

Uso de aprendizado de máquina probabilístico para avaliação da deformação permanente em misturas asfálticas

Use of probabilistic machine learning for assessing rutting in asphalt mixtures

Pedro Luiz Ribeiro Rocha¹, Jorge Barbosa Soares², Iuri Sidney Bessa³, César Lincoln Cavalcante Mattos⁴, e Carlos David Rodrigues Melo⁵

¹Universidade Federal do Ceará, Fortaleza, Ceará, Brasil, pedro.luz.rocha@det.ufc.br, 0009-0004-6154-2302

²Universidade Federal do Ceará, Fortaleza, Ceará, Brasil, jsoares@det.ufc.br, 0000-0002-2940-6309

³Universidade Federal do Ceará, Fortaleza, Ceará, Brasil, iuri@det.ufc.br, 0000-0002-4711-0552

⁴Universidade Federal do Ceará, Fortaleza, Ceará, Brasil, cesarlincoln@dc.ufc.br, 0000-0002-2404-3625

⁵Universidade Federal do Ceará, Fortaleza, Ceará, Brasil, cdavidmelo@ufc.br, 0000-0003-3334-2309

Recebido: 13 de agosto de 2025

Revisão final: 9 de abril de 2026

Aceito: 10 de abril de 2026

Publicado: 9 de junho de 2026

Editor Associado:

Francisco Thiago Sacramento Aragão
Universidade Federal do Rio de Janeiro,
RJ, Brasil

Palavras-chave:

Misturas asfálticas;
Deformação permanente;
Aprendizado de máquina probabilístico.

Keywords:

Asphalt mixtures;
Permanent deformation;
Probabilistic machine learning.

DOI: [10.58922/transportes.v34.e3165](https://doi.org/10.58922/transportes.v34.e3165)



RESUMO

A deformação permanente é um defeito que afeta a durabilidade e a qualidade dos pavimentos asfálticos, sendo influenciada pelas propriedades do ligante e dos agregados que compõem as misturas asfálticas usadas em revestimentos. O ensaio uniaxial de carga repetida, cujo resultado é o *Flow Number*, é usado para avaliar a resistência da mistura à deformação permanente. Uma alternativa a esse ensaio é a aplicação de algoritmos de aprendizado de máquina, uma abordagem que permite prever o nível de tráfego equivalente ao *Flow Number* a partir de dados de parâmetros das misturas. Outra opção é a aplicação de modelos probabilísticos, que consideram a incerteza das variáveis previstas. Dentre eles, destacam-se os Processos Gaussianos, que se sobressaem com uma pequena quantidade de dados. Este trabalho tem como objetivo avaliar o desempenho de diferentes modelos probabilísticos na classificação de misturas asfálticas com relação à deformação permanente, auxiliando na seleção de materiais e no dimensionamento de pavimentos. Construiu-se um banco de dados com 252 misturas, por meio da coleta de informações de ensaios laboratoriais. Aplicaram-se modelos de regressão e classificação para prever o nível de tráfego equivalente, obtendo resultados que demonstram a viabilidade dessa abordagem, alcançando uma acurácia de até 90%.

ABSTRACT

Permanent deformation is a distress that affects the durability and quality of asphalt pavements, being influenced by the properties of the binder and aggregates in asphalt mixtures used on surface courses. The uniaxial repeated load test, which provides the Flow Number, is used to evaluate the mixture rutting resistance. An alternative to this test is the use of machine learning algorithms, an approach that enables the prediction of the traffic level equivalent to the Flow Number using data from various parameters of asphalt mixtures. Another option is the use of probabilistic models, which consider the uncertainty of the predicted variables. Among these, Gaussian Processes stand out, excelling with small amounts of data. The study herein aims to assess the performance of various probabilistic models in classifying asphalt mixtures concerning permanent deformation. The goal is to assist in material selection and pavement design. A database, containing 252 mixtures, was constructed by gathering information from various laboratory tests. Using regression and classification models to predict the equivalent traffic level, the results obtained demonstrate the feasibility of this approach, achieving an accuracy of up to 90%.

1. Introdução

Um dos problemas mais comuns na estrutura de pavimentos asfálticos é a deformação permanente (DP). Este defeito depende das propriedades dos componentes da mistura asfáltica usada em revestimentos, o ligante asfáltico e os agregados, e da interação entre ambos (Portela Neto, 2018; Gomes, Soares & Bastos, 2019; Ferreira *et al.*, 2020), o que torna necessária a avaliação da resistência a essa patologia.

A resistência à deformação permanente pode ser avaliada de duas maneiras: previsão do defeito ao longo do tempo por meio de modelos viscoplásticos (Kim, 2015) ou estabelecimento de um critério de falha, como medida de prevenção. Esse critério pode ser determinado por meio do ensaio uniaxial de

carga repetida (DNIT, 2018), que fornece o *Flow Number* (FN). Quanto maior for esse valor, melhor é a capacidade da mistura de resistir à deformação permanente. Assim, é possível classificar o FN de acordo com o nível de tráfego equivalente que o revestimento poderá suportar (Bastos, Soares & Nascimento, 2017). Como esse ensaio demanda tempo e recursos, buscaram-se alternativas mais econômicas para avaliar a deformação permanente de misturas, sobretudo para fases preliminares do projeto (dosagem das misturas) ou para anteprojetos.

Uma dessas alternativas é o uso de modelos de aprendizado de máquina (*machine learning*, ML), uma metodologia que vem sendo cada vez mais difundida nas mais diversas áreas de estudo. Os modelos de ML têm sido usados para previsão de diversos parâmetros das misturas asfálticas. Dentre os parâmetros previstos, destacam-se o módulo dinâmico (El-Badawy, El-Hakim & Awed, 2018), os parâmetros volumétricos de dosagem (Ozturk & Kutay, 2014) e a vida de fadiga (Ahmed, Green & Khalid, 2017).

Para prever o nível de tráfego equivalente ao FN, Mariano (2023) construiu um banco de dados contendo propriedades de 251 misturas asfálticas. Selecionaram-se parâmetros do ligante, dos agregados e da mistura como variáveis de entrada de modelos de aprendizado baseados em Redes Neurais e Árvores de Decisão, dois dos algoritmos mais comumente utilizados. Esse banco de dados foi modificado e expandido (atingindo 333 misturas) para o propósito da presente pesquisa, usando informações adicionais coletadas por Carvalho (2025) e pelos próprios autores.

Mariano (2023) usou uma abordagem determinística, a qual considera que as saídas do modelo assumem um valor específico. Uma outra opção é a aplicação de modelos de aprendizado de máquina probabilístico. Essa metodologia considera que as saídas do modelo são distribuições de probabilidade, que possuem uma média e um desvio padrão, gerando uma faixa de valores possíveis.

Modelos probabilísticos como os Processos Gaussianos (*Gaussian Processes*, GP) se destacam especialmente em cenários de dados limitados, como é o caso do banco de dados avaliado. Além disso, a possibilidade de ajuste direto dos hiperparâmetros (variáveis de configuração dos modelos) via máxima verossimilhança marginal (Williams & Rasmussen, 2006) diminui a necessidade de busca por valores apropriados para a configuração do modelo, uma etapa que demanda bastante tempo em geral. Os trabalhos mais recentes sobre misturas asfálticas que empregam GP focaram na previsão dos parâmetros volumétricos de dosagem (Zhao et al., 2022) e do módulo dinâmico (Liu et al., 2023).

