ciclos de carregamento é complexo e a literatura apresenta alguns baseados em análise de regressão, como os de Sweere (1990) e Wolff e Visser (1994). Segundo Alnedawi *et al.* (2019), modelos de regressão carecem de uma generalização, sendo adequados ao conjunto de condições a partir das quais foram produzidos.

Atualmente, alguns estudos para a compreensão das misturas asfálticas em relação à deformação permanente (DP) trazem abordagens alternativas com o uso de aprendizado de máquina. No entanto, há um ceticismo sobre esses algoritmos, que são vistos como “caixa-

preta” devido à dificuldade de se visualizar o funcionamento dos códigos. Um dos desafios no uso da inteligência artificial (IA) é a aceitação e a confiança do público nas respostas fornecidas, sendo que o comportamento incompreensível de alguns modelos torna a desconfiança com relação aos mesmos ainda maior. Em diversas aplicações, compreender a previsão do modelo pode ser tão relevante quanto à precisão dessa previsão. No entanto, os melhores resultados em aplicações de aprendizado de máquina são alcançados com a utilização de modelos mais complexos e de difícil interpretação, o que abre uma lacuna entre a precisão e a interpretabilidade.

Uma resposta para esse problema pode estar no uso da chamada IA interpretável, que pode ser definida como uma ferramenta de IA aplicada a algoritmos que não fornecem uma explicação clara das decisões tomadas. Uma dessas técnicas é o método *SHAP* (*SHAP*ley Additive exPlanations), usado para explicar a relevância de cada variável e como essas afetam o modelo. O *SHAP* permite realizar uma análise local e global do conjunto de dados do fenômeno estudado. Sendo assim, o presente trabalho tem por objetivo a utilização do aprendizado de máquina para desenvolver modelos de classificação para misturas asfálticas utilizando uma ferramenta de aprendizado de máquina explicável.

1.3 Problema de pesquisa

A deformação permanente é um defeito recorrente nas estruturas de pavimentação asfáltica, mas que pode ser mitigado na etapa de dosagem da mistura aplicada no revestimento. Dessa forma, torna-se interessante uma ferramenta de aprendizado de máquina que possa fornecer informações prévias sobre o comportamento da mistura em relação a esse parâmetro. Apesar do avanço do aprendizado de máquina em diversas áreas de estudo, ainda existem dúvidas em relação ao uso de algoritmos devido a sua complexidade e a dificuldade em compreender seu funcionamento.

A dificuldade no entendimento de algoritmos de aprendizado de máquina, com foco em misturas asfálticas, revela-se como o problema de pesquisa desta dissertação. Os algoritmos preditivos são treinados em grande quantidade de dados e podem inadvertidamente transmitir algum viés, a julgar pela qualidade e variação desses dados, o que pode resultar em resultados tendenciosos ou injustos, dificultando o uso e a aceitação dessas técnicas. A proposição do aprendizado de máquina interpretável torna-se uma alternativa interessante para mitigar esse problema. Aprender mais sobre essas técnicas e compreender melhor o funcionamento de algoritmos de previsão é fundamental para trazer confiança e apoio em tomada de decisões.

A utilização de ferramentas de aprendizado de máquina interpretável pode ajudar a ampliar o conhecimento e a aplicação de técnicas de ciência de dados na engenharia viária, além de fornecer uma visão alternativa e trazer novos *insights* a respeito do efeito de variáveis dos constituintes, ou seja, do ligante, dos agregados e da própria volumetria da mistura.

Sendo assim, entende-se como relevante a proposição de modelos de aprendizado de máquina interpretável para classificação de misturas asfálticas em relação a defeitos como deformação permanente. Para além das métricas de desempenho que esses modelos podem abordar, é necessária uma visão sobre como os algoritmos utilizam as variáveis para chegar a suas previsões, jogando luz sobre as variáveis que mais impactam no modelo. Essa compreensão pode nortear a obtenção de novos dados na área de infraestrutura, com foco nos parâmetros mais relevantes para a previsão de um parâmetro como o FN de misturas asfálticas, o que pode tornar os algoritmos mais robustos e confiáveis.

1.4 Objetivos

1.4.1 Objetivo Geral

A presente pesquisa tem como objetivo geral implementar modelos de aprendizado de máquina interpretável para o desenvolvimento de algoritmos precisos e intuitivos de classificação de misturas asfálticas em relação à deformação permanente e compatibilidade com classes de tráfego correspondentes.

1.4.2 Objetivos Específicos

A partir do objetivo geral, foram traçados os seguintes objetivos específicos:

- a) Avaliar quais dados possuem maior poder de contribuição em um modelo preditivo de classificação de misturas asfálticas com foco na deformação permanente;
- b) Analisar modelos de classificação de misturas asfálticas a partir dos dados obtidos, tais como: ensaios de caracterização do ligante, granulometria dos agregados, volumetria e rigidez da mistura;
- c) Investigar uma ferramenta de aprendizado interpretável para avaliar a importância individual e global das variáveis em relação ao fenômeno da deformação permanente.

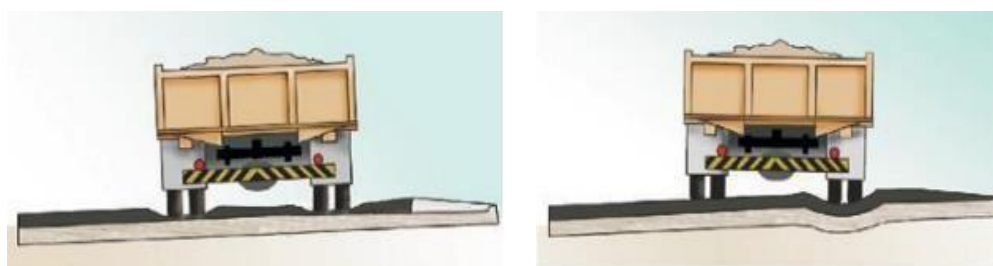
2 REVISÃO BIBLIOGRÁFICA

A revisão da literatura aborda uma breve visão geral sobre a deformação permanente em pavimentos asfálticos, o comportamento dos componentes da mistura asfáltica em relação a esse defeito e o ensaio laboratorial para estimar um parâmetro indicativo da resistência ao mesmo, no caso o ensaio uniaxial de carga repetida que fornece o *flow number*. Em seguida, será mostrado um panorama a respeito do aprendizado de máquina, detalhando algoritmos de previsão e sobre ferramentas de interpretação de modelos preditivos.

2.1 Deformação permanente em pavimentos asfálticos

Nos estudos relacionados à mecânica dos pavimentos, a deformação permanente é um dos principais defeitos observados na estrutura do sistema em camadas, especialmente em locais com altas temperaturas. Essa patologia é caracterizada por uma deformação plástica nos eixos de passagem dos veículos (ATR ou afundamento da trilha de roda). O ATR, como anteriormente mencionado, pode ocorrer apenas na camada de revestimento e também na estrutura do pavimento como ilustra a Figura 2.

Figura 2 – Trilha de roda em pavimentos



a) Trilha de roda no revestimento

b) Trilha de roda no pavimento

Fonte: Bernucci *et al.* (2022)

Quando o pavimento é submetido a uma carga, a estrutura se deforma e logo após a aplicação da carga, parte da deformação é restaurada imediatamente (parte elástica), e parte dela será recuperada em função do tempo (parte viscosa). Isso se dá porque as misturas asfálticas que frequentemente são usadas nos revestimentos apresentam um comportamento viscoelástico. Há ainda uma outra parte dessa deformação que não pode ser recuperada e se acumula com o tempo, chamada de deformação plástica.

A depender do tipo e do teor de ligante, as misturas asfálticas se tornam mais ou menos vulneráveis à variação de temperatura, à frequência e tempo de aplicação de carga (Specht *et al.*, 2017), resultando em uma maior chance de ocorrer deformações permanentes em temperaturas mais altas e em frequências mais baixas. As deformações permanentes resultam em danos na rodovia, gerando insegurança, atrasos, aumento no consumo de combustível, aumento no tempo de viagem e podendo também gerar acúmulo de água no pavimento, que podem implicar no fenômeno de aquaplanagem.

Um material não acumula deformação permanente quando sua estrutura interna é capaz de absorver os esforços solicitantes. Para Masad *et al.* (1999), os mecanismos de resistência à deformação permanente das misturas asfálticas estão ligados a estrutura interna desse material, o que inclui a distribuição do ligante asfáltico, dos agregados e dos vazios. Pode-se mencionar a transmissão interna de cargas como um dos mecanismos de resistência à deformação permanente, sendo essa caracterizada pela interação força-deslocamento produzida pelo contato entre as partículas adjacentes (Zhu e Nades, 2000).

2.2 Impacto do ligante na deformação permanente

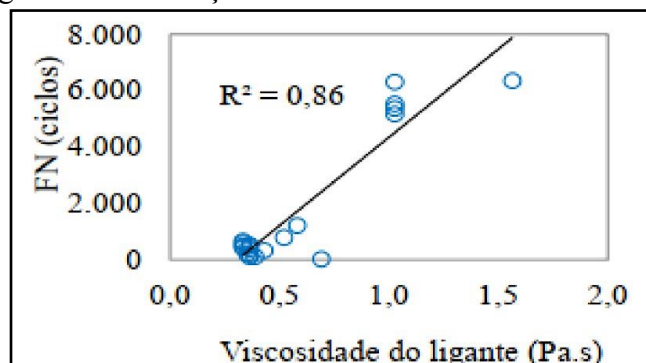
O teor de ligante é um dos fatores que influencia diretamente na deformação permanente das misturas asfálticas. Uma dosagem imprópria, ou seja, um teor elevado de ligante, pode produzir um material com maior suscetibilidade ao acúmulo de deformação (Roberts *et al.*, 1996). De acordo com Sybilski (1996), a deformação acumulada no ligante asfáltico, em razão do tráfego de veículos, é um dos principais responsáveis pela deformação permanente de misturas e, conseqüentemente, dos pavimentos asfálticos.

Outra característica importante das misturas asfálticas é a sua termosensibilidade e suas propriedades viscoelásticas derivadas dos ligantes. Segundo Bernucci *et al.* (2022), a maioria dos países utiliza características físicas do ligante como, resistência ao fluxo, medida por ensaios de viscosidade. Os autores também afirmam que as propriedades físicas do asfalto estão associadas à temperatura, entre elas a resistência ao escoamento.

Devido à dependência da viscosidade da mistura asfáltica em relação à temperatura, em países tropicais como o Brasil, a mistura asfáltica torna-se menos viscosa, produzindo uma lubrificação maior no esqueleto pétreo e prejudicando o intertravamento dos agregados, tornando a mistura menos rígida e mais susceptível à deformação permanente. Ferreira (2017) analisou a relação entre a viscosidade e a deformação permanente, utilizando o *flow number*

como parâmetro. A Figura 3 ilustra a correlação entre esses dois parâmetros.

Figura 3 – Correlação entre viscosidade rotacional e FN



Fonte: Ferreira (2017)

Os ligantes asfálticos que apresentam maiores resistências à penetração ou mais elevado *Performance Grade* (PG) tendem a gerar misturas asfálticas mais resistentes à deformação permanente. O PG, segundo a norma *Superpave*, é representado por dois números: o primeiro correspondente à temperatura mais alta do pavimento, correspondente aos ensaios que avaliam a resistência ao acúmulo de deformação permanente devem ser realizados, o segundo número corresponde à temperatura mínima do pavimento, aquela em que os ensaios que analisam a resistência às trincas térmicas devem ser realizados. O ligante asfáltico é classificado em classes de temperaturas, o intervalo entre as classes de temperatura varia de 6°C em 6°C, tanto para máxima como mínima temperatura.

A depender da região onde for empregada a mistura, poderá haver mudanças na rigidez na camada asfáltica em relação à variação de temperatura ao longo do dia (Vassoler *et al.*, 2011). Uma mudança de temperatura de 5°C pode apresentar uma variação de rigidez de até 70%.

Vários estudos indicam a importância dos ligantes asfálticos na contribuição da resistência à deformação permanente. Viscosidade, teor de ligante e ligantes modificados são alguns dos fatores relevantes. Asfaltos menos viscosos tendem a tornar a mistura asfáltica menos rígida e conseqüentemente mais sucessível à DP.

Almeida Júnior (2015) comparou misturas asfálticas utilizando ligantes modificados por polímeros (AMP 60/85) e ligantes não modificados (CAP 50/70). As misturas que continham ligante modificado apresentaram valores de *flow number* até 8 vezes maiores em relação às misturas com o ligante não modificado. A melhoria do ligante modificado no

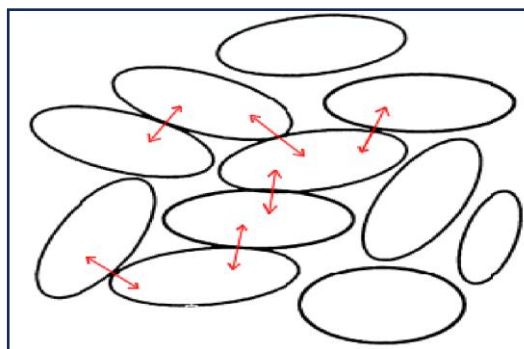
desempenho das misturas em relação à deformação permanente também foi observada nos trabalhos de Onofre (2012), Borges (2014), Bohn (2016) e Bastos (2016).

2.3 Influência do agregado na deformação permanente

Com o avanço do crescimento da frota de veículos no Brasil, os antigos pavimentos passaram a não suportar de maneira satisfatória a demanda do tráfego cada vez mais severo. Desde então, novas abordagens são consideradas no dimensionamento dos pavimentos. Segundo Kennedy *et al.* (1994), as propriedades e o uso dos agregados em misturas asfálticas podem afetar até 80% da resistência à deformação permanente.

A distribuição granulométrica das partículas que compõem as misturas é feita de modo a promover o intertravamento dos agregados, onde os agregados de menor dimensão preenchem os espaços vazios deixados pelos agregados maiores (Momm e Domingues, 1998), conforme ilustrado na Figura 4.

Figura 4 – Contato entre grãos em um conjunto de partículas



Fonte: Adaptado Zhu e Nodas (2000)

Segundo Kim (2006), a seleção granulométrica com um intertravamento mais adequado, pode garantir um melhor comportamento da mistura em relação à resistência à deformação permanente, independente da qualidade de outras variáveis. A granulometria, bem como a seleção dos agregados minerais, é geralmente definida através de faixas especificadas por órgãos rodoviários, que carecem de revisões com o passar do tempo, para garantir um desempenho mais satisfatório.

Entre os métodos sistemáticos de seleção granulométrica, encontra-se na literatura o método Bailey e a metodologia Faixa de Agregados Dominantes (FAD), resultantes de

estudos focados em deformação permanente.

O método Bailey foi desenvolvido na década de 1980 pelo engenheiro Robert Bailey do Departamento de Transportes de Illinois norte-americano, como um procedimento sistemático para definir a composição granulométrica que conferisse trabalhabilidade e resistência para a mistura asfáltica. Posteriormente, o método foi refinado por Vavrik *et al.* (2002) para melhorar sua durabilidade através de uma composição mineral com características densas e descontínuas que garantisse um maior intertravamento.

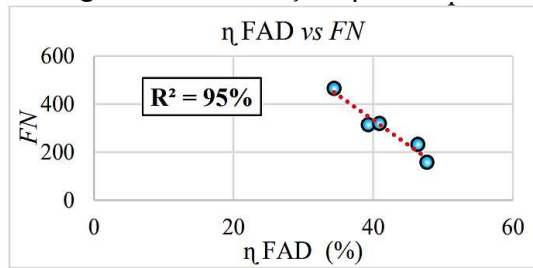
Kim (2006) propôs a metodologia FAD, uma ferramenta para verificar a estrutura granulométrica das misturas asfálticas por meio de um esqueleto pétreo estável, capaz de resistir a deformação permanente por meio do contato entre partículas graúdas, acima de 1,18mm, denominados agregados dominantes. Os agregados mantêm um espaçamento constante e seu contato efetivo pode garantir um esqueleto pétreo com boa resistência a danos. As partículas são avaliadas por porosidade, medida adimensional que indica o comportamento de deformação permanente da mistura (Ferreira *et al.*, 2015; Bastos, 2015; Ferreira *et al.*, 2016).

Portela Neto (2018) relacionou a influência da origem do agregado e sua composição granulométrica a partir dos parâmetros FAD e dos componentes intersticiais (CI), que são agregados com diâmetros menores que a última partícula graúda dominante, bem como do tipo de ligante e da dosagem da mistura, nos resultados de FN. Ele ressaltou que o método FAD pode avaliar e especificar granulometrias de misturas asfálticas para melhor resistir à deformação permanente e que parâmetros da FAD estão satisfatoriamente associados ao FN, desde que tenham matriz granular e composição e características do ligante semelhantes.

Oliveira (2019) avaliou o impacto da metodologia FAD-CI (FAD acrescida dos Componentes Intersticiais) na vida de fadiga de misturas asfálticas, concluindo que misturas que atendiam a mais parâmetros dessa metodologia eram mais resistentes à fadiga.

Ferreira (2015) citou as vantagens de usar esse método de seleção granulométrica, pois pode ser avaliado e facilitado para modificar as graduações para garantir que a mistura tenha boas propriedades de intertravamento. Neto (2018) mostrou que as misturas com mesmo tipo de agregado, porém com diferentes proporções entre os agregados FAD, possuíam boa correlação com o FN, permitindo dessa maneira ranquear as misturas que apresentassem essa relação. A Figura 5 ilustra essa correlação entre a porosidade FAD e o FN.

Figura 5 – Correlação $\eta_{FAD} \times FN$



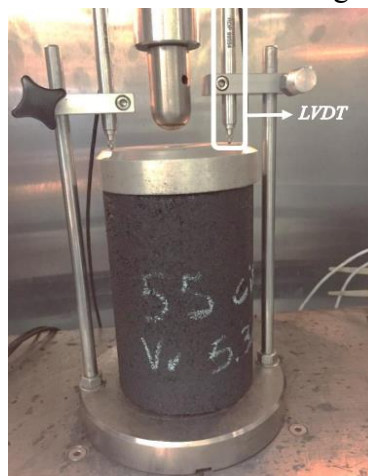
Fonte: Neto (2018)

2.4 Ensaio uniaxial de carga repetida

A NBR 16505 (2016) é a norma brasileira que rege a avaliação de laboratório com relação à deformação permanente de misturas asfálticas, além da atual norma do DNIT 184/2018 – ME. Nos Estados Unidos o procedimento é padronizado pela norma AASHTO T 378-17. O *flow number* é o parâmetro obtido por esse ensaio e atualmente é o mais comum quantificar a deformação permanente acumulada sob cargas repetidas em temperatura alta (60°C). Esse critério é utilizado no atual método de dimensionamento nacional de pavimentos flexíveis. Diversos trabalhos nacionais adotaram esse ensaio (Nascimento, 2008; Onofre, 2012; Borges, 2014; Pazos, 2015; Ferreira, 2015; Bastos, 2016; Vale, 2020).

O corpo de prova é acondicionado à temperatura de 60°C durante um tempo mínimo de 3 horas. Os deslocamentos verticais são medidos por meio de transdutores de deslocamento eletrônicos (LVDTs, do inglês *Linear Variable Differential Transducer*), posicionados na superfície do corpo de prova (CP), conforme Figura 6. A deformação permanente acumulada (ϵ_p) é obtida em função do número de ciclos de carregamento.

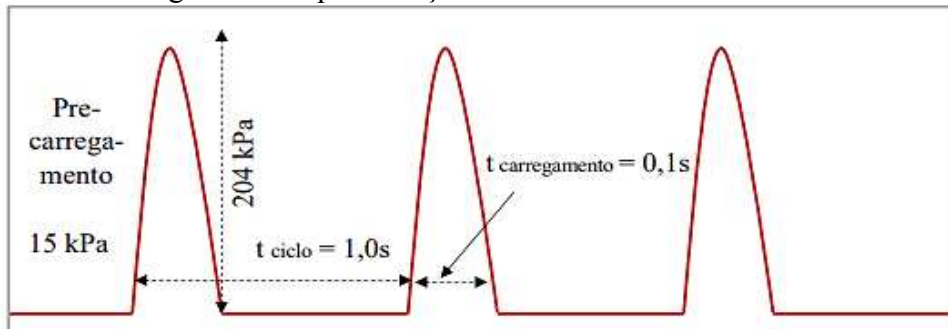
Figura 6 – Ensaio uniaxial de carga repetida



Fonte: Santos (2017)

No início do ensaio a mistura é pré-carregada para acomodação do CP, seguido de um ciclo que corresponde a uma compressão de 204kPa durante 0,1s e com repouso de 0,9s. Durante o repouso a mistura é submetida a um carregamento de 10,2kPa, correspondente a 5% da tensão principal e a deformação permanente acumulada em função do tempo é encontrada pelo número de ciclos. A Figura 7 mostra o carregamento realizado durante o ensaio.

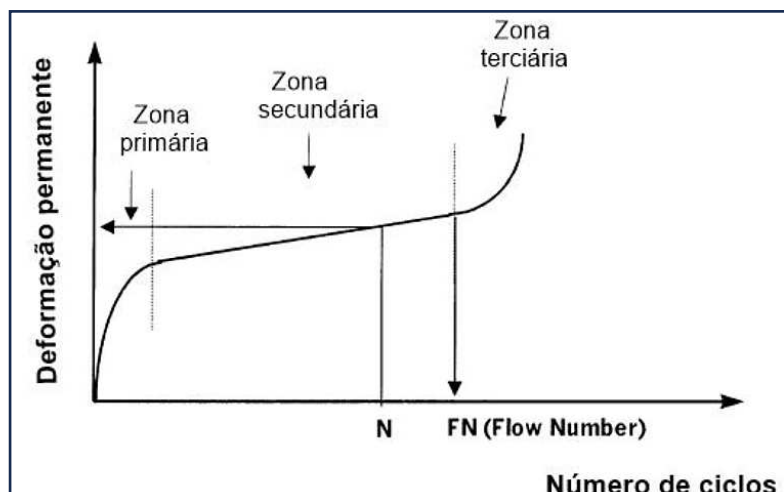
Figura 7 – Representação dos ciclos durante o ensaio



Fonte: Ferreira (2015)

A deformação permanente acumulada é dividida em 3 zonas: primária, secundária e terciária. Na zona primária ocorre uma deformação com uma taxa elevada e variável, que se mantém constante durante a zona secundária e então a deformação aumenta rapidamente na zona terciária, o que descreve a falha do material. A Figura 8 ilustra as zonas de deformação no ensaio uniaxial.

Figura 8 – Zonas da curva de deformação plástica e *flow number*



Fonte: Ferreira (2015)

O FN é o resultado do número de ciclos correspondente a passagem da zona secundária para a zona terciária e serve para a análise da deformação permanente. A norma do DNIT estabelece outros dois critérios para o fim do ensaio: um deles corresponde a 7.200 ciclos que a mistura atinge e o outro quando a mistura atingir 5% de deformação em ensaios confinados e 2% de deformação para ensaios não confinados.

Na literatura, existem divergências acerca de quais valores de FN poderiam ser associados a determinado tipo de tráfego. Nascimento (2008) propôs valores de FN obtidos em laboratório de acordo com uma adaptação do ensaio uniaxial de carga repetida inicialmente proposto por Witczak *et al.* (2002), considerando 9 misturas. Bastos *et al.* (2017) compararam a metodologia dos ensaios laboratoriais propostos por Witczak *et al.* (2002) e aplicados por Nascimento (2008). Posteriormente, comparou os dados laboratoriais com valores de deformação permanente em campo medidos em 12 trechos experimentais. O relatório 673 do *Advanced Asphalt Technologies* (AAT) do programa do NCHRP apresentou valores de FN para diferentes tipos de tráfego considerando 27 misturas em laboratório testadas em 6 trechos da *Federal Highway Administration* (FHWA). Bonaquist (2012) utilizou misturas aplicadas em trechos submetidos a cargas em baixa velocidade e diferentes níveis de tráfego, associados às características volumétricas da mistura, temperatura do ensaio e a tensão de aplicação utilizada na mistura em laboratório. Os valores propostos para o FN de acordo com o tipo de tráfego dos autores citados estão dispostos na Tabela 1.

Tabela 1 – Critérios de FN para diferentes níveis de tráfego

Pesquisas / Níveis de Tráfego (Número N)	Vv (%)	Temp. (°C)	Tensão (kPa)	Leve (3x10⁶)	Médio (3.10⁶ a 1.10⁷)	Pesado (1.10⁷ a 3.10⁷)	Extremamente pesado (> 3.10⁷)
Nascimento (2008)	6,5 a 7,5	60	204	--	300	750	--
AAT (2011)	4,7 a 8,2	31,3 a 54,3	600	--	53	190	740
Bonaquist (2012)	6,0 a 7,3	49,6	600	15	50	135	415
Bastos <i>et al.</i> (2017)	5,0 a 7,5	60	204	--	100	300	> 1000

Fonte: adaptada de Bastos *et al.* (2017)

2.5 Modelo de Francken

A necessidade de um modelo que possa reduzir a subjetividade do FN foi justificada por Ameri *et al.* (2014), avaliando que o modelo de Francken, recomendado por Biligiri *et al.* (2007), é apropriado no tratamento de dados para misturas com ligante modificado ou não. O modelo proposto é formado por uma função de potência e uma função exponencial, de acordo com a Equação 1.

$$\varepsilon_p(N) = AN^B + C(e^{DN} - 1) \quad (1)$$

Onde: $\varepsilon_p(N)$ é a deformação permanente; N é o ciclo de carregamento; e A , B , C e D são coeficientes de regressão.

Os parâmetros A e B representam as zonas primária e secundária da curva de deformação permanente e os parâmetros C e D ajustam o modelo na zona terciária. A inclinação da curva de deformação permanente é encontrada ao derivar a Equação 1 em relação a N , obtendo-se a Equação 2.

$$\frac{d\varepsilon_p}{dN} = A \times B \times N^{(B-1)} + (C \times D \times e^{DN}) \quad (2)$$

Para encontrar a curva ajustada ou o gradiente da inclinação da curva de deformação e também para determinar o FN deve-se calcular a segunda derivada da Equação 2, expressa pela Equação 3. O FN ocorre no ciclo de carregamento em que a segunda derivada é igual a zero.

$$\frac{d^2\varepsilon_p}{dN^2} = A \times B \times (B-1) \times N^{(B-2)} + (C \times D^2 \times e^{DN}) \quad (3)$$

2.6 Aprendizado de máquina

A aprendizagem de máquina é um campo da inteligência artificial que habilita os computadores a aprenderem com dados e aprimorarem seu desempenho em uma tarefa específica ao longo do tempo. Allen-dencid (2019) define-a como uma área de pesquisa computacional que tem como objetivo explorar um conjunto de métodos que reconhecem padrões em um conjunto de dados e os empregam para prever eventos futuros.

Para Russel e Norvig (2013), a aprendizagem de máquina pode ser dividida em três categorias principais: aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço. Na primeira, o computador é alimentado com um conjunto de dados

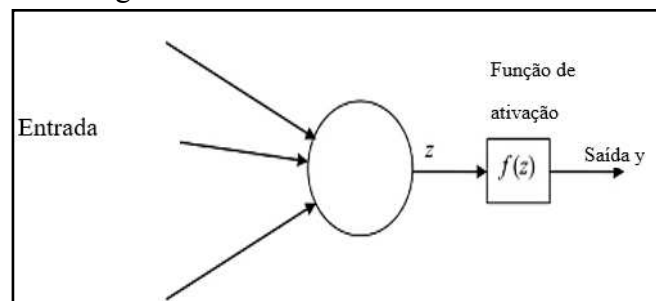
rotulados, onde cada linha de dados está associada a uma variável alvo. O objetivo é aprender um mapeamento das variáveis de entrada para a variável de saída. Na aprendizagem não supervisionada, o computador busca padrões em uma estrutura de dados não rotulada. Na aprendizagem por reforço, o computador aprende por meio de uma interação com um ambiente e recebe recompensas ou penalidades por suas decisões.

Algoritmos de classificação são algoritmos de aprendizagem supervisionada onde o objetivo é prever uma classe ou um rótulo associado com uma variável de entrada contendo determinados atributos. Inicialmente, o algoritmo é treinado com um conjunto de dados com classes conhecidas, podendo esses dados estar divididos em somente duas (classificação binária) ou em várias classes (classificação multiclasse).

2.7 Redes Neurais Artificiais (RNAs)

McCulloch e Pitts (1943) introduziram as primeiras redes neurais artificiais e propuseram um modelo computacional simples que demonstrava como os neurônios biológicos trabalham juntos para realizar cálculos complexos usando a lógica proposicional. O modelo com uma ou mais entradas binárias (ativadas e desativadas) e saídas binárias (mais tarde chamadas de neurônios artificiais). O modelo neural de McCulloch está ilustrado na Figura 9.

Figura 9 – Modelo neural McCulloch

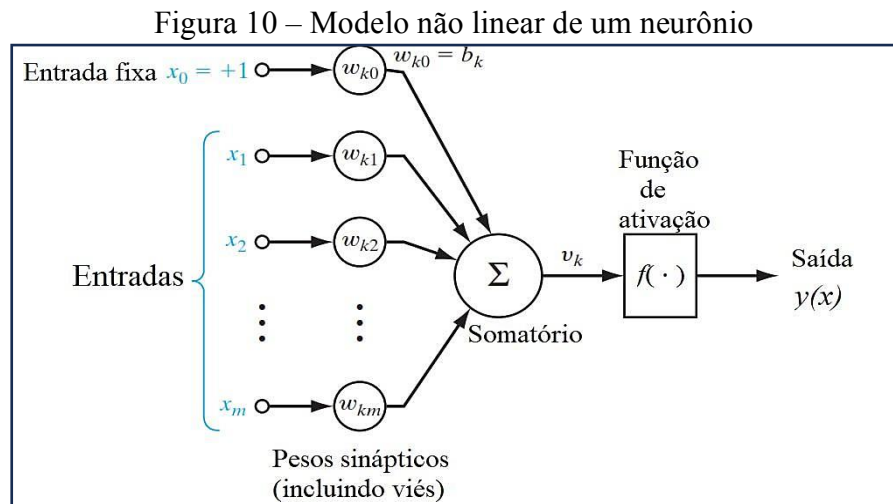


Fonte: Fernandes (1999)

Rosenblatt (1958) aprimorou o modelo, que foi batizado de *Perceptron*. Do ponto de vista neurofisiológico, esse neurônio representa um método mais simplificado, mas retém várias características básicas dos neurônios biológicos e permite cálculos computacionais. Haykin (2001) descreveu uma rede neural artificial como um processador paralelo a uma unidade de processamento simples distribuída, por meio da qual o conhecimento experimental

pode ser armazenado e fornecido para uso posterior.

