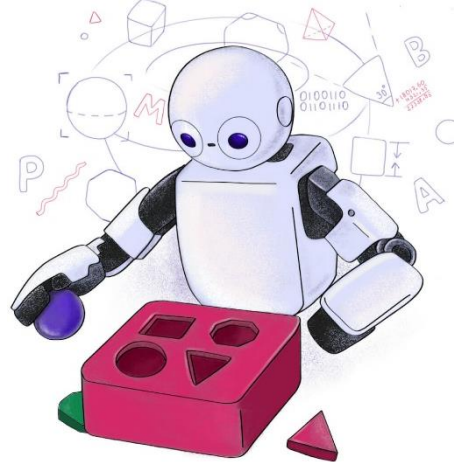


T320 - Introdução ao Aprendizado de Máquina II: *Classificação (Parte V)*



Recapitulando

- Anteriormente, vimos como lidar com problemas de classificação que envolvem mais de 2 classes, também chamados de classificação multi-classes.
- Nesta aula, veremos várias métricas utilizadas para medir o desempenho de classificadores.

Métricas de avaliação de classificadores

- As métricas para avaliação do desempenho de classificadores que estudaremos são:
 - Taxa de erro e acurácia
 - Matriz de confusão
 - Pontuação-F (*F-score*)
 - Curva Característica Operacional do Receptor (*Receiver Operating Characteristic - ROC*)

Métricas de avaliação de classificadores

Taxa de erro e acurácia

- A **taxa de erro**, é intuitivamente, a métrica mais direta para se avaliar o desempenho de um classificador.
- Ela corresponde à porcentagem de exemplos classificados incorretamente considerando o conjunto de dados disponíveis para **validação**.
- A **taxa de erro** é dada por

$$p_e(\hat{y}(\mathbf{x})) = \frac{1}{N} \sum_{i=0}^{N-1} (1 - \delta(y(i), \hat{y}(\mathbf{x}(i)))) ,$$

onde $\delta(i, j) = \begin{cases} 0, & \text{se } i \neq j \\ 1, & \text{se } i = j \end{cases}$ é o delta de Kronecker. Observe que $p_e(\hat{y}(\mathbf{x})) \in [0, 1]$.

- O complemento da **taxa de erro** é conhecido como **acurácia**, e é definida por

$$\text{acc}(\hat{y}(\mathbf{x})) = 1 - p_e(\hat{y}(\mathbf{x})).$$

Métricas de avaliação de classificadores

Matriz de Confusão

- O nome, **matriz de confusão**, deriva do fato de que ela torna fácil verificar se o classificador está confundindo classes (ou seja, geralmente rotulando incorretamente uma como a outra).
- A **matriz de confusão** $C \in \mathbb{R}^{Q \times Q}$ contabiliza o número de classificações corretas e incorretas para cada uma das Q classes existentes.

- A **matriz de confusão** C é definida como

Quantidade de exemplos pertencentes à classe 1.

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1Q} \\ C_{21} & C_{22} & \dots & C_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ C_{Q1} & C_{Q2} & \dots & C_{QQ} \end{bmatrix}.$$

Exemplos classificados como pertencentes à classe 1.

- C_{11} indica quantos padrões da classe 1 foram corretamente atribuídos à classe 1.
- C_{12} indica quantos padrões da classe 2 foram atribuídos à classe 1.

- A diagonal de C fornece o número de classificações corretas.
- Cada **linha** representa os exemplos que foram classificados como pertencentes a uma dada classe.
- Cada **coluna** representa os exemplos realmente pertencentes a uma dada classe.
- A informação apresentada nesta matriz permite verificar quais classes o **classificador** tem maior dificuldade em classificar.

Métricas de avaliação de classificadores

Matriz de Confusão

Exemplo para $Q = 2$.

Classes Estimadas	+	Verdadeiro Positivo (TP)	Falso Positivo (FP)
	C_2		
	-	Falso Negativo (FN)	Verdadeiro Negativo (TN)
	C_1		
		+	-
		C_2	C_1
		Classes Verdadeiras	

- **Verdadeiro Positivo** (TP): número de exemplos da classe positiva (+), C_2 , classificados corretamente.
 - **Verdadeiro Negativo** (TN): número de exemplos da classe negativa (-), C_1 , classificados corretamente.
 - **Falso Positivo** (FP): número de exemplos classificados como positivos (+), mas que, na verdade, pertencem à classe negativa (-).
 - **Falso Negativo** (FN): número de exemplos atribuídos à classe negativa (-), mas que, na verdade, pertencem à classe positiva (+).
- Algumas definições que vamos precisar a seguir:
- N_+ define o número de exemplos pertencentes à classe positiva = TP + FN (coluna de C_2).
 - N_- define o número de exemplos pertencentes à classe negativa = FP + TN (coluna de C_1).
 - N define o número total de exemplos = TP + FN + FP + TN.