Utilizando diferentes modelos baseados em algoritmos de aprendizado de máquina probabilístico, espera-se que o presente trabalho possa ser usado de forma alternativa à realização de ensaios laboratoriais de resistência à deformação permanente. Para prever o nível de tráfego equivalente ao FN, aplicaram-se modelos probabilísticos para classificação, variando os atributos de entrada e as saídas do banco de dados. Trata-se de uma ferramenta auxiliar ao projeto de misturas asfálticas, reduzindo ou mesmo eliminando o ensaio de obtenção do FN, nesse último caso sobretudo para fins de anteprojetos.

2. Revisão da literatura

2.1. Ensaio uniaxial de carga repetida

O ensaio uniaxial de carga repetida tem sido usado frequentemente por sua simplicidade e por também representar razoavelmente o comportamento mecânico de misturas asfálticas após a incorporação de agentes modificadores (Onofre, 2012). Esse ensaio consiste na aplicação de um carregamento de compressão repetido, sem confinamento, em um corpo de prova cilíndrico, em que pulsos de carga são seguidos por períodos de repouso (DNIT, 2018). Ao longo do ensaio, são observadas três zonas de deformação distintas (Witczak, 2007), como mostra a Figura 1:

- Zona Primária: elevada deformação permanente inicial, com uma taxa de deformação plástica decrescente. Associada predominantemente à mudança de volume;
- Zona Secundária: taxa de deformação se mantém constante em cada ciclo. Também associada à mudança de volume;
- Zona Terciária: ocorre no ponto de inflexão da curva, quando a taxa de deformação é mínima. Associada às deformações de cisalhamento, sem mudança de volume.

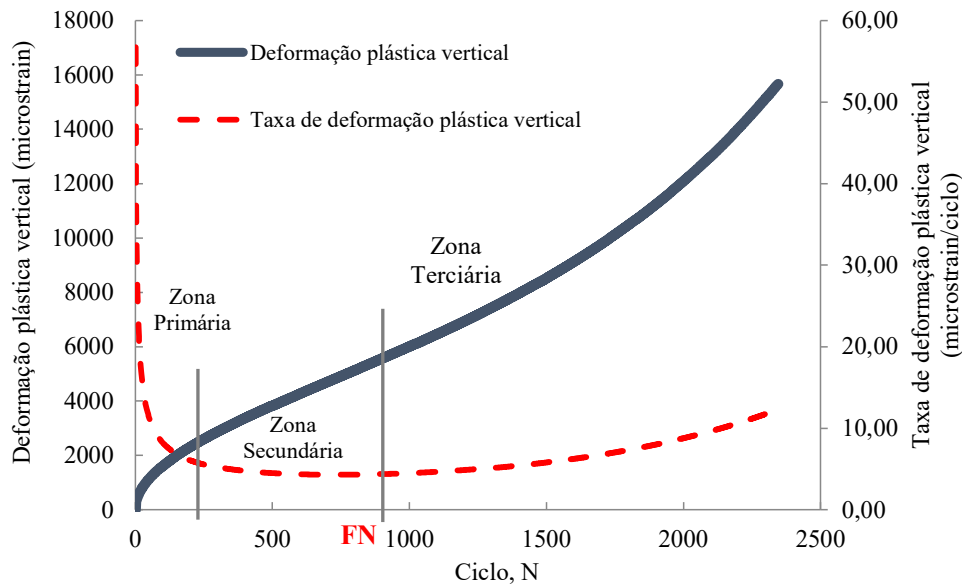


Figura 1 • Curva típica do ensaio uniaxial de carga repetida (Barros, 2022).

Esse procedimento fornece o FN, que corresponde ao número de ciclos de carga em que a taxa de deformação uniaxial do corpo de prova é mínima, e representa o início do fluxo de cisalhamento com volume constante. O ensaio é realizado até a ruptura do corpo de prova (início da zona terciária) ou até o número de ciclos atingir 7.200, o que equivale a aproximadamente 2 horas de carregamento (DNIT, 2018).

2.2. Aplicação de modelos para caracterização da deformação permanente

Diversas abordagens foram utilizadas para avaliar a deformação permanente de misturas asfálticas. Ouni, Dony & Colin (2014) aplicaram uma abordagem estocástica para modelar a incerteza do afundamento de trilha de roda (ATR) gerado por um simulador de tráfego, usando amostras de asfalto virgem e envelhecido. Singh & Swamy (2020) usaram distribuições de probabilidade normal e lognormal para prever o ATR obtido em um simulador de tráfego.

Várias pesquisas foram realizadas com o intuito de determinar o FN por meio de modelos estatísticos que o representem por equações. Apeageyi (2011) construiu modelos de regressão múltipla de 1ª. ordem, usando o módulo dinâmico da mistura e a granulometria dos agregados como variáveis de entrada para prever o FN. Gandomi et al. (2011) utilizaram programação genética (PG), uma corrente dos algoritmos genéticos, para correlacionar parâmetros de dosagem, dos agregados e da mistura ao FN.

Com o passar dos anos, o uso de técnicas baseadas em aprendizado de máquina se popularizou, graças ao aumento do poder computacional e ao desenvolvimento de algoritmos cada vez mais eficazes e de fácil aplicação. Mirzahosseini et al. (2015) usaram parâmetros de dosagem, dos agregados e da mistura para prever o FN. Foram aplicados modelos de programação genética e Redes Neurais Artificiais (RNA), resultando em um R^2 de 0,94. Mariano (2023) transformou as saídas, originalmente valores de FN, no nível de tráfego equivalente, para aplicação de modelos de classificação. Usando algoritmos baseados em RNAs e árvores de decisão, obteve uma acurácia de 83%.

2.3. Aprendizado probabilístico Bayesiano

Um modelo de aprendizado de máquina é um algoritmo que estabelece a relação entre as variáveis de entrada e as variáveis de saída, sendo análogo a uma função matemática (Equação 1):

$$Y = f(X, W) \quad (1)$$

em que: \mathbf{X} são as variáveis de entrada; \mathbf{Y} são as variáveis de saída; e \mathbf{W} são os parâmetros do modelo, que correlacionam \mathbf{X} a \mathbf{Y} . O aprendizado de máquina pode ser aplicado por meio de uma perspectiva probabilística, ou seja, tratando os valores desconhecidos (os parâmetros \mathbf{W} e as saídas \mathbf{Y} de um modelo, por exemplo) como variáveis aleatórias, representadas por distribuições de probabilidade que descrevem um conjunto de valores possíveis que a variável pode assumir (Murphy, 2022).

O aprendizado Bayesiano é uma subdivisão do aprendizado probabilístico. É uma abordagem baseada na interpretação Bayesiana da probabilidade, em que a incerteza de um evento pode ser atualizada ao incorporar algum conhecimento sobre esse evento por meio do teorema de Bayes (Bishop, 2006).