Para Bishop (2006), o termo “rede neural” nasceu da tentativa de encontrar uma representação matemática do processamento da informação em sistemas biológicos. Desde então, tem sido amplamente utilizado para representar uma variedade de modelos diferentes. Para Haykin (2009), o neurônio é a unidade de processamento de informações necessária para o funcionamento da rede neural. Na Figura 10 são ilustrados os 3 elementos básicos do modelo neural, que são descritos em seguida.



Fonte: Haykin (2009)

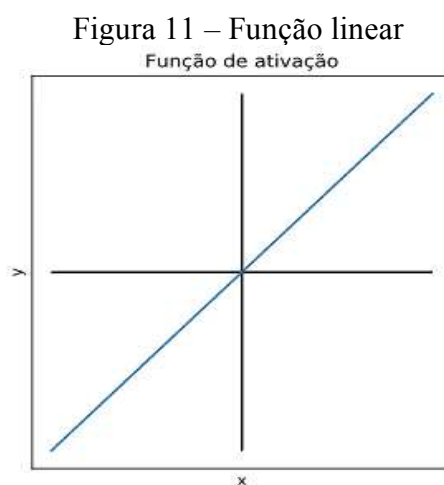
- a) Um conjunto de sinapses, links conectados e dotados de um peso. Especificamente, o sinal x_j na entrada da sinapse j conectado ao neurônio k é multiplicado pelo peso da sinapse w_{kj} . O primeiro índice em w_k aponta para o neurônio em questão, e o segundo índice aponta para o final da entrada da sinapse apontada pelo peso. Ao contrário do peso das sinapses no cérebro, o peso sináptico dos neurônios artificiais pode estar em uma faixa que inclui valores positivos e negativos;
- b) A soma do sinal de entrada é ponderada pelo peso sináptico de cada neurônio, sendo que essas operações formam uma combinação linear;
- c) O último elemento é a função de ativação, que é usada para limitar a amplitude da saída do neurônio. A função de ativação limita a faixa de amplitude permitida do sinal de saída a um determinado valor finito.

2.7.1 Funções de ativação

Por mais complexa que seja a rede neural, ela só pode captar relações lineares entre as variáveis de entrada e a variável de saída. Para torná-las capazes de modelar também relações não-lineares, os resultados de saída de cada camada passaram a ser processados pelas chamadas funções de ativação. Isso é especialmente importante nas camadas escondidas da rede neural. Nas camadas de saída, as funções de ativação podem ter finalidades específicas, a depender do problema que a rede neural resolvendo. Dentre as várias funções, as mais comuns são as seguintes: linear, sigmoide, tangente hiperbólica e ReLU (Ribeiro, 2013).

2.7.1.1 Função linear

A função linear é limitada em sua capacidade de compreender relações mais complexas entre os dados, justamente por conta da sua linearidade. Além disso, sua derivada é constante, o que faz com que o gradiente a cada etapa de *backpropagation* seja constante. Assim, a etapa de descida do gradiente não tende a convergir para produzir um erro estável próximo de zero. Na camada de saída, a função de ativação linear pode ser utilizada em problemas de regressão, já que produz resultados em todo o domínio dos números reais. A Figura 11 ilustra a função linear.



Fonte: Ceccon (2020)

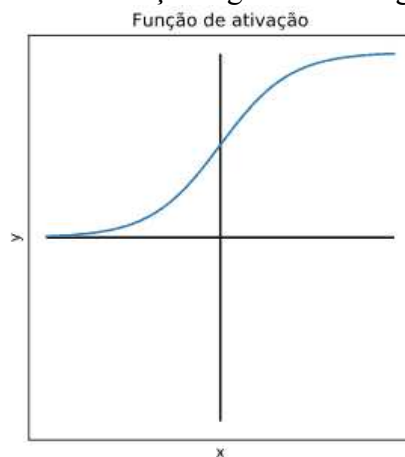
2.7.1.2 Função sigmoide ou logística

A função sigmoide é vantajosa na questão da não linearidade, e o valor da derivada é máximo quando x está próximo de 0, o que tende a “empurrar” seu resultado para as

extremidades do intervalo [0, 1] ao longo do treinamento, o que é uma característica desejável, por exemplo, em problemas de classificação. Entretanto, sua característica não linear aumenta o custo computacional.

Na camada de saída, a função sigmoide é útil para produzir probabilidades em problemas de classificação binária, já que seus resultados, na faixa de [0, 1], podem ser interpretados como a probabilidade de determinada instância pertencer ou não a determinada classe. A Figura 12 ilustra a curva da função sigmoide.

Figura 12 – Função sigmoide ou logística



Fonte: Ceccon (2020)

A função sigmoide ou logística é dada pela Equação 4:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (4)$$

2.7.1.3 Função tangente hiperbólica

Outra importante função de ativação é a tangente hiperbólica, sendo recomendada como substituta da logística (Lecun *et al.*, 2015). A função tangente hiperbólica pode ser definida através da divisão entre seno hiperbólico e cosseno hiperbólico. Diferente da sigmoide, essa função atua no intervalo [-1,1]. Na Figura 13, tem-se a ilustração do comportamento dessa função.

Figura 13 – Função tangente hiperbólica



Fonte: Ceccon (2020)

Essa simples flexão do intervalo em relação à função sigmoide, resulta em uma melhor convergência da rede neural e é representada pela Equação 5.

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} \quad (5)$$

2.7.1.4 Função ReLU

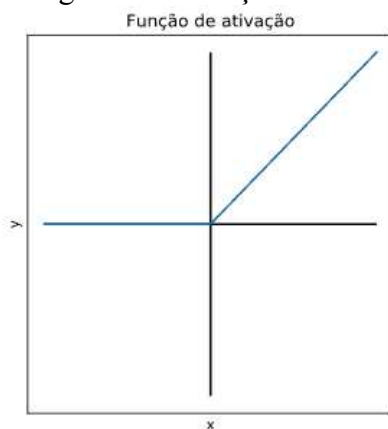
A função Linear Retificada (*Rectified Linear Unit, ReLU*) foi apresentada por Hahnloser *et al.* (2000), em uma rede dinâmica. É a função de ativação mais utilizada nos projetos de redes neurais modernos (*Data Science Academy, 2019*). Essa função objetiva reduzir o tempo de treinamento de redes neurais artificiais profundas e é aplicada aos neurônios das camadas ocultas.

Quando a entrada do neurônio é negativa ele não é ativado o que torna o processamento mais eficiente. O gradiente 0 para $x < 0$, apesar de reduzir tempo de processamento, pode ser um problema para o treinamento, sem dar um direcionamento para o processo de treinamento. A função *ReLU* é representada pela Equação 6.

$$\text{ReLU}(x) = \max(0, x) \quad (6)$$

Na Figura 14 pode-se notar o comportamento da função *ReLU*, que retorna um valor igual a zero para valores negativos de x . Essa função não costuma ser utilizada na camada de saída de uma RNA.

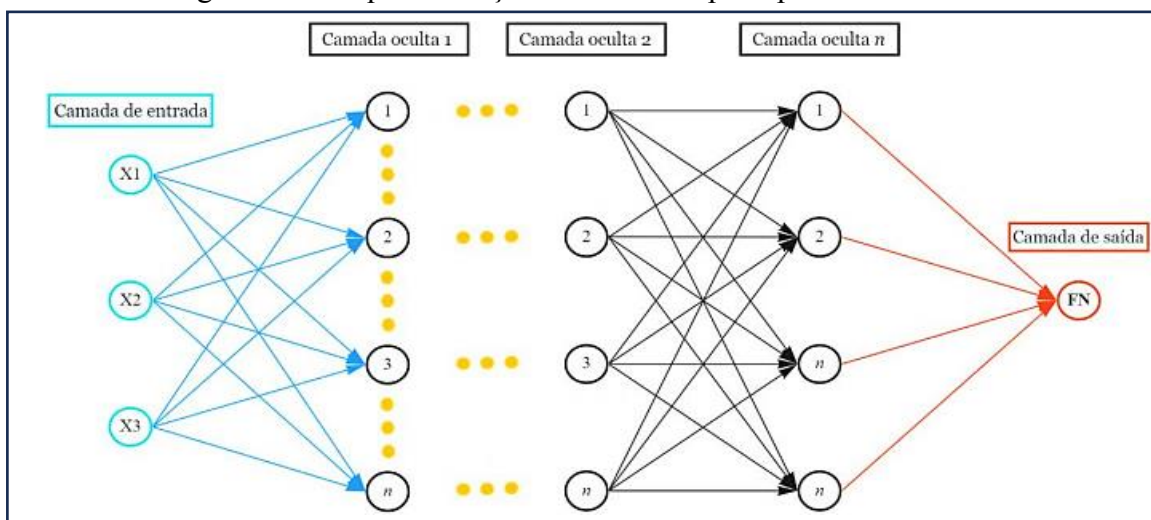
Figura 14 – Função ReLU



2.7.2 Perceptron Multicamadas

As RNAs do tipo *perceptrons* de múltiplas camadas (*Multilayer Perceptron – MLP*) são comumente utilizadas, consistindo em um conjunto de unidades de processamento com unidades de entrada, unidades intermediárias (ou camadas ocultas) e camada de saída. Os estímulos de entrada são propagados pelas camadas intermediárias até a camada de saída (Bocanegra, 2002). A saída de cada neurônio é calculada a partir de uma função de ativação, que pode ser linear ou não-linear. O uso de redes com mais de uma camada oculta é recomendada quando envolve altos graus de não linearidade (Bocanegra, 2002). O *MLP* é capaz de aprender a partir de exemplos e pode ser utilizado em diversas aplicações, como: reconhecimento de padrões, classificação, previsão, entre outras. Uma RNA do tipo *MLP* é ilustrada na Figura 15.

Figura 15 – Esquemática de uma rede perceptron multicamadas



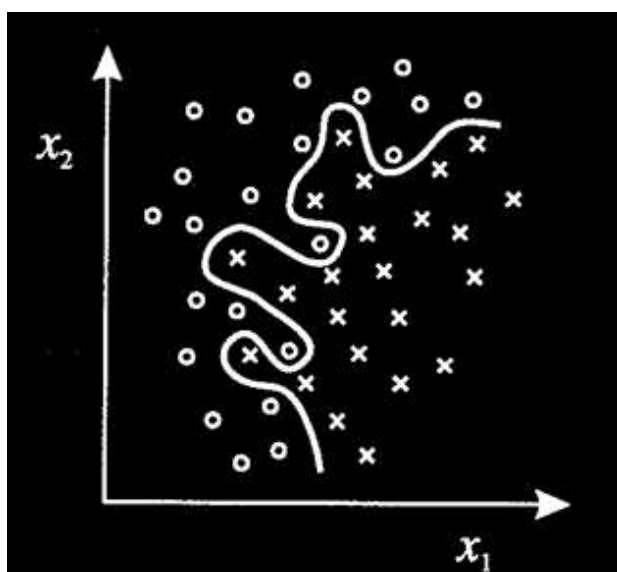
Fonte: O autor

2.7.3 Algoritmo de retropropagação do erro (*Backpropagation*)

Haykin (2001) afirma que o algoritmo conhecido por *error backpropagation* é o mais utilizado em aplicações como reconhecimento de padrões, previsão e classificação. O *backpropagation* é um processo envolvido no treinamento de uma rede neural.

A principal vantagem desse algoritmo é capacidade de lidar com problemas não linearmente, ou seja, aquele onde não é possível separar 2 classes distintas no eixo cartesiano de duas dimensões apenas traçando uma reta, como ilustrado na Figura 16.

Figura 16 – Ilustração de um problema não-linearmente separável



Fonte: Devmedia (2013)

O funcionamento do *backpropagation* baseia-se em, a partir do erro obtido na camada de saída da rede neural, calcular o valor dos pesos do vetor da última camada de neurônios e assim proceder para as camadas anteriores, de trás para a frente, ou seja, atualizar todos os pesos das camadas a partir da última até atingir a camada de entrada da RNA, para isso realizando a retropropagação do erro obtido pela rede.

2.8 XGBoost (eXtreme Gradient Boosting)

Entre os algoritmos de aprendizado supervisionado, existem os algoritmos de aprendizado em conjunto ou *ensemble learning*. A ideia foi proposta originalmente para tarefas de classificação em forma de aprendizagem supervisionada (Nilsson, 1965). O conceito básico do aprendizado de conjunto é treinar vários modelos base como membros do conjunto e

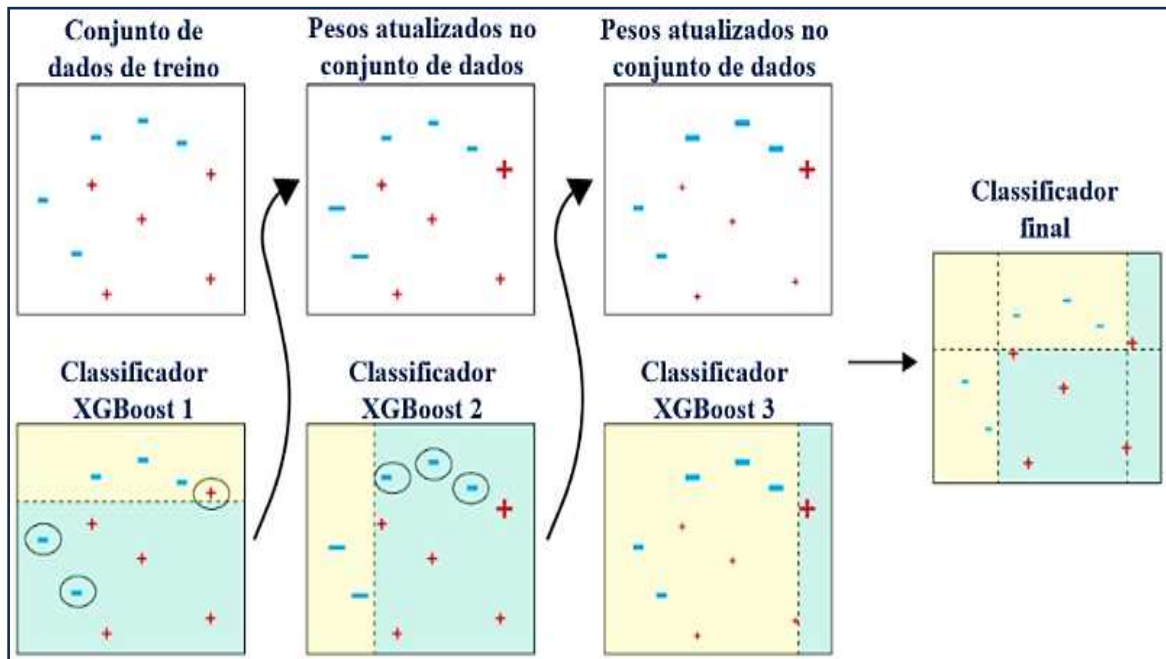
combinar suas previsões em uma única saída que deve ter um desempenho melhor do que qualquer outro membro do conjunto com erro não correlacionado nos conjuntos de dados alvo.

O *XGBoost* (*eXtreme Gradient Boosting*) é um algoritmo de aprendizado de máquina baseado em árvores de decisão que melhora o desempenho, necessitando de menos recursos computacionais e produzindo resultados superiores a outras técnicas em menor tempo. Utiliza uma formalização de modelo mais regularizada para controlar o *overfitting* (quando o modelo apresenta um ótimo desempenho nos dados de treino, mas apresenta resultados ruins nos dados de teste). Podemos interpretar o *overfitting* como uma super adaptação do modelo ao conjunto de treino, em que ele não aprendeu bem as relações existentes entre as variáveis, apenas “decorou” o que deveria ser feito e por isso não possui uma boa capacidade de generalização, apresentando um desempenho ruim com a adição de novos dados. O modelo foi apresentado por Chen e Guestrin (2016) e desde então tem sido amplamente utilizado, principalmente em casos com dados estruturados de pequeno a médio porte.

O *XGBoost* é um tipo de aprendizado em conjunto que consiste na criação de uma sequência de previsores onde o modelo sucessor corrige o erro do modelo antecessor e assim por diante. Dessa forma, o modelo cria um classificador “fraco” e então melhora os classificadores seguintes com o objetivo de diminuir os erros residuais, para capturar e tratar qualquer padrão, ainda que aleatório, nos erros encontrados. O algoritmo então combinará as previsões das árvores criadas para gerar um classificador “forte”.

No exemplo ilustrado na Figura 17, o conjunto de dados do treino é passado para o classificador 1. Nos quadrados coloridos (verde e amarelo), estão em amarelo as previsões com sinal de menos e em verde as previsões para o sinal de mais. O modelo do classificador 1 prevê incorretamente dois sinais de menos e um sinal de mais, destacados com um círculo. Então, o peso desses pontos de dados previstos erroneamente é aumentado e enviado para o próximo classificador. O classificador 2 prevê corretamente os dois sinais de menos, algo que o primeiro classificador não conseguiu. Entretanto, o classificador 2 comete erros de classificação. O processo continua até chegar a um classificador final combinado que prevê todos os pontos de dados corretamente.

Figura 17 – Processo de aprendizado Classificador *XGBoost*



Fonte: Shah, 2020 (adaptado)

2.9 Relevância do atributo

A *Feature Importance* (FI) é amplamente utilizada e vem integrada à maioria das bibliotecas de aprendizado de máquina. É comumente associada a técnicas de *ensembles* de árvores de decisão e pode ser medida por diferentes métricas. A métrica *Split* é o padrão para essa técnica e pode ser obtida contando o número de vezes que a variável foi usada no modelo. Essa métrica indica importância, pois presume que quanto mais vezes a variável foi usada no modelo ou quanto maior o ganho da quebra que usou aquela variável, maior foi o impacto daquela variável na predição (Zhou e Hooker, 2021).

2.10 Interpretabilidade em aprendizado de máquina

Os métodos de aprendizado podem ser chamados de “caixa-preta” devido à dificuldade de interpretação dos resultados obtidos, ainda que forneçam respostas adequadas. Esses modelos não revelam seus mecanismos internos e não possuem mecanismos muito complexos que impedem seu entendimento. Apesar de não haver uma definição matemática, alguns autores como Miller (2019) definem interpretabilidade como o grau no qual um ser humano consegue entender a causa que levou a uma determinada decisão.

Para Doshi-Velez e Kim (2018), não é possível avaliar problemas reais apenas sob o aspecto de uma única métrica, como por exemplo a acurácia. Entender o porquê de certas decisões terem sido tomadas leva a uma maior compreensão do fenômeno e pode levar a um melhor direcionamento da atenção e melhorias nos métodos de coleta dos dados. Os modelos são obtidos após o treinamento de uma base de dados, e os resultados podem estar acompanhados de vieses. A interpretabilidade pode ser uma ferramenta para detectar e evitar esse tipo de viés em modelos de aprendizado de máquina (Molnar, 2022).

Entender a saída de um modelo de previsão pode resultar em maior confiança do usuário e melhor compreensão do fenômeno estudado. Lundberg e Lee (2017) propuseram o *SHAP* (*Shapley Additive exPlanations*) para a interpretação da saída dos modelos. O método é baseado na teoria dos jogos *Shapley* (Strumbeij e Kononenko, 2014) e na abordagem de explicação local LIME (Ribeiro *et al.*, 2016) e permite estimar a contribuição de cada variável no modelo.

Atualmente o *SHAP* é um dos métodos mais populares entre cientistas de dados (Linardatos, Papastefanopoulos e Kotsiantis, 2021), fornecendo a relevância dos atributos, com foco na visualização gráfica das informações. O valor *SHAP* permite análises que abrangem o escopo local e global, o que pode não ser encontrado quando são combinados métodos distintos, como por exemplo o LIME (Ludenberg e Lee, 2017).

O conceito dos valores *SHAP* estabelece que todas as combinações possíveis devem ser consideradas para estimar o valor de cada variável. Para obter o número total de combinações possíveis, deve-se considerar também conjuntos em que uma ou mais variáveis não estão presentes.

Por exemplo, para uma base fictícia com 3 variáveis A, B, C, há o total de 8 combinações possíveis. Dessa forma, pode-se expressar o número total de combinações por 2^N , onde N é o número de variáveis distintas. A contribuição marginal (CM) de cada atributo é calculada pela diferença entre a predição de dois modelos, para uma instância qualquer x_0 . Partindo do modelo sem atributos, a predição seria simplesmente a média da variável resposta em toda a base de dados.

O próximo modelo treinado teria apenas uma variável, como por exemplo a variável A. A diferença entre a primeira predição e a predição do modelo que tem apenas essa variável resulta na contribuição marginal da variável A, como mostra a Equação 7.

$$CM_{A,\{A\}}(x_0) = \text{Predição}_{\{A\}}(x_0) - \text{Predição}_{\emptyset}(x_0) \quad (7)$$

Para obter a contribuição marginal geral da variável, é preciso considerar o impacto dela em todos os modelos, calculando a diferença entre os modelos subsequentes em que essa variável está presente, com aqueles em que ela não está presente. Nesse exemplo, com o total de 8 modelos, essa situação ocorre 4 vezes para a variável A. É preciso utilizar uma média ponderada por w , para cada vez que uma contribuição marginal acontece, de modo que $w_1 + w_2 + w_3 + w_4 = 1$. A contribuição marginal geral de A pode ser calculada pela Equação 8.

$$SHAP_A(x_0) = w_1 \times CM_{A,\{A\}}(x_0) + w_2 \times CM_{A,\{A,B\}}(x_0) + w_3 \times CM_{A,\{A,C\}}(x_0) + w_4 \times CM_{A,\{A,B,C\}}(x_0) \quad (8)$$

2.11 Uso de IA em infraestrutura de transportes

As redes neurais já tem sido utilizadas na infraestrutura viária. Pode-se destacar trabalhos com aplicações de RNA em estudos geotécnicos, previsão do comportamento de misturas asfálticas e também na gerência de pavimentos.

Ferreira (2008) usou ferramentas de *data mining* e RNA para prever valores de Módulo de Resiliência de solos estabilizados e britas, utilizando um banco de dados do laboratório de pavimentação da UFRJ. Seus resultados apresentaram R^2 próximos de 0,90.

Thube (2012) utilizou modelos de RNA para auxiliar na tomada de decisão em serviços de manutenção de pavimentos, usando parâmetros relacionados às condições de pavimento como área trincada e o Índice Internacional de Irregularidade (IRI), para construir modelos de previsão de defeitos. Seus resultados obtiveram um coeficiente de determinação R^2 com valores de até 0,98.

Ribeiro, da Silva e Barroso (2015) fizeram o uso da RNA para prever e classificar solos da Região Metropolitana de Fortaleza (RMF). Foram usados parâmetros como pedologia, vegetação, geomorfologia, geologia e altimetria. A precisão dos acertos do modelo gerado foi de 92,6%. Seguindo a linha de pesquisa, Ribeiro, da Silva e Barroso (2018) elaboraram modelos de previsão e de mapeamento de Índice de Suporte Califórnia (CBR), com a finalidade de criar mapas temáticos para auxiliar na tomada de decisões durante o planejamento de obras rodoviárias. Os resultados obtidos pelos modelos apresentaram R^2 de 0,98.

Ribeiro (2017) usou redes neurais para criar um modelo de previsão para módulo de resiliência dos solos no estado do Ceará. O autor utilizou parâmetros obtidos em ensaios

como variáveis de entrada, entre elas: umidade ótima, massa específica aparente seca máxima, limite de liquidez, limite de plasticidade, tensão confinante, tensão desvio, entre outras. A melhor configuração dos modelos propostos chegou a atingir R^2 acima de 0,98.

Mirzahosseini *et al.* (2015) desenvolveram modelos não lineares para a previsão do FN utilizando duas técnicas: a programação genética linear e RNAs. Esse modelo foi realizado em termos de percentagem de agregado graúdo, filler, ligante asfáltico, volume de vazios, vazios no agregado mineral e quociente de Marshall, que é um parâmetro indicador do comportamento de misturas asfálticas em relação à deformação permanente.

Paz *et al.* (2019) utilizaram redes neurais para detecção automática de defeitos do tipo panela em pavimentos. A técnica utilizada foi a Rede Neural Convolutiva (RNC), esse tipo de rede neural é utilizado com o objetivo de processar e analisar imagens. O banco de dados para treino e teste foi obtido por meio de fotografias cedidas pelo órgão rodoviário do Ceará em parceria com uma empresa privada, que contribuiu com imagens tiradas metro a metro das rodovias do estado. Os resultados obtidos ultrapassaram 80% de precisão no reconhecimento dessa patologia. Os autores ainda obtiveram cerca de 80% de previsão quando utilizaram imagens de *smartphone*.

Celeste e Oliveira (2019) usaram a técnica de RNA para apresentar a retroanálise do Módulo de Resiliência, a partir das espessuras das camadas e das deflexões medidas na superfície do pavimento, como alternativa à retroanálise tradicional. No referido estudo o modelo computacional obteve R^2 de 99%.

Gong *et al.* (2020) desenvolveram modelos de rede neural para estimar valores do módulo dinâmico $|E^*|$ de misturas asfálticas a quente utilizando parâmetros do ligante, da volumetria e da granulometria dos agregados. Os dados foram extraídos da *National Cooperative Highway Research Program* (NCHRP) e obtiveram valores de R^2 acima de 0,97.

Ullah, Tanyu e Zainab (2020) desenvolveram um modelo de previsão baseado em RNA para analisar o comportamento de misturas asfálticas em relação à deformação permanente com diferentes teores de *Reclaimed Asphalt Pavement* (RAP), utilizando teores de 20%, 40% e 60%. Com resultados de R^2 de 0,98, esse estudo mostrou que as redes neurais são capazes de lidar com a heterogeneidade dos dados, resultando em valores precisos de previsão.

Liu *et al.* (2022) utilizaram *XGBoost* para otimizar economicamente a dosagem de misturas asfálticas. O banco de dados contou com 27 variáveis de entrada, entre elas: condições climáticas, tráfego e propriedades dos materiais. Os autores também utilizaram esses dados para

estimar valores de afundamento de trilha de roda, obtendo um R^2 de 0,87 em relação ao afundamento de trilha medido em campo.

3 MATERIAIS E MÉTODOS

Esta pesquisa foi motivada pela busca de formas alternativas de examinar o efeito de variáveis diversas em um defeito relevante da área de infraestrutura viária através do uso do aprendizado de máquina. Para isso, realizou-se uma extensa pesquisa para a aquisição de dados de misturas asfálticas submetidas ao ensaio uniaxial de carga repetida. Neste capítulo, são detalhadas todas as etapas metodológicas para o desenvolvimento da pesquisa, incluindo as características dos materiais considerados e as ferramentas usadas na análise.

3.1 Materiais

Para a realização do presente trabalho foram obtidas 251 misturas asfálticas de diferentes fontes: projetos do Centro de Pesquisas Leopoldo Américo Miguez de Mello (Cenpes), relatório do Centro de Pesquisas Rodoviárias da CCR em parceria com Agência Nacional de Transportes Terrestres (ANTT) e revisão de trabalhos acadêmicos (Onofre, 2012; Ferreira, 2015; Almeida Jr, 2016; Barros, 2017; Faccin, 2018; Possebon, 2018). Dividiram-se os ligantes entre modificados e não modificados. Os ligantes não modificados (CAP 30/45 e CAP 50/70) foram colocados em uma só classificação, enquanto os demais foram divididos em asfalto modificado por polímero 60/85 (AMP), asfalto modificado por borracha moída de pneus tipo AB-08 (AB) e ligante de alto módulo (AM).

As variáveis de entrada escolhidas para a composição do banco de dados foram: ponto de amolecimento (ABNT NBR 6560), penetração (ABNT NBR 6576) e viscosidade *Brookfield* (ABNT NBR 15184) do ligante, módulo dinâmico e volume de vazios da mistura, tamanho máximo nominal do agregado, teor de ligante, porosidade FAD (faixa de agregado dominante); percentual passante nas peneiras e tipo de ligante. Buscaram-se outras variáveis, como por exemplo parâmetros obtidos no ensaio MSCR como: o percentual de recuperação e compliância não recuperável (J_{nr}), porém nem todas as fontes de dados disponibilizavam tal propriedade.

Com os dados adquiridos, realizou-se uma análise quantitativa e descritiva para detectar variáveis com valores discrepantes, como volumes de vazios e teores de ligante excessivamente baixos em relação aos observados normalmente. No total, foram descartados 8 dados de misturas asfálticas. Após a remoção desses dados, foram selecionadas 251 misturas, compostas por 16 variáveis de entrada, sendo 12 variáveis numéricas e 4 variáveis categóricas; há 1 variável de saída dividida em 3 classes de tráfego: médio, pesado e extremamente pesado,

ou seja, a mistura avaliada, à luz de suas características ou de seus constituintes é tida como adequadamente correspondente a uma dessas 3 categorias de tráfego. Uma mistura melhor seria adequada ao tráfego mais elevado (extremamente pesado) e, em contraposição, uma mistura pior seria indicada para o tráfego de menor magnitude, ou seja, no caso o tráfego médio. As variáveis que compõem o banco de dados analisados constam na Tabela 2, sendo os 12 primeiros atributos correspondentes às variáveis numéricas e os 4 últimos correspondentes às variáveis categóricas.

Tabela 2 – Variáveis numéricas e categóricas de entrada do banco de dados desenvolvido

Variável	Descrição
PEN	Penetração (0,1mm)
PA	Ponto de amolecimento (°C)
VIS 135	Viscosidade a 135°C (cp)
TL	Teor de ligante (%)
E*	Módulo dinâmico da mistura asfáltica 10Hz 21°C (MPa)
TMN	Tamanho máximo nominal (mm)
FAD	Porosidade FAD (%)
Vv	Volume de vazios (%)
P38	Passante peneira 3/8" (%)
P34	Passante peneira 3/4" (%)
P4	Passante peneira #4 (%)
P200	Passante peneira #200 (%)
CAP	Ligante puro
AB	Asfalto borracha
AMP	Asfalto modificado por polímero
AM	Ligante asfáltico de alto módulo

Fonte: O autor

A penetração do ligante foi escolhida como variável de entrada devido a sua relação com a deformação permanente, onde menores valores de penetração correspondem a uma maior resistência a esse defeito (Onofre, 2013). A termosensibilidade das propriedades é uma característica fundamental transmitida dos ligantes às misturas asfálticas correspondentes. O ponto de amolecimento é importante por permitir uma aplicação das misturas em condições de serviço em temperaturas mais elevadas, sem perda das propriedades mecânicas. Quanto maior o ponto de amolecimento, menor a penetração e, por conseguinte, maior a tendência de maior

resistência à deformação do ligante e, por consequência, da mistura.

