

Relatório 25 - Métricas e Validação de Modelos de Aprendizado de Máquina

Guilherme Loan Schneider

Descrição da atividade

Primeiramente, o autor comenta sobre as diferenças entre modelos preditivos voltados para problemas de classificação e regressão.

O primeiro tipo geralmente está atrelado a rotular alguma coisa, ou seja, dizer se algo é bom ou ruim, problemas multiclass, dentre outros. Já os problemas de regressão são geralmente utilizados para analisar dados e gerar previsões sobre eles, ou seja, prever se determinada ação na bolsa de valores irá aumentar ou diminuir na próxima semana, ou também prever o número de vendas de uma loja, dentre muitas outras aplicações.

Outro ponto importante é a contextualização do que é uma matriz de confusão e como podemos compreender o que ela nos mostra. A análise dessa matriz é imprescindível quando vamos tomar uma decisão sobre o desempenho de um algoritmo, funcionando como um norte para tomar uma decisão sobre o que fazer.

		Classe Predita	
		0	1
Classe Original	0	TN	FP
	1	FN	TP

Em primeiro plano, é importante salientar que ela serve de base para os cálculos de acurácia, precisão, recall, ROC, dentre outros.

Taxas de Verdadeiro Positivo e Falso Positivo

A TPR, também conhecida como sensibilidade ou recall, calcula a proporção de casos positivos que foram corretamente identificados pelo modelo. Matematicamente, é expressa por:

$$TPR = \frac{VP}{VP + FN}$$

onde VP representa os verdadeiros positivos e FN os falsos negativos. Essa métrica é particularmente importante em aplicações onde o custo de deixar passar um caso positivo é elevado, como no diagnóstico de doenças graves.

Por outro lado, a FPR mede a proporção de casos negativos que foram erroneamente classificados como positivos:

$$FPR = \frac{FP}{FP + TN}$$

sendo FP os falsos positivos e VN os verdadeiros negativos. Em sistemas de detecção de fraudes, por exemplo, uma FPR alta indicaria um número excessivo de alarmes falsos, o que poderia sobrecarregar o sistema com verificações desnecessárias.

Precisão do Modelo

A precisão tem o objetivo de mensurar a confiabilidade das previsões positivas feitas pelo modelo:

$$\text{Precisão} = \frac{VP}{VP + FP}$$

Diferentemente da TPR, que foca em capturar a maioria dos casos positivos, a precisão avalia quantos dos itens classificados como positivos realmente pertencem a essa classe. Em filtros de spam, por exemplo, uma alta precisão significa que poucas mensagens legítimas estão sendo marcadas incorretamente como spam.

Recall

O Recall é uma métrica que mede a capacidade de um modelo de classificação em identificar corretamente todas as instâncias positivas de um conjunto de dados. Ou seja, dentro do seu conjunto de dados que são positivos, quais o modelo realmente acertou que eram positivos.

$$\text{Recall} = \frac{VP}{VP + FN}$$

Especificidade

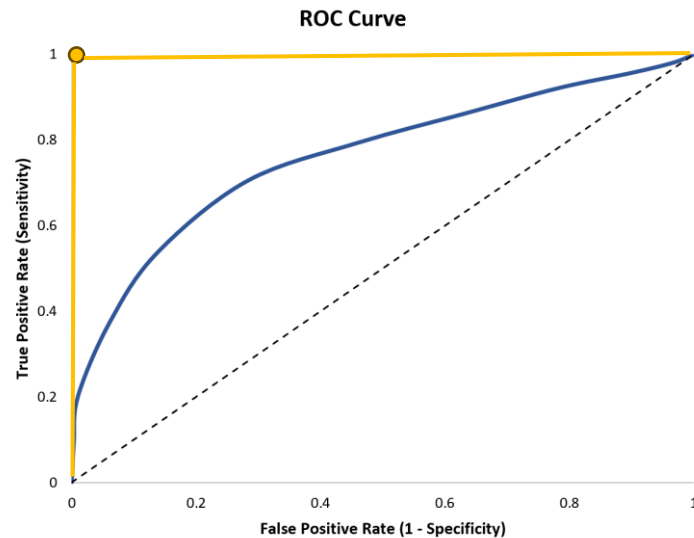
A Especificidade funciona similarmente ao Recall, mas a ideia aqui é ao invés de pegar as instâncias positivas de um conjunto de dados, ele acessará a parte negativa.