Para um vetor de parâmetros \mathbf{w} e um conjunto de dados $D = (\mathbf{X}, \mathbf{y})$, têm-se:

$$p(\mathbf{w}|D) = \frac{p(D|\mathbf{w})p(\mathbf{w})}{p(D)} \quad (2)$$

em que: $p(\mathbf{w})$ é a probabilidade a priori (incerteza inicial sobre os parâmetros antes de observar os dados); $p(D|\mathbf{w})$ é a verossimilhança (descreve a relação entre os parâmetros e os dados); $p(\mathbf{w}|D)$ é a probabilidade a posteriori (incerteza sobre os parâmetros após observar os dados); e $p(D)$ é a evidência (termo de normalização). Por exemplo, inicialmente pode-se assumir que a média e o desvio padrão que representam a incerteza sobre um parâmetro são, respectivamente, 0 e 1. Por meio da verossimilhança, a incerteza sobre o parâmetro é atualizada, obtendo-se novos valores de média e desvio padrão, após considerar como os parâmetros e os dados se correlacionam.

2.4. Processos Gaussianos

Uma das principais distribuições de probabilidade é a distribuição Gaussiana, também conhecida como normal. É representada por $N(x|\mu, \sigma^2)$, sendo descrita pela Equação 3:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} \quad (3)$$

em que: $p(x)$ é a distribuição Gaussiana univariada; μ é a média; e σ^2 é a variância. A raiz quadrada da variância é o desvio padrão (σ). Ao considerar um vetor n -dimensional de múltiplas variáveis \mathbf{x} , tem-se a Equação 4:

$$p(\mathbf{x}) = \frac{1}{|\Sigma|^{1/2}(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} \quad (4)$$

em que: $p(\mathbf{x})$ é a distribuição Gaussiana multivariada; $\boldsymbol{\mu}$ é o vetor média; e Σ é a matriz de covariância, sendo $|\Sigma|$ o determinante dessa matriz.

Os Processos Gaussianos são modelos não paramétricos, cujas saídas são representadas por uma distribuição Gaussiana multivariada (Murphy, 2012). Nesses modelos, uma função desconhecida, que representa as saídas sem a presença de ruído, é representada por uma distribuição a priori (Equação 5):

$$\mathbf{f} = f(\mathbf{X}) \sim N(\mathbf{f}|\mathbf{0}, \mathbf{K}) \quad (5)$$

em que: $\mathbf{K} = k(\mathbf{x}, \mathbf{x}')$ é a matriz de covariância, obtida por uma função de *kernel* (ou covariância).

Diferentes funções de *kernel* apresentam graus distintos de suavidade, que são mais ou menos adequados dependendo do conjunto de dados. Além disso, essas funções podem ser combinadas, aumentando a expressividade do modelo. São essas funções que permitem representar relações complexas e não-lineares entre as variáveis. Um exemplo de função de *kernel* bastante utilizada é a Matérn, dada pela Equação 6:

$$k_{\text{Matérn}}(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)} (\sqrt{2\nu} d)^\nu K_\nu(\sqrt{2\nu} d) \quad (6)$$

em que: $d = (\mathbf{x} - \mathbf{x}')^T \Theta^{-2}(\mathbf{x} - \mathbf{x}')$ é a distância Euclidiana entre \mathbf{x} e \mathbf{x}' , sendo Θ a escala de comprimento; ν é um parâmetro de suavidade da função; e K_ν é uma função de Bessel modificada.

Por serem modelos não paramétricos, a complexidade dos GP cresce com o tamanho do conjunto de dados. Na Equação 6, quanto maior é a quantidade de dados, maior é o número de cálculos realizados dentro da função de *kernel*, aumentando a demanda de processamento. Logo, os GP são ideais para conjuntos de dados pequenos, em que o custo computacional não limita sua utilização.

Os hiperparâmetros do GP são as variáveis de entrada da função de *kernel* escolhida e o ruído de observação, sendo possível otimizá-los para determinar a distribuição preditiva. Evita-se assim a necessidade de buscar por hiperparâmetros adequados via *random-search* ou *grid-search*. A Equação 2 é adaptada da seguinte maneira, calculando-se a posteriori das saídas sem ruído \mathbf{f} (Equação 7):

$$p(\mathbf{f}|D) = \frac{p(\mathbf{y}|\mathbf{f}) p(\mathbf{f}|\mathbf{X})}{p(\mathbf{y}|\mathbf{X})}, \quad (7)$$

em que: $p(\mathbf{y}|\mathbf{f})$ é a verossimilhança, a qual descreve a relação entre as saídas observadas \mathbf{y} e as saídas não observadas (latentes) \mathbf{f} ; $p(\mathbf{f}|\mathbf{X})$ é a distribuição a priori; $p(\mathbf{f}|D)$ é a distribuição a posteriori; e $p(\mathbf{y}|\mathbf{X})$ é a evidência.

A distribuição a priori é dita conjugada à verossimilhança se a distribuição a posteriori resultante tem o mesmo formato da distribuição a priori. Se a priori e a verossimilhança são Gaussianas, a posteriori também é Gaussiana. Nesse caso, o modelo é definido como um GP exato (Williams & Rasmussen, 2006). Esse modelo é mais indicado para problemas de regressão, em que os conjuntos de dados possuem saídas contínuas.

Em problemas de classificação, é possível usar outras verossimilhanças. Os GP baseados em Dirichlet permitem transformar um problema de classificação em um problema de regressão, em que as saídas preditas são as probabilidades de cada classe (Milios et al., 2018). A verossimilhança do modelo $p(\mathbf{y}|\boldsymbol{\alpha})$ é categórica, sendo expressa pela Equação 8:

$$p(\mathbf{y}|\boldsymbol{\alpha}) = \text{Cat}(\boldsymbol{\pi}), \text{ onde } \boldsymbol{\pi} \sim \text{Dir}(\boldsymbol{\alpha}) \quad (8)$$

em que: $\boldsymbol{\alpha}$ são os parâmetros de concentração, iguais a α_ϵ ou $(1 + \alpha_\epsilon)$; e $\boldsymbol{\pi}$ são as probabilidades das classes da distribuição categórica $\text{Cat}(\boldsymbol{\pi})$, inferidas a partir de uma distribuição de Dirichlet $\text{Dir}(\boldsymbol{\alpha})$. Essa distribuição pode ser expressa como uma combinação de distribuições lognormais. Como parâmetros $\boldsymbol{\alpha}$ nulos convergiriam para o infinito após a transformação log, soma-se uma pequena quantidade α_ϵ a cada um dos termos de $\boldsymbol{\alpha}$, evitando instabilidades numéricas. O hiperparâmetro α_ϵ deve ser ajustado a fim de minimizar o erro do modelo.

Uma outra alternativa é usar a verossimilhança Softmax (Chai, 2012), tratando o problema como uma tarefa de classificação. Como essa verossimilhança não é Gaussiana, a posteriori não possui solução analítica. Realiza-se então uma aproximação Gaussiana da posteriori por meio de inferência variacional esparsa (Titsias, 2009), que possui o formato a seguir (Equação 9):

$$q(\mathbf{f}, \mathbf{u}) = p(\mathbf{f}|\mathbf{u})q(\mathbf{u}), \quad (9)$$

em que: $q(\mathbf{f}, \mathbf{u})$ é a posteriori aproximada; $p(\mathbf{f}|\mathbf{u})$ é uma distribuição condicional obtida a partir da priori do GP; e $q(\mathbf{u})$ é uma distribuição variacional Gaussiana. Conjuntos de pontos de indução \mathbf{u} , de tamanho inferior ao conjunto de dados original, são usados para construir o modelo aproximado. Um conjunto \mathbf{u} é definido para cada uma das classes de saída.

