

# Relatório LAMIA 25

## Prática: Métricas e Validação de Modelos de Aprendizado de Máquina (III)

Kaique Medeiros Lima

### 1 Introdução

Este card abordará o uso do SHAP (Shapley Additive Explanations) como uma ferramenta para interpretar modelos de machine learning, explicando como ele auxilia na compreensão das contribuições individuais de cada variável para a previsão do modelo. Além disso, será feita uma análise comparativa entre diferentes modelos de aprendizado de máquina, destacando suas vantagens, desvantagens e aplicações em diferentes contextos.

### 2 Descrição da atividade

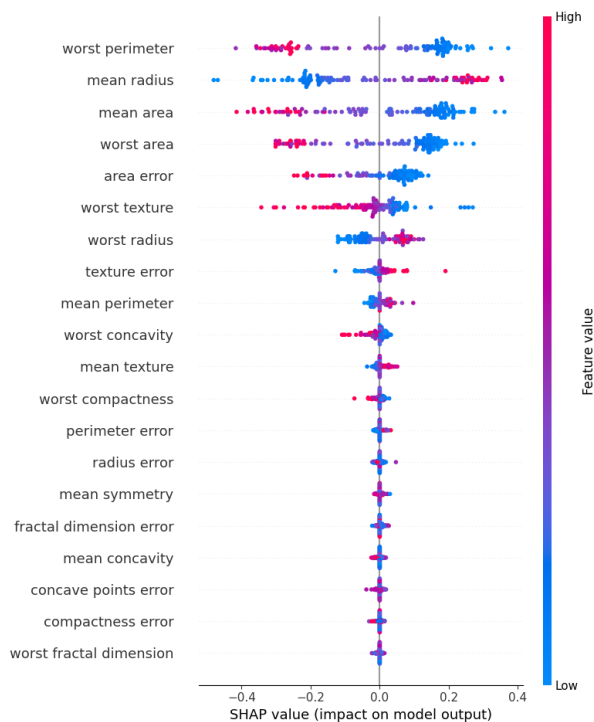
#### 2.1 Explain Machine Learning Models with SHAP in Python

##### 2.1.1 Explicação de Modelos com SHAP

O vídeo ensina a utilizar o SHAP (Shapley Additive Explanations) para interpretar decisões de modelos de machine learning. O SHAP é baseado na Teoria dos Jogos e oferece uma abordagem para explicar modelos que não possuem interpretabilidade nativa, como Support Vector Machines (SVM). O autor menciona que alguns modelos, como árvores de decisão, possuem explicabilidade inerente, enquanto outros apenas geram previsões sem justificar as decisões tomadas. Métodos como SHAP e LIME são úteis nesses casos, pois permitem visualizar a influência de cada variável no resultado.

##### 2.1.2 Exemplo Prático: Uso do SHAP com SVM

No exemplo prático, é utilizado o conjunto de dados breast cancer do pacote `sklearn.datasets`, onde os dados são divididos entre treino e teste. Um modelo de SVM linear (`LinearSVC`) é treinado e avaliado, atingindo uma precisão de 92,9%. No entanto, por não fornecer explicações sobre suas decisões, o SHAP é empregado para interpretar os resultados. O `KernelExplainer` do SHAP é aplicado para calcular os valores SHAP e gerar um `summary_plot`, que exibe a importância de cada variável na previsão. Segue o `summary_plot` gerado:



### 2.1.3 Visualização das Decisões com SHAP

Além do `summary_plot`, o autor demonstra como visualizar decisões individuais utilizando o `force_plot`, que ilustra como cada característica influencia a classificação de um exemplo específico. Ele destaca que algumas variáveis, como a `worst area`, exercem maior impacto na decisão final, deslocando a previsão para um lado ou outro da escala de classificação. Essa análise permite compreender melhor o funcionamento do modelo em nível granular.

### 2.1.4 Importância das Características

Por fim, o vídeo mostra como calcular a importância das características somando os valores absolutos dos SHAP values e ordenando os atributos do maior para o menor impacto. Essa análise fornece insights sobre quais características exercem maior influência nas decisões do modelo, mesmo que ele não forneça essa informação diretamente.