Outra variável selecionada para o modelo foi a viscosidade. Domingos *et al.* (2012) compararam as viscosidades rotacionais de ligantes asfálticos utilizados por Onofre (2012). Notou-se uma tendência dos ligantes com menores viscosidades levarem a misturas com maiores percentuais de deformação permanente ao fim do ensaio uniaxial de carga repetida. Ferreira (2017) também demonstrou uma forte correlação entre a viscosidade e o *flow number*.

A escolha do teor de ligante justifica-se pelo possível impacto no comportamento da mistura em relação à deformação permanente. Segundo Moura (2010), não há uma proporcionalidade entre o aumento do teor de ligante e o aumento da deformação permanente em misturas asfálticas, no entanto, há maiores chances de acontecer essa deformação em trilha de roda com o aumento do teor, ocasionados pela elevação da espessura do filme asfáltico no entorno dos agregados. Isso pode interferir na resistência ao cisalhamento resultante do atrito entre os grãos, devido ao aumento da mobilidade resultante de uma maior lubrificação do contato.

Para avaliar a rigidez da mistura foi escolhido como variável o módulo dinâmico. Alguns autores buscaram correlacionar a resistência à deformação permanente de misturas asfálticas com o módulo dinâmico, como Witzack *et al.* (2002) e Nascimento (2008). Apeagiey (2011) encontrou boas correlações entre o módulo dinâmico e valores de FN na temperatura de 38°C. No entanto, Zhang *et al.* (2013), Barroso e Rezende (2019), ao correlacionarem resultados de MD com resultados de FN, não verificaram fortes correlações. Por meio do ensaio de MD pode-se verificar o comportamento viscoelástico da mistura, obtendo-se as chamadas propriedades viscoelásticas no domínio da frequência, quais sejam, o módulo complexo ($|E^*|$) e o ângulo de fase (δ). O ângulo de fase representa o atraso, ou seja, a defasagem de tempo, da resposta elástica do material em relação à aplicação de carga. O MD apresenta uma medida da rigidez da mistura asfáltica e é o valor absoluto do módulo complexo. Quanto maior o valor do módulo dinâmico, maior a rigidez da mistura, que tende a apresentar uma menor deformação do material quando submetido a aplicação de uma carga (Asphalt Institute, 2001). A temperatura de 21,1°C com a frequência de 10Hz foi usada como referência, uma vez que essa foi uma variável comumente encontrada durante as pesquisas nos diversos bancos de dados.

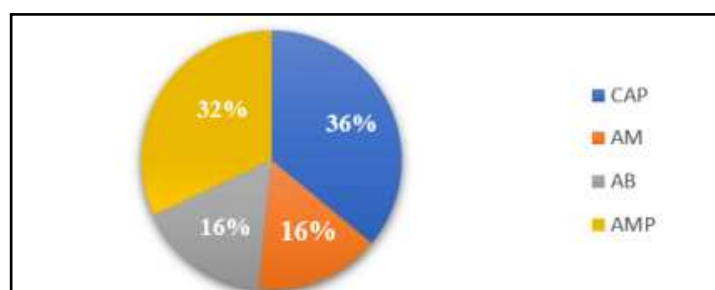
A porosidade FAD foi uma das variáveis granulométricas selecionadas. Como já mencionado, a metodologia FAD se baseia na teoria do empacotamento de partículas esféricas de vários tamanhos e não consideram as características de formas, de textura e da angularidade

dos agregados. É a partir da determinação da porosidade FAD, que é estimado o potencial de resistência. Uma porosidade FAD menor do que 48% indica composições granulométricas com grande potencial de resistência à deformação permanente. Valores entre 48 e 52% são chamadas de marginais e indicam composições com contato granular questionável, não sendo possível estimar com precisão a resistência à deformação permanente (Portela Neto, 2018). Estima-se uma baixa resistência à deformação permanente de misturas com porosidade FAD superior a 52% (Ferreira; Bastos; Soares, 2015). Faccin (2018) cita que a porosidade FAD não se mostra adequada analisando misturas de diferentes tipos de ligantes e de diferentes britagens.

Foram selecionadas variáveis granulométricas dos agregados e também volumétricas da mistura. Segundo Lira *et al.* (2013), a granulometria adequada para uma mistura asfáltica é aquela que confere o maior grau de empacotamento, garante uma maior estabilidade e contato intergranular, reduzindo os vazios de ar. É necessário atentar-se ao volume de vazios necessário para conferir flexibilidade e garantir a devida incorporação de ligante asfáltico na mistura (Brown *et al.*, 2009). Segundo Yu (2012), o baixo teor de ligante asfáltico e o baixo volume de vazios, podem ocasionar uma baixa resistência da mistura em relação à deformação permanente. Por essa razão, o volume de vazios foi escolhido como uma das variáveis do presente estudo, assim como os percentuais passantes nas peneiras 3/4", 3/8", #4 e #200, além do tamanho nominal máximo (TMN).

Segundo Onofre (2013), modificações feitas nos ligantes asfálticos proporcionam aumento da resistência à deformação permanente das misturas asfálticas. Por essa razão, optou-se em juntar os ligantes não modificados (CAP 30/45 e CAP 50/70) em apenas uma categoria. Os demais tipos de ligante são: asfalto modificado por polímero 60/85 (AMP), asfalto modificado por borracha moída de pneus tipo AB-08 (AB) e ligante asfáltico de alto módulo (AM). A Figura 18 ilustra a distribuição percentual das misturas no conjunto de dados contendo os diferentes tipos de ligante.

Figura 18 – Distribuição das misturas do banco de dados em relação ao tipo de ligante

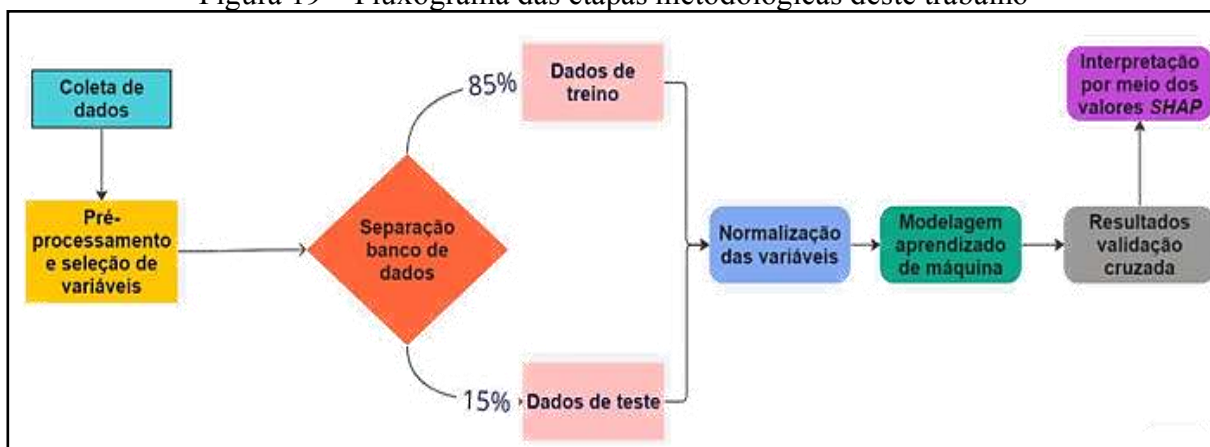


Fonte: O autor

3.2 Método empregado

As etapas metodológicas para a realização do presente trabalho estão ilustradas no fluxograma indicado na Figura 19 e é composto pelas etapas: coleta de dados, pré-processamento e seleção de variáveis, separação do banco de dados de treino e teste, normalização das variáveis, modelagem dos algoritmos de aprendizado de máquina, validação cruzada e interpretação por meio dos valores *SHAP*. Nas próximas seções serão detalhados os processos de cada etapa.

Figura 19 – Fluxograma das etapas metodológicas deste trabalho



Fonte: O autor

3.2.1 Coleta de Dados

A obtenção dos dados para esta pesquisa foi realizada por meio de revisões de trabalhos acadêmicos, publicações científicas e pesquisas em projetos e relatórios rodoviários. A coleta e organização dos dados corresponde a uma etapa longa e meticulosa em um trabalho que envolve ciência de dados, levando meses para ficar devidamente estruturada. Após a coleta dos dados, iniciou-se um processo para organizar todas as informações em planilhas, definindo as variáveis de interesse a serem utilizadas no trabalho. A extração dos dados para a planilha foi realizada de maneira cuidadosa para evitar erros durante o preenchimento. Os dados coletados estão anexados ao final deste trabalho no Apêndice A e encontram-se disponíveis para futuras pesquisas.

3.2.2 Pré-processamento e seleção das variáveis

Nesta etapa os dados foram submetidos a um pré-processamento para se adequarem às limitações da modelagem computacional. Alguns algoritmos conseguem trabalhar diretamente com variáveis categóricas, entretanto, em sua maioria, os modelos necessitam que sejam transformados em valores numéricos. Nesse caso, usou-se codificação *one-hot* para transformar as variáveis categóricas em colunas com variáveis binárias, onde o número 1 representa o valor afirmativo e o 0 negativo. A Figura 20 apresenta a conversão de uma variável categórica para colunas com variáveis binárias.

Figura 20 – Conversão da variável categórica em variáveis numéricas

Original	Codificação <i>one-hot</i>			
Tipo	CAP	AB	AMP	AM
CAP	1	0	0	0
CAP	1	0	0	0
AMP	0	0	1	0
AMP	0	0	1	0
AMP	0	0	1	0
AMP	0	0	1	0
AB	0	1	0	0
AB	0	1	0	0
AB	0	1	0	0

Fonte: O autor

Um modelo computacional envolvendo aprendizado de máquina pode ser modelado de diversas maneiras. Apesar de existirem ferramentas para seleção de melhores hiperparâmetros, há um trabalho empírico no início do processo, seja com a seleção das variáveis na composição do banco de dados, seja na adição ou remoção de dados ou na definição da proporção do conjunto de treino e teste.

Ainda que todas as variáveis escolhidas tenham influência sobre a deformação permanente de misturas asfálticas, esses parâmetros não são capazes, isoladamente, de representar tal comportamento em sua totalidade. Neste sentido, atributos redundantes podem ser removidos, com a finalidade de melhorar o desempenho dos modelos preditivos. Para verificar a relação das variáveis de entrada e de saída, elaborou-se a matriz de correlação, que foi realizada por meio da correlação de Pearson, comumente usada para avaliar a dependência linear entre duas variáveis contínuas.

Para a construção de uma melhor arquitetura para a modelagem, buscou-se encontrar um modelo simplificado com o menor número de parâmetros envolvidos, mas que fosse capaz de oferecer um desempenho satisfatório na representação do fenômeno em estudo. Para isso, foram empregadas 3 configurações de variáveis de entrada (C1, C2 e C3). A configuração C1 possui todas as variáveis do banco de dados, a configuração C2 eliminou as variáveis de percentuais passante nas peneiras e na configuração C3 foram removidas as mesmas da configuração C2, e eliminou-se ainda as variáveis referentes à porosidade FAD e ao tamanho máximo nominal. As variáveis utilizadas em cada configuração estão descritas na Tabela 3.

Tabela 3 – Configurações consideradas para as variáveis de entrada

Configuração	Quantidade de variáveis	Variáveis de entrada consideradas
C1	16	PA, PEN, VIS 135, TL, E* , FAD, Vv%, TMN, P34, P38, P4, P200, CAP, AB, AMP e AM
C2	12	PA, PEN, VIS 135, TL, E* , FAD, Vv%, TMN, CAP, AB, AMP e AM
C3	10	PA, PEN, VIS 135, TL, E* , Vv%, CAP, AB, AMP e AM

Fonte: O autor

3.2.3 Separação banco de dados, Treino e Teste

A literatura estabelece como uma boa regra prática a separação do conjunto em 75% de treino e 25% teste (Muller e Guido, 2016). Esse percentual pode variar, podendo aumentar para até 90% o percentual do conjunto de treinamento ao utilizar a validação cruzada na avaliação do modelo, segundo Muller e Guido (2016). Para a modelagem do algoritmo foi realizada a divisão do conjunto de dados, sendo 85% para treinamento e 15% para teste. Utilizou-se a validação cruzada *k-fold* dividida em 5 subconjuntos para medição do desempenho do algoritmo.

Para a modelagem dos algoritmos, utilizou-se o *Jupyter lab*, ferramenta gratuita e muito utilizada na área de aprendizado de máquina e ciência de dados. O código foi escrito em Python e utilizou-se a biblioteca *scikit-learn* para auxiliar o desenvolvimento do algoritmo.

3.2.4 Normalização das variáveis

Após a etapa de separação dos dados de treino e teste, procedeu-se com a normalização dos dados, que consiste em colocar as variáveis em uma mesma escala para evitar o viés em direção às variáveis com uma ordem de grandeza maior. A normalização objetiva colocar as variáveis dentro de um intervalo de 0 e 1 ou -1 e 1. Nesse trabalho, foi utilizado a normalização min-max, técnica usada para colocar as variáveis no intervalo de 0 e 1. A Equação 9 foi empregada para a normalização.

$$X_{\text{norm}} = \frac{X - X_{\text{Min}}}{X_{\text{Max}} - X_{\text{Min}}} \quad (9)$$

Onde:

X_{norm} = Variável resultante da normalização;

X = Variável a ser normalizada;

X_{Min} = Valor mínimo da variável a ser normalizada;

X_{Max} = Valor máximo da variável a ser normalizada.

A Tabela 4 ilustra os valores de uma mistura após o processo de normalização das variáveis

Tabela 4 – Exemplo de normalização dos dados

Variável	Valor original	Valor normalizado
PEN	54	0,71
PA	58,8	0,48
VIS 135	1258	0,82
TL	5	0,21
E*	6860	0,26
FAD	42	0,52
Vv%	4	0,49
TMN	19	0,19
P34	97,4	1,00
P38	72,6	0,59
P4	54,5	0,39
P200	3,9	0,23

Fonte: O autor

3.2.5 *SMOTE (Técnica de Super Amostragem Sintética da Minoria)*

Desenvolvida por Chawla *et al.* (2002), a técnica SMOTE efetua uma super amostragem da classe alvo, criando novas amostras sintéticas, ao invés de selecionar amostras. Em trabalhos de modelagem supervisionada com foco na classificação, é preferível que haja um equilíbrio na quantidade das classes de saída de modo a evitar que o modelo seja enviesado em direção a uma classe específica, o que pode ocorrer quando uma determinada classe tem uma quantidade muito maior de dados que outra. No entanto, alguns conjuntos de dados podem conter um desequilíbrio em relação às classes de saída. O treinamento de um modelo de classificação com a variável desbalanceada pode resultar em uma previsão com viés, ou seja, a classe dominante vai se sobressair em relação às demais classes de menor frequência (Chawla *et al.*, 2002). O modelo construído usando dados desbalanceados pode apresentar acurácia elevada, porém não prever corretamente uma observação para a classe com menor frequência. Como resultado, pode-se interpretar erroneamente que o modelo possui uma boa performance.

A literatura acerca do tema aborda a questão da desproporção de classes. Uma das formas de lidar com um conjunto de dados desbalanceado é produzir uma amostra a partir do conjunto de dados original, seja por *oversampling* ou *undersampling* da classe majoritária (Japkowicz, 2000). O *oversampling* é uma técnica que consiste em aumentar a quantidade de dados de treino da classe de frequência para que a base de dados possua uma quantidade equilibrada entre as classes de variáveis de saída. Para evitar que existam dados idênticos, pode ser usada a técnica *SMOTE*, que consiste em sintetizar novas informações com base nos dados já existentes. Esses dados sintéticos podem ser muito próximos aos dados reais, mas não são iguais.

3.2.6 *Modelagem Via RNA e XGBOOST*

3.2.6.1 *Ajuste de hiperparâmetros do XGBOOST*

A modelagem utilizando o algoritmo *XGBoost* envolve a definição dos hiperparâmetros do modelo e o treinamento do *ensemble* de árvores de decisão. Hiperparâmetros são atributos do modelo que devem ser definidos antes do processo de treinamento. Existem diferentes técnicas que buscam otimizá-los para encontrar uma melhor

acurácia para o modelo. No presente trabalho foi usada a ferramenta *GridSearchCV*, contida na biblioteca *scikit-learn*. A grade de hiperparâmetros foi escolhida entre intervalos para cada um dos hiperparâmetros. Buscou-se escolher valores próximos dos valores padrões do *XGBoost*, a fim de não causar um *overfitting* ou um *underfitting*, mas que fossem valores que fizessem a diferença em relação ao padrão. Foram escolhidos os seguintes valores para cada um:

- *booster*: esse hiperparâmetro determina se o algoritmo utilizará árvores de decisão ou um modelo linear generalizado como base para o *ensemble*. A escolha depende do conjunto de dados e do problema específico que está sendo abordado. Foram escolhidos “*gbtree*” e “*dart*”. O padrão do algoritmo é “*gbtree*”.
- *gamma*: é um parâmetro de regularização que penaliza a criação de novos nós na árvore de decisão, reduzindo a sua profundidade. Esse parâmetro controla a quantidade mínima de redução de erro necessária para criar um novo nó na árvore. O intervalo escolhido foi entre 0 e 0,5;
- *learning_rate*: controla a taxa de aprendizado do modelo durante o treinamento. Ajustar esse hiperparâmetro é importante para encontrar um equilíbrio entre a velocidade de convergência e o risco de *overfitting*. Foram utilizados valores entre 0,1 e 0,5, sendo 0,5 o predefinido no *XGBoost*;
- *Max_depth*: é usado para controlar a profundidade máxima de cada árvore no *ensemble*. Ele determina o número máximo de camadas que uma árvore pode ter, ou seja, quantas perguntas ou divisões podem ser feitas antes de tomar uma decisão final. Um valor maior desse hiperparâmetro permite que as árvores sejam mais profundas e mais complexas, pois podem fazer mais perguntas e capturar relações mais detalhadas nos dados, mas aumenta o risco de *overfitting*. Foi utilizado o intervalo de 3 a 9. O valor padrão do *XGBoost* é utilizar o *Max_depth* como 6;
- *n_estimators*: controla o número de árvores no *ensemble*. É utilizado para controlar a complexidade do modelo e evitar o *overfitting*. A escolha do valor ideal envolve um equilíbrio entre a capacidade de capturar padrões complexos nos dados e a capacidade de generalizar para novos dados. O intervalo escolhido foi entre 100 e 300, sendo 100 o valor padrão.

Outros hiperparâmetros foram modificados manualmente, porém não foram observadas alterações significativas no resultado do modelo e por essa razão optou-se por não colocá-los na grade de hiperparâmetros a serem variados e testados. Para a modelagem do

algoritmo foi realizada a divisão do conjunto de dados de maneira aleatória e sem repetição, sendo 85% para treinamento e 15% para teste. Após a divisão do conjunto de dados em treino e teste, foi iniciado o treinamento utilizando o *XGBOOST* para modelar um algoritmo de classificação. A configuração dos hiperparâmetros está ilustrada na Tabela 5.

Tabela 5 – Hiperparâmetros utilizados no modelo *XGBoost*

Hiperparâmetros	Valores testados	Opção utilizada
<i>booster</i>	<i>gbtree, dart</i>	<i>gbtree</i>
<i>gamma</i>	0 - 0,5	0
<i>learning_rate</i>	0,1 - 0,5	0,1
<i>max_depth</i>	3 - 9	3
<i>n_estimators</i>	100 - 300	100

Fonte: O autor

3.2.6.2 Ajuste de hiperparâmetros da RNA

O código usado na montagem da RNA foi escrito em *Python* e para auxiliar nas etapas de treinamento e teste do modelo. Utilizou-se as bibliotecas de aprendizado de máquina *Scikit-Learn*. Para tanto optou-se pelo uso de uma RNA do tipo *perceptrons* de múltiplas camadas (MLP).

- *activation*: função de ativação aplicadas aos neurônios em cada camada da rede neural, incluindo a camada de entrada. A função de ativação introduz a não-linearidade na rede e ajudar a modelar relações complexas nos dados;
- *learning_rate*: controla a taxa de aprendizado da rede neural;
- *solver*: utilizado para otimizar os pesos da rede durante o treinamento;
- *batch_size*: É o número de amostras de treinamento que são utilizadas em cada passo de atualização dos pesos da rede;
- Número de camadas ocultas: número de camadas intermediárias entre a camada de entrada e a camada de saída da rede. As camadas ocultas são responsáveis por aprender representações intermediárias dos dados;
- Quantidade de neurônios em cada camada oculta: determina a capacidade da rede de aprender e modelar relações complexas nos dados;
- *epochs*: representa o número de vezes que todos os dados de treinamento são apresentados à rede durante o treinamento. Cada época consiste em um ciclo completo

de passar os dados pela rede, calcular o erro e ajustar os pesos.

Os hiperparâmetros obtidos pelo *GridSearchCV* usados na modelagem da RNA, estão detalhados na Tabela 6.

Tabela 6 – Hiperparâmetros utilizados na modelagem da RNA

Hiperparâmetros	Valores testados	Opção utilizada
<i>activation</i> (camada de entrada e camadas ocultas)	<i>linear, sigmoid e ReLU</i>	<i>ReLU</i>
<i>learning_rate</i>	<i>constant, adaptive</i>	ReLU
<i>solver</i>	sgd, Adam	<i>Adam</i>
<i>batch_size</i>	8 - 32	16
Número de camadas ocultas	2 - 3	2
Quantidade de neurônios em cada camada oculta	32 - 256	256
<i>epochs</i>	100 - 500	150

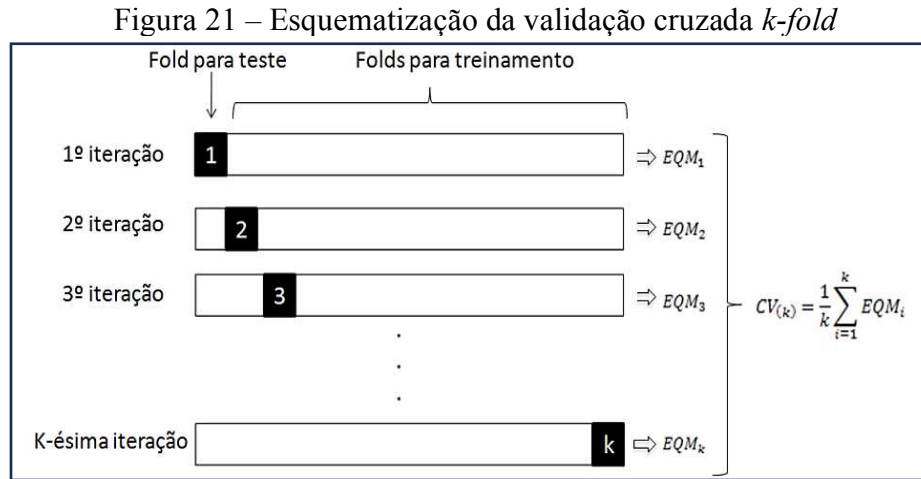
Fonte: O autor

3.2.7 Validação cruzada *k-fold*

É vantajoso usar a validação cruzada em vez de uma única divisão aleatória em um conjunto de treinamento e um conjunto de teste. Na divisão dos dados em treino e teste, pode ocorrer que dados mais “difíceis” de serem classificados caiam no conjunto de treino. Dessa forma, o conjunto de testes seria composto por exemplos mais “fáceis” de classificar, resultando em uma precisão falsamente alta. O contrário também poderia ocorrer resultando em um modelo de baixa precisão (Muller e Guido, 2016). No entanto, ao usar a validação cruzada, cada exemplo estará no subconjunto de treinamento exatamente uma vez. Esse subconjunto é usualmente chamado de dobra. Dessa forma, cada exemplo está em uma das dobras e cada dobra é o conjunto de teste uma vez.

A validação cruzada é uma técnica bastante usada para estimar o erro de teste associado ao ajuste de um determinado método de aprendizagem em um conjunto de observações. Ajuda na avaliação do modelo e é útil para garantir que esse não esteja em *overfitting*. A validação cruzada *k-fold* é utilizada para encontrar os melhores valores de

hiperparâmetros que produzem um desempenho de generalização satisfatório. A Figura 21 ilustra a divisão dos subconjuntos de treino e teste na validação cruzada *k-fold*.



3.2.8 Análise dos resultados

Optou-se por modelar algoritmos de classificação, com intervalos de FN divididos em tipos de tráfego. A decisão foi justificada devido a elevada variância no valor do FN dentro do conjunto de dados estruturado no presente trabalho. Além disso, o valor do FN pode variar significativamente em ensaios utilizando misturas idênticas. Portanto, dividiu-se os resultados em intervalos correspondentes ao tipo de tráfego empregado. Conforme já mencionado, foram escolhidas 3 classes correspondentes às misturas conforme a resistência das mesmas quanto à deformação permanente, todavia divididas por tipo de tráfego que as mesmas suportariam para não atingirem níveis críticos desse defeito específico. Os níveis de tráfego considerados foram os seguintes: médio, pesado e extremamente pesado. Foi adotado o critério de correspondência entre FN e tráfego definido por Bastos *et al.* (2017).

O coeficiente de correlação Pearson foi utilizado para avaliar o grau de relação linear entre cada par de variáveis. Os valores variam no intervalo de -1 a +1. Valores positivos ilustram uma relação direta entre os dados analisados, enquanto valores negativos indicam uma relação inversa entre o par analisado. O desempenho do modelo é ilustrado por meio de uma matriz de confusão, que é uma tabela que ilustra as frequências de classificação para cada classe do modelo. Essa matriz ilustra de maneira sucinta as seguintes frequências:

- Verdadeiro positivo (VP): quando no conjunto real, a classe buscada foi prevista corretamente;

- Falso positivo (FP): quando no conjunto real, a classe alvo foi prevista incorretamente;
- Verdadeiro negativo (VN): quando no conjunto real, a classe que não se quer obter foi prevista corretamente;
- Falso negativo (FNe): quando no conjunto real, a classe que não se está buscando foi prevista incorretamente.

A acurácia da modelagem diz respeito ao acerto de todas as previsões possíveis. É dada pela razão entre as previsões corretas e todas as previsões realizadas, conforme Equação 10.

$$Acurácia = \frac{VP+VFN}{VP+FP+VN+FNe} = \frac{Previsões\ verdadeiras}{Todas\ as\ previsões} \quad (10)$$

Outra métrica utilizada foi o *recall*, que corresponde ao quão bom o modelo é em prever resultados positivos, ou seja, a capacidade de prever positivamente a classe desejada. O resultado é a razão entre os verdadeiros positivos e a soma de verdadeiros positivos com falsos negativos, conforme mostra a Equação 11.

$$Recall = \frac{VP}{VP+FNe} \quad (11)$$

A precisão do modelo refere-se à proporção de identificações corretamente previstas. Ela corresponde a razão entre verdadeiros positivos e a soma de verdadeiros positivos e falsos positivos, como indica a Equação 12.

$$Precisão = \frac{VP}{VP+FP} \quad (12)$$

A métrica F1 ou *F1-score* considera tanto a precisão quanto o *recall*, sendo definida pela média harmônica entre as duas, como aponta a Equação 13.

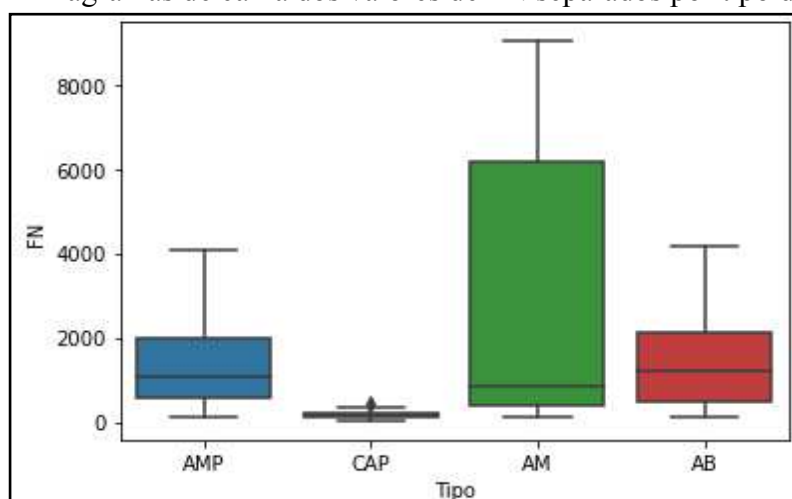
$$F1 = 2 * \frac{Precisão * Recall}{Precisão + Recall} \quad (13)$$

Pode-se interpretar que um modelo que apresenta um bom *F1-score* é capaz tanto de acertar suas predições (precisão alta), quanto de recuperar os exemplos da classe de interesse (*recall* alta). Portanto, essa métrica tende a ser um resumo melhor da qualidade do modelo.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo foi realizada uma análise exploratória do banco de dados por meio da estatística descritiva e análise de correlação de variáveis. Foram apresentados histogramas e diagramas de caixa para melhor visualização da composição e ocorrência das variáveis do banco de dados. A Figura 22 ilustra um diagrama de caixa que descreve a variação dos valores de FN para cada um dos 4 tipos de ligante considerados na base de dados.

Figura 22 – Diagramas de caixa dos valores de FN separados por tipo de ligante



Fonte: O autor

Além de ilustrar a distribuição dos valores de FN, o diagrama de caixas da Figura 22 mostrou os valores atípicos encontrados em cada ligante. O asfalto de alto módulo (AM) foi o único que não apresentou *outliers*. Destaca-se também que o ligante puro (CAP) corresponde às misturas que apresentaram os menores valores de FN.

4.1 Análise Descritiva dos Dados

Conforme já mencionado, o banco de dados utilizado na presente pesquisa é composto por 251 amostras de misturas, com 16 variáveis numéricas e 1 variável de saída, o *flow number*. Na análise descritiva, as variáveis relacionadas ao tipo de ligante não foram utilizadas. A Tabela 7 traz a estatística descritiva das variáveis numéricas. Dessa forma, pôde-se extrair informações de valores mínimos, valores máximos, média, desvio padrão, variância e coeficiente de variação (CV). A saída do modelo corresponde a 1 variável categórica dividida

em 3 classes de tráfego, de acordo com o critério anteriormente citado na Tabela 1.