3. Metodologia

Construiu-se um banco de dados como base nos trabalhos desenvolvidos por Mariano (2023) e Carvalho (2025), cujos atributos são descritos na Tabela 1. Selecionaram-se apenas misturas ensaiadas conforme o protocolo estabelecido pelo DNIT (2018).

Dentre os atributos escolhidos, quatro são categóricos (CAP, AMP, AB e AM), correspondendo ao tipo de ligante utilizado na mistura. Ao avaliar as misturas com ligante AMP do banco de dados, verifica-se que a adição de diferentes teores e tipos de polímero gera ligantes modificados com consistências diversas, como demonstram as propriedades medidas em laboratório. Assim, há indicativo de que misturas com

Tabela 1 • Descrição dos atributos

Atributo	Descrição
CAP	Ligante asfáltico puro
AMP	Asfalto modificado por polímero
AB	Asfalto borracha
AM	Asfalto de alto módulo
PEN	Penetração (0,1 mm)
VIS	Viscosidade a 135 °C ou 175 °C (cp)
PA	Ponto de amolecimento (°C)
P1_2	Passante na peneira 1/2" (%)
P3_8	Passante na peneira 3/8" (%)
P4	Passante na peneira n.º 4 (%)
P200	Passante na peneira n.º. 200 (%)
TMN	Tamanho máximo nominal (mm)
Vv	Volume de vazios (%)
TL	Teor de ligante (%)

ligante de maior consistência estão associadas à maior capacidade de resistir à DP. A adição de novas variáveis de entrada que diminuam a heterogeneidade do subconjunto AMP, como o tipo de polímero adicionado, deve melhorar o desempenho dos modelos. Para o presente trabalho, decidiu-se utilizar apenas o atributo AMP, visto que as misturas do banco de dados são modificadas com uma gama de tipos e combinações de polímeros, o que aumentaria o número de atributos de forma significativa.

Os outros dez atributos são numéricos, correspondendo a parâmetros do ligante (PEN, VIS, PA), que indicam sua consistência; dos agregados (P1_2, P3_8, P4, P200, TMN), obtidos a partir da análise granulométrica; e da mistura (Vv, TL), determinados após sua dosagem. Os valores estatísticos e histogramas das variáveis são apresentados na Tabela 2 e na Figura 2. Como o ensaio de FN não representa bem o comportamento do asfalto borracha com relação à DP (Faccin, 2018), as misturas com AB foram retiradas do conjunto de dados para as análises realizadas, gerando um total de 252 misturas.

Tabela 2 • Valores estatísticos das variáveis do banco de dados

Métrica	PEN	VIS	PA	P1_2	P3_8	P4	P200	TMN	Vv	TL	FN
Média	44,8	1081,2	59,3	90,4	82,5	49,7	6,3	14,9	4,2	5,0	1851
DP	14,2	689,8	8,1	7,7	10,7	8,7	2,7	4,5	0,4	0,6	2801
CV (%)	31,6	63,8	13,7	8,5	13,0	17,5	43,1	30,0	10,3	11,7	151
Mínimo	20,0	282,0	41,0	64,2	53,6	34,0	3,0	9,5	2,2	3,7	27
Máximo	80,0	3070,0	88,1	100,0	100,0	75,0	23,3	25,0	5,7	7,3	10000

3.1. Seleção e combinação de atributos

Uma das principais preocupações em relação aos modelos desenvolvidos é verificar se é necessário utilizar todas as propriedades dos materiais da mistura para sua aplicação. Por exemplo, é possível que usar apenas as propriedades do ligante asfáltico represente melhor o comportamento da mistura com relação à deformação permanente. Ou seja, o uso das propriedades dos agregados e da mistura tornaria o processo de aprendizagem menos eficaz. Para resolver essa questão, inicialmente montou-se a matriz de correlação entre os atributos numéricos, apresentada na Figura 3. Observa-se que os atributos VIS e PA, do ligante, e P1_2, P3_8 e TMN, dos agregados, apresentam correlação significativa entre si. No entanto, como nenhum dos coeficientes de Pearson foi superior a 0,90, optou-se por não descartar essas variáveis, visto que a quantidade de atributos não é tão elevada.

Em seguida, avaliaram-se diferentes combinações de atributos numéricos. Essas combinações foram feitas com base no tipo de material que os atributos caracterizam (ligante, agregados ou mistura), como mostra a Tabela 3. Os atributos categóricos foram usados junto com os atributos numéricos em todas as combinações, já que as misturas modificadas apresentam um aumento considerável da resistência à DP