## 2.2 COMPARANDO MODELOS DE MACHINE LEARNING!!

### 2.2.1 Modelos Preditivos

Os modelos preditivos podem ser classificados em duas categorias principais: regressão e classificação. Os modelos de regressão são utilizados para estimar valores numéricos contínuos, como prever o preço de um imóvel com base em suas características ou a variação do preço de uma ação ao longo do tempo. Já os modelos de classificação têm o objetivo de categorizar dados em classes distintas, como identificar se um e-mail é spam ou não, ou determinar se uma fruta em uma imagem é uma laranja ou uma maçã. Enquanto a regressão busca prever quantidades, a classificação foca na atribuição de rótulos a partir de padrões nos dados.

### 2.2.2 Matriz de Confusão

Para medir o desempenho de um modelo, são utilizadas algumas métricas baseadas em cálculos matemáticos, que ajudam a verificar se ele está ajustado corretamente. A principal métrica a se considerar inicialmente é a matriz de confusão, uma tabela que organiza os resultados previstos pelo modelo e os compara com os valores reais.

Por exemplo, imagine um modelo treinado para reconhecer maçãs em um conjunto de 100 imagens. Se o modelo previu corretamente a presença de uma maçã, esse caso é registrado como verdadeiro positivo. Se apontou uma maçã onde não havia uma, é um falso positivo. Um falso negativo ocorre quando o modelo não identificou uma maçã que estava presente, enquanto o verdadeiro negativo representa uma previsão correta de que não havia uma maçã. No exemplo abaixo, a matriz de confusão de um modelo de classificação binária é apresentada utilizando classificação de Churns:

Matriz de confusão		Modelo	
		Não Churn	Churn
VERDADEIRO	Não Churn	1	1
	Churn	1	5

A análise dessa matriz possibilita entender as dificuldades do modelo e identificar formas de melhorar sua taxa de acertos, como aumentar a precisão de 65% para um valor mais elevado.

### 2.2.3 Acurácia e Precisão

A taxa de acertos do modelo, chamada de acurácia, é determinada pela razão entre o número de previsões corretas e o total de previsões realizadas. Esse indicador fornece uma visão geral do desempenho do modelo, mas pode ser insuficiente em contextos com classes desbalanceadas.

Por outro lado, a precisão mede a confiabilidade das previsões positivas, ou seja, quantas das amostras classificadas como positivas realmente pertencem à classe positiva. Ela é calculada pela seguinte fórmula:

$$\frac{\text{Verdadeiro Positivo}}{\text{Verdadeiro Positivo} + \text{Falso Positivo}}.$$

### 2.2.4 Recall e Especificidade

O recall, ou sensibilidade, é calculado a partir de uma única coluna da matriz. Suponha que a classe correta seja "positivo" e que o modelo tenha acertado 50 previsões e errado outras 50. O recall, então, é dado por:

$$\frac{\text{Verdadeiro Positivo}}{\text{Verdadeiro Positivo} + \text{Falso Negativo}} = 50\%.$$

Outra métrica relevante é a especificidade, utilizada no cálculo da curva ROC. Ela é obtida dividindo o número de verdadeiros negativos pelo total de verdadeiros negativos somados aos falsos positivos.

### 2.2.5 Curva ROC

A curva ROC é construída ao comparar a taxa de verdadeiros positivos (recall) com a taxa de falsos positivos (1 - especificidade) para diferentes valores de limiar do modelo. Esses limiares representam os pontos de corte que determinam a partir de qual probabilidade uma amostra é classificada como positiva. Por exemplo, um modelo pode considerar uma amostra como positiva apenas se a probabilidade prevista for maior que 50%, mas esse valor pode ser ajustado para priorizar maior sensibilidade ou maior especificidade. Ao testar vários limiares e registrar os respectivos valores de recall e taxa de falsos positivos, é possível traçar a curva ROC, que fornece uma visão ampla do desempenho do modelo sob diferentes configurações, auxiliando na escolha do melhor equilíbrio entre sensibilidade e especificidade.

## 3 Conclusão

Nesse card, foi discutido o uso do SHAP para interpretar modelos de machine learning e a comparação de diferentes modelos preditivos. O SHAP é uma ferramenta útil para explicar decisões de modelos complexos, como SVM, e visualizar a importância de cada variável nas previsões. Já a comparação de modelos envolve a análise de métricas como matriz de confusão, precisão, recall, especificidade e curva ROC, que permitem avaliar o desempenho e a confiabilidade dos algoritmos. Essas técnicas são essenciais para garantir a qualidade das previsões e identificar possíveis melhorias nos modelos de machine learning.