Tabela 7 – Estatística descritiva do banco de dados

Variável	Mínimo	Máximo	Média	Desvio Padrão	Variância	CV (%)
PEN	20	69	43,3	12,7	162,2	29,3
PA	47,1	88,1	59,2	7,3	53,2	12,3
VIS 135	282	2300	976,9	502,8	252808	51,5
TL	4	7,3	5,1	0,6	0,4	11,8
E*	5480	17976	10284,2	2181,1	4757225	21,2
FAD	30	77,7	48,9	10,7	114,7	21,9
Vv%	2,2	5,7	4,2	0,4	0,2	9,5
TMN	9,5	25	13,4	4,1	16,5	30,6
P34	83,3	100	99,7	1,3	1,7	1,3
P38	53,6	100	83,2	11,1	122,3	13,3
P4	30	77,7	49,1	10,54	111,1	21,5
P200	3,2	8,5	6,1	1,1	1,2	18,0
FN	27	9027	1388,3	1944,2	3779771	140,0

Fonte: O autor

As variáveis granulométricas apresentaram menores valores de desvio padrão, o que pode ser explicado pelo fato de diferentes misturas terem sido dosadas com granulometrias iguais. O módulo dinâmico da mistura (|E*|) foi a variável que apresentou maior variância, possivelmente pelos valores correspondentes ao ligante de alto módulo (AM) no conjunto de dados. Outra propriedade desse ligante é a baixa penetração, que deve variar entre 15-25 décimos de milímetro (ABNT NBR 6576).

Para ilustrar a distribuição da frequência das variáveis, foram gerados os histogramas das variáveis numéricas do conjunto de dados. A Figura 23 resume a distribuição das variáveis.

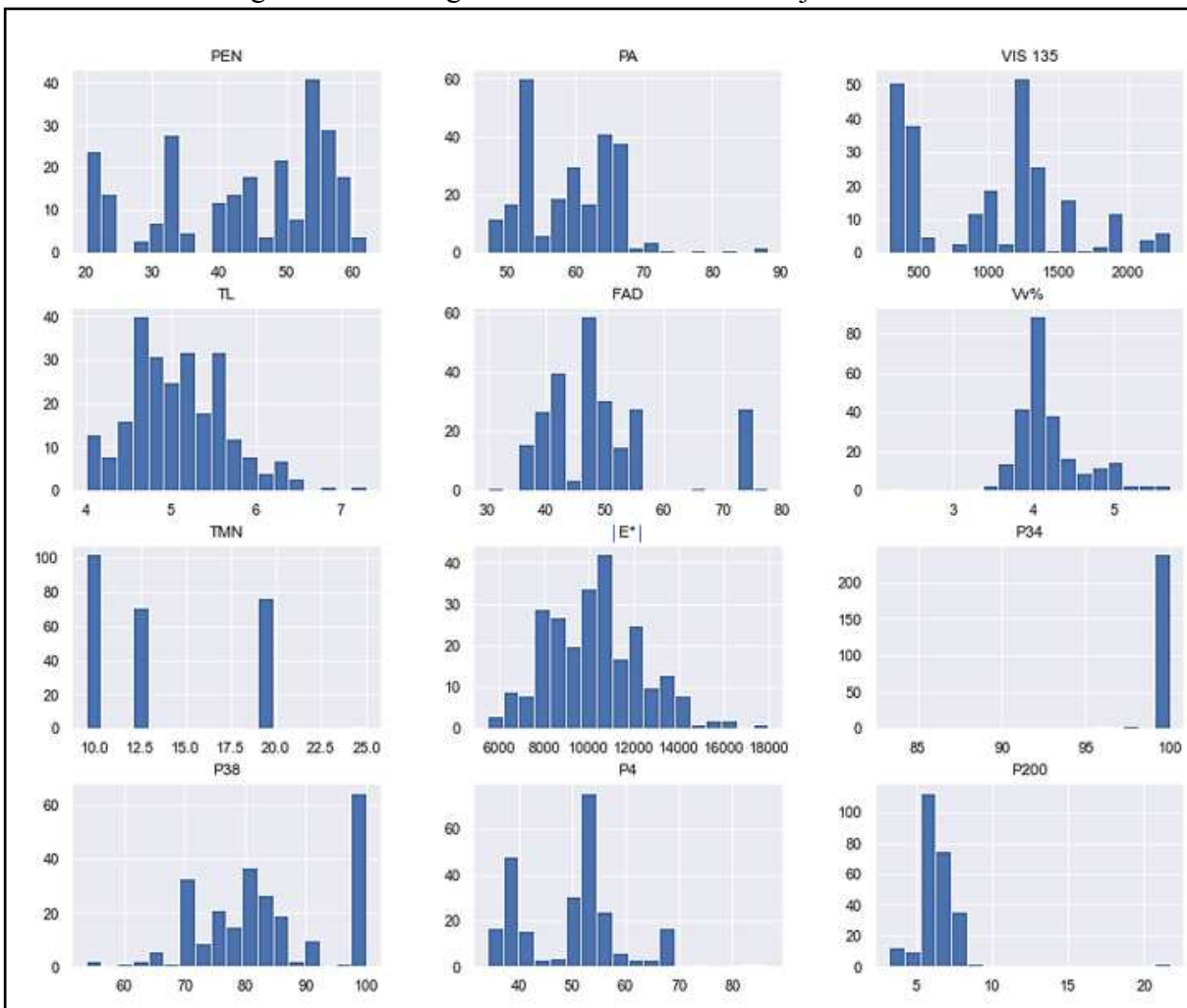
Algumas observações podem ser feitas a partir da análise dos histogramas:

- a) As variáveis TMN, P34 e P200 apresentaram uma menor variação em relação à distribuição. Isso pode indicar que essas variáveis não exercem um impacto significativo para a modelagem por meio de aprendizado de máquina.
- b) As variáveis PEN, VIS 135, TL e |E*| apresentaram uma maior amplitude

de variação. As variáveis com frequências distribuídas de maneira mais equilibrada, tendem a impactar significativamente no desempenho do modelo computacional.

- c) A variável P34 apresenta a menor variação. Isso sugere que sua remoção pode ser realizada sem gerar um impacto significativo nos modelos.

Figura 23 – Histogramas das variáveis no conjunto de dados



Fonte: O autor

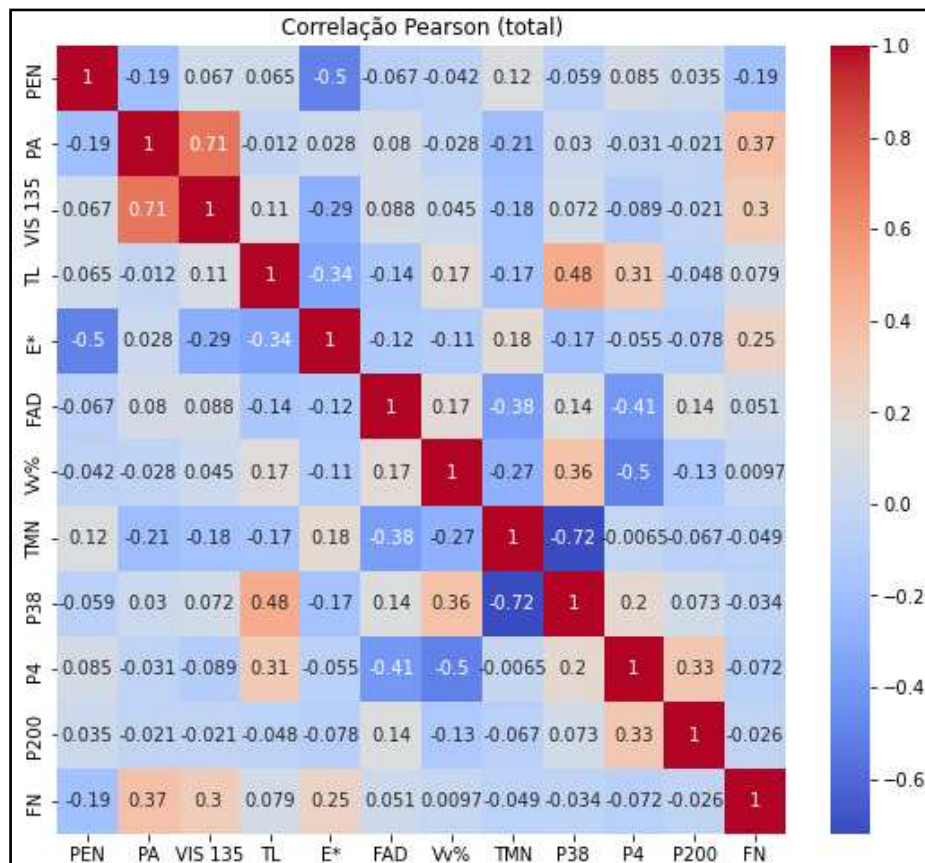
Além da distribuição das variáveis, os histogramas permitem avaliar a variabilidade dos valores das variáveis dentro do banco de dados. Dessa forma, pode-se identificar possíveis atributos que não irão contribuir significativamente na modelagem do algoritmo preditivo. A apresentação dos histogramas é uma etapa importante do trabalho, que ajuda de maneira prévia a selecionar as melhores variáveis que irão compor os modelos de previsão.

4.1.1 Relações Estatísticas entre as Variáveis

4.1.1.1 Relações estatísticas para todo o conjunto de dados

Para analisar o nível de correlação entre as variáveis consideradas neste estudo, procedeu-se com a elaboração de mapas de calor que ilustram a correlação de Pearson entre as variáveis de entrada com a variável de interesse FN. Para Cohen (1988), valores entre 0,10 e 0,29 indicam correlações pequenas entre as variáveis, se os valores estiverem entre 0,30 e 0,49 podem ser considerados como correlações médias e se os resultados estiverem entre 0,50 e 1 podem ser interpretados como fortes correlações. No entanto, recomenda-se utilizar todos os dados, até mesmo os com menores correlações, para verificar a precisão da modelagem. A Figura 24 mostra a correlação considerando todo o banco de dados.

Figura 24 – Matriz de correlação Pearson das variáveis do banco de dados



Fonte: O autor

Nesse mapa de calor, as variáveis: PA, VIS 135 e |E*| apresentam as maiores

correlações em relação ao FN. Destaca-se novamente que o conjunto de 251 amostras contém diferentes tipos de ligante e que misturas com ligantes modificados apresentam comportamentos diferentes das misturas com ligante puro. Portanto, a Figura 24 ilustra o comportamento das variáveis sob o ponto de vista geral.

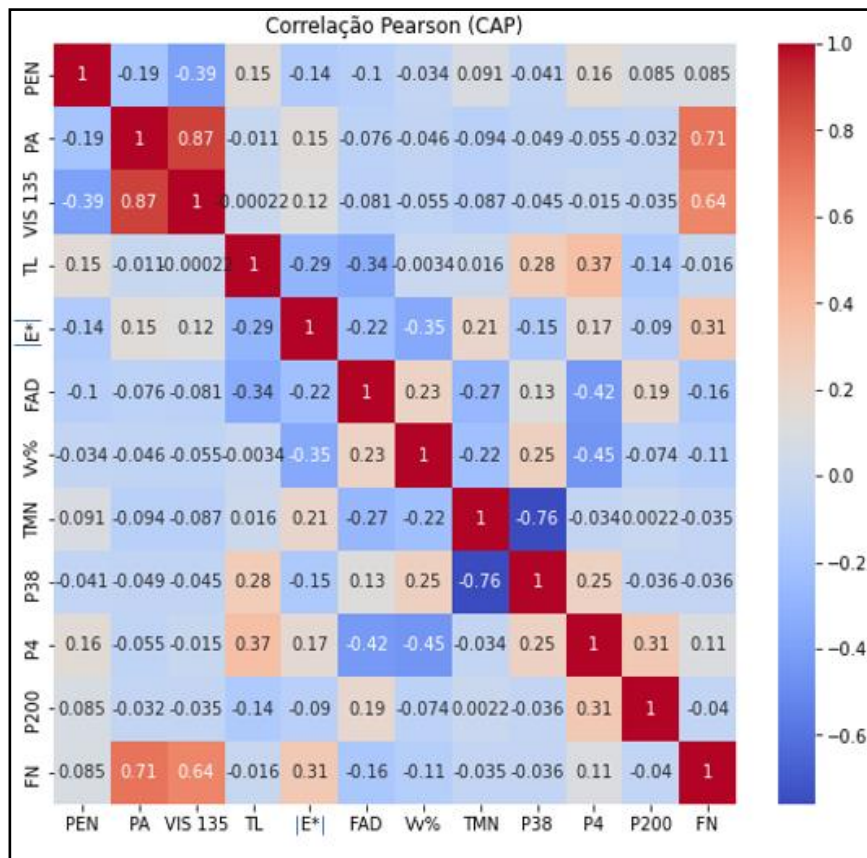
Nota-se a influência da viscosidade e do ponto de amolecimento do ligante em relação à deformação permanente da mistura. As misturas com ligantes menos viscosos podem produzir uma lubrificação maior no esqueleto pétreo e prejudicar o intertravamento dos agregados, tornando a mistura menos rígida e mais sujeita à deformação permanente. A análise de correlação mostrou também que ligantes mais resistentes à penetração (menor PEN) tendem a gerar misturas mais resistentes à deformação permanente. Observou-se que as variáveis que conferem propriedades mais rígidas à mistura, tenderam a elevar a resistência à deformação, o que era esperado.

Por outro lado, em um conjunto de dados com diversos ligantes, as variáveis granulométricas mostraram uma correlação baixa em relação ao valor de FN. Todavia, essas variáveis não devem ser desprezadas. Neto (2018) ressaltou que o método FAD pode avaliar e especificar granulometrias de misturas asfálticas para melhor resistir à deformação permanente e está satisfatoriamente associado ao FN desde que tenham matriz granular, composição e características do ligante semelhantes.

4.1.1.2 Relações estatísticas para ligante puro (CAP)

Para analisar as correlações apenas com o CAP, foi gerado um mapa de calor com os valores de correlação das misturas que contém apenas o ligante puro. No banco de dados estruturado há 89 misturas asfálticas dosadas com esse tipo de ligante. Na figura abaixo é possível notar que há uma forte correlação das variáveis ponto de amolecimento e viscosidade do ligante a 135°C, o que evidencia a grande influência das propriedades do ligante asfáltico na previsão do comportamento de uma mistura asfáltica contendo ligante puro. Com relação as variáveis referentes a granulometria como: TMN, P38, P34 e P200, observou-se uma baixa correlação com a variável FN. A Figura 25 ilustra o mapa de calor com os valores de correlação de Pearson das variáveis das misturas apenas com CAP.

Figura 25 – Matriz de correlação Pearson das variáveis com ligante CAP



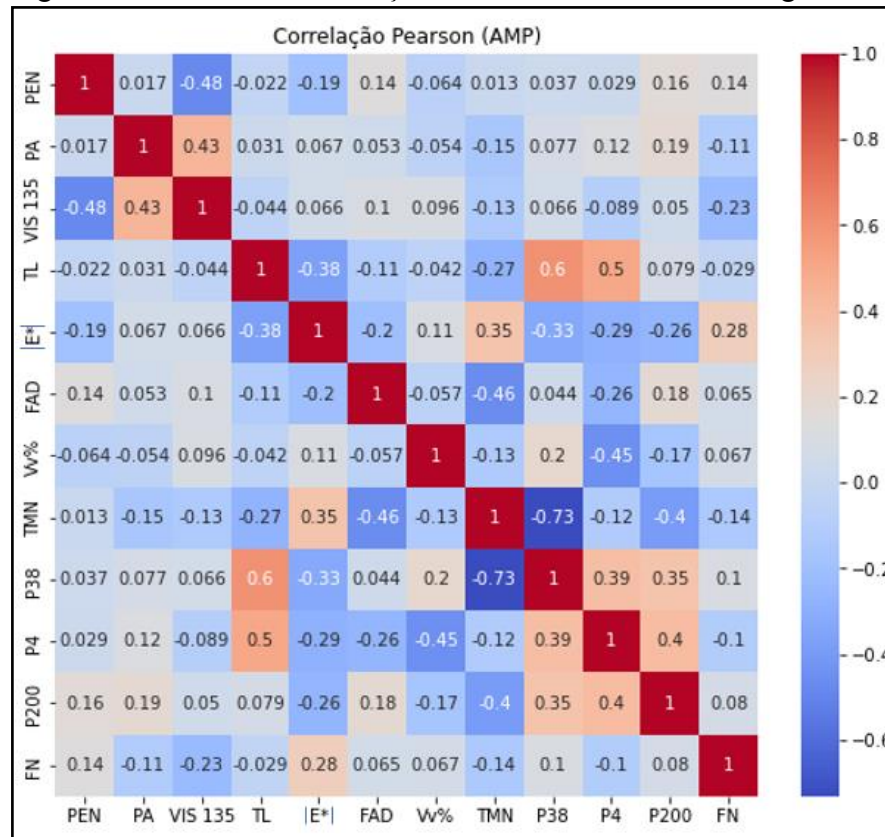
Fonte: O autor

Embora tenha apresentado um valor mais baixo, o módulo dinâmico exerceu uma grande influência em relação ao desempenho da mistura à deformação permanente. Destaca-se também a leve correlação negativa do volume de vazios, o que implica que menores teores de vazios geram maiores valores de FN.

4.1.1.3 Relações estatísticas para asfalto modificado por polímero

O mapa de calor para as 76 misturas asfálticas compostas por ligante modificado por polímero está ilustrado na Figura 26. A variável que mais exerceu influência positiva nos valores de FN para esse tipo de ligante específico foi o módulo dinâmico. Destaca-se também que o valor da viscosidade obteve a maior correlação negativa, junto com o tamanho máximo nominal dos agregados. Uma possível explicação seria que um ligante menos viscoso e com agregados menores, facilitaria o intertravamento da mistura, favorecendo um melhor desempenho em relação à deformação permanente.

Figura 26 – Matriz de correlação Pearson das variáveis do ligante AMP



Fonte: O autor

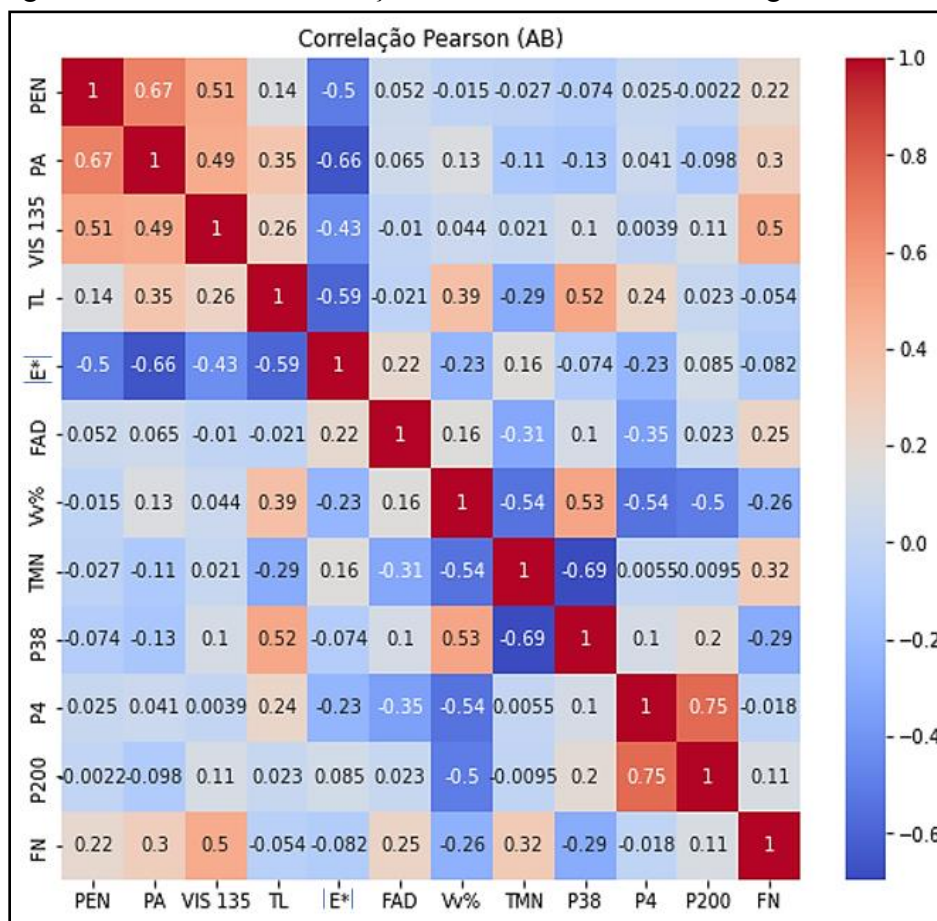
4.1.1.4 Relações estatísticas para o asfalto borracha

Posteriormente, gerou-se o mapa de calor contendo as correlações entre as variáveis para as 48 misturas asfálticas contendo asfalto borracha. No mapa de calor da Figura 27, referente às misturas com asfalto borracha, as variáveis viscosidade, ponto de amolecimento e tamanho máximo nominal exerceram as maiores influências positivas, destacando-se também a penetração como uma variável influente. As variáveis volume de vazios e material passante na peneira 3/8 obtiveram as maiores correlações negativas.

Uma possível explicação para esses valores seria que a alta viscosidade do asfalto borracha exerceria um papel fundamental na resistência da mistura à deformação permanente. Por outro lado, essa viscosidade elevada pode dificultar o preenchimento do ligante no esqueleto pétreo e dificultaria o intertravamento e consequentemente aumentaria os vazios da mistura. Esse comportamento justificaria a correlação negativa da variável volume de vazios nos valores de FN e o leve valor positivo nos valores de % passante na peneira 200. O bom desempenho

dessa mistura ocorre quando há uma alta viscosidade do ligante, aliada a um bom intertravamento entre os agregados.

Figura 27 – Matriz de correlação Pearson das variáveis do ligante AB



Fonte: O autor

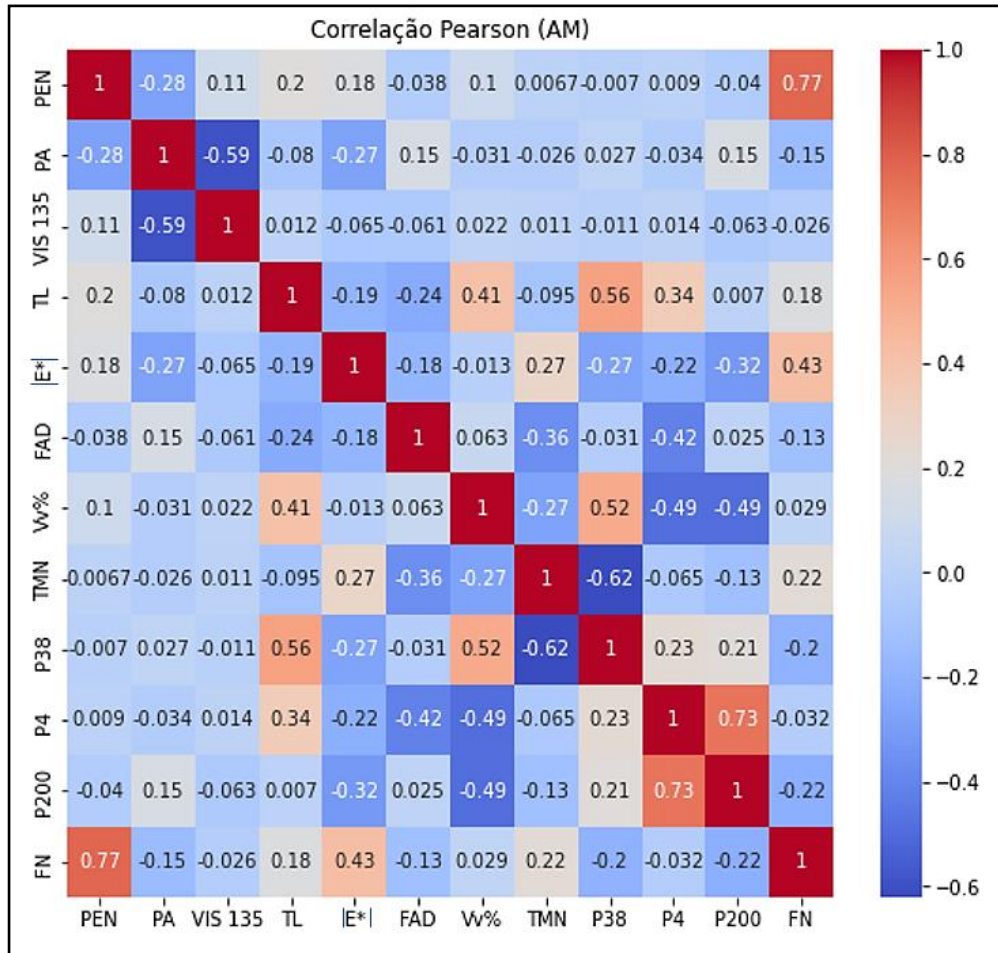
4.1.1.5 Relações estatísticas para o ligante de alto módulo

O quarto tipo de ligante utilizado na pesquisa é o AM e o mapa de calor das variáveis das 38 misturas contidas no banco de dados estruturado está apresentado na Figura 28. As variáveis que exerceram as maiores influências positivas foram a penetração, o módulo dinâmico e o tamanho máximo nominal dos agregados. Enquanto as variáveis % passante na peneira 3/8 e % passante na peneira 200 exerceram as maiores correlações negativas.

O asfalto de alto módulo é caracterizado, como o próprio nome indica, por possuir valores elevados de módulo dinâmico, o que confere uma maior rigidez ao ligante e, por conseguinte, à mistura correspondente. Apesar da forte correlação positiva da variável

penetração, o ligante de alto módulo possui um curto intervalo de valores de penetração que varia entre 20 e 24 decímetros de milímetro, o que dificultaria tirar conclusões acerca dessa variável.

Figura 28 – Matriz de correlação Pearson das variáveis do ligante AM



Fonte: O autor

4.2 Desempenho da Rede Neural Artificial

Foram desenvolvidos 3 modelos de RNA para realizar a classificação do FN das misturas de acordo com o tipo de tráfego recomendado para que as mesmas não desenvolvam o tipo de defeito analisado, qual seja, a deformação permanente. Os resultados das topologias são apresentados alterando-se as variáveis de entrada em busca de uma relação otimizada entre o número de parâmetros de entrada e o desempenho do algoritmo. A divisão do conjunto de dados foi realizada por meio da validação cruzada, sendo 85% para treinamento e 15% para

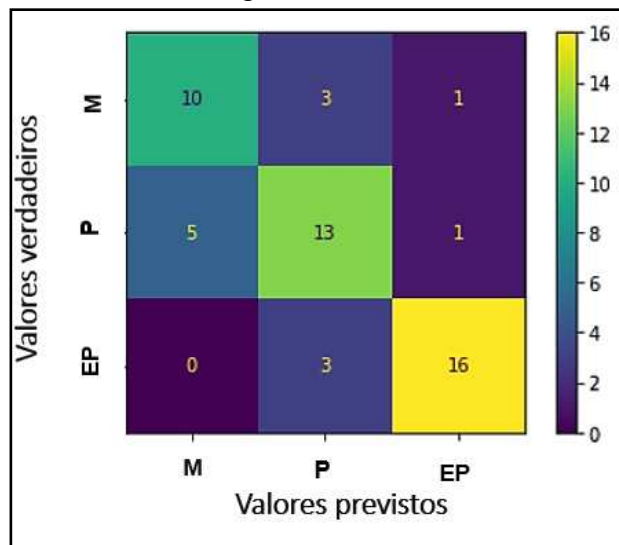
teste. O desempenho do modelo foi avaliado por meio de uma matriz de confusão que é uma tabela resumo com as métricas citadas na metodologia.

Primeiramente gerou-se uma matriz de confusão, que é uma tabela bastante utilizada em verificação de algoritmos de classificação. Por meio dela é possível avaliar as classificações que foram realizadas corretamente, mas também permite verificar as classificações incorretas. Gerou-se uma tabela resumo com as métricas: precisão, *recall* e *F1-score*. E por fim avaliou-se o modelo por meio da validação cruzada *k-fold* dividida em 5 partições.

4.2.1 Modelo de RNA para a configuração C1

Conforme informado no capítulo anterior, a configuração C1 possui todas as variáveis do banco de dados, sendo assim o modelo neural para a configuração C1 é uma RNA do tipo MLP. As variáveis referentes a configuração C1 são: PA, PEN, VIS 135, TL, |E*|, FAD, Vv%, TMN, P34, P38, P4, P200, CAP, AB, AMP e AM. Os resultados para a configuração C1 podem ser observados na matriz de confusão ilustrada na Figura 29.

Figura 29 – Matriz de confusão para o modelo de RNA na configuração C1



Fonte: O autor

Nessa matriz de confusão, as linhas correspondem ao número de amostras de referência contidas no conjunto de teste. Na Figura 29, a primeira linha, correspondente à classe de tráfego M (Médio), tem o total de 14 misturas (10 + 3 + 1). Na linha abaixo, que corresponde

à classe P (Pesado), há um total de 19 misturas no conjunto de teste ($5 + 13 + 1$). Na última linha foram registradas 19 ($3 + 16$) misturas da classe EP (Extremamente Pesado).

As colunas representam os valores do total de misturas classificadas. Na primeira coluna da matriz da Figura 29, referente à classe M, pode-se contar 15 amostras ($10 + 5$). Na segunda coluna, referente à classe P, a quantidade de misturas classificadas foi de 19 ($3 + 13 + 3$). E na terceira coluna, que responde à classe EP, o número total de previsões foi 18 ($1 + 1 + 16$).

Deve-se atentar principalmente para a diagonal da matriz de confusão, pois nela estão contidos os valores previstos corretamente pelo modelo. A precisão de cada classe é definida pela razão entre o valor previsto corretamente e pelo número total de amostras classificadas. No primeiro número da diagonal, em relação à classe M, tem-se o valor 10. Conforme descrito anteriormente, cada coluna representa o total de amostras usada no teste. Portanto, a precisão da classe M é a razão entre o número de amostras previstas corretamente (10), dividida pelo número total de amostras (15), que resulta em 0,67 ou 67% de precisão na classe M.

No segundo número da diagonal, pode-se analisar a precisão do modelo referente à classe P. A precisão é dada pela razão entre o número de misturas classificadas corretamente (13) e o total de misturas previstas para essa classe, 19, o que resulta no valor de 0,68 ($13/19$). Para finalizar, na classe EP, há 16 amostras previstas corretamente em um total de 18, resultando em uma precisão de 0,89.

Outra métrica que pode ser obtida na análise da matriz de confusão é o *recall*, que permite avaliar o quão bom o modelo prevê resultados positivos corretamente. Considera-se como resultados positivos as previsões corretas (verdadeiro positivo) e as previsões que foram classificadas como negativas de maneira incorreta (falso negativo). O cálculo do *recall* para a classe M é dado pela razão entre as misturas classificadas corretamente (10), pelo total de amostras reais contidas no conjunto de teste que nesse caso são 14 ($10 + 3 + 1$), que resulta em 71%.

