

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

Higor Miller Grassi

**Descrição da atividade:** No primeiro vídeo, aprendemos mais conceitualmente as métricas para os modelos de aprendizado de máquina, vendo sobre matriz de confusão, accuracy, precisão, recall, especificidade e a AUC, explicando detalhadamente como funciona. Já no segundo vídeo, vemos sobre SHAP no machine learning

Os modelos preditivos são amplamente utilizados em aprendizado de máquina para fazer previsões com base em dados históricos, sendo divididos em duas categorias principais, classificação e regressão.

Na classificação, o objetivo é prever uma classe ou rótulo, exemplos de aplicações incluem a detecção de fraude, previsão de churn (cancelamento de serviço), comportamento de compra e diagnósticos médicos. As métricas comuns para avaliar modelos de classificação incluem acurácia, precisão, sensibilidade, especificidade. Essas métricas ajudam a determinar a eficácia do modelo em identificar corretamente as classes.

Por outro lado, os modelos de regressão são usados para prever números ou valores contínuos, incluem a previsão de demanda, valor de compra, preços de ações e outros indicadores quantitativos. As métricas para avaliar modelos de regressão geralmente incluem o erro quadrático médio (MSE) que penaliza erros maiores mais severamente, visto nos cards anteriores, ajudando a entender o quão próximas as previsões estão dos valores reais.

A Matriz de Confusão, como dado o exemplo no vídeo, prevê a probabilidade do cliente adquirir um produto novo nos próximos 3 meses de acordo com os dados, o verdadeiro funcionando para saber se ele comprou recentemente ou não, já o Predito será o modelo a respeito do dado depois de ensinar o algoritmo, aparecendo um volume de pessoas em cada retângulo.

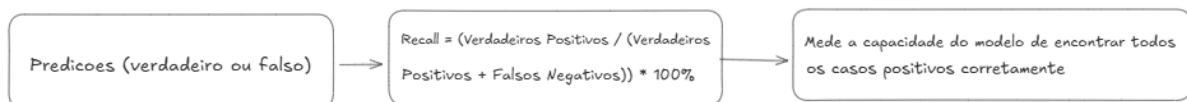
		Predito	
		0	1
Verdadeiro / Real	0	Verdadeiro Negativo	Falso Positivo
	1	Falso Negativo	Verdadeiro Positivo

A Accuracy será a quantidade de acertos pelo total para demonstrar qual a probabilidade de estar certo o resultado, tendo um modelo acurado, acertando bastante os resultados, ele é fundamental para avaliar o desempenho de modelos de classificação, pois indica o quão frequentemente o modelo faz previsões corretas.

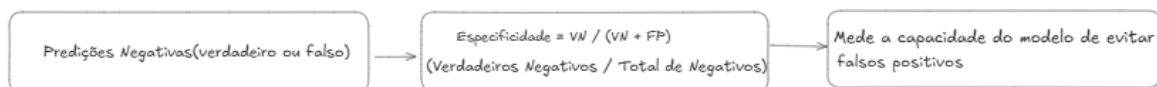


Ser preciso, como exemplo a tabela anterior, seria pegar a coluna do Predito em que está o 1, ou seja, o falso positivo e verdadeiro positivo, para sabermos a precisão correta de acordo com o número total, buscando sempre um modelo mais preciso, indicando a proporção de resultados positivos corretamente identificados em relação ao total de resultados classificados como positivos.

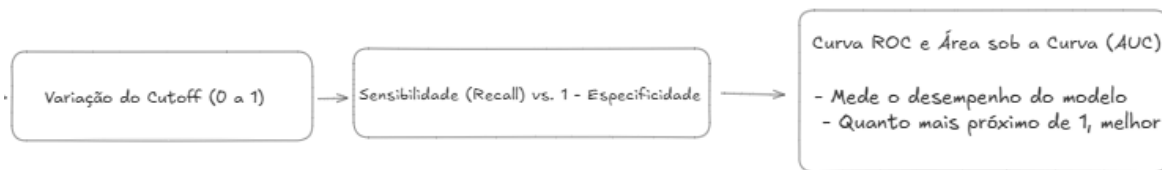
Já o recall/sensibilidade, novamente olhando para a tabela anterior, de tudo que o modelo falou que ia acontecer, quanto realmente aconteceu, conseguindo a porcentagem da resposta capturada. O recall varia de 0 a 1 e é essencial para avaliar a capacidade do modelo em identificar corretamente todos os casos positivos.



A especificidade seria o quanto os verdadeiros negativos estão sendo capturados de acordo com o falso positivo, funcionando de maneira parecida com o recall, mas focando nos resultados negativos, isso ajuda a entender a capacidade do modelo de evitar falsos positivos.



A Área sob a Curva ROC (AUC) seria, de uma forma resumida, a relação entre sensibilidade e especificidade, colocando em um gráfico com os dois pontos, permitindo visualizar o desempenho do modelo em diferentes limiares de classificação, sendo uma métrica bem completa para avaliar a qualidade geral do modelo, tudo isso com base no cutoff, que seria valor de limite usado em modelos de classificação para decidir a qual classe uma observação pertence, costumeiramente sendo definido como 0,5, ou seja, caso for menor, será avaliado como negativo e caso maior ou igual, como positivo.



O SHAP (SHapley Additive exPlanations) atribui a cada feature a contribuição que ela tem para uma predição específica, ajudando a entender o impacto das variáveis no modelo, sendo útil para tornar modelos complexos mais interpretáveis, fornecendo explicações locais (para cada previsão) e globais (para o modelo como um todo), e é aplicável a diversos tipos de algoritmos, onde na imagem a seguir podemos visualizar melhor.

