

Relatório 25 - Prática: Métricas e Validação de Modelos de Aprendizado de Máquina (III)

Lucas Augusto Nunes de Barros

Descrição das atividades

O primeiro vídeo proposto traz uma abordagem prática acerca das métricas de validação, ferramentas essenciais para avaliar o desempenho de modelos de classificação. Uma das ferramentas apresentadas é a matriz confusão, que organiza as previsões do modelo em quatro categorias: verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. Essa matriz é muitas vezes referenciada como base para o cálculo de diversos parâmetros durante a validação e avaliação das métricas dos erros do modelo e serve como base para o cálculo de outras métricas.

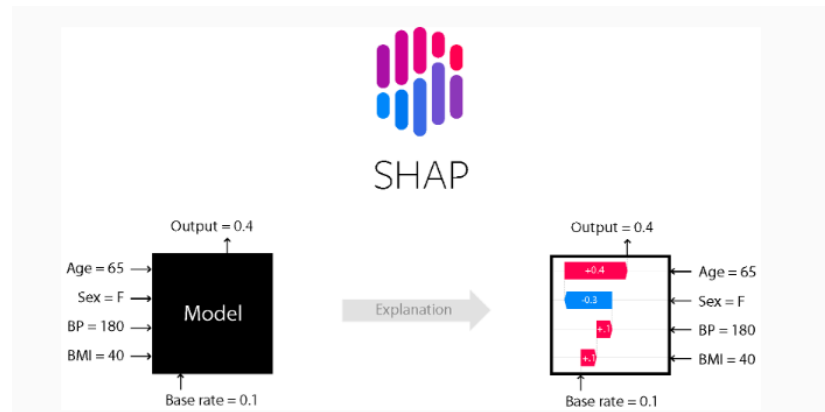
		Predito	
		0	1
Verdadeiro / Real	0	Verdadeiro Negativo	Falso Positivo
	1	Falso Negativo	Verdadeiro Positivo

A partir da matriz confusão, é possível calcular métricas como acurácia, precisão e recall. A acurácia representa a proporção de previsões corretas em relação ao total de previsões. A precisão mede a proporção de previsões corretas entre todas as previsões de verdadeiros positivos e falsos positivos. Já o *recall*, ou sensibilidade, indica a proporção de casos positivos que foram corretamente identificados.

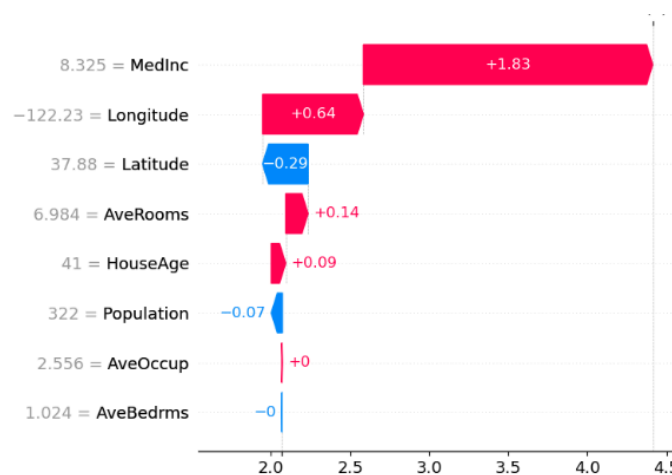
Outra métrica importante que é abordada é a curva ROC e a área sob a curva (AUC). A curva ROC plota a taxa de verdadeiros positivos (*recall*) contra a taxa de falsos positivos em diferentes limiares de classificação, oferecendo uma visão abrangente do desempenho do modelo. Característica que enquadra a curva ROC na classe das curvas empíricas.

A AUC quantifica o desempenho geral do modelo: um valor próximo de 1 indica excelente capacidade de discriminação entre as classes, enquanto um valor próximo de 0,5 sugere um modelo quase que aleatório. Essas métricas, quando combinadas, fornecem uma avaliação robusta do modelo, fornecendo o *feedback* do que ainda deve ser ajustado e quais otimizações podem ser implementadas para melhorar a performance em cenários reais.

O segundo vídeo traz o conceito de SHAP (*SHapley Additive exPlanations*), uma metodologia que oferece uma abordagem para explicar as previsões de modelos de *machine learning*, atribuindo a cada característica um valor que indica sua contribuição para o resultado final. Suas principais funcionalidades incluem a geração de valores SHAP, que quantificam a importância de cada variável para cada previsão específica ou ainda global, realizando o cálculo da média das importâncias de cada variável em todo o dataset.



A ideia dessa abordagem é tirar um pouco o perfil de “caixa preta” do modelo, onde a decisão é difícil de ser avaliada devido a quantidade enorme de *features*, o que torna particularmente difícil compreender o motivo de cada valor de saída do modelo. É abordado um método para visualizar os gráficos de resumo e dependência, o que auxilia na interpretação das contribuições. As aplicações do SHAP abrangem desde a validação de modelos em áreas críticas até a melhoria de modelos ao identificar variáveis irrelevantes ou redundantes.



O gráfico acima mostra a contribuição de cada variável na saída de um modelo. Os recursos que aumentam a previsão são mostrados em vermelho e os que diminuem a previsão são mostrados em azul.

Conclusão

Ao longo dos vídeos propostos, foram aplicadas técnicas e métricas estudadas anteriormente, porém sob uma nova ótica é possível perceber que o conhecimento se solidifica e assim os conceitos se tornam mais notáveis durante a implementação. Utilizando métricas de validação como acurácia, precisão, *recall* e AUC (*Area Under Curve*) para avaliar o desempenho dos modelos. Um tópico muito interessante foi o uso do SHAP (*SHapley Additive exPlanations*) para interpretar as previsões de modelos. A combinação de ferramentas, aliada ao SHAP, garante uma avaliação robusta do modelo e também aumenta a confiabilidade nos resultados.

Referências

[1] DATA CAMP.COM. Uma introdução aos valores SHAP e à interpretabilidade do aprendizado de máquina. Disponível em: <<https://www.datacamp.com/pt/tutorial/introduction-to-shap-values-machine-learning-interpretability>>. Acesso em 03 de junho de 2025

[2] YOUTUBE.COM . COMPARANDO MODELOS DE MACHINE LEARNING!!. Disponível em: <<https://www.youtube.com/watch?v=Y1SWuKr4Tg>>. Acesso em 02 de junho de 2025

[3] YOUTUBE.COM. Explain Machine Learning Models with SHAP in Python. Disponível em: <https://www.youtube.com/watch?v=zyXapWWGG4&ab_channel=NeuralNine>. Acesso em 03 de junho de 2025

[4] GITHUB.COM. Shap: A game theoretic approach to explain the output of any machine learning model. Disponível em: <<https://github.com/shap/shap>>. Acesso em 04 de junho de 2025