Para a classe P, o valor do *recall* é a razão entre previsões corretas (13) e o número total de misturas de classe P contidas no conjunto de testes, ou seja, 19 misturas ($5 + 13 + 1$), que resulta em 0,68. Para a classe EP, o *recall* corresponde às 16 misturas classificadas corretamente e o número total de misturas no conjunto de teste (19). Para a classe EP, portanto, o *recall* foi de 0,84 ($16/19$).

A Tabela 8, resume os valores das métricas de precisão, *recall* e *F1-score*. A validação cruzada *k-fold* dividida em 5 partições foi utilizada na etapa de validação.

Tabela 8 – Tabela de desempenho do modelo de RNA para a configuração C1

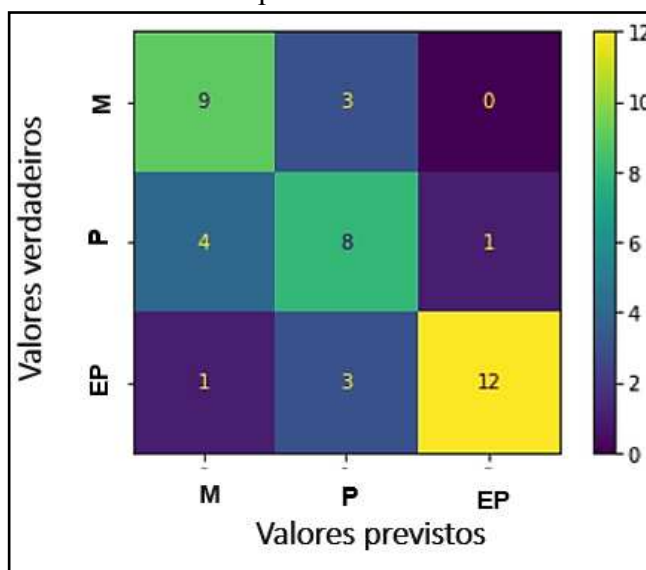
Classe de tráfego	Precisão	Recall	F1-score
M	0,67	0,71	0,69
P	0,68	0,68	0,68
EP	0,89	0,84	0,86
Acurácia			0,75
Validação			0,73

Fonte: O autor

4.2.2 Modelo de RNA para a configuração C2

Conforme informado no capítulo anterior, o modelo neural para a configuração C2 é uma RNA do tipo MLP com as seguintes variáveis na camada de entrada: PA, PEN, VIS 135, TL, |E*|, FAD, Vv%, TMN, CAP, AB, AMP e AM. Os resultados para a configuração C2 podem ser observados na matriz de confusão ilustrada na Figura 30.

Figura 30 – Matriz de confusão para o modelo de RNA na configuração C2



Fonte: O autor

A Tabela 9 contém o resumo dos resultados das métricas obtidas na matriz de confusão da configuração C2. Da mesma forma anterior, os índices de acertos foram estimados

pelo conjunto de teste e o desempenho medido através das métricas de precisão, *recall* e *F1-score*. Para a etapa de validação, usou-se a validação cruzada *k-fold* dividida em 5 subconjuntos.

Tabela 9 – Tabela de desempenho do modelo de RNA para a configuração C2

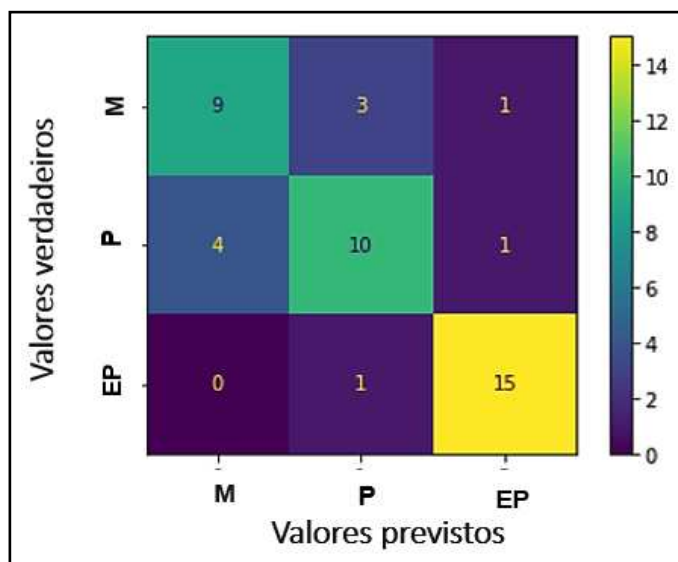
Classe de tráfego	Precisão	<i>Recall</i>	<i>F1-score</i>
M	0,64	0,75	0,69
P	0,57	0,62	0,59
EP	0,92	0,75	0,83
Acurácia			0,71
Validação			0,75

Fonte: O autor

4.2.3 Modelo de RNA para o modelo C3

Conforme informado no capítulo anterior, o modelo neural para a configuração C3 é uma RNA do tipo MLP com as seguintes variáveis na camada de entrada: PA, PEN, VIS 135, TL, |E*|, Vv%, CAP, AB, AMP e AM. Os resultados para a configuração C3 podem ser observados na matriz de confusão ilustrada na Figura 31 e o desempenho do modelo está resumido na Tabela 10. Novamente, os índices de acertos foram estimados pelo conjunto de teste e o desempenho foi medido através das métricas *recall* e *F1-score*. A validação cruzada *k-fold* dividida em 5 partições foi usada na etapa de validação.

Figura 31 – Matriz de confusão para o modelo de RNA na configuração C3



Fonte: O autor

Tabela 10 – Tabela de desempenho do modelo de RNA para a configuração C3

Classe de tráfego	Precisão	Recall	F1-score
M	0,69	0,69	0,69
P	0,71	0,67	0,69
EP	0,88	0,94	0,91
Acurácia			0,77
Validação			0,77

Fonte: O autor

4.3 Desempenho dos modelos utilizando o *XGBoost*

Foram testadas 3 diferentes configurações utilizando o *XGBoost* para realizar a classificação do FN de acordo com o tipo de tráfego recomendado. Os resultados das topologias são apresentados modificando-se as variáveis de entrada em busca de uma relação otimizada entre o número de parâmetros de entrada e o desempenho do algoritmo.

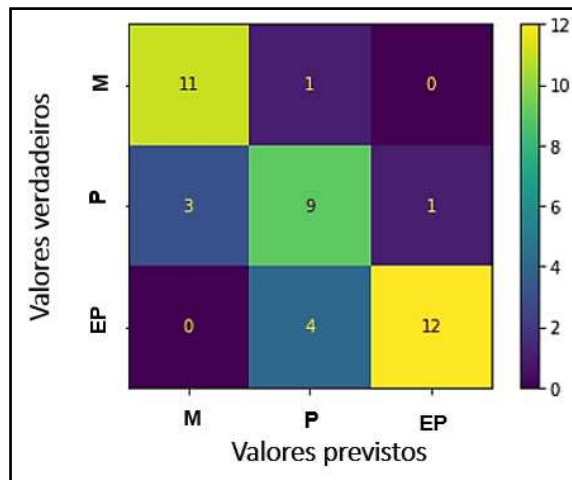
O conjunto de dados foi dividido, sendo 85% para treinamento e 15% para teste. O desempenho do modelo foi avaliado por meio de uma matriz de confusão, tabela resumo com as métricas citadas na metodologia. A validação do modelo foi feita por meio da validação cruzada *k-fold* dividida em 5 subconjuntos.

Além das métricas de avaliação do modelo, a biblioteca *XGBoost* fornece também um gráfico que ilustra a importância das variáveis. A métrica *Split* é o padrão para essa técnica e pode ser obtida contando o número de vezes que a variável foi usada no modelo. Essa métrica mede o número de vezes que uma característica é usada para dividir os dados nas árvores do modelo. Variáveis que são usadas com mais frequência para dividir os dados são consideradas mais importantes (Zhou e Hooker, 2021).

4.3.1 Modelo utilizando o *XGBoost* para a configuração C1

Primeiramente foi desenvolvido um modelo para a configuração C1 com as seguintes variáveis na camada de entrada: PA, PEN, VIS 135, TL, |E*|, FAD, Vv%, TMN, P34, P38, P4, P200, CAP, AB, AMP e AM. Os resultados do modelo podem ser observados na matriz de confusão ilustrada na Figura 32.

Figura 32 – Matriz de confusão para o modelo *XGBoost* na configuração C1



Fonte: O autor

A Tabela 11 contém as métricas referentes ao modelo *XGBoost* na configuração C1. Validou-se o modelo por meio da validação cruzada *k-fold*, particionado em 5 subconjuntos.

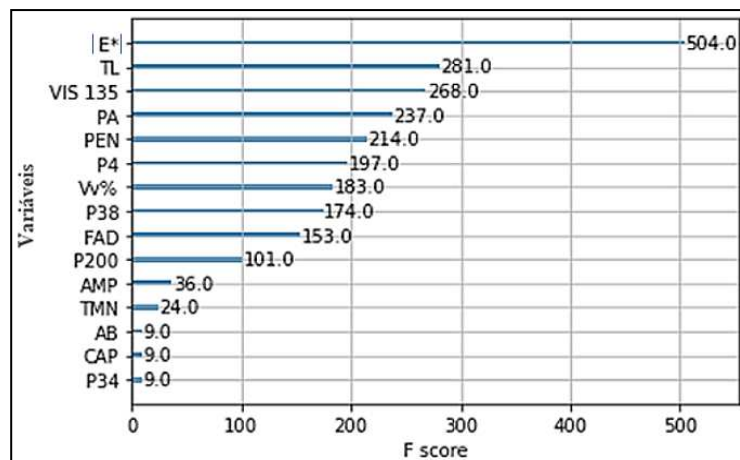
Tabela 11 – Tabela de desempenho do modelo *XGBoost* para a configuração C1

Classe de tráfego	Precisão	Recall	F1-score
M	0,74	1,00	0,85
P	0,63	0,63	0,63
EP	0,86	0,63	0,73
Acurácia			0,73
Validação			0,79

Fonte: O autor

A Figura 33 ilustra o gráfico de relevância das variáveis obtido no modelo *XGBoost* usado na configuração C1. Esse gráfico indica o número de vezes que as variáveis da configuração C1 foram usadas na predição.

Figura 33 – Relevância das variáveis obtido no modelo *XGBoost* usado na configuração C1



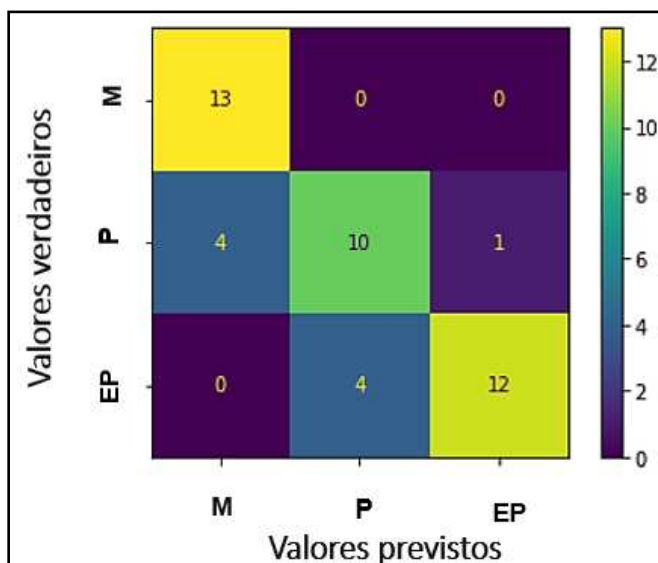
Fonte: O autor

O gráfico de relevância das variáveis da configuração C1 indica que o módulo dinâmico da mistura ($|E^*|$) foi a variável mais relevante no desempenho do algoritmo. As variáveis que aparecem em sequência são referentes à caracterização do ligante asfáltico, como: teor de ligante, viscosidade, ponto de amolecimento e penetração. Entre as variáveis afetadas pela granulometria dos agregados, o volume de vazios foi a que mais contribuiu para a previsão do modelo. O percentual passante na peneira 3/4" e o tipo de CAP foram as variáveis consideradas menos relevantes pelo algoritmo para realizar a classificação.

4.3.2 Modelo utilizando o XGBoost para a configuração C2

Objetivando um código com menos variáveis de entrada, foi desenvolvido um modelo para a configuração C2 com as seguintes variáveis na camada de entrada PA, PEN, VIS 135, TL, $|E^*|$, FAD, Vv%, TMN, CAP, AB, AMP e AM. Os resultados do modelo podem ser observados na matriz de confusão ilustrada na Figura 34.

Figura 34 – Matriz de confusão para o modelo XGBoost na configuração C2



Fonte: O autor

O desempenho matriz de confusão ilustrada na Figura 34, está resumido na Tabela 12. Uma vez mais os índices de acertos foram estimados pelo conjunto de teste, o desempenho medido através das métricas *recall* e *F1-score*, e a validação cruzada *k-fold* dividida em 5 partições.

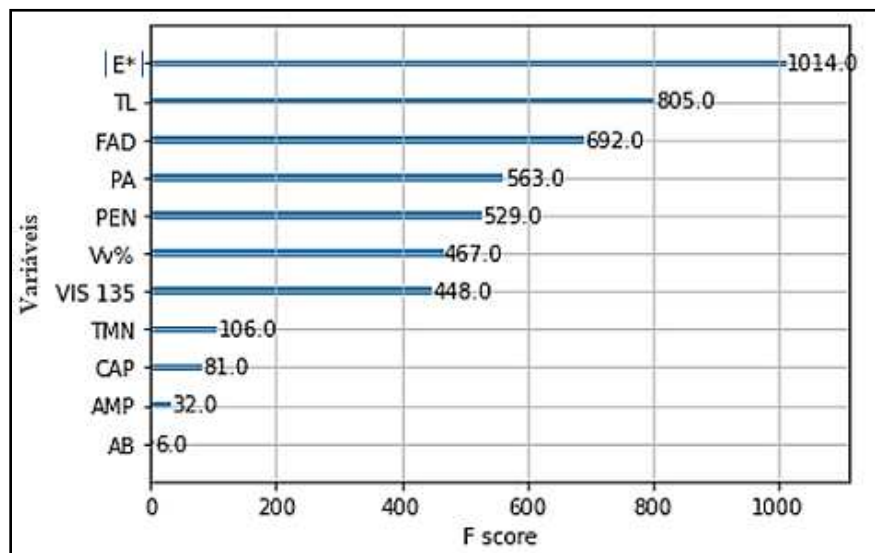
Tabela 12 – Tabela de desempenho do modelo *XGBoost* para a configuração C2

Classe de tráfego	Precisão	Recall	F1-score
M	0,76	1,00	0,87
P	0,71	0,67	0,69
EP	0,92	0,75	0,83
Acurácia			0,80
Validação			0,80

Fonte: O autor

A Figura 35 traz o gráfico de relevância das variáveis conforme obtido no modelo *XGBoost* usado na configuração C2. Esse gráfico indica o número de vezes que as variáveis da configuração C2 foram usadas na predição. A variável |E*| foi a mais relevante para as previsões. As variáveis que aparecem em sequência são: teor de ligante, porosidade FAD, ponto de amolecimento e penetração. Nessa configuração específica, as variáveis correspondentes ao tipo de ligante foram as menos utilizadas pelo modelo.

Figura 35 – Relevância das variáveis obtido no modelo *XGBoost* usado na configuração C2

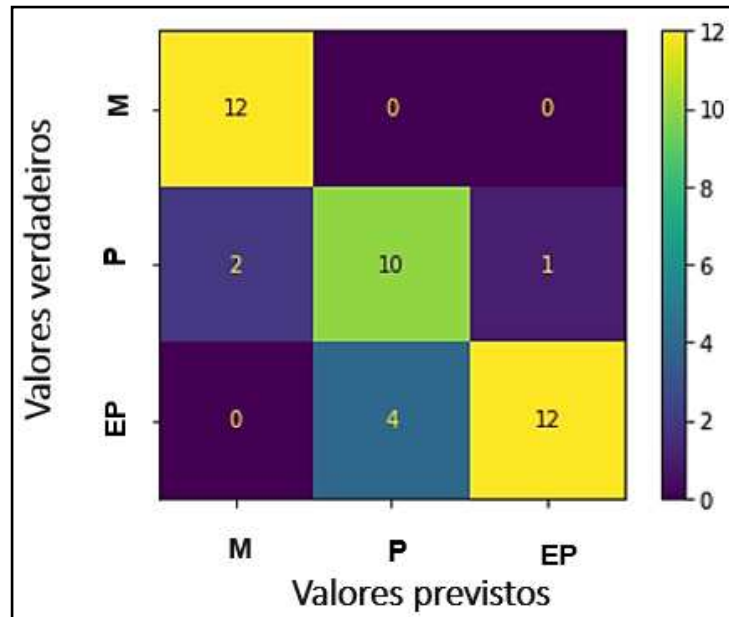


Fonte: O Autor

4.3.3 Modelo utilizando o *XGBoost* para a configuração C3

Por último, foi desenvolvido um modelo para a configuração C3 com as seguintes variáveis na camada de entrada PA, PEN, VIS 135, TL, |E*|, FAD, CAP, AB, AMP e AM. Os resultados podem ser observados na matriz de confusão ilustrada na Figura 36.

Figura 36 – Matriz de confusão para o modelo *XGBoost* na configuração C3



Fonte: O autor

O desempenho do modelo *XGBoost* na configuração C3 está resumido na Tabela 13. Os índices de acertos foram estimados pelo conjunto de teste, o desempenho medido através novamente de recall e F1-score, bem como a validação cruzada *k-fold* dividida em 5 partições foi usada na etapa de validação.

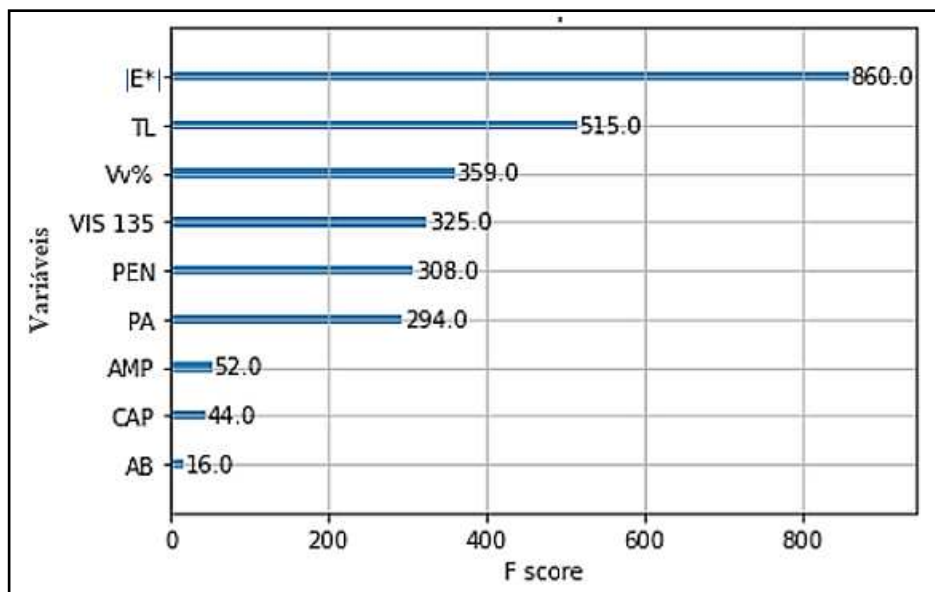
Tabela 13 – Tabela de desempenho do modelo *XGBoost* para a configuração C2

Classe de tráfego	Precisão	Recall	F1-score
M	0,86	1,00	0,92
P	0,71	0,77	0,74
EP	0,92	0,75	0,83
Acurácia do modelo			0,83
Validação			0,84

Fonte: O autor

A Figura 37 ilustra o gráfico de relevância das variáveis obtido no modelo *XGBoost* usado na configuração C3, apontando o número de vezes que as variáveis da configuração C3 foram usadas na predição.

Figura 37 – Relevância das variáveis obtido no modelo *XGBoost* usado na configuração C3



Fonte: O autor

A observação do gráfico referente a relevância das variáveis na configuração C3 evidencia que variável $|E^*|$ também foi a mais relevante para as previsões realizadas no modelo. As variáveis que aparecem em sequência são: teor de ligante, volume de vazios, viscosidade, penetração e ponto de amolecimento. Na configuração C3, as variáveis correspondentes ao tipo de ligante foram as menos usadas pelo modelo. No entanto, há limitações ao visualizar a contribuição das variáveis através do gráfico de relevância das variáveis. Por meio dessa análise não é possível distinguir se a contribuição dessas variáveis impactou positivamente ou negativamente a previsão. Também não é possível distinguir o comportamento de cada variável em relação a classe de saída. Além disso, essa métrica pode ser tendenciosa em relação às variáveis com um grande número de valores únicos e pode não levar em conta a melhoria real no desempenho do modelo proporcionada por cada variável.

4.4 Comparação entre os modelos

Conforme já mencionado, foram consideradas 3 diferentes configurações para estimar a classificação de misturas asfálticas em relação à deformação permanente seguindo o critério de tipo de tráfego. Usaram-se as redes neurais artificiais e também o algoritmo *XGBoost*.

O *GridSearch* foi utilizado para encontrar os melhores hiperparâmetros de cada modelo. A Tabela 14 ilustra o resumo do desempenho de cada configuração nos diferentes algoritmos.

Tabela 14 – Comparação entre os desempenhos dos modelos para as 3 configurações

Configuração	Modelo	Acurácia teste	Validação
C1	RNA	0,75	0,73
	XGBOOST	0,73	0,79
C2	RNA	0,71	0,75
	<i>XGBOOST</i>	0,8	0,8
C3	RNA	0,77	0,77
	<i>XGBOOST</i>	0,83	0,84

Fonte: O autor

Durante a fase de teste, a RNA nas configurações C1, C2 e C3 obteve, respectivamente, valores de: 0,75; 0,71; e 0,77. Já nos modelos utilizando o *XGBoost* os valores referentes a acurácia foram respectivamente de: 0,73; 0,80 e 0,83. Tendo em vista esses valores, observa-se que com exceção da configuração C2 da RNA, a redução do número de variáveis contribuiu para melhorar a acurácia dos modelos desenvolvidos

Na fase de validação, a RNA nas configurações, C1, C2 e C3 obteve valores de 0,73; 0,75; e 0,77, respectivamente. Os modelos desenvolvidos utilizando o *XGBoost* obtiveram os valores 0,79; 0,80 e 0,84. Esses valores evidenciam mais uma vez a melhora do modelo ao reduzir as variáveis, principalmente retirando as variáveis relacionadas com a granulometria.

Pode-se constatar que o modelo desenvolvido com o algoritmo *XGBoost* obteve resultados mais expressivos que as redes neurais artificiais, exceto na fase de teste da configuração C1. Entre todas as combinações testadas, a configuração C3 com as variáveis: PA, PEN, VIS 135, TL, |E*|, Vv%, CAP, AB, AMP e AM, foi a que obteve os melhores resultados em ambos os modelos.

Assim, à luz dos resultados da presente pesquisa, pode-se afirmar que a configuração C3 é a preferível para classificar as misturas em relação à deformação permanente seguindo o critério de tráfego escolhido. Deve-se ressaltar que essa afirmação vale apenas para o banco de dados estruturado durante a pesquisa.

4.5 Gráficos *SHAP*

Embora os modelos de aprendizado de máquina possam ser muito precisos na previsão de resultados, eles nem sempre possuem uma fácil interpretação. Isso pode dificultar a compreensão de como as previsões são realizadas e quais as variáveis são mais importantes. Por meio dos gráficos *SHAP* é possível interpretar a contribuição de cada variável para a saída de um modelo de aprendizado de máquina.

A análise por meio dos valores *SHAP* permite ilustrar a contribuição da variável individualmente em cada mistura. Também é possível interpretar de forma geral o comportamento de todas as variáveis em relação à classe de saída por meio de uma análise global.

A utilização da abordagem *SHAP* pode ajudar a aumentar a transparência e a confiança nos modelos de aprendizado de máquina, permitindo que os usuários entendam como as previsões são feitas e quais variáveis são mais importantes. Isso pode ser especialmente útil em domínios onde a interpretabilidade é importante. Os gráficos gerados pela ferramenta *SHAP* foram obtidos a partir do treinamento de um modelo utilizando o *XGBoost* com todas as variáveis disponíveis no banco de dados.

4.5.1 *Interpretação local – gráficos de força e gráfico cachoeira*

Para uma interpretação local utilizou-se o gráfico de força, que é uma ferramenta de visualização utilizada na interpretação dos valores *SHAP* e que permite avaliar a influência das variáveis em uma única linha de dados, ou seja, uma análise do comportamento das variáveis de uma única mistura asfáltica. Seu uso é muito conveniente para uma análise de erros ou uma melhor compreensão de um caso particular.

O gráfico de força fornece os valores *SHAP* das variáveis e sua influência, positiva ou negativa, para a previsão da classe em questão. As barras são coloridas de acordo com o valor da variável correspondente. Os atributos em verde indicam uma contribuição de forma a aumentar o valor da predição da classe alvo, enquanto os atributos em vermelho sinalizam que aquelas variáveis contribuem de forma a diminuem a predição da classe analisada. Cada mistura asfáltica possui seu próprio gráfico de força.

Foi selecionada uma mistura dentro do banco de dados como exemplo para ilustrar o gráfico de força utilizado para realizar a interpretabilidade local com os valores *SHAP*. As

variáveis dessa mistura e seus respectivos valores, estão descritos na Tabela 15.

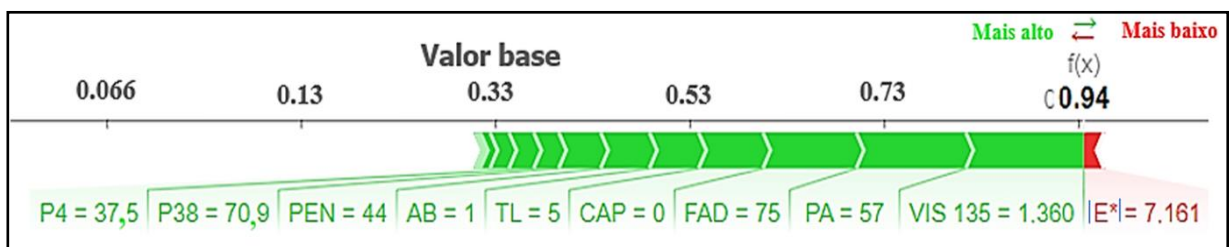
Tabela 15 – Valores das variáveis da mistura utilizada no gráfico de força

Variável	Valor correspondente
PEN (0,1mm)	44
PA (°C)	57
VIS 135 (cp)	1360
TL (%)	5,0
E* (Mpa)	7161
FAD (%)	75
Vv% (%)	4
TMN (mm)	9,5
P34 (%)	100
P38 (%)	70,9
P4 (%)	37,5
P200 (%)	5,6
CAP	0
AB	1
AMP	0
AM	0

Fonte: O autor

A Figura 38 mostra um gráfico de força para essa mistura específica. A classe alvo para a saída é a classe de tráfego EP, ou seja, Extremamente Pesado. Importante notar que essas observações valem apenas para essa única mistura selecionada. Outras amostras terão outros valores *SHAP*.

Figura 38 – Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (EP), definida como alvo



Fonte: O autor

A partir desse gráfico de força pode-se interpretar como a IA atribuiu a essa mistura específica a classe de tráfego Extremamente Pesado (EP), por meio das seguintes observações:

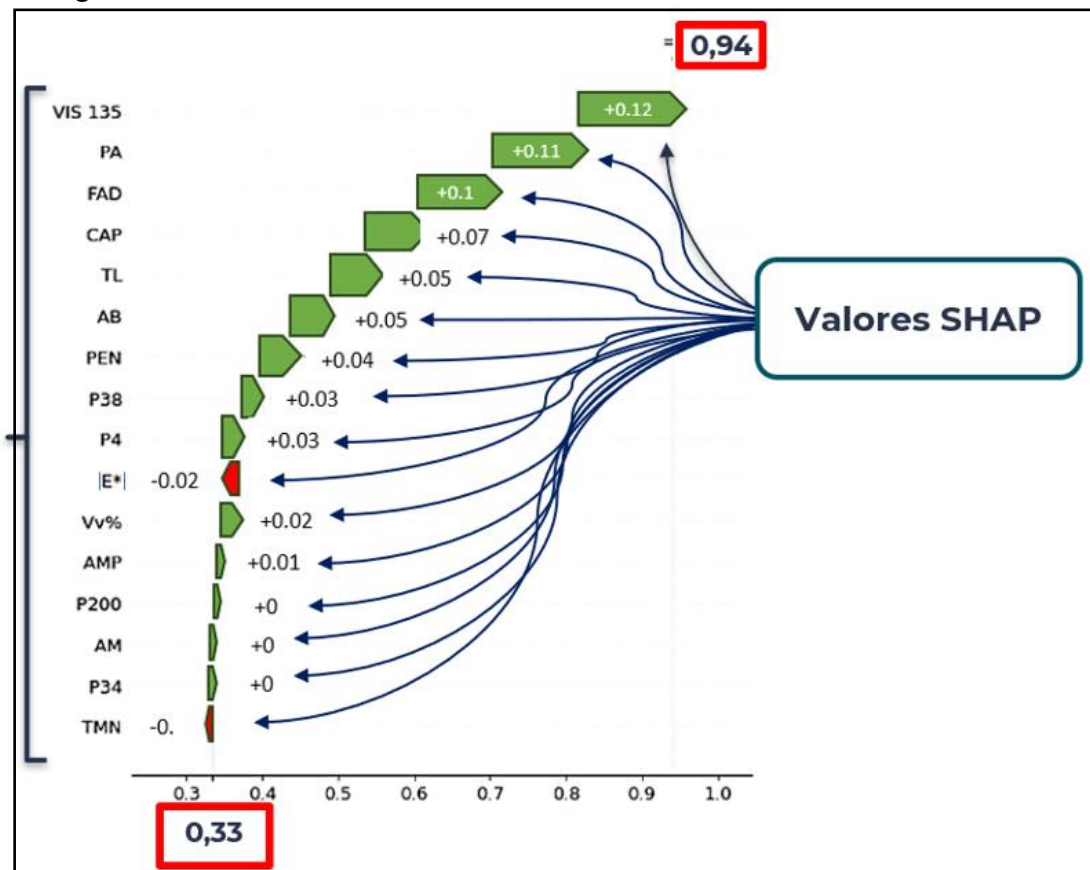
- a) O valor base corresponde a probabilidade de selecionar uma classe ao acaso. Como as 3 classes de saída (3 classes de tráfego consideradas) estão divididas em 3 partes iguais, o valor base para uma classe de saída corresponde a aproximadamente $1/3$, ou seja, 0,33.
- b) Em verde estão as variáveis que influenciaram positivamente para o modelo interpretar essa mistura como classe EP e em vermelho as variáveis que influenciaram para que o algoritmo não interpretasse a mistura em questão como EP. Os números ao lado de cada variável correspondem aos valores delas na mistura em questão.
- c) O tamanho das setas é proporcional à relevância da contribuição das variáveis na observação dessa mistura.
- d) Pode-se interpretar a partir do gráfico que as variáveis contribuíram para que a probabilidade do modelo interpretar a mistura como classe EP saísse do valor base de 0,33 para 0,94.
- e) A viscosidade elevada, o ponto de amolecimento, a porosidade FAD e a ausência do ligante puro (CAP) contribuíram positivamente para o modelo interpretar essa mistura como sendo da classe (EP).
- f) A mistura em questão possui ligante modificado, no caso o asfalto borracha. Observa-se que o modelo associou o ligante modificado a uma mistura mais resistente à deformação permanente.
- g) O baixo valor do módulo dinâmico da mistura ($|E^*|$) apresentou uma pequena influência para que o modelo não interpretasse essa mistura como sendo da classe (EP).