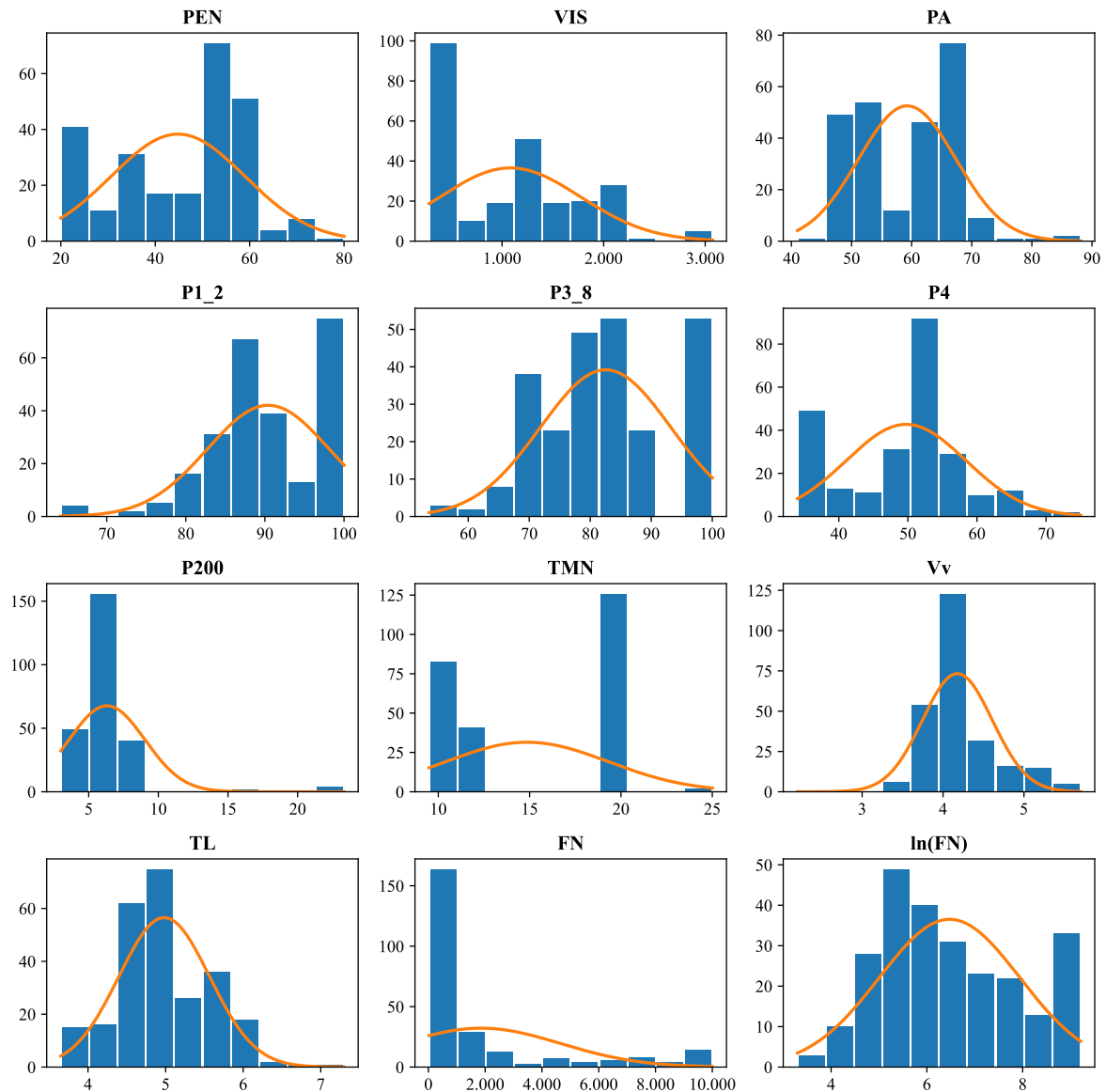


Figura 2 • Histogramas das variáveis do banco de dados.

(Onofre, 2012). Para isso, o atributo tipo de ligante foi transformado em três novas variáveis de entrada (CAP, AMP e AM) por meio de *one-hot-encoding*.

3.2. Divisão do banco de dados pelo nível de tráfego

Para aplicar os modelos de ML, transformaram-se os valores de FN no nível de tráfego equivalente, com base nos critérios de Bastos, Soares & Nascimento (2017). Assim, uma mistura com maior capacidade de resistir à DP é capaz de suportar um tráfego de maior intensidade. Os critérios estabelecem quatro níveis de tráfego: Leve (L), Médio (M), Pesado (P) e Extremamente Pesado (EP). Esses níveis foram agrupados em duas classes, como mostra a Tabela 4, para criar um problema de classificação binária entre as misturas mais e menos suscetíveis à deformação permanente. Avaliaram-se três diferentes divisões, variando-se o limiar de FN. Observa-se que o banco de dados é heterogêneo quanto ao tipo de ligante, visto que as proporções das classes variam de forma significativa, o que pode limitar o desempenho dos modelos.

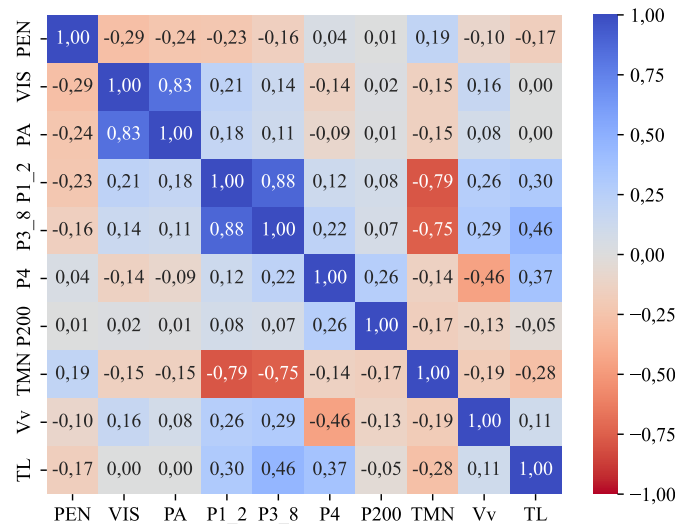


Figura 3 • Matriz de correlação dos atributos numéricos.

Tabela 3 • Combinações de atributos numéricos avaliadas

Combinação	Número de atributos	Atributos numéricos
Lig. + Agr. + Mist.	10	PEN, VIS, PA, P1_2, P3_8, P4, P200, TMN, Vv, TL
Ligante	3	PEN, VIS, PA
Agregados	5	P1_2, P3_8, P4, P200, TMN
Mistura	2	Vv, TL
Lig. + Agr.	8	PEN, VIS, PA, P1_2, P3_8, P4, P200, TMN
Lig. + Mist.	5	PEN, VIS, PA, Vv, TL
Agr. + Mist.	7	P1_2, P3_8, P4, P200, TMN, Vv, TL

Tabela 4 • Divisões dos dados por nível de tráfego e tipo de ligante

Classe	Flow Number	Número de misturas (%)			Total
		CAP	AMP	AM	
Divisão 1					
L	FN < 100	17 (16%)	0 (0%)	0 (0%)	17 (7%)
M/P/EP	FN ≥ 100	90 (84%)	107 (100%)	38 (100%)	235 (93%)
Divisão 2					
L/M	FN < 300	80 (75%)	6 (6%)	7 (18%)	93 (37%)
P/EP	FN ≥ 300	27 (25%)	101 (94%)	31 (82%)	159 (63%)
Divisão 3					
L/M/P	FN < 1000	106 (99%)	38 (36%)	20 (53%)	164 (65%)
EP	FN ≥ 1000	1 (1%)	69 (64%)	18 (47%)	88 (35%)

3.3. Seleção e aplicação dos modelos

O treinamento e a avaliação dos modelos foram feitos na linguagem Python, usando as bibliotecas pandas (McKinney, 2010) e scikit-learn (Pedregosa et al., 2011) para o pré-processamento dos dados. Para avaliar a capacidade de generalização dos modelos de maneira mais robusta, decidiu-se usar validação cruzada K-Fold estratificada com 10 folds.

Escolheram-se três modelos de GP, da biblioteca GPyTorch (Gardner et al., 2018), para previsão dos resultados: GP exato, para previsão do FN (regressão); GP baseado em Dirichlet e GP com verossimi-

lhança Softmax, para previsão do nível de tráfego equivalente (classificação). Utilizou-se a função de *kernel* Matérn, considerando três valores usualmente adotados para o parâmetro ν : 1/2, 3/2 e 5/2. O uso de valores além desses três aumenta consideravelmente o tempo de processamento. No caso do GP baseado em Dirichlet, avaliaram-se quatro valores possíveis para o hiperparâmetro α_ϵ : 10^{-1} , 10^{-2} , 10^{-3} e 10^{-4} . Ao implementar o GP com verossimilhança Softmax, utilizaram-se 16 pontos de indução.

Ao montar o histograma dos valores de FN (Figura 2), observou-se que a distribuição de frequência era enviesada para a direita. Assim, antes de empregar os modelos de regressão, aplicou-se uma transformação log nas saídas. Também foi feita a normalização *Z-score* dos atributos numéricos e do FN, para garantir que as diferenças de suas escalas não interferissem na modelagem. O modelo GP exato gera um valor de FN médio e um desvio padrão para cada uma das misturas, representando a distribuição prevista. Em seguida, a transformação log e a normalização aplicadas foram revertidas, obtendo-se os valores na escala original. A partir da distribuição preditiva, calculou-se a probabilidade do FN assumir um valor maior que o limiar definido para cada um dos critérios, por meio do cálculo da função de distribuição acumulada. Dessa forma, é possível avaliar o desempenho do modelo GP exato por métricas de classificação.