$$\text{Especificidade} = \frac{VN}{VN + FP}$$

Curva ROC

A curva ROC é uma forma de analisar um algoritmo a partir de duas métricas fundamentais, o recall sobre a especificidade. A ideia dessa curva é analisar os diferentes valores de cutoff (threshold de um algoritmo, ou seja, a partir desse valor o algoritmo classifica como spam, e abaixo do valor é não spam) e plotar os valores retornados pela função de recall sobre especificidade.

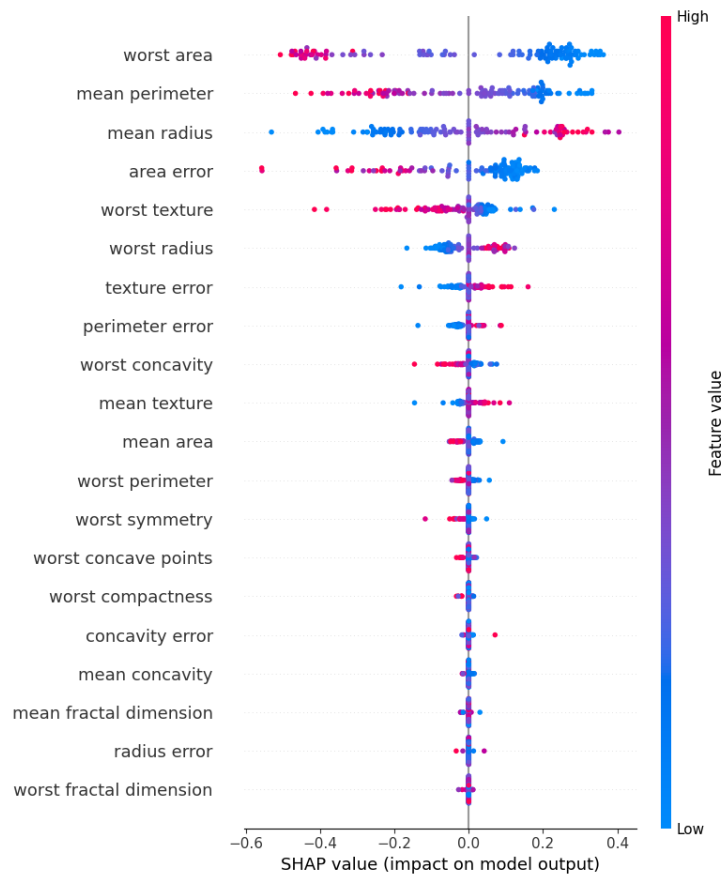
É interessante salientar que essa métrica não retorna um único valor como as métricas citadas anteriormente, permitindo uma análise mais representativa do atual estado do algoritmo. Um valor perfeito dessa métrica é quando formamos um ângulo de 90 graus com o canto superior esquerdo do gráfico, representado pela linha amarela.



SHAP

O SHAP (SHapley Additive exPlanations) é uma biblioteca capaz de explicar para nós como um modelo está tomando as decisões de classificação. No exemplo de aula, utilizamos a base “breast cancer” que classifica tumores como malignos ou benignos.

A divisão do dataset foi realizado e o treinamento também, tudo de forma transparente, não possuindo uma visualização clara de como o modelo toma as decisões. Para isso, utilizamos o SHAP a fim de visualizar as features que mais importam no modelo.



Além dessa visualização total, podemos observar também a análise individual de cada instância, ou seja, quais foram as features que importaram na decisão do modelo para uma única instância classificada como maligna ou benigna.

Na imagem abaixo nós conseguimos analisar cada um dos parâmetros que implicaram na classificação do tumor como maligno.



Nesse outro exemplo, é possível verificar que a classificação é benigna e o impacto de cada valor presente nas features.



Referencias

<https://www.youtube.com/watch?v=Y1SWuKr4Tg>

[Explain Machine Learning Models with SHAP in Python](#)