Outra forma de interpretação local fornecida pela biblioteca *SHAP* é o chamado gráfico cachoeira, também conhecido como gráfico de cascata. Esse gráfico mostra um total cumulativo à medida que valores são adicionados ou subtraídos. O gráfico cachoeira complementa o gráfico de força, anteriormente ilustrado na Figura 38. A principal vantagem desse gráfico consiste em mostrar o valor exato da contribuição de cada variável para a predição e ajuda a entender como um valor inicial é afetado por uma série de valores positivos e

negativos.

A Figura 39 ilustra um gráfico cachoeira para a mistura anteriormente descrita na Tabela 15. Nesse gráfico, a probabilidade inicial da predição está ilustrada na parte inferior, correspondente a 0,33 e sobe até o resultado de 0,94. Em verde estão as variáveis que contribuíram positivamente para a previsão e em vermelho a variável que contribuiu negativamente. As variáveis correspondentes a viscosidade e ao ponto de amolecimento foram as de maior impacto positivo, com valores *SHAP* de +0,12 e +0,11 respectivamente.

Figura 39 – Gráfico cachoeira relacionado à mistura descrita na Tabela 15



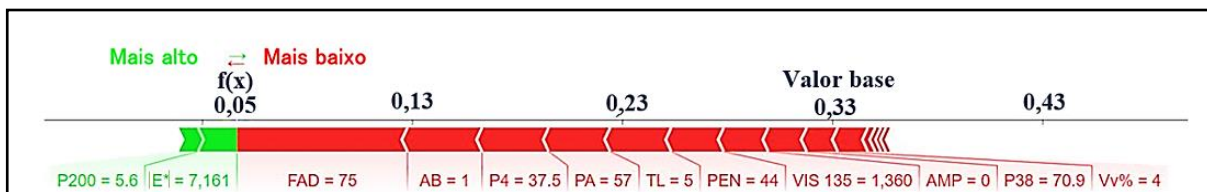
Fonte: O autor

A partir da observação da Figura 39, interpreta-se que o valor *SHAP* de +0,12 atribuído a variável VIS 135 na mistura em questão, elevou em 12% as chances de o algoritmo interpretar essa mistura como a classe alvo, nesse caso uma mistura de classe (EP). O mesmo raciocínio é considerado para as demais variáveis, lembrando que as variáveis com valores *SHAP* negativos contribuem reduzindo essas chances.

A mistura em questão apresentou uma probabilidade de 0,94 de ser uma mistura de

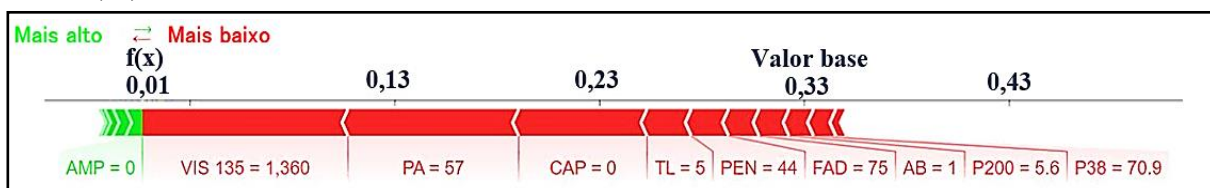
classe (EP), o que justifica a classificação do algoritmo. Utilizou-se a mesma mistura para avaliar a probabilidade dessa mistura ser da classe (P) e da classe (M). As Figuras 40 e 41 ilustram os valores *SHAP* obtidos nas análises.

Figura 40 – Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (P), definida como alvo



Fonte: O autor

Figura 41 – Gráfico de força correspondente a mistura descrita na Tabela 15, com a classe (M), definida como alvo



Fonte: O autor

Na Figura 40, percebe-se que o modelo atribuiu a chance dessa mistura ser de classe (P) é de apenas 0,05. No gráfico de força, é possível observar que as variáveis: FAD, AB, P4 e PA; foram as que mais influenciaram para a mistura não ser classificada como destinada a tráfego pesado. Na Figura 41, a VIS 135, o PA e o CAP, foram as variáveis que apresentaram os valores *SHAP*, contribuindo para uma diminuição do algoritmo classificar essa mistura como tráfego médio. Os valores *SHAP* na Figura 41, baixaram para apenas 0,01 a probabilidade dessa mistura ser classificada como (M).

Nota-se que as probabilidades das 3 classes se complementam, sendo 0,94 para a classe (EP), 0,05 para a classe (P) e 0,01 para a classe (M).

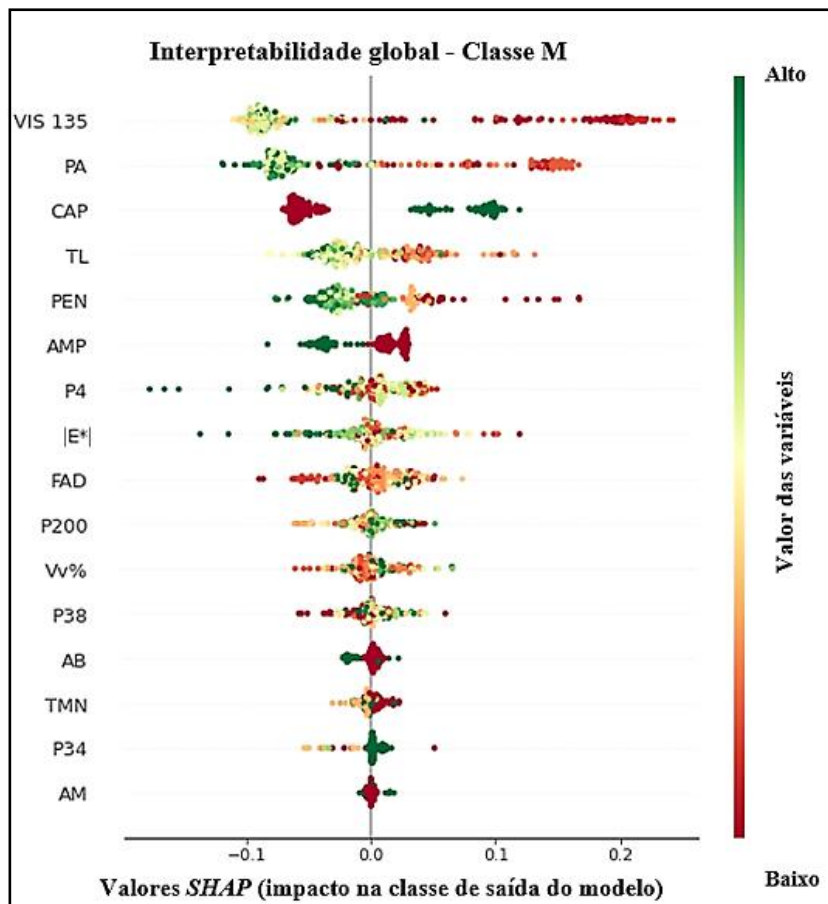
4.5.2 Interpretabilidade global

Na interpretabilidade local os valores correspondem a uma análise para uma única mistura. Para entender a relevância das variáveis de uma maneira geral em todo o conjunto de dados, pode-se usar o gráfico resumo de interpretabilidade global. Nesse gráfico, em cada linha onde se encontra cada uma das variáveis, cada ponto representa uma mistura do banco de dados.

Se o ponto está vermelho, significa que para aquela mistura, aquela variável possui um valor baixo (em relação aos valores que aquela variável assume no conjunto de dados). Quanto mais verde o ponto, mais alto é o valor. A posição do ponto no eixo horizontal indica o valor *SHAP*, daquela variável, sendo que quanto mais a direita, mais positiva é a contribuição daquela variável naquele ponto.

A Figura 42 mostra a interpretação global do modelo referente às misturas da classe (M). Os valores referentes às classes categóricas como o tipo de ligante são variáveis *dummy* e assumem apenas valores de 0 ou 1.

Figura 42 – Interpretação global do modelo referente as misturas da classe (M)



Fonte: O autor

Na Figura 42, os valores *SHAP* positivos, localizados à direita do gráfico, indicam a contribuição favorável à classe desejada, enquanto os valores à esquerda indicam que esses dados contribuem negativamente para a previsão da classe alvo.

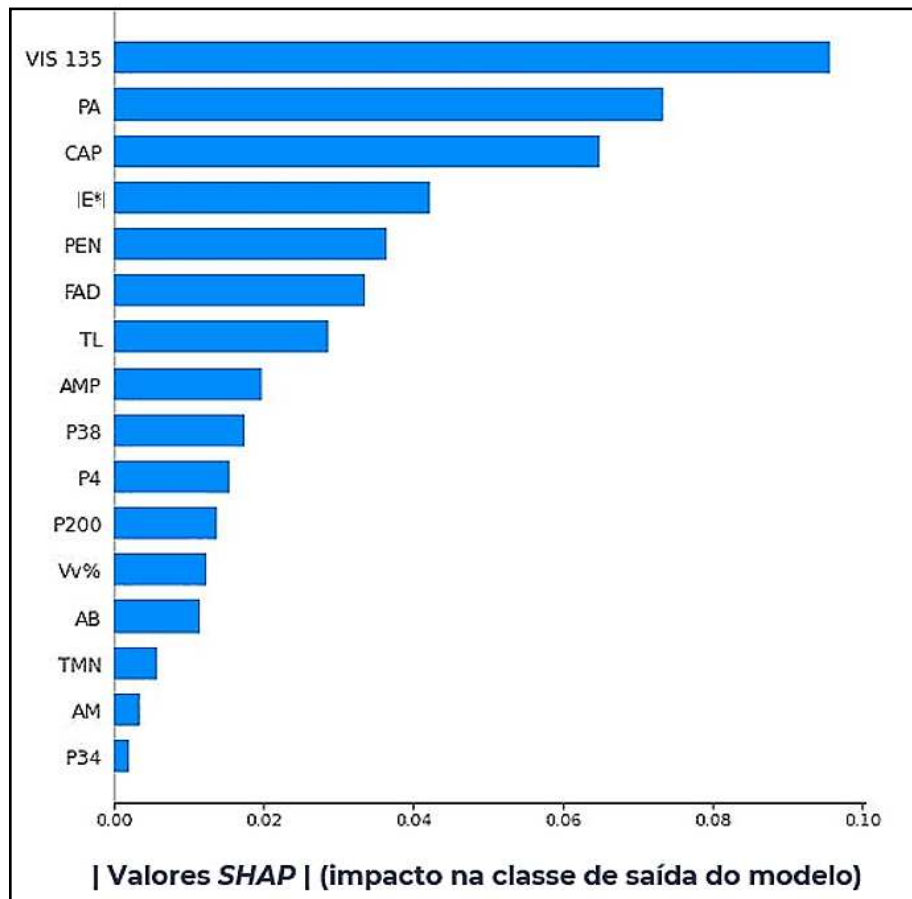
A classe M, no caso desta pesquisa, pertence as misturas menos resistentes à deformação permanente e consequentemente indicadas para tráfego médio. Verifica-se que o

modelo interpretou que misturas com valores baixos (em vermelho) de viscosidade, de ponto de amolecimento e de penetração contribuem para maiores valores *SHAP* referentes à classe M. Valores mais altos indicam (em verde) uma contribuição negativa dessas variáveis na previsão de uma mistura de classe M.

Como citado anteriormente, cada ponto do gráfico global significa um valor *SHAP* de uma determinada variável em uma mistura específica. Para cada recurso, pode-se calcular o valor médio de todas as observações, mais precisamente a média dos valores absolutos, para que os valores positivos e negativos não se compensem. As variáveis que apresentaram mais contribuições positivas ou negativas terão maiores valores *SHAP* médios.

Dessa forma o gráfico de barras com a média absoluta dos valores *SHAP* pode ser indicado para fornecer um gráfico de importância das variáveis, conforme ilustrado na Figura 43.

Figura 43 – Gráfico da média absoluta dos valores *SHAP* para classe (M)



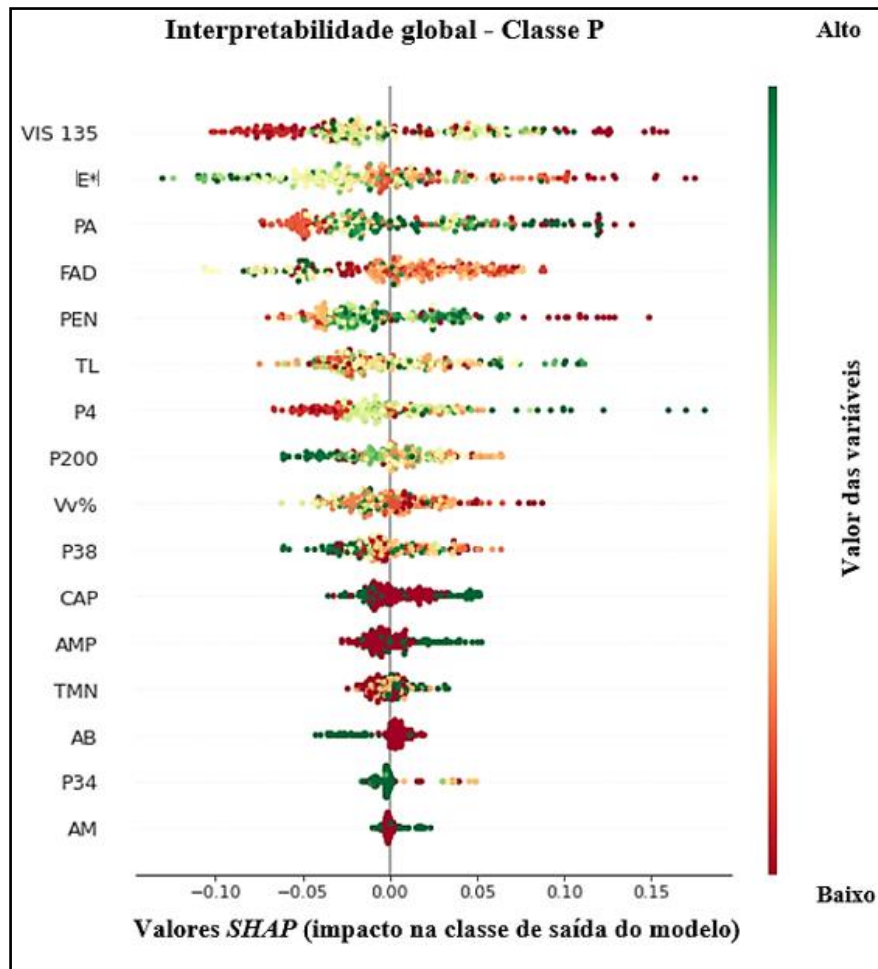
Fonte: O autor

A visualização global da importância das variáveis torna-se mais fácil de

compreender por meio do gráfico dos valores médios absolutos. Pode-se interpretar que a contribuição média da variável VIS 135 está em torno de 10%, o mesmo raciocínio é aplicado para as demais variáveis. Na Figura 43, além das VIS 135, as variáveis PA, CAP e E*; apresentaram contribuições significativas, enquanto as variáveis TMN, AM e P34 foram as que apresentaram menores influências.

A Figura 44 ilustra os valores *SHAP* de interpretação global envolvendo misturas da classe P (tráfego pesado), que é uma classe intermediária entre as classes de tráfego médio (M) e extremamente pesado (EP).

Figura 44 – Interpretação global do modelo referente as misturas da classe (P)



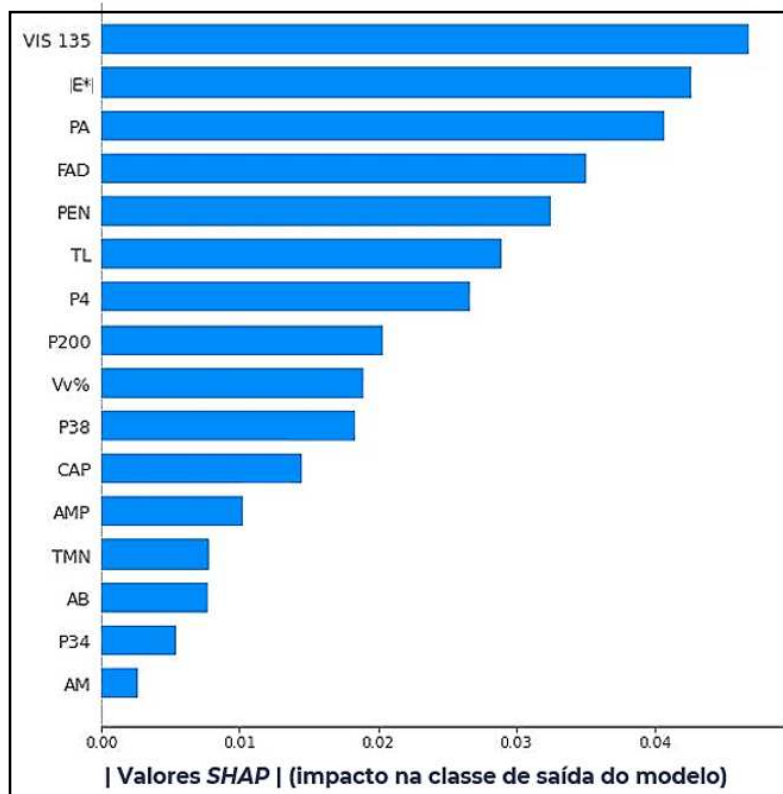
Fonte: O autor

É possível verificar que valores altos de $|E^*|$ implicam em uma contribuição negativa do modelo em relação à classe P e misturas com $|E^*|$ mais baixo apresentaram valores *SHAP* maiores para prever a classe de tráfego Pesado (P). Em relação à penetração do ligante,

o modelo interpretou que valores baixos contribuem positivamente para uma previsão da classe alvo. Valores altos na variável de percentual passante na peneira #4 contribuiram positivamente na previsão dessa classe. Outra variável relevante na interpretação foi o teor de ligante, onde os maiores valores *SHAP* estão relacionados a um teor de ligante mais elevado.

Calculando a média absoluta dos valores *SHAP* obtidos na análise global referente a misturas da classe (P), plotou-se o gráfico de barras, ilustrado na Figura 45, para visualizar a importância das variáveis dessa classe.

Figura 45 – Gráfico da média absoluta dos valores SHAP para classe (P)



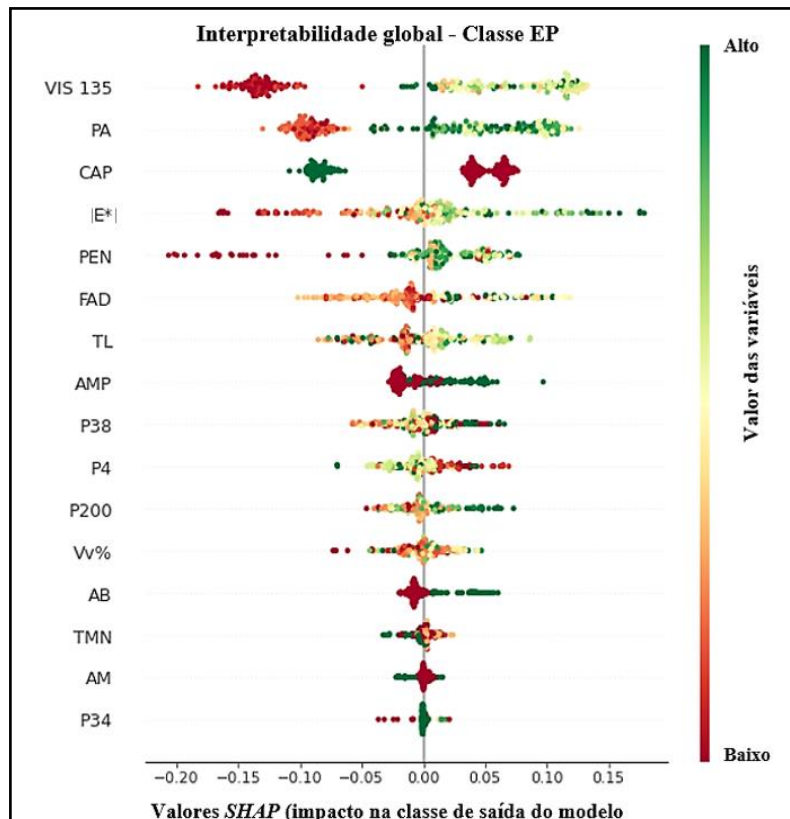
Fonte: O autor

No gráfico da Figura 45, a variável VIS 135 novamente se apresenta como o atributo que mais contribuiu para a interpretação do algoritmo. No entanto, para a classe (P), observa-se valores *SHAP* mais baixos para as variáveis, quando comparados com os valores obtidos para classe (M). Esses valores mais baixos podem indicar uma dificuldade do modelo em encontrar uma variável mais explicativa para a análise dessa classe, no entanto pode-se observar que as variáveis VIS 135, |E*| e PA; continuam aparecendo nos primeiros lugares.

Definindo a classe EP como alvo, é possível verificar que valores baixos de viscosidade do ligante contribuem negativamente e que valores mais altos contribuem

positivamente para a previsão de uma mistura do tipo EP. O $|E^*|$ da mistura apresentou resultado semelhante ao da viscosidade do ligante, indicando que o algoritmo interpretou um comportamento já esperado, ou seja, valores elevados de $|E^*|$ contribuem para misturas mais resistentes à deformação permanente. O gráfico resumo para classe EP está ilustrado na Figura 46.

Figura 46 – Interpretação global do modelo referente as misturas da classe (EP)

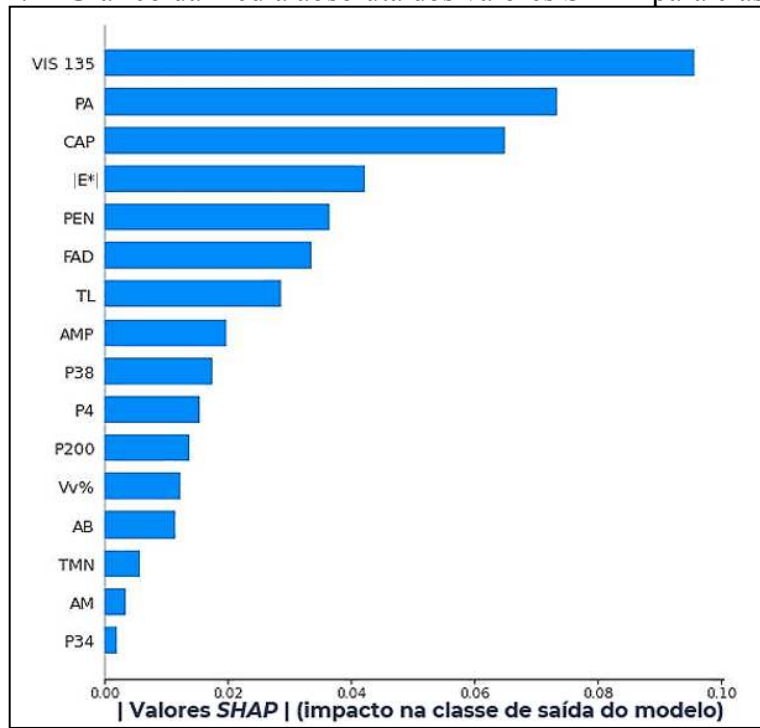


Fonte: O autor

É possível notar que a presença de ligante puro (CAP) diminui as chances do modelo de interpretar a classe como (EP), e a presença de ligantes modificados, com exceção do AM, indica que a mistura resistirá bem à DP. A variável relacionada à porosidade FAD, do ponto de vista computacional não se mostrou promissora para indicar a resistência de uma mistura asfáltica para as técnicas e dados aqui utilizados. Observa-se que valores baixos de ponto de amolecimento do ligante contribuem negativamente para previsão de tráfego EP.

Foi plotado o gráfico com a média absoluta dos valores SHAP das variáveis referentes à classe (EP), conforme observado na Figura 47. Nesse gráfico pode-se observar que a variável VIS 135 novamente aparece como a mais relevante para o modelo, com valores próximos a 0,10, seguidos das variáveis: PA, CAP e $|E^*|$.

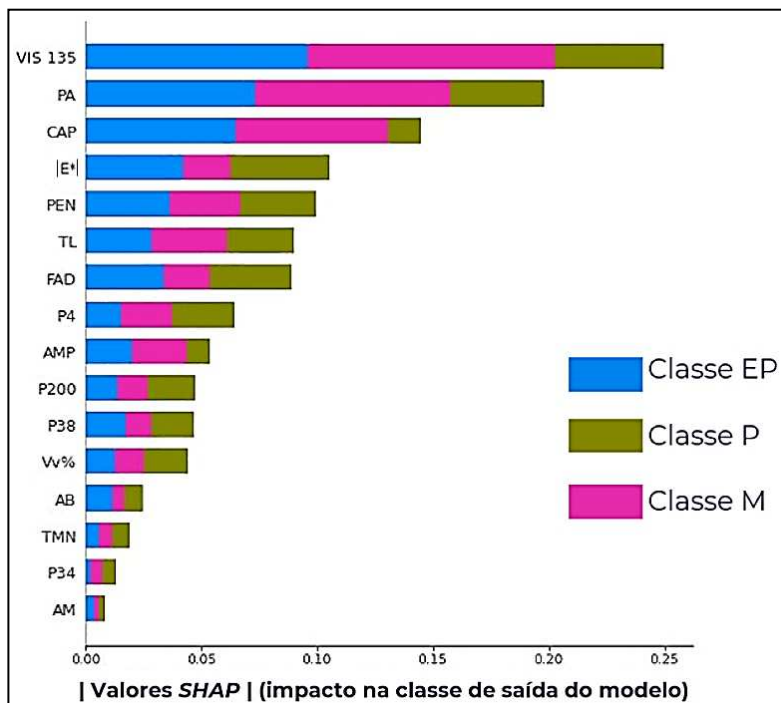
Figura 47 – Gráfico da média absoluta dos valores SHAP para classe (EP)



Fonte: O autor

Por último, foi gerado um gráfico de barras com a soma de todos os valores médios absolutos obtidos em cada classe. A Figura 48 ilustra o resultado dessa soma.

Figura 48 – Gráfico da soma das médias absolutas dos valores SHAP de todas as classes



Fonte: O autor

O gráfico da Figura 48 fornece uma visão geral do impacto das variáveis ranqueadas pelos valores *SHAP*. De maneira geral, observa-se que a variável VIS 135 foi a que apresentou a maior influência no modelo, com valores próximos a 0,25. As variáveis PA, CAP e E* também se mostraram como atributos relevantes para classificação de uma mistura asfáltica em relação a deformação permanente. Pode-se interpretar que essas variáveis são as mais importantes dentro o modelo desenvolvido utilizando as 251 misturas do banco de dados.

A análise do impacto das variáveis pode fornecer informações relevantes para a compreensão do fenômeno da deformação permanente. Conhecer os atributos que mais irão influenciar um modelo preditivo pode ajudar a orientar a coleta de novos dados, uma vez que a escassez deles dificulta a obtenção de todas as informações relativas à mistura asfáltica.

O uso de ciência de dados pode fornecer uma visão alternativa a partir de variáveis de fácil obtenção, como por exemplo a viscosidade, que é um atributo referente a trabalhabilidade do ligante, mas que apresentou-se como uma variável de grande impacto na classificação de misturas asfálticas com foco na deformação permanente.

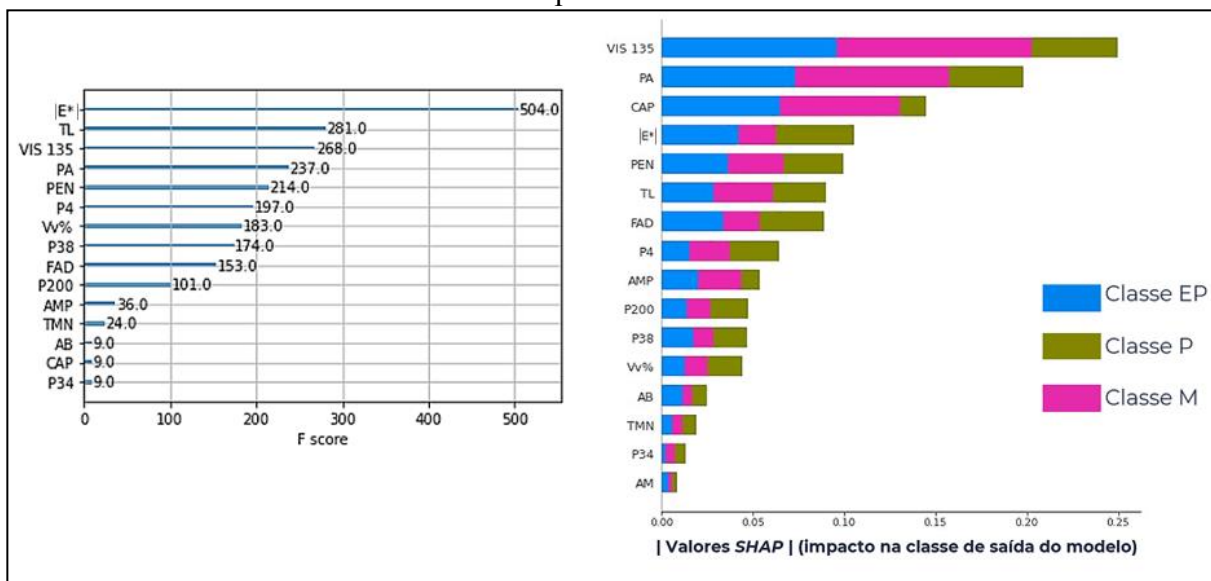
4.6 Comparação entre a relevância dos atributos e os valores *SHAP*

O algoritmo *XGBoost* fornece um gráfico resumo contendo a relevância das variáveis. Essa medida é estimada ao registrar o número de vezes que cada variável é usada na divisão dos dados nas árvores do modelo. Essa é uma métrica comumente utilizada em trabalhos de aprendizado de máquina para ilustrar a relevância das variáveis. No entanto, essa métrica é relativa apenas ao desempenho do algoritmo em si e não reflete necessariamente que aquela variável é a mais importante na previsão do fenômeno em estudo.