3.4. Métricas de avaliação utilizadas

Foram usadas 4 medidas para avaliação dos modelos: verdadeiros positivos (*VP*); verdadeiros negativos (*VN*); falsos positivos (*FP*); e falsos negativos (*FNe*). A partir dessas medidas, definem-se as quatro principais métricas de avaliação de modelos mostradas nas Equações 10, 11, 12 e 13 (Murphy, 2012):

- Acurácia – representa a porcentagem de acertos do modelo:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FNe}. \quad (10)$$

- Revocação – aponta, das saídas positivas reais, quantas o modelo conseguiu classificar corretamente:

$$\text{Revocação} = \frac{VP}{VP + FNe}. \quad (11)$$

É essencial para evitar que misturas que suportam um menor nível de tráfego (classe positiva) sejam previstas como sendo capazes de aguentar um nível de tráfego elevado (classe negativa), resultando em problemas estruturais futuros.

- Precisão – indica, das classificações positivas preditas, quantas foram acertadas:

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

É usada quando é importante que não haja presença elevada de falsos positivos. O aumento da precisão tende a reduzir a revocação. Essa métrica foi obtida para calcular o *F1 score*.

- *F1 score* – é a média harmônica entre precisão e revocação:

$$\text{F1 score} = 2 \frac{\text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}}. \quad (13)$$

É uma métrica que balanceia a importância da precisão e da revocação. Ao determinar o *F1 score* para cada uma das classes e calcular a média desses valores, obtém-se o *F1 score* médio, que representa o desempenho geral do modelo. O modelo com as maiores métricas em geral foi escolhido como a opção a ser utilizada. Também se construiu a matriz de confusão dos modelos selecionados, para visualização do desempenho. Essa matriz fornece a relação entre as previsões feitas pelo modelo e os valores reais das saídas, possibilitando a verificação do erro de cada uma das classes. Como foi utilizado *K-Fold* estratificado em 10 partes, calculou-se a matriz de confusão média para cada um dos modelos empregados.

A importância das variáveis foi obtida a partir do vetor de escalas de comprimento (Θ) da função de *kernel* Matérn. Cada variável do modelo possui um valor de associado. Atributos com valores baixos de possuem maior importância na resposta do modelo, e vice-versa. Assim, ao inverter e normalizar os valores do vetor Θ , calculou-se a importância relativa das variáveis em porcentagem.

Para verificar a acurácia das probabilidades previstas pelo modelo selecionado, utilizou-se a métrica *Brier score*, definida como (Equação 14):

$$\text{Brier score} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{p}_i)^2, \tag{14}$$

em que: N é o número de dados; y_i é a classe real do dado i (0, negativa ou 1, positiva); e \hat{p}_i é a probabilidade do dado i pertencer à classe 1. O *Brier score* assume um valor entre 0 e 1, e quanto menor for essa medida, maior é a acurácia das probabilidades previstas.

4. Resultados

A seguir, são apresentados os resultados obtidos pelos modelos GP exato, Dirichlet e Softmax para a divisão de classes L/M/P e EP, variando os atributos de entrada com base no tipo de material (ligante asfáltico, agregados e mistura asfáltica) como detalhado na Tabela 3, gerando um total de 21 modelos. A Figura 4 apresenta as métricas de avaliação determinadas. Os três modelos apresentam valores similares de acurácia e *F1 score* médio, enquanto os modelos de Dirichlet e Softmax possuem os maiores valores de revocação.

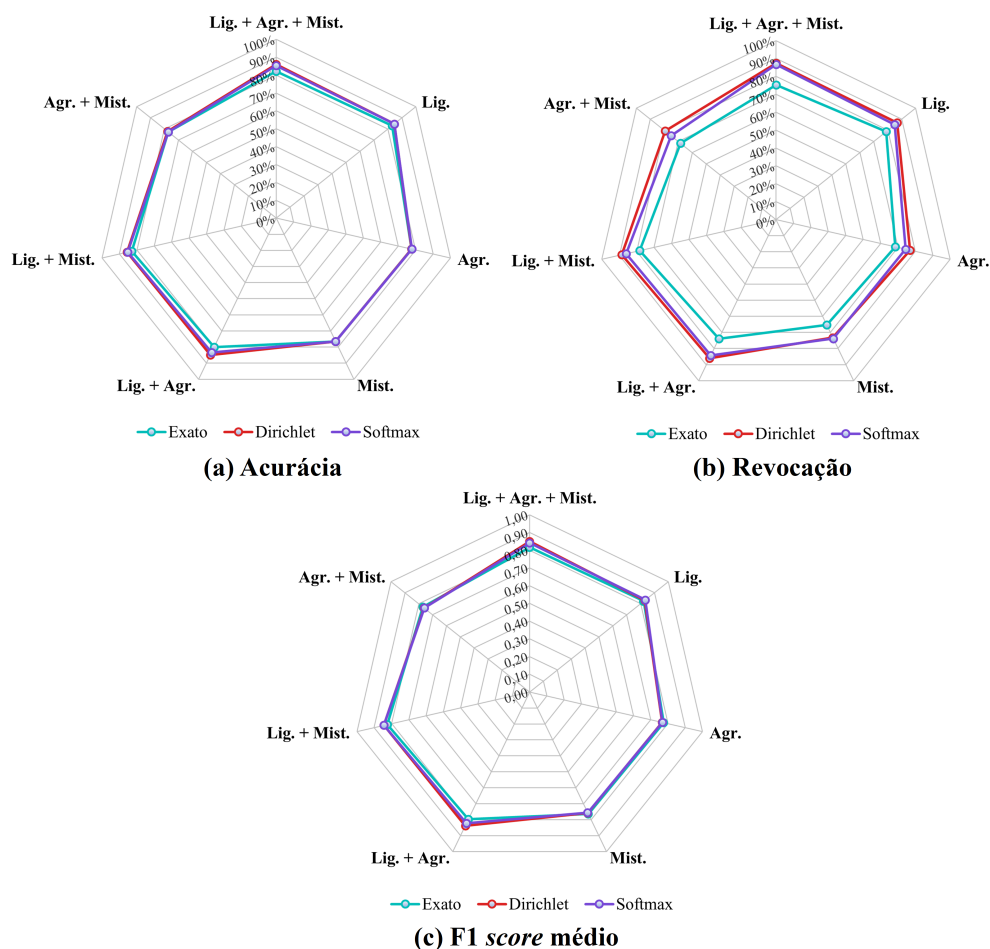


Figura 4 • Métricas de avaliação dos modelos – Divisão L/M/P e EP.

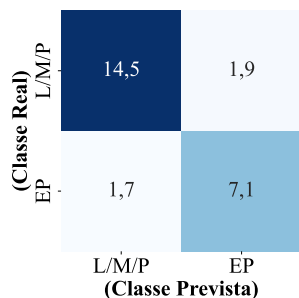


Figura 5 • Matriz de confusão média do melhor modelo – Divisão L/M/P e EP.

