

# TAIL — CONTEÚDOS EXTRAS

Diretoria de Matemática - Tail

11/09/2020

## 1 Métricas de avaliação de modelos de Classificação

Agora já que as métricas de avaliação de modelos de regressão foram vistas, vamos abordar as métricas de avaliação de modelos de classificação.

### 1.1 Matriz de confusão

O mais simples dentre os meios de avaliação de modelos de classificação, porém a base para todas as outras métricas. A matriz de confusão é composta por todas as informações que nosso modelo obteve, Falso Negativo (FN), Falso Positivo (FP), Verdadeiro Negativo (VN) e Verdadeiro Positivo (VP). Além disso, esse meio de avaliação nos permite a visualização **direta** se nosso modelo está classificando um grupo melhor que o outro, ou seja, favorecendo um dos grupos sujeitos.

		Predito	
		Positivo	Negativo
Real	Positivo	<b>VP</b>	<b>FN</b>
	Negativo	<b>FP</b>	<b>VF</b>

### 1.2 Acurácia

A acurácia é uma métrica muito usada para avaliar o desempenho de modelos de classificação, identificando quantas observações de fato foram classificadas de forma correta, seja ela da classe negativa, positiva ou qualquer outra. Porém, a acurácia apresenta alguns problemas como o não ponderamento, digamos que temos uma amostra de 5000 observações, sendo 4890 da classe positiva e apenas 110 da classe negativa, logo nosso modelo classificará melhor o a classe positiva, por fim, poderemos ver que nosso modelo não estará trabalhando bem.

$$\frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}}$$

### 1.3 Precisão

Perceba que a precisão trabalha em cima das observações classificadas como positivas, ou seja, essa métrica nos ajuda a identificar: Dentre os exemplos classificados como positivos, quantos são realmente **positivos**?

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

### 1.4 Recall

O recall ou sensibilidade é a quantidade de observações positivas (reais) sobre a quantidade de classificações positivas encontradas, tanto verdadeiras quanto falsas. Quanto mais próximo de 1, menos erros o modelo cometeu em marcar valores como verdadeiro.

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

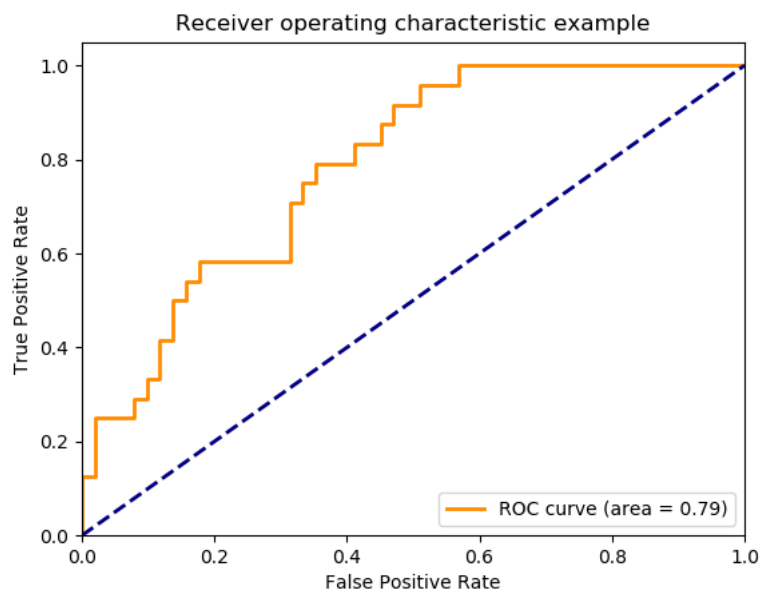
## 1.5 F1-Score

Perceba uma dependência dessa métrica com o **Recall** e da **Precisão**, assim sendo, quanto maior nosso F1-score, podemos afirmar que nosso modelo está tendo uma precisão mais alta (Capacidade de classificar melhor as classes) e maior será meu recall, evidenciando que o nosso modelo erra menos as classes classificadas como positivas

$$2 \cdot \frac{\text{Preciso} \cdot \text{Recall}}{\text{Preciso} + \text{Recall}}$$

### 1.5.1 Curva ROC

Receiver Operating Characteristic Curve ou simplesmente Curva ROC, é um meio de avaliação diferente dos outros, justamente pelo fato de não ser sensível ao desbalanço entre classes diferentes. Podemos analisar a curva ROC da seguinte forma:



Quanto maior a área da curva ROC e mais próxima da parte superior esquerda do gráfico, melhor será nosso modelo. Para o cálculo da curva ROC utilizamos dois valores bem importantes: Taxa de Falso Positivo (FPR) e a Taxa de Verdadeiro Positivo (TPR) que auxiliam na construção de um limiar para o gráfico, onde tudo que estiver superior a ele será classificado como positivo, caso esteja abaixo, classificado como negativo. Sendo o valor perfeito o (0, 1).

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

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

A reta azul tracejada representa um modelo que apenas classifica os dados de forma aleatória, evidenciada para comparação com seu modelo.