Os valores *SHAP* medem o impacto das variáveis na previsão, levando em consideração a interação com outras variáveis. Eles calculam a importância de uma variável para uma previsão específica, comparando o valor previsto com o valor que seria previsto caso a variável não estivesse presente. É importante ressaltar que a ordem na qual as variáveis são apresentadas ao modelo pode afetar suas previsões, por isso os valores *SHAP* são calculados em todas as ordens possíveis para que as variáveis sejam comparadas de maneira justa.

A Figura 49 ilustra a comparação entre a relevância dos atributos obtidos com o uso do algoritmo *XGBoost* e o impacto das variáveis ranqueadas pelos valores *SHAP*.

Figura 49 – Comparação entre os resultados obtidos pelo *XGBoost* e pela biblioteca *SHAP*, referente a importância das variáveis



Fonte: O autor

Inicialmente, percebe-se que há diferenças na ordem das variáveis que cada ferramenta ranqueou. Enquanto o E* foi o atributo mais relevante para o *XGBoost*, seguido pelas variáveis TL, VIS 135 e PA. Já na análise com valores *SHAP*, a variável VIS 135 se apresentou como a variável mais relevante. Ao comparar os valores *SHAP* e a relevância das variáveis fornecida pelo algoritmo *XGBoost*, pode-se concluir que os valores *SHAP* são uma medida mais precisa para estimar o impacto das variáveis na previsão, pois levam em consideração a interação entre as variáveis. Adicionalmente, os valores *SHAP* fornecem uma explicação individualizada para cada predição, com essa informação é possível interpretar o comportamento das variáveis em cada mistura.

A utilização dos valores SHAP se mostrou uma ferramenta útil para a análise da relevância das variáveis em modelos de aprendizado de máquina. A análise do impacto das variáveis é realizada de maneira justa, pois distribui a contribuição de cada variável para a predição e levam em consideração a interação entre todas elas.

5 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Neste capítulo final da dissertação são realizadas considerações finais referentes ao uso de aprendizado de máquina interpretável para classificar as misturas asfálticas em relação à deformação permanente seguindo um critério de tráfego. São apresentadas as conclusões obtidas a partir do desenvolvimento desta pesquisa, e também indicadas sugestões para trabalhos futuros.

5.1 Principal Contribuição da Pesquisa

Os resultados obtidos na realização desse trabalho mostraram que a IA pode ser uma importante ferramenta para o desenvolvimento de pesquisas na engenharia rodoviária. O modelo utilizado obteve um ótimo desempenho, que não objetiva substituir o ensaio uniaxial, mas sim orientar a dosagem. É possível obter um conhecimento prévio do comportamento da mistura em relação à deformação permanente.

As contribuições dessa pesquisa vão além dos resultados obtidos ao final da dissertação. Uma das maiores dificuldades em um trabalho de ciência de dados é a carência de informações. Para auxiliar em pesquisas futuras, os dados estruturados durante a realização desse trabalho estão organizados e disponibilizados para futuros pesquisadores.

Um dos motivos para a utilização de uma ferramenta de aprendizado de máquina interpretável é buscar trazer um melhor entendimento dos resultados e do potencial dos algoritmos preditivos. Espera-se que esse trabalho possa estimular futuras pesquisas na área do aprendizado de máquina, trazendo essa poderosa ferramenta para a área da infraestrutura viária.

5.2 Principais Limitações da Pesquisa

O desenvolvimento de uma pesquisa envolvendo aprendizado de máquina, necessita da disponibilidade de uma quantidade razoável de dados. A escassez de informações apresentou-se como uma das principais dificuldades durante a realização dessa pesquisa. Essa escassez ocorre não somente na ausência de dados, mas também quando há dados existentes contendo informações incompletas.

Outro fator limitante na pesquisa foi a ausência de dados estruturados. Estruturar os dados encontrados durante a pesquisa foi uma etapa demorada devido o processo de

organização e tratamento desses dados. Essa etapa poderia ter sido otimizada caso houvesse uma melhor disponibilidade de dados.

5.3 Principais Constatações da Pesquisa

Este trabalho buscou desenvolver ferramentas de aprendizado de máquina interpretável para classificar as misturas asfálticas em relação à deformação permanente de acordo com o tipo de tráfego.

Primeiramente, foi realizado um levantamento de dados de trabalhos acadêmicos, publicações científicas e relatórios técnicos. A obtenção dos dados se revelou em uma extensa etapa necessária para a realização do tipo de pesquisa desenvolvida. Esse esforço resultou em um banco de dados de algumas misturas asfálticas, organizado em planilha e disponível para futuras pesquisas. Acredita-se que o mesmo será um ponto de partida capaz de economizar tempo, material e recurso humano de outras pesquisas que possam fazer uso de ferramentas como as apresentadas aqui, entre outras.

Foi realizada uma análise descritiva das variáveis a fim de buscar um entendimento do banco de dados que deu suporte a este trabalho. Evidencia-se que as misturas apresentam comportamentos distintos dependendo do tipo de ligante asfáltico usado. Os resultados da análise das correlações utilizando todo o banco de dados não reflete de maneira adequada o comportamento das misturas, devido à heterogeneidade do banco de dados com 4 tipos de ligantes diferentes (CAP, AMP, AB e AM). Dessa forma, o estudo das correlações mostrou-se mais eficaz ao separar os ligantes, sendo então possível obter uma melhor compreensão do comportamento dos dados de entrada em relação aos valores do parâmetro usado como indicador da resistência à deformação permanente, no caso o *flow number* (FN).

Para realização dos modelos preditivos para classificação das misturas asfálticas, foram usadas redes neurais artificiais e também modelos com o algoritmo *XGBOOST*. O banco de dados do trabalho foi composto com dados de 251 misturas asfálticas, com 16 variáveis de entrada (PA, PEN, VIS 135, TL, |E*|, FAD, Vv%, TMN, P34, P38, P4, P200, CAP, AB, AMP e AM) e 1 variável de saída dividida em 3 classes.

Com o objetivo de encontrar um modelo simplificado com o menor número de parâmetros, mas que fosse capaz de oferecer um desempenho adequado, foram criadas 3 diferentes configurações (C1, C2 e C3). A configuração C3 com as variáveis: PA, PEN, VIS 135, TL, |E*|, Vv%, CAP, AB, AMP e AM, apresentou o melhor desempenho. Após a etapa

de treino e teste, obteve-se 84% de precisão na etapa de validação. Os hiperparâmetros podem diferir a depender de fatores como variação na quantidade de dados, escolha das variáveis de entrada, alteração na divisão de treino e teste, entre outros.

No entanto, esses modelos possuem limitações referentes à compreensão do comportamento das variáveis em relação ao fenômeno em estudo, pois não é possível interpretar como o algoritmo executa as previsões. Para contornar a lacuna da falta de interpretabilidade dos modelos de aprendizado de máquina, utilizou-se a ferramenta de inteligência artificial interpretável (*SHAP*), a fim de compreender o que o modelo de aprendizado de máquina leva em conta durante as previsões.

Para a interpretação individual de uma mistura, gerou-se o gráfico de força, que apresenta de maneira intuitiva quais as variáveis que mais contribuíram para a previsão do modelo. Além disso, esse gráfico indica se o impacto foi positivo ou negativo na análise individual. O gráfico cachoeira complementa o gráfico de força e fornece o valor *SHAP* exato da variável numa dada mistura.

O gráfico resumo permitiu a análise global das variáveis, ou seja, não se restringe a uma análise individual. Foram gerados 3 gráficos resumos, sendo um para cada tipo de tráfego. Considerando a classe de tráfego Médio (M), foi observado que o modelo interpretou valores menores de viscosidade e ponto de amolecimento do ligante, como fortes contribuições para chegar à conclusão de que era uma mistura asfáltica compatível com a classe M de tráfego. Verificou-se também que a ferramenta concluiu que o ligante puro contribuiu para uma mistura destinada a um tráfego menos intenso.

Na classe de tráfego Pesado (P), destaca-se a viscosidade do ligante, o módulo da mistura (IE^*) e o ponto de amolecimento do ligante com as variáveis que obtiveram os maiores valores *SHAP*, o que evidencia a relevância dessas variáveis no modelo de classificação. Para a classe EP de tráfego, foi possível observar que valores altos de viscosidade e ponto de amolecimento do ligante e módulo da mistura asfáltica apresentaram uma influência positiva para o modelo classificar a mistura para tráfego Extremamente Pesado (EP).

O número de dados considerado no presente estudo pode não ser suficientemente representativo. No entanto, a pesquisa contribui trazendo novas ferramentas com vistas a uma melhor compreensão de como variáveis diversas afetam a resistência de misturas à deformação permanente.

O uso de inteligência artificial tem sido cada vez mais difundido em diversos contextos e grupos de pesquisa. No entanto, o seu uso ainda é visto com reserva devido à dificuldade em se entender o funcionamento dos algoritmos de aprendizado de máquina. Espera-se com esse trabalho que a apresentação de uma ferramenta para interpretar modelos preditivos diminua as desconfianças em relação ao aprendizado de máquina e que sirva como um catalisador para a adoção do aprendizado de máquina em futuras pesquisas.

5.4 Sugestões para pesquisas futuras

No decorrer do desenvolvimento desta dissertação, foi possível identificar oportunidades diversas de ramificações deste trabalho por meio de pesquisas futuras. Assim, algumas sugestões são apresentadas a seguir:

- a) Criação de uma plataforma para facilitar a pesquisa de dados de misturas asfálticas e armazenamento de novos dados. Se no presente trabalho foi possível se trabalhar com 251 misturas, acredita-se que uma plataforma poderia elevar este número consideravelmente;
- b) Alimentação imediata do banco de dados com novas misturas, aumentando o número de exemplos e incluindo valores de variáveis que estão fora do intervalo dos dados utilizados neste trabalho;
- c) Criação de um *software* treinado com o modelo aqui desenvolvido para auxiliar o processo de dosagem de misturas asfálticas;
- d) Desenvolvimento de um modelo para estimar a evolução da deformação permanente a partir do treino com dados de resultados obtidos pelo ensaio triaxial *Stress Sweep Rutting* (SSR);
- e) Aplicação de técnicas de aprendizado de máquina ativo para seleção de variáveis em modelos preditivos;
- f) Criação de um modelo de aprendizado de máquina utilizando dados de corpos de provas virtuais.

REFERÊNCIAS

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 16505**: Misturas Asfálticas – Resistência à Deformação Permanente Utilizando o Ensaio Uniaxial de Carga Repetida. Rio de Janeiro, 2016.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 6560**: Ligantes asfálticos – Determinação do ponto de amolecimento – Método do anel e bola. Rio de Janeiro, 2016.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 6576**: Materiais asfálticos – Determinação da penetração. Rio de Janeiro, 2007.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 15184**: Materiais betuminosos – Determinação da viscosidade em temperaturas elevadas usando um viscosímetro rotacional. Rio de Janeiro, 2021.
- ALLEN-DENCID, H.; ZAMORA, J.; ALFARO, P. e ALONSO, M. F. **A machine learning approach for the automatic classification of schizophrenic discourse**. <https://doi.org/10.1109/ACCESS.2019.29086>, 2019.
- ASPHALT INSTITUTE. Superpave Mix Design. **Asphalt Institute Research Center**. Series, no. 02 (SP-2), Lexington, 2001.
- APEAGYEI, A. K. Rutting as a function of dynamic modulus and gradation. **Journal Of Materials In Civil Engineering**, Charlottesville. v.23, p. 1302-1310, 2011.
- ALAVI, A. H. A new approach for modeling of flow number of asphalt mixtures. **Archives of Civil and Mechanical Engineering**, Pitsburgo, v. 17, n. 2, p. 326–335, 2017.
- ALMEIDA JR, P. O. Avaliação laboratorial do tipo e teor de ligante e da granulometria na deformação permanente de misturas asfálticas. **Transportes**, Santa Maria, v. 26, n. 2, p. 1–15, 2018.
- AL-AMERI, R. ALNEDAWI, A. e NEPAL, K. New shakedown criterion and permanent deformation properties of unbound granular materials. **Journal of Modern Transportation**. v.27, p. 108–119, 2019.
- BARROSO, L. e REZENDE, L. **Resultados de ensaios de módulo dinâmico para avaliação preliminar do desempenho de misturas asfálticas**, In: CONGRESSO NACIONAL DE PESQUISA E ENSINO EM TRANSPORTE, 33º, 2019, Balneário Camboriú. Anais [...] Balneário Camboriú: ANPET, 2019.
- BASTOS, J. B. S. **Considerações sobre a deformação permanente de pavimentos asfálticos no dimensionamento mecânico-empírico**. 2016. Tese (Doutorado em Engenharia de Transportes) - Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2016.
- BASTOS, J. B. S.; SOARES, J. B. e NASCIMENTO, L. A. H. Critérios para os resultados do ensaio uniaxial de carga repetida de misturas asfálticas em laboratório a partir do desempenho em campo. **Revista Transportes**, v. 25, n. 2, p. 29-40, 2017.
- BERNUCCI, L. L. B.; MOTTA, L. M. G.; CERATTI, J. A. P. e SOARES, J. B. **Pavimentação Asfáltica – Formação Básica para Engenheiros**. Rio de Janeiro: Petrobras; ABEDA, 2022.
- BAGHERI, A.; NAZARI, A. e SANJAYAN, J. The use of machine learning in boron-based geopolymers: Function approximation of compressive strength by ANN and GP. **Measurement: Journal of the International Measurement Confederation**, v. 141, p. 241–249, 2019.
- BILIGIRI, K. P.; KALOUSH, K.; MAMLOUK, M. e WITCZAK, M. Rational Modeling of Tertiary Flow for Asphalt Mixtures. **Transportation Research Record**. v. 2001, p. 63-72, 2007.
- BISHOP, C.M. **Pattern Recognition and Machine Learning**. Berlin: Springer, 2006.
- BOCANEGRA, C.W.R. **Procedimentos para tornar mais efetivo o uso das redes neurais artificiais em**

planejamento de transportes. 2002. Dissertação (Mestrado em Engenharia Civil – Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2003.

BOHN, K. A. **Avaliação de misturas asfálticas recicladas mornas com uso de ligantes convencional e modificado por polímero**. 2017. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Santa Maria, Santa Maria, 2017.

BOHN K. A. e SPECHT, L. P. **Avaliação do comportamento viscoelástico e da deformação permanente de mistura asfálticas recicladas mornas com uso de ligantes convencional e modificado por polímero**. In: CONGRESSO IBERO-LATINOAMERICANO DEL ASFALTO – CILA, 19º, 2017, Medellín. Anais [...] Medellín: CILA, 2017.

BONAQUIST, R. Evaluation of flow number (FN) as a discriminating HMA mixture property. **Wisconsin Highway Research Program**. Disponível em: <https://rosap.ntl.bts.gov/view/dot/25898>. Acesso em: 20 de nov. 2023.

BORGES, R. L. **Deformação permanente em misturas asfálticas a partir do Shift Model viscoplastico e ensaios triaxiais de carga repetida**. 2014. Dissertação (Mestrado em Engenharia de Transportes) – Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2014.

BRASIL. Ministério dos transportes. Departamento Nacional de Infraestrutura de Transportes. **DNIT 184 – ME: Pavimentação - Misturas asfálticas - Ensaio uniaxial de carga repetida para determinação da resistência à deformação permanente – Método de ensaio**, Rio de Janeiro, RJ. 2018.

BROWN, E. R.; KANDHAL, P. S.; ROBERTS, F. L.; KIM, Y. R.; LEE, D. Y. e KENNEDY, T. W. **Hot mix asphalt materials, mixture design, and construction**. 3rd ed. Lanham: NAPA Research and Education Foundation, 2009.

CELESTE, A. B. e OLIVEIRA, F. H. L. DE. Study of retroanalysis of asphaltic pavements resilience modules with the use of artificial neural networks. **Transportes**, v. 27, n. 4, p. 123–133, 2019.

CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O. e KEGELMEYER, W. P. SMOTE: synthetic minority over-sampling technique. **J. Artif. Intell. Res. JAIR**, v.16, p. 321–357. 2022.

CHEN, T. e GUESTRIN, C. **XGBoost: a scalable tree boosting system**. New York: ACM, p.785-794, 2016.

CONFEDERAÇÃO NACIONAL DO TRANSPORTE - CNT. **Pesquisa CNT de Rodovias 2022: SEST: SENAT**. - Relatório gerencial. – Brasília. 2022.

COHEN, J. **Statistical power analysis for the behavioral sciences**. New York: Lawrence Erlbaum Pub. 1988.

DACOREGIO, F. A. **Estimativa Preliminar De Custos De Obras Utilizando Redes Neurais Artificiais**. 2017. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal de Santa Catarina. Florianópolis. 2017.

DE PINA, A. C. ANN-based surrogate models for the analysis of mooring lines and risers. **Applied Ocean Research**, v. 41, p. 76–86, 2013.

DOMINGOS, M. D. I.; PAMPLONA, T. F.; FAXINA, A. L. e GIGANTE, A. C. Viscosidade rotacional de ligantes asfálticos modificados de mesmo grau de desempenho. **Revista Transportes**, v. 20, n. 2, p. 15-22, 2012.

DOSHI-VELEZ, F. e KIM, B. **Towards a Rigorous Science of Interpretable Machine Learning**. arXiv: Machine Learning. 2017.

FACCIN, C. **Concretos asfálticos em utilização no Rio Grande do Sul: comportamento mecânico e desempenho em campo quanto à deformação permanente**. 2018. Dissertação (Mestrado em Engenharia Civil). Programa de Pós-Graduação em Engenharia Civil. Universidade Federal de Santa Maria. Santa Maria. 2018.

FERREIRA, J.G.H. De M. **Tratamento de dados geotécnicos para predição de módulos de resiliência de solos**

e britas utilizando ferramentas de data mining. Tese de Doutorado em Engenharia Civil, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2008.

FERREIRA, J. L. S.; BASTOS, J. B. DOS S. e SOARES, J. B. **Validação da metodologia de Faixa de Agregados Dominantes para avaliação e especificação da granulometria de misturas asfálticas densas.** In: CONGRESSO NACIONAL DE PESQUISA E ENSINO EM TRANSPORTE, 29, 2015, Ouro Preto. Anais [...] Ouro Preto: ANPET, 2015.

FERREIRA, J. L. S. **Novas diretrizes para a seleção granulométrica de misturas asfálticas densas com foco na deformação permanente.** 2016. Dissertação (Mestrado em Engenharia de Transportes) - Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2017.

FRANCO A. C. P. e MOTTA, L. M. G. **Análise elástica de múltiplas camadas: manual de utilização.** Versão 2.4. Rio de Janeiro, 2018.

FUHG, J. N.; MARINO, M. e BOUKLAS, N. Local approximate Gaussian process regression for data-driven constitutive laws: Development and comparison with neural networks. **Computer Methods in Applied Mechanics and Engineering**, Ithaca, v. 388, p 1-23, 2021.

GOMES, J. C. B. **Estimação não-paramétrica para função de covariância de processos gaussianos espaciais.** 2009. Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matematica, Estatística e Computação Científica, Campinas, São Paulo. 2009.

GOUVEIA, B. C. S. **Métodos de previsão e determinação de módulo dinâmico de misturas asfálticas aplicadas em rodovias brasileiras.** 2016. Dissertação (Mestrado em Engenharia de Transportes) – Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2016.

GONG, H., SUN, Y.; DONG, Y.; HAN, B.; POLACZYK, P.; HU, W. e HUANG, B. Improved estimation of dynamic modulus for hot mix asphalt using deep learning. **Construction and Building Materials**. v. 263. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0950061820319176?via%3Dihub>. Acesso em: 20 de nov. 2020.

GROSK, A. P e DIAS, J. S. **Algoritmo Genético: Revisão Histórica e Exemplificação.** Pontifícia Universidade Católica do Paraná. 2006.

GUIMARÃES, A. C. R. **Um método mecanístico-empírico para a previsão da deformação permanente em solos tropicais constituintes de pavimentos.** 2009. Tese (Doutorado em Engenharia Civil) – Programa de Engenharia Civil da COPPE/UFRJ. Rio de Janeiro, Rio de Janeiro. 2009.

HAHNLOSER, R.; SARPESHKAR, R.; MAHOWALD, M.; DOUGLAS, R. e SEUNG, H. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. **Nature**. v. 405, p. 947-951, 2000.

HAYKIN, S. **Redes Neurais: Princípios e prática.** 2. ed. Ontário, Canadá: ARTMED, 2008.

HOU, Y. The State-of-the-Art Review on Applications of Intrusive Sensing, Image Processing Techniques, and Machine Learning Methods in Pavement Monitoring and Analysis. **Engineering**, 2020.

KENNEDY, T.W.; HUBER, G. A.; HARRIGAN, E. T.; COMINSKY, R. J.; HUGHES, C.S.; QUINTUS, H.V. e JAMES S. M. **Superior performing asphalt pavements (Superpave): the product of the SHRP asphalt research program.** Washington: SHRP, 1994.

KIM, S. Identification and assessment of the dominant aggregate size range (DASR) of asphalt mixture, **Asphalt Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions**. v.75, p. 789-814, 2006.

KRONHEIM, B. S.; KUCHERA, M. P. e PROSPER H. B. TensorBNN: Bayesian inference for neural networks using TensorFlow. **Computer Physics Communications**, Davidson, v. 270, p. 108-116, 2022.

- LECUN, Y. e BENGIO, Y. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.
- LINARDATOS, P.; PAPASTEFANOPOULOS, V. e KOTSIANTIS, S. Explainable AI: A Review of Machine Learning Interpretability Methods. **Entropy**. v. 23. Disponível em: <https://www.mdpi.com/1099-4300/23/1/18>. Acesso em: 20 de nov de 2023.
- LIRA, B.; JELAGIN, D. e BIRGISSON, B. Gradation-based framework for asphalt mixture. **Materials and Structures**, Estocolmo, v. 46, p. 1401-1414, 2013.
- LIU, J.; LIU, F.; ZHENG, C.; ZHOU, D. e WANG, L. Optimizing asphalt mix design through predicting the rut depth of asphalt pavement using machine learning, **Construction and Building Materials**, Blacksburg, v. 356, 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0950061822028677>. Acesso em: 20 de nov. 2023
- LUNDBERG, S.M. e LEE, S.I. **A unified approach to interpreting model predictions**. In: INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS. 31., 2017, Red Hook. Anais [...] New York: Curran Associates Inc, p. 4768–4777, 2017.
- MAJIDIFARD, H.; ADU-GYAMFI, Y. e BUTTLAR, W.G. Deep machine learning approach to develop a new asphalt pavement condition index. **Construction and Building Materials**, Columbia, v. 247, 2020.
- MAJIDIFARD, H. New machine learning-based prediction models for fracture energy of asphalt mixtures. **Measurement: Journal of the International Measurement Confederation**, Columbia, v. 135, p. 438–451, 2019.
- MASAD, E. Internal structure characterization of asphalt concrete using image analysis. **Journal of Computing in Civil Engineering**, Pullman, v. 13, n. 2, p. 88–95, 1999.
- MCCULLOCH, W. S. e PITTS, W. H. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, Illinois, v. 7, p. 115–133, 1943.
- MILLER, T. Explanation in artificial intelligence: Insights from the social sciences. **Artificial Intelligence**, Melbourne, v. 267, p. 1-38, 2019.
- MIRZAHOSSEINI, M.; GANDOMI, A.; ALAVI, A. e NEJAD, F. Nonlinear genetic-based models for prediction of flow number of asphalt mixtures. **Journal of Materials in Civil Engineering**, Tehran, v.23, p. 248-263, 2011.
- MOLNAR, C. **Interpretable Machine Learning: a guide for making black box models explainable**. 2. ed. [S. l.: s. n.], Disponível em: <https://christophm.github.io/interpretable-ml-book>. Acesso em: 20 de nov. 2023.
- MOMM, L. E DOMINGUES, F. A. A. **Estudo dos efeitos da granulometria sobre a macrotextura superficial do concreto asfáltico e seu comportamento mecânico**. São Paulo: Universidade de São Paulo, 1998.
- MOURA, E. **Estudo de deformação permanente em trilha de roda de misturas asfálticas em pista e em laboratório**. 2010. 299f. Tese (Doutorado em Engenharia de Transportes) - Universidade de São Paulo, São Paulo, SP, 2010.
- MÜLLER, A. C. e GUIDO, S. **Introduction to machine learning with python: a guide for data scientists**. 1st. ed. Sebastopol: O'Reilly Media, 2017.
- MURPHY, K. **Machine learning: a probabilistic perspective**. 1st. ed. Cambridge: MIT Press, 2012.
- PORTELA NETO, M. G. **Avaliação da faixa de agregados dominantes e componentes intersticiais em misturas asfálticas**. 2018. Dissertação (Mestrado em Engenharia de Transportes) – Pró Reitoria de Pesquisa e Pós- Graduação, Universidade Federal do Ceará, 2018.
- NILSSON, N. **Learning machines: foundations of trainable pattern-classifying systems**. New York: McGraw-Hill, 1965.

OLIVEIRA, J. A. **Avaliação do impacto da metodologia FAD-CI de seleção granulométrica na vida de fadiga de misturas asfálticas**. 2019. Tese (Doutorado em Engenharia de Transportes) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Transportes, Fortaleza. 2019.

ONOFRE, F. C. **Avaliação do efeito da modificação de ligantes asfálticos por ácido poli fosfórico e polímeros na resistência à deformação permanente de misturas asfálticas**. 2012. Dissertação (Mestrado em Engenharia de Transportes) – Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade Federal do Ceará, Fortaleza, 2012.

ONYELOWE, K. C.; EBID, A. M. e NWOBIA, L. I. Predictive models of volumetric stability (durability) and erodibility of lateritic soil treated with different nanotextured bio-ashes with application of loss of strength on immersion, GP, ANN and EPR performance study. **Cleaner Materials**, v. 1, p. 100-106, 2021.

PACHECO, M. A. C. **Algoritmos Genéticos: Princípios e Aplicações. Laboratório de Inteligência Computacional Aplicada**. Pontifícia Universidade Católica do Rio de Janeiro. 1999.

PAZ, K.B.; ALENCAR, J. P.; ALMEIDA, L. C. de; LIMA, C. L. F. A.; GOMES, J. P. P. e SOARES, J.B. **Identificação de defeitos do tipo “panela” em pavimento asfáltico por meio de redes neurais convolucionais**. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES. 34^o, 100% Digital, 2020. Anais [...], [s.l]: Editora: Associação Nacional de Pesquisa e Ensino em Transportes, 2020.

PIOVEZAN, R e ANDRADE JUNIOR, P.P. Machine learning method for return direction forecasting of Exchange Traded Funds using classification and regression models. **Computational Economics**, p. 1-26, 2023

POSSEBON, E. P. **Estudo laboratorial das propriedades mecânicas de misturas com diferentes cimentos asfálticos de petróleo produzidos no Brasil**. 2018. Dissertação (Mestrado em Engenharia Civil) - Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Santa Maria, Santa Maria, 2018.

RASMUSSEN, C. E.; WILLIAMS, C. K. I. **Gaussian processes for machine learning**. [S.l.]: MIT Press, 2006.

RIBEIRO, A. J. A.; SILVA, C. A. U. DA e BARROSO, S. H. DE A. Metodologia de baixo custo para mapeamento geotécnico aplicado à pavimentação. **Transportes**, Fortaleza, v. 26, n. 2, p. 84–100, 2018.

ROBERTS, F.L.; KANDHAL, P.S.; BROWN, R.E.; LEE, D. e KENNEDY, T.W. **Hot Mix Asphalt Materials: Mixture Design, and Construction**. Maryland: National Asphalt Pavement Association. 1996.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, v. 65, p.386–408,1958.

RUSSEL, S. e NORVIG, P. **Artificial Intelligence - A Modern Approach**. Prentice Hall, Englewood Cliffs. 2013.

SANTANA, A. F. **Treinamento de redes neurais artificiais utilizando algoritmo genético em plataforma distribuída**. 2012. Trabalho de conclusão de curso (Graduação em Sistemas de Informação) - Universidade Federal de Lavras, Lavras, 2012.

SCHUSTER, S. L. **Estudo da influência da areia natural nas propriedades mecânicas de uma mistura asfáltica**. 2016. Trabalho de Conclusão de Curso (Graduação em Engenharia Civil) - Universidade Federal de Santa Maria, Santa Maria, 2016.

SEBAALY, H.; VARMA, S. e MAINA, J. W. Optimizing asphalt mix design process using artificial neural network and genetic algorithm. **Construction and Building Materials**, Pretoria, v.168, p. 660–670, 2018.

SHAN, L. Analysis of linear viscoelastic response function model for asphalt binders. **Journal of materials in civil engineering**, v. 28, n. 6, p. 04016010, 2016.