Tabela 5 • Melhores modelos para cada uma das divisões de classes

Divisão de classes	Modelo	Atributos numéricos	Acurácia (%)	Revocação (%)	F1 score médio	Brier score
L e M/P/EP	Softmax	Lig.	93	25	0,58	0,047
L/M e P/EP	Dirichlet	Agr.	90	83	0,90	0,097
L/M/P e EP	Dirichlet	Lig. + Mist.	86	89	0,84	0,113

Dentre os 21 modelos gerados, escolhe-se aquele que apresenta o melhor desempenho, ou seja, que possui as maiores métricas de avaliação. Destaca-se o modelo de Dirichlet com atributos numéricos do ligante e da mistura, que possui os maiores valores de acurácia (86%) e revocação (89%). O Brier score desse modelo é 0,113, o que indica que a acurácia das probabilidades previstas é elevada. A Figura 5 apresenta a matriz de confusão média do modelo destacado. Verifica-se que o modelo é capaz de diferenciar as duas classes, apresentando melhor desempenho na previsão da classe L/M/P.

As variáveis mais relevantes para o modelo, cujas importâncias estão indicadas entre parênteses, são os atributos categóricos CAP (28%) e AMP (45%), e o atributo numérico VIS (15%). Observa-se que o tipo de ligante possui um peso considerável, visto que quase todas as misturas com ligante CAP pertencem à classe L/M/P, e a maioria das misturas com ligante AMP pertence à classe EP. Também é possível que ligantes mais consistentes tendam a apresentar valores mais elevados de viscosidade, conferindo maior rigidez à mistura asfáltica e, conseqüentemente, aumentando sua capacidade de resistir à deformação permanente. A importância do atributo VIS é constatada ao avaliar a Figura 4, em que os modelos que não contêm essa variável possuem menores valores de revocação.

Da mesma forma, obtêm-se os melhores modelos para outras duas divisões de classes, escolhidos dentre outros 42 modelos, cujos resultados são apresentados na Tabela 5. O modelo da divisão L e M/P/EP apresenta baixos valores de revocação e F1 score médio, devido ao elevado desequilíbrio entre as classes (Tabela 4). Logo, o uso dessa classificação é descartado. No caso da divisão L/M e P/EP, o modelo apresenta métricas de avaliação elevadas e baixo valor de Brier score. Ao avaliar a importância das variáveis, verifica-se que o atributo categórico CAP tem uma importância de 73% na resposta do modelo, muito maior que a dos demais. Como a importância das variáveis AMP (4%) e AM (4%) não é elevada, pode-se inferir que é mais relevante, para o modelo, diferenciar se as misturas possuem ou não ligante modificado. Isso é corroborado pela divisão de classes da Tabela 4, em que a maioria das misturas das classes L/M e P/EP contém ligante puro e modificado, respectivamente.

5. Conclusões

Com base em critérios de transformação do FN no nível de tráfego equivalente, definiram-se três problemas de classificação binária entre as misturas asfálticas de menor resistência à DP (valores menores de FN) e maior resistência à DP (valores maiores de FN). Duas das divisões de classes avaliadas (L/M e P/EP; L/M/P e EP) resultaram em modelos de ML capazes de prever se uma mistura possui um valor de FN abaixo ou acima de um valor definido, indicando se o revestimento é capaz ou não de resistir a um tráfego de maior intensidade. Para essas classificações, obtiveram-se valores de acurácia, revocação

e F1 score médio de até 90%, 89% e 0,90, respectivamente. O modelo de Dirichlet gerou os maiores valores de revocação em ambos os casos, demonstrando sua capacidade de distinguir corretamente as misturas com menor resistência à DP. O atributo CAP é o mais importante na resposta do modelo para a divisão L/M e P/EP, enquanto os atributos CAP, AMP e VIS são os mais importantes para a divisão L/M/P e EP. Os modelos selecionados apresentam baixos valores de Brier score, o que indica alta acurácia das probabilidades previstas.

É importante ressaltar que, caso ocorram alterações tanto no banco de dados quanto na metodologia empregada, outro modelo poderá ser escolhido, com valores diferentes de importância dos atributos. Apesar dos resultados obtidos serem satisfatórios, deve-se considerar que o banco de dados é heterogêneo, contendo informações de misturas com diferentes tipos de ligante, o que afeta a performance dos modelos. Além disso, os dados de misturas com asfalto borracha não foram avaliados, já que o FN não representa bem seu comportamento com relação à DP. Futuramente, pretende-se avaliar o desempenho dos modelos considerando cada tipo de ligante de forma isolada, a fim de aperfeiçoá-los. Serão avaliados 3 conjuntos de dados para os ligantes CAP, AMP e AM. Além disso, será analisada a adoção de um valor mínimo para a probabilidade das misturas pertencerem à classe negativa (maior resistência à DP), de modo a evitar o uso de misturas com maior risco de terem sido classificadas incorretamente.

Agradecimentos

Os autores agradecem a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), o Departamento Nacional de Infraestrutura de Transportes (DNIT) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento das bolsas e pelo Instituto Nacional de Ciência e Tecnologia em Caracterização de Materiais para Infraestrutura (INCT-Infra).

Declaração de contribuição de autoria CRediT

Pedro Luiz Ribeiro Rocha: Curadoria de dados, Análise Formal, Investigação, Metodologia, Software, Validação, Visualização, Escrita (versão inicial); *Jorge Barbosa Soares*: Conceituação, Metodologia, Supervisão, Escrita (revisão e edição); *Iuri Sidney Bessa*: Metodologia, Supervisão, Escrita (revisão e edição); *César Lincoln Cavalcante Mattos*: Metodologia, Supervisão, Validação, Escrita (revisão e edição); *Carlos David Rodrigues Melo*: Metodologia, Supervisão, Escrita (revisão e edição).

Uso de tecnologia assistida por inteligência artificial

Foram utilizadas ferramentas de inteligência artificial na produção da pesquisa, com a finalidade de prever variáveis. Para isso, construíram-se modelos de aprendizado de máquina usando a linguagem de programação Python e suas bibliotecas. Os autores assumem total responsabilidade pelo uso de tais ferramentas.

Declaração de conflitos de interesse

Os autores declaram que não há conflito de interesses.

Declaração de disponibilidade de dados

Os dados, modelos e códigos que suportam os resultados deste estudo estão disponíveis mediante solicitação ao autor correspondente, Pedro Luiz Ribeiro Rocha.

References

- Ahmed, T. M., P. L. Green, & H. A. Khalid (2017). Predicting fatigue performance of hot mix asphalt using artificial neural networks. *Road Materials and Pavement Design* 18, 141–154. DOI: 10.1080/14680629.2017.1306928.
- Apeageyi, A. K. (2011). Rutting as a function of dynamic modulus and gradation. *Journal of Materials in Civil Engineering* 23(9), 1302–1310. DOI: 10.1061/(ASCE)MT.1943-5533.0000309.