- SITTON, J. D. e ZEINALI, B.A. Story Rapid soil classification using artificial neural networks for use in constructing compressed earth blocks. **Construction and Building Materials**, Dallas v. 138, p. 214-221, DOI: 10.1016/j.conbuildmat.2017.02.006, 2017. Acesso em: 20 de nov 2023.
- SPECHT, L.P. Application of the theory of viscoelasticity to evaluate the resilient modulus test in asphalt mixes. **Construction and Building Materials**, Santa Maria, v.149, p.648–658, 2017.
- ŠTRUMBELJ, E. e KONONENKO, I. Explaining prediction models and individual predictions with feature contributions. **Knowl Inf Syst**, Ljubljana, v. 41, p. 647–665, 2014.
- SYBILSKI, D. Zero-shear viscosity of bituminous binder and its relation to bituminous mixture's rutting resistance. **Transportation Research Record Journal of the Transportation Research Board**, v. 1535, p. 15-21, 1996.
- SWEERE, G. T. H. **Unbound Granular Bases for Roads**. Delft, Netherlands: University of Delft, PhD Thesis. 1990.
- THUBE, D. T. Artificial Neural Network (ANN) based pavement deterioration models for low volume roads in India. **International Journal of Pavement Research and Technology**, Roorkee, v. 5, n. 2, p. 115–120, 2012.
- ULLAH, S.; TANYU, B. F. e ZAINAB, B. Development of an artificial neural network (ANN)-based model to predict permanent deformation of base course containing reclaimed asphalt pavement (RAP). **Road Materials and Pavement Design**, Fairfax, v. 1, p. 1–19, 2020.
- VALE, A. C. F.; SILVA, L. S. V.; BASTOS, J. B. S.; BARROSO, E. S.; BABADOPULOS, L. F. A. L.; PARENTE JUNIOR, E. e SOARES, J. B. **Acúmulo de deformação permanente em camadas asfálticas no dimensionamento de pavimentos a partir de modelo viscoplástico**. In: CONGRESSO NACIONAL DE PESQUISA E ENSINO EM TRANSPORTES, 28., Fortaleza, CE. 2022. Anais [...] Fortaleza: ANPET, 2020.
- VAVRIK, W. R.; HUBER, G. e PINE, W. Bailey method for gradation selection in HMA mixture design. **Transportation Research Board: Transportation Research Circular Number**, Washington, v. 70, p. 132–175, 2002.
- WOLFF, H. e ALEX T. VISSER. Incorporating elasto-plasticity in granular layer pavement design. [s. l.]: **ARRB**, v.105, p. 259-272, 1994.
- WITCZAK, M. W.; KALOUSH, K.; PELLINEN, T.; EL-BASYOUNY, M. e VON QUINTUS, H. NCHRP Report 465: Simple performance test for Superpave mix design. **Transportation Research Board**. Washington, DC. 2002.
- YU, H. **Design and characterization of asphalt mixtures based on particle packing and mechanical modeling**. 2012. Dissertation (PhD) – Washington State University, Washington D.C., 2012.
- ZHANG, J.; ALVAREZ, A. E.; LEE, S. I.; TORRES, S. A. e WALUBITA, L. Comparison of flow number, dynamic modulus, and repeated load tests for evaluation of HMA permanent deformation. **Construction and Building Materials**, Houston, v. 44, p.391-398, 2013.
- ZHOU, Z. e HOOKER, G. Unbiased measurement of feature importance in tree-based methods. **ACM transactions on knowledge discovery from data (TKDD)**, New York, v. 15, n. 2, p. 1–21. 2021.
- ZHU, H. e NODES, J. E. Contact based analysis of asphalt pavement with the effect of aggregate angularity. **Mechanics of Materials**, Tempe, p. 193-202, 2000.

APÊNDICE A – Banco de dados das 251 misturas utilizadas na pesquisa

ID da Mistura	Tipo	PEN	PA	VIS 135	TL	E*	FAD	Vv%	TMN	P34	P38	P4	P200	FN
1	AMP	54	58.8	1258	5.0	6860	42	4.00	19	97.4	72.6	54.5	3.9	406
2	CAP	28	53.2	587.5	5.0	14000	42	4.00	19	97.4	72.6	54.5	3.9	169
3	CAP	28	53.2	587.5	5.5	12424	38	4.00	12.5	99.1	90.6	54.4	3.2	213
4	CAP	43	48.7	445	5.5	10800	38	4.00	12.5	99.1	90.6	54.4	3.2	140
5	CAP	43	48.7	445	5.0	10700	42	4.00	19	97.4	72.6	54.5	3.9	95
6	AMP	51	55.1	810	4.6	13400	42	4.00	19	96.5	76.1	50.8	4.7	339
7	AMP	44	59.6	1048.7	4.5	14192	42	4.9	19	96.5	76.1	50.8	4.7	4237
8	AMP	53	60.85	1229	4.7	8750	37	4.00	19	95.84	64.69	46.58	4.3	3016
9	AMP	53	60.85	1229	7.3	5480	53.1	2.2	9.5	100	98.12	74.65	6.44	448
10	AMP	47	52.6	460.8	4.6	12270	36	4.00	19	95.84	64.69	46.58	4.3	2288
11	CAP	46	50.2	458.33	6.0	7830	42.8	3.7	12.5	99	77.65	54.20	6.46	27
12	CAP	46	50.2	458.33	5.0	9009	48.8	4.8	25	83.27	62.16	50.62	8.52	49
13	CAP	54	52.1	350	4.2	9226	30	4.00	19	100	64.3	42.9	7.3	354
14	AMP	50	57.6	995.83	4.51	9410	36.5	5.48	19	100	64.4	42.10	7.4	1685
15	AMP	50	57.6	995.83	4.8	10115	36.5	4.57	19	99.77	68.08	40.54	4.29	1318
16	CAP	31	53.6	458.33	5.0	13292	42.4	4.1	19	98	65	43	5	315
17	CAP	30	55.4	522.5	5.5	9954	38	4.00	12.5	99.1	90.6	54.4	3.2	215
18	CAP	33	55.4	477.5	5.5	10388	38	4.00	12.5	99.1	90.6	54.4	3.2	166
19	CAP	32	54.4	438.3	5.5	11148	38	4.00	12.5	99.1	90.6	54.4	3.2	220
20	CAP	33	53.6	482.5	5.5	11930	38	4.00	12.5	99.1	90.6	54.4	3.2	177
21	CAP	55	50.1	387.8	5.5	11166	38	4.00	12.5	99.1	90.6	54.4	3.2	133
22	CAP	55	50	310	5.5	9952	38	4.00	12.5	99.1	90.6	54.4	3.2	107
23	CAP	58	50	387.5	5.5	11168	38	4.00	12.5	99.1	90.6	54.4	3.2	110
24	AMP	47.5	54.1	1030	5.5	9070	40	4.22	12.5	99.4	82.9	50	5.4	1953
25	CAP	57.6	47.1	370	5.5	9287	40	4.22	12.5	99.4	82.9	50	5.4	448
26	CAP	31	52.1	585	5.5	11227	40	4.22	12.5	99.4	82.9	50	5.4	895
27	AM	22.6	61.3	1700	5.5	13540	40	4.22	12.5	99.4	82.9	50	5.4	7200
28	AM	23	58.4	737.5	5.5	17976	40	4.22	12.5	99.4	82.9	50	5.4	7200
29	AMP	31	88.1	1790	5.5	11672	40	4.22	12.5	99.4	82.9	50	5.4	7200
30	AMP	50	62	1040	5.1	9136	38.6	3.42	12.5	100	87.9	58.7	6.2	556
31	AMP	50	62	1040	5.8	6848	49.6	4.00	9.5	100	98.3	65.1	6.3	1588
32	AMP	29.6	65	2300	4.6	12726	41.8	3.9	19	100	63	38.2	4.2	972
33	AMP	29.6	65	2300	4.8	11400	38.4	4.5	12.5	100	87.9	58.7	6.2	495
34	AMP	32.6	82.6	740	4.9	13658	41.3	3.8	12.5	99.9	85.5	60.1	6.8	7200
35	AMP	48	77.1	1535	4.9	10182	35.6	3.8	12.5	99.9	85.5	60.1	6.8	7200
36	CAP	62	48.2	321	5.5	8918	40	4.22	12.5	99.4	82.9	50	5.4	972
37	CAP	51	50.5	353	5.5	9753	40	4.22	12.5	99.4	82.9	50	5.4	735
38	CAP	50	50.5	336.7	4.1	9046	50	5.00	19	99.4	53.6	38.5	7.5	152
39	AMP	59	63.4	1163	4.7	9616	48	3.32	19	99.4	53.6	38.5	7.4	2001
40	AMP	61	67.8	1093	4.7	8352	40	4.00	12.5	100	82.4	56.1	7.6	1669
41	CAP	44	52.1	377.7	6.2	10576	77.7	4.3	19	99.7	86.2	68.52	5.87	497
42	CAP	55	49.7	345.8	6.9	7359	41.8	4.00	19	100	83.9	54.1	6.8	158
43	CAP	54	49.5	360	4.7	9776	53.6	4.00	12.5	99.85	84.92	66.84	21.67	119
44	CAP	35	54.4	543.3	4.9	11160	53.7	4.00	12.5	99.85	84.92	66.84	21.67	212
45	CAP	55	50	348	5.0	11971	45.4	4.00	9.5	100	97.4	65.1	8.4	132
46	CAP	60	49.9	335	5.1	12957	39.1	4.00	19	100	83.18	63.69	6	262
47	AMP	56	68	991	5.2	12369	36.2	4.00	19	100	74.8	51.1	5.9	7200
48	AMP	50	87.5	1448	5.2	11353	36.5	4.00	12.5	100	74.8	51.1	5.9	1033
49	CAP	57	50	350	5.1	11480	39.9	3.7	19	100	81.5	55	6.3	284
50	CAP	58	51	349	4.7	11834	42.9	3.49	19	99.32	80.38	59.13	7.8	564

ID da Mistura	Tipo	PEN	PA	VIS 135	TL	E*	FAD	Vv%	TMN	P34	P38	P4	P200	FN
51	AMP	51	65.2	1850	4.8	10994	44.6	4.00	19	98	65.26	48.38	5.7	175
52	AMP	48	66	1395	5.4	10869	35.5	4.5	12.5	99.88	81.74	60.87	4.86	7200
53	AMP	48	66	1395	5.4	9595	39	4.5	19	99.88	81.74	60.87	4.86	7200
54	CAP	51	50.8	282	4.6	15838	37.8	05.03	19	100	76.5	51.4	5.5	779
55	CAP	57	48.2	335	5.0	12044	40.4	04.05	12.5	100	90.4	81.8	5.7	359
56	AMP	56.5	72	1115	5.2	10208	38.7	3.57	19	100	75.70	86.69	5.8	647
57	AMP	29	70.6	1528	5.5	12123	38.1	3.47	12.5	100	75.7	59.4	5.8	2001
58	CAP	57	48.2	335	4.8	13240	41.5	3.71	12.5	100	81	56.5	5.7	485
59	CAP	56	49.5	343	5.6	10423	45.6	4.00	9.5	100	98.4	61.7	6.3	226
60	AMP	52	72	1610	5.0	8226	48.1	4.00	19	100	80.5	56	6	923
61	AMP	52	72	1610	5.2	8851	37.4	4.00	19	95.9	70.7	54.3	4.5	986
62	CAP	57	48.2	355	4.8	10510	48.4	4.8	19	100	71.5	46.2	5.5	336
63	CAP	52	48.6	328	5.0	9156	46.6	04.07	12.5	100	82.55	54.39	6.82	822
64	CAP	58	49.4	295	4.6	11641	39.8	4.31	19	94.04	58.84	42.96	5.12	549
65	AMP	56	73	1252	5.0	10633.5	44.9	3.93	19	100	82.55	54.39	6.82	7200
66	CAP	53	52	360	4.6	9465	56	3.6	19	100	80.8	54.8	7.6	81
67	CAP	32	52	470	4.6	12304	55	3.6	19	100	80.8	54.8	7.6	98
68	CAP	32	52	470	4.4	10700	75	3.9	19	100	70.9	37.5	5.6	193
69	CAP	53	52	360	4.5	10132	75	3.7	19	100	70.9	37.5	5.6	184
70	CAP	32	52	470	4.6	11788	43	3.9	9.5	100	99.3	52	7	384
71	CAP	53	52	360	4.6	10776	43	3.9	9.5	100	99.3	52	7	384
72	CAP	32	52	470	4.9	10286	73	5.00	9.5	100	100	34	6.2	215
73	CAP	53	52	360	5.1	8605	73	4.9	9.5	100	100	34	6.2	92
74	CAP	32	52	470	4.9	10571	49	4.4	19	100	78	53.3	6.8	87
75	CAP	53	52	360	4.7	9673	49	4.1	19	100	78	53.3	6.8	48
76	CAP	32	52	470	4.4	13329	46	4.2	19	100	69.5	37.2	5.3	173
77	CAP	53	52	360	4.4	10985	46	4.3	19	100	69.5	37.2	5.3	68
78	CAP	32	52	470	4.9	10571	49	3.8	12.5	100	86.3	53.7	7.1	384
79	CAP	52	52	360	5.0	8704	50	4.00	12.5	100	86.3	53.7	7.1	384
80	CAP	32	52	470	4.0	10700	46	4.6	12.5	100	80.4	37.5	5.3	214
81	CAP	53	52	360	4.0	11480	46	4.4	12.5	100	80.4	37.5	5.3	96
82	CAP	32	52	470	4.8	12298	54	4.2	19	100	83.3	54	7.5	252
83	CAP	53	52	360	5.0	10683	54	4.1	19	100	83.3	54	7.5	126
84	CAP	32	52	470	5.1	13908	41	3.9	19	100	75.6	48.9	6.2	307
85	CAP	53	52	360	5.2	12446	41	4.5	19	100	75.6	48.9	6.2	44
86	CAP	32	52	470	5.1	11937	47	3.9	9.5	100	98.5	66.6	6.8	427
87	CAP	53	52	360	5.2	10803	47	4.5	9.5	100	98.5	66.6	6.8	170
88	CAP	32	52	470	4.0	9583	46	4.6	9.5	100	98.7	40.9	5.3	214
89	CAP	53	52	360	5.8	6775	66	5.4	9.5	100	98.7	40.9	5.3	70
90	CAP	30	63.2	430	5.5	12000	36.2	4.00	19	100	72.1	51.7	6.2	214
91	CAP	62	48.2	369	5.4	9700	36.1	4.00	19	100	72.2	51.8	6.3	128
92	CAP	57	48.6	300	5.7	9800	36.3	4.00	19	100	72.3	51.9	6.4	176
93	CAP	59	48.1	314	5.9	8200	36.2	4.00	19	100	72.4	51.10	6.5	70
94	CAP	54	49	392	5.7	9100	36.4	4.00	19	100	72.5	51.11	6.6	165
95	CAP	50	50.2	315	5.6	10000	36.3	4.00	19	100	72.6	51.12	6.7	132
96	CAP	53	52	360	4.6	12304	55	3.6	12.5	100	80.8	54.8	7.6	98
97	CAP	32	52	470	4.6	9465	56	3.9	12.5	100	80.8	54.8	7.6	81
98	AMP	53	64	1290	4.6	8530	55	3.9	12.5	100	80.8	54.8	7.6	2390
99	AMP	50	67	1952	4.7	8684	55	3.9	12.5	100	80.8	54.8	7.6	915
100	AMP	40	65	1560	4.8	11230	55	4	12.5	100	80.8	54.8	7.6	394

ID da Mistura	Tipo	PEN	PA	VIS 135	TL	E*	FAD	Vv%	TMN	P34	P38	P4	P200	FN
101	AMP	57	66	1240	4.7	10175	55	3.9	12.5	100	80.8	54.8	7.6	1582
102	AB	44	57	1360	5.5	7303	55	3.9	12.5	100	80.8	54.8	7.6	5034
103	AB	42	59	1190	5.3	7707	55	4	12.5	100	80.8	54.8	7.6	2125
104	AB	58	59	1320	5.1	6947	55	3.9	12.5	100	80.8	54.8	7.6	1931
105	AM	20	67	900	4.7	11232	55	3.9	12.5	100	80.8	54.8	7.6	399
106	AM	21	63	1200	4.7	11406	55	3.9	12.5	100	80.8	54.8	7.6	284
107	AM	24	65	1020	4.5	8439	55	3.7	12.5	100	80.8	54.8	7.6	7030
108	CAP	53	52	360	4.4	10700	75	3.9	9.5	100	70.9	37.5	5.6	193
109	CAP	32	52	470	4.5	10132	75	3.7	9.5	100	70.9	37.5	5.6	184
110	AMP	53	64	1290	4.3	8023	75	3.6	9.5	100	70.9	37.5	5.6	2944
111	AMP	50	67	1952	4.5	8232	75	3.9	9.5	100	70.9	37.5	5.6	512
112	AMP	40	65	1560	4.6	10089	75	4.3	9.5	100	70.9	37.5	5.6	2424
113	AMP	57	66	1240	4.5	10960	75	3.9	9.5	100	70.9	37.5	5.6	1978
114	AB	44	57	1360	5.0	7161	75	4	9.5	100	70.9	37.5	5.6	6286
115	AB	42	59	1190	5.4	8600	75	4.2	9.5	100	70.9	37.5	5.6	2576
116	AB	58	59	1320	5.2	7709	75	4.1	9.5	100	70.9	37.5	5.6	2123
117	AM	20	67	900	4.4	13727	75	3.9	9.5	100	70.9	37.5	5.6	1120
118	AM	21	63	1200	4.5	12237	75	3.8	9.5	100	70.9	37.5	5.6	583
119	AM	24	65	1020	4.8	10830	75	3.9	9.5	100	70.9	37.5	5.6	7818
120	CAP	53	52	360	4.6	11788	43	3.9	9.5	100	99.3	52	7	179
121	CAP	32	52	470	4.6	10776	43	3.9	9.5	100	99.3	52	7	113
122	AMP	53	64	1290	4.8	8336	43	3.7	9.5	100	99.3	52	7	1621
123	AMP	50	67	1952	4.9	8607	43	4.1	9.5	100	99.3	52	7	752
124	AMP	40	65	1560	4.9	9723	43	4	9.5	100	99.3	52	7	3169
125	AMP	57	66	1240	5.0	12106	43	4	9.5	100	99.3	52	7	4071
126	AB	44	57	1360	5.3	9197	43	4	9.5	100	99.3	52	7	1436
127	AB	42	59	1190	5.5	9847	43	4	9.5	100	99.3	52	7	6855
128	AB	58	59	1320	5.6	8844	43	4.1	9.5	100	99.3	52	7	7808
129	AM	20	67	900	5.0	10751	43	4	9.5	100	99.3	52	7	544
130	AM	21	63	1200	4.7	12515	43	4.1	9.5	100	99.3	52	7	129
131	AM	24	65	1020	5.2	12611	43	3.9	9.5	100	99.3	52	7	1931
132	CAP	53	52	360	4.9	10286	73	5	9.5	100	100	34	6.2	215
133	CAP	32	52	470	5.1	8605	73	4.9	9.5	100	100	34	6.2	92
134	AMP	53	64	1290	5.2	8569	73	5.1	9.5	100	100	34	6.2	2816
135	AMP	50	67	1952	5.2	9372	73	4.9	9.5	100	100	34	6.2	608
136	AMP	40	65	1560	5.2	10271	73	4.9	9.5	100	100	34	6.2	1077
137	AMP	57	66	1240	5.4	9984	73	5.1	9.5	100	100	34	6.2	1416
138	AB	44	57	1360	5.9	8773	73	4.8	9.5	100	100	34	6.2	1762
139	AB	42	59	1190	5.9	9750	73	4.9	9.5	100	100	34	6.2	1617
140	AB	58	59	1320	5.5	8187	73	5.1	9.5	100	100	34	6.2	1387
141	AM	20	67	900	5.1	13464	73	5	9.5	100	100	34	6.2	1140
142	AM	21	63	1200	5.1	12355	73	5	9.5	100	100	34	6.2	388
143	AM	24	65	1020	5.3	10585	73	5	9.5	100	100	34	6.2	1931
144	CAP	53	52	360	4.9	10571	49	4.4	12.5	100	78	53.3	6.8	87
145	CAP	32	52	470	4.7	8704	49	4.1	12.5	100	78	53.3	6.8	48
146	AMP	53	64	1290	4.7	6950	49	4.1	12.5	100	78	53.3	6.8	312
147	AMP	50	67	1952	4.7	7622	49	4	12.5	100	78	53.3	6.8	115
148	AMP	40	65	1560	4.7	8619	49	4.2	12.5	100	78	53.3	6.8	496
149	AMP	57	66	1240	4.6	7936	49	4.3	12.5	100	78	53.3	6.8	297
150	AB	44	57	1360	5.0	8381	49	4.3	12.5	100	78	53.3	6.8	3516
151	AB	42	59	1190	5.2	6292	49	4.1	12.5	100	78	53.3	6.8	340

ID da Mistura	Tipo	PEN	PA	VIS 135	TL	E*	FAD	Vv%	TMN	P34	P38	P4	P200	FN
152	AB	58	59	1320	5.2	6802	49	4.3	12.5	100	78	53.3	6.8	1543
153	AM	20	67	900	4.8	10915	49	4.0	12.5	100	78	53.3	6.8	362
154	AM	21	63	1200	4.7	10055	49	4.0	12.5	100	78	53.3	6.8	123
155	AM	24	65	1020	5.3	9371	49	4.1	12.5	100	78	53.3	6.8	626
156	CAP	53	52	360	4.4	13329	46	4.2	19	100	69.5	37.2	5.3	173
157	CAP	32	52	470	4.4	10985	46	4.3	19	100	69.5	37.2	5.3	68
158	AMP	53	64	1290	4.3	9782	46	4.0	19	100	69.5	37.2	5.3	1704
159	AMP	50	67	1952	4.2	11722	46	4.0	19	100	69.5	37.2	5.3	948
160	AMP	40	65	1560	4.3	9973	46	4.1	19	100	69.5	37.2	5.3	584
161	AMP	57	66	1240	4.3	10268	46	4.1	19	100	69.5	37.2	5.3	622
162	AB	44	57	1360	4.9	8227	46	4.2	19	100	69.5	37.2	5.3	1860
163	AB	42	59	1190	4.9	9172	46	4.2	19	100	69.5	37.2	5.3	3700
164	AB	58	59	1320	4.8	7757	46	4.0	19	100	69.5	37.2	5.3	2706
165	AM	20	67	900	4.3	15543	46	4.1	19	100	69.5	37.2	5.3	4645
166	AM	21	63	1200	4.3	15198	46	4.0	19	100	69.5	37.2	5.3	1251
167	AM	24	65	1020	4.7	13961	46	4.2	19	100	69.5	37.2	5.3	9027
168	CAP	53	52	360	4.9	10571	50	3.8	9.5	100	86.3	53.7	7.1	81
169	CAP	32	52	470	5.0	8704	50	4.0	9.5	100	86.3	53.7	7.1	62
170	AMP	53	64	1290	5.0	6950	50	4.1	9.5	100	86.3	53.7	7.1	1502
171	AMP	50	67	1952	5.0	7622	50	4.0	9.5	100	86.3	53.7	7.1	932
172	AMP	40	65	1560	5.0	8619	50	4.0	9.5	100	86.3	53.7	7.1	1197
173	AMP	57	66	1240	5.0	7936	50	4.1	9.5	100	86.3	53.7	7.1	455
174	AB	44	57	1360	5.4	8381	50	4.2	9.5	100	86.3	53.7	7.1	146
175	AB	42	59	1190	5.7	8811	50	4.3	9.5	100	86.3	53.7	7.1	458
176	AB	58	59	1320	5.5	7840	50	4.3	9.5	100	86.3	53.7	7.1	323
177	AM	20	67	900	5.0	10505	50	4.0	9.5	100	86.3	53.7	7.1	173
178	AM	21	63	1200	4.9	13697	50	4.2	9.5	100	86.3	53.7	7.1	466
179	AM	24	65	1020	5.3	16096	50	4.1	9.5	100	86.3	53.7	7.1	6915
180	CAP	53	52	360	4.0	10700	46	4.6	9.5	100	80.4	37.5	5.3	214
181	CAP	32	52	470	4.0	11480	46	4.4	9.5	100	80.4	37.5	5.3	96
182	AMP	53	64	1290	4.0	10395	46	4.2	9.5	100	80.4	37.5	5.3	4622
183	AMP	50	67	1952	4.0	12462	46	4.6	9.5	100	80.4	37.5	5.3	2744
184	AMP	40	65	1560	4.0	9838	46	4.5	9.5	100	80.4	37.5	5.3	567
185	AMP	57	66	1240	4.0	10827	46	4.4	9.5	100	80.4	37.5	5.3	1058
186	AB	44	57	1360	4.7	8659	46	4.6	9.5	100	80.4	37.5	5.3	542
187	AB	42	59	1190	4.6	8385	46	4.6	9.5	100	80.4	37.5	5.3	272
188	AB	58	59	1320	4.7	7928	46	4.7	9.5	100	80.4	37.5	5.3	269
189	AM	20	67	900	4.1	13626	46	4.4	9.5	100	80.4	37.5	5.3	228
190	AM	21	63	1200	4.0	12828	46	4.6	9.5	100	80.4	37.5	5.3	401
191	AM	24	65	1020	4.0	15944	46	4.5	9.5	100	80.4	37.5	5.3	5844
192	CAP	53	52	360	4.8	12298	54	4.2	12.5	100	83.3	54	7.5	252
193	CAP	32	52	470	5.0	10683	54	4.1	12.5	100	83.3	54	7.5	126
194	AMP	53	64	1290	5.0	10566	54	4.5	12.5	100	83.3	54	7.5	6885
195	AMP	50	67	1952	4.9	12054	54	4.3	12.5	100	83.3	54	7.5	8630
196	AMP	40	65	1560	5.1	10936	54	3.9	12.5	100	83.3	54	7.5	6627
197	AMP	57	66	1240	4.7	11590	54	3.9	12.5	100	83.3	54	7.5	3445
198	AB	44	57	1360	5.6	11985	54	3.9	12.5	100	83.3	54	7.5	8856
199	AB	42	59	1190	5.4	9788	54	4.1	12.5	100	83.3	54	7.5	5554
200	AB	58	59	1320	5.4	7681	54	4.0	12.5	100	83.3	54	7.5	2252

ID da Mistura	Tipo	PEN	PA	VIS 135	TL	E*	FAD	Vv%	TMN	P34	P38	P4	P200	FN
201	AM	20	67	900	4.8	13311	54	3.9	12.5	100	83.3	54	7.5	846
202	AM	21	63	1200	4.7	13392	54	3.8	12.5	100	83.3	54	7.5	252
203	AM	24	65	1020	5.2	14355	54	4.1	12.5	100	83.3	54	7.5	6847
204	CAP	53	52	360	5.1	13908	41	3.9	19	100	75.6	48.9	6.2	307
205	CAP	32	52	470	5.2	12446	41	4.5	19	100	75.6	48.9	6.2	44
206	AMP	53	64	1290	5.3	10655	41	4.3	19	100	75.6	48.9	6.2	3174
207	AMP	50	67	1952	5.4	11823	41	4.1	19	100	75.6	48.9	6.2	2027
208	AMP	40	65	1560	5.1	10046	41	3.9	19	100	75.6	48.9	6.2	483
209	AMP	57	66	1240	4.8	12004	41	3.9	19	100	75.6	48.9	6.2	1478
210	AB	44	57	1360	6.3	7973	41	4.2	19	100	75.6	48.9	6.2	921
211	AB	42	59	1190	5.6	7912	41	4.3	19	100	75.6	48.9	6.2	876
212	AB	58	59	1320	5.8	7852	41	3.7	19	100	75.6	48.9	6.2	831
213	AM	20	67	900	4.9	12215	41	4.1	19	100	75.6	48.9	6.2	595
214	AM	21	63	1200	5.8	13014	41	4.1	19	100	75.6	48.9	6.2	863
215	AM	24	65	1020	6.0	14403	41	4.2	19	100	75.6	48.9	6.2	8658
216	CAP	53	52	360	5.1	11937	47	3.9	9.5	100	98.5	66.6	6.8	427
217	CAP	32	52	470	5.2	10803	47	4.5	9.5	100	98.5	66.6	6.8	170
218	AMP	53	64	1290	5.3	9729	47	4.3	9.5	100	98.5	66.6	6.8	2856
219	AMP	50	67	1952	5.4	9258	47	4.1	9.5	100	98.5	66.6	6.8	1506
220	AMP	40	65	1560	5.9	8104	47	4.3	9.5	100	98.5	66.6	6.8	2309
221	AMP	57	66	1240	5.8	8321	47	3.9	9.5	100	98.5	66.6	6.8	708
222	AB	44	57	1360	6.3	6686	47	4.2	9.5	100	98.5	66.6	6.8	1205
223	AB	42	59	1190	6.3	7095	47	4.5	9.5	100	98.5	66.6	6.8	1051
224	AB	58	59	1320	6.2	7503	47	4.1	9.5	100	98.5	66.6	6.8	897
225	AM	20	67	900	6.1	9917	47	4.0	9.5	100	98.5	66.6	6.8	298
226	AM	21	63	1200	5.8	10906	47	4.1	9.5	100	98.5	66.6	6.8	463
227	AM	24	65	1020	6.0	14029	47	4.2	9.5	100	98.5	66.6	6.8	8178
228	CAP	53	52	360	5.7	9583	46	5.6	9.5	100	98.7	40.9	5.3	186
229	CAP	32	52	470	5.8	6775	46	5.4	9.5	100	98.7	40.9	5.3	70
230	AMP	53	64	1290	5.6	10135	46	5.1	9.5	100	98.7	40.9	5.3	4025
231	AMP	50	67	1952	6.2	10021	46	5.6	9.5	100	98.7	40.9	5.3	849
232	AMP	40	65	1560	6.1	6544	46	5.2	9.5	100	98.7	40.9	5.3	221
233	AMP	57	66	1240	5.7	8411	46	5.0	9.5	100	98.7	40.9	5.3	889
234	AB	44	57	1360	6.5	5542	46	5.3	9.5	100	98.7	40.9	5.3	419
235	AB	42	59	1190	6.5	5897	46	5.1	9.5	100	98.7	40.9	5.3	569
236	AB	58	59	1320	6.1	6251	46	4.8	9.5	100	98.7	40.9	5.3	718
237	AM	20	67	900	6.4	10592	46	4.9	9.5	100	98.7	40.9	5.3	1070
238	AM	21	63	1200	6.1	9956	46	5.2	9.5	100	98.7	40.9	5.3	444
239	AM	24	65	1020	6.3	12104	46	5.7	9.5	100	98.7	40.9	5.3	6269
240	AB	57	59	2250	5.1	7934	52	3.9	19	100	80.8	54.8	7.6	2179
241	AB	57	59	2250	5.2	7934	52	4.1	19	100	70.9	37.5	5.6	2482
242	AB	57	59	2250	5.6	7934	52	4.1	9.5	100	99.3	52	7	2174
243	AB	57	59	2250	5.5	7934	52	5.1	9.5	100	100	34	6.2	1096
244	AB	44	57	2140	5.5	7934	52	3.9	19	100	80.8	54.8	7.6	4185
245	AB	44	57	2140	5	7934	52	4	19	100	70.9	37.5	5.6	6039
246	AB	44	57	2140	5.3	7934	52	4	9.5	100	99.3	52	7	2130
247	AB	44	57	2140	5.9	7934	52	4.8	9.5	100	100	34	6.2	1783
248	AB	34	53	490	4.6	10708	52	3.6	19	100	80.8	54.8	7.6	132
249	AB	34	53	490	4.4	10708	52	3.9	19	100	70.9	37.5	5.6	193
250	AB	34	53	490	4.6	10708	52	3.9	9.5	100	99.3	52	7	179
251	AB	34	53	490	4.9	10708	52	5	9.5	100	100	34	6.2	215