- Barros, L. M. (2022). *Implementação do Ensaio Stress Sweep Rutting e do Shift Model para a Previsão da Deformação Permanente de Misturas Asfálticas Brasileiras*. Doutorado (tese), Universidade Federal do Rio de Janeiro, Rio de Janeiro. URL: <http://hdl.handle.net/11422/28334> [visitado 13/08/2025].
- Bastos, J. B. S., J. B. Soares, & L. A. H. Nascimento (2017). Critérios para os resultados do ensaio uniaxial de carga repetida de misturas asfálticas em laboratório a partir do desempenho em campo. *Transportes* 25(2), 29–40. DOI:10.14295/transportes.v25i2.1284.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Cambridge: Springer.
- Carvalho, P. H. J. (2025). Coleta, estruturação e análise de um banco de dados de ensaios sobre misturas asfálticas e seus componentes. Monografia (bacharelado), Universidade Federal do Ceará, Crateús. URL: <http://repositorio.ufc.br/handle/riufc/80054> [visitado 13/08/2025].
- DNIT (2018). *DNIT 184/2018-ME: Pavimentação - Misturas Asfálticas - Ensaio Uniaxial de Carga Repetida para Determinação da Resistência à Deformação Permanente – Método de Ensaio*. Rio de Janeiro: Departamento Nacional de Infraestrutura de Transportes (DNIT).
- El-Badawy, S., R. A. El-Hakim, & A. Awed (2018). Comparing artificial neural networks with regression models for hot-mix asphalt dynamic modulus prediction. *Journal of Materials in Civil Engineering* 30(7), 04018128. DOI:10.1061/(ASCE)MT.1943-5533.0002282.
- Faccin, C. (2018). Concretos asfálticos em utilização no Rio Grande do Sul: Comportamento mecânico e desempenho em campo quanto à deformação permanente. Mestrado (dissertação), Universidade Federal de Santa Maria, Santa Maria. URL: <http://repositorio.ufsm.br/handle/1/15888> [visitado 13/08/2025].
- Ferreira, J., L. Babadopulos, J. Bastos, & J. Soares (2020). A tool to design rutting resistant asphalt mixes through aggregate gradation selection. *Construction & Building Materials* 236, 117531. DOI:10.1016/j.conbuildmat.2019.117531.
- Gandomi, A. H., A. H. Alavi, M. R. Mirzahosseini, & F. M. Nejad (2011). Nonlinear genetic-based models for prediction of flow number of asphalt mixtures. *Journal of Materials in Civil Engineering* 23(3), 248–263. DOI:10.1061/(ASCE)MT.1943-5533.0000154.
- Gardner, J. R., G. Pleiss, D. Bindel, K. Q. Weinberger, & A. G. Wilson (2018). GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. In *Advances in Neural Information Processing Systems*. arXiv:1809.11165. DOI:10.48550/arXiv.1809.11165.
- Gomes, O. J. F., J. B. Soares, & J. B. S. Bastos (2019). Relação das diferentes escalas da mistura asfáltica com a sua resistência à deformação permanente. In *20º Congresso Ibero-Latino-Americano do Asfalto (CILA)*, Guadalajara, México. CILA.
- Kim, D. (2015). *Modulus and Permanent Deformation Characterization of Asphalt Mixtures and Pavements*. Doutorado (tese), North Carolina State University, Raleigh. URL: <http://www.lib.ncsu.edu/resolver/1840.16/10587> [visitado 13/08/2025].
- Liu, J., F. Liu, Z. Wang, E. O. Fanijo, & L. Wang (2023). Involving prediction of dynamic modulus in asphalt mix design with machine learning and mechanical-empirical analysis. *Construction & Building Materials* 407, 133610. DOI:10.1016/j.conbuildmat.2023.133610.
- Mariano, A. L. G. (2023). Uso de aprendizado de máquina interpretável para avaliação da deformação permanente em misturas asfálticas. Mestrado (dissertação), Universidade Federal do Ceará, Fortaleza. URL: <http://repositorio.ufc.br/handle/riufc/75193> [visitado 13/08/2025].
- McKinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference*, Austin, TX, pp. 56–61. SciPy Proceedings. DOI:10.25080/Majora-92bf1922-00a.
- Milios, D., R. Camoriano, P. Michiardi, L. Rosasco, & M. Filippone (2018). Dirichlet-based Gaussian processes for large-scale calibrated classification. In *Advances in Neural Information Processing Systems*. arXiv:1805.10915. DOI:10.48550/arXiv.1805.10915.
- Mirzahosseini, M., Y. M. Najjar, A. H. Alavi, & A. H. Gandomi (2015). Next-generation models for evaluation of the flow number of asphalt mixtures. *International Journal of Geomechanics* 15(6), 04015009. DOI:10.1061/(ASCE)GM.1943-5622.0000483.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Cambridge: The MIT Press.
- Murphy, K. P. (2022). *Probabilistic Machine Learning: An Introduction*. Cambridge: The MIT Press.
- Onofre, F. C. (2012). Avaliação do comportamento mecânico de misturas asfálticas produzidas com ligantes asfálticos modificados por ácido polifosfórico e aditivos poliméricos, enfatizando a resistência à deformação. Mestrado (dissertação), Universidade Federal do Ceará, Fortaleza. URL: <http://repositorio.ufc.br/handle/riufc/11166> [visitado 13/08/2025].
- Ouni, A. E., A. Dony, & J. Colin (2014). Probabilistic parametric approach for rutting evaluation: Application to hot and warm asphalt. *The International Journal of Pavement Engineering* 15(1), 58–65. DOI:10.1080/10298436.2012.725473.
- Ozturk, H. I. & M. E. Kutay (2014). An artificial neural network model for virtual Superpave asphalt mixture design. *The International Journal of Pavement Engineering* 15(2), 151–162. DOI:10.1080/10298436.2013.808341.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, & B. Thirion (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830. URL: <https://www.jmlr.org/papers/v12/pedregosa11a.html> [visitado 21/5/2026].
- Portela Neto, M. G. (2018). Avaliação da faixa de agregados dominantes e componentes intersticiais em misturas asfálticas. Mestrado (dissertação), Universidade Federal do Ceará, Fortaleza. URL: <http://repositorio.ufc.br/handle/riufc/>

- 49901 [visitado 13/08/2025].
- Singh, P. & A. K. Swamy (2020). Probabilistic approach to characterise laboratory rutting behaviour of asphalt concrete mixtures. *The International Journal of Pavement Engineering* 21(3), 384–396. DOI:10.1080/10298436.2018.1480780.
- Titsias, M. K. (2009). Variational learning of inducing variables in sparse Gaussian processes. *Journal of Machine Learning Research* 5, 567–574.
- Williams, C. K. I. & C. E. Rasmussen (2006). *Gaussian Processes for Machine Learning*. Cambridge: The MIT Press.
- Witczak, M. W. (2007). Specification Criteria for Simple Performance Tests for Rutting, Volume II: Flow Number and Flow Time. NCHRP Report 580, Transportation Research Board, Washington, D.C. URL: <https://trid.trb.org/View/844079> [visitado 21/5/2026].
- Zhao, Y., K. Zhang, Y. Zhang, Y. Luo, & S. Wang (2022). Prediction of air voids of asphalt layers by intelligent algorithm. *Construction & Building Materials* 317, 125908. DOI:10.1016/j.conbuildmat.2021.125908.