Métricas de avaliação de classificadores

Matriz de Confusão

Nós podemos calcular diversas métricas de desempenho a partir das informações contidas na **matriz de confusão**:

- **Taxa de falso negativo**: é a proporção de exemplos da classe positiva (+) classificados incorretamente.

$$\text{Taxa de falso negativo} = p_e^+(\hat{y}(\mathbf{x})) = \frac{\text{FN}}{\text{TP} + \text{FN}} = \frac{\text{FN}}{N_+}.$$

- **Taxa de falso positivo**: é a proporção de exemplos da classe negativa (-) classificados incorretamente.

$$\text{Taxa de falso positivo} = p_e^-(\hat{y}(\mathbf{x})) = \frac{\text{FP}}{\text{TF} + \text{FP}} = \frac{\text{FP}}{N_-}.$$

- **Taxa de erro**:

$$p_e(\hat{y}(\mathbf{x})) = \frac{\text{FP} + \text{FN}}{N}.$$

- **Acurácia**:

$$\text{acc}(\hat{y}(\mathbf{x})) = \frac{\text{TP} + \text{TN}}{N}.$$

Métricas de avaliação de classificadores

Matriz de Confusão

- **Precisão**: corresponde à proporção de exemplos da classe positiva (+) corretamente classificados em relação a todos os exemplos atribuídos à classe positiva (+).

$$\text{precisão}(\hat{y}(x)) = \frac{TP}{TP+FP}.$$

- **Sensibilidade** (ou *recall*): também conhecida como ***taxa de verdadeiros positivos***. Corresponde à proporção de exemplos da classe positiva (+) corretamente classificados.

$$\text{recall}(\hat{y}(x)) = \frac{TP}{TP+FN} = 1 - p_e^+(\hat{y}(x)).$$

- **Especificidade**: também conhecida como ***taxa de verdadeiros negativos***. Corresponde à proporção de exemplos da classe negativa (-) corretamente classificados.

$$\text{especificidade}(\hat{y}(x)) = \frac{TN}{TN+FP} = 1 - p_e^-(\hat{y}(x)).$$

Observações importantes quanto à matriz de confusão

- É possível estender as métricas obtidas com a **matriz de confusão** para o cenário multi-classes (i.e., $Q > 2$):
 - Para isto, basta selecionar, uma vez, cada classe C_q , $q = 1, \dots, Q$ como sendo a classe positiva, enquanto todas as demais classes formam a classe negativa. Assim, obtem-se os valores das métricas para cada classe.
- Veja o exemplo abaixo para $Q = 3$.

Classes Estimadas	+	Verdadeiro Positivo (TP)	Falso Positivo (FP)	Falso Positivo (FP)
	-	Falso Negativo (FN)	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)
	-	Falso Negativo (FN)	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)
		+	-	-
		Classes Verdadeiras		

Classes Estimadas	-	Verdadeiro Negativo (TN)	Falso Negativo (FN)	Verdadeiro Negativo (TN)
	+	Falso Positivo (FP)	Verdadeiro Positivo (TP)	Falso Positivo (FP)
	-	Verdadeiro Negativo (TN)	Falso Negativo (FN)	Verdadeiro Negativo (TN)
		-	+	-
		Classes Verdadeiras		

Classes Estimadas	-	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)	Falso Negativo (FN)
	-	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)	Falso Negativo (FN)
	+	Falso Positivo (FP)	Falso Positivo (FP)	Verdadeiro Positivo (TP)
		-	-	+
		Classes Verdadeiras		

Observações importantes quanto à matriz de confusão

- **Precisão** diz o quão preciso é o modelo em relação à todos os exemplos classificados como positivos, quantos deles são realmente positivos.

$$\text{Precisão} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

- A **precisão** é uma boa medida para determinar a qualidade do classificador quando os custos de **falsos positivos** são altos.
 - Por exemplo, na classificação de **spams**, um **falso positivo** significa que um **ham** (verdadeiro negativo) foi classificado como **spam**. O usuário de email pode perder emails importantes se a **precisão** não for alta.
- **Recall** calcula quantos exemplos realmente positivos o classificador captura, rotulando-os como positivos verdadeiros.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

- O **recall** é uma boa medida para determinar a qualidade de um classificador quando houver um alto custo associado à **falsos negativos**.
 - Por exemplo, na classificação de doenças, se um paciente doente (**positivo verdadeiro**) for classificado como não doente (**falso negativo**). O custo associado ao **falso negativo** será extremamente alto se a doença for contagiosa.

Observações importantes quanto à matriz de confusão

- Uma **precisão** = 1 significa que todo exemplo classificado como pertencente à classe **positiva**, realmente pertence à ela, ou seja, o número de **falsos positivos** é igual a 0.
 - Entretanto, essa métrica não dá informações a respeito de quantos exemplos desta classe foram classificados de forma incorreta, ou seja, quantidade de **falsos negativos**.
- Por outro lado, um **recall** = 1 indica que todos os exemplos da classe positiva foram classificados como sendo pertencentes a ela, ou seja, o número de **falsos negativos** é igual a 0.
 - Porém, essa métrica não traz informações a respeito de quantos exemplos da classe negativa foram classificados como sendo pertencentes à classe positiva, ou seja, a quantidade de **falsos positivos**.
- Portanto, para analisarmos melhor o desempenho de um classificador, precisamos usar uma métrica que combine essas duas métricas.

Métricas de avaliação de classificadores

Pontuação-F

- As métricas de **precisão** e **recall** costumam ser analisadas conjuntamente através de uma métrica que combina ambas métricas, chamada de **pontuação-F** (ou **F-score**), denotada por F_m , que combina as duas métricas através de uma **média harmônica ponderada** dada pela equação abaixo:

$$F_m = \frac{(m+1) \times \text{recall}(\hat{y}(x)) \times \text{precisão}(\hat{y}(x))}{\text{recall}(\hat{y}(x)) + m \times \text{precisão}(\hat{y}(x))},$$

onde m é o **fator de ponderação**.

- Quando $m = 1$, a mesma importância é dada para a **precisão** e para o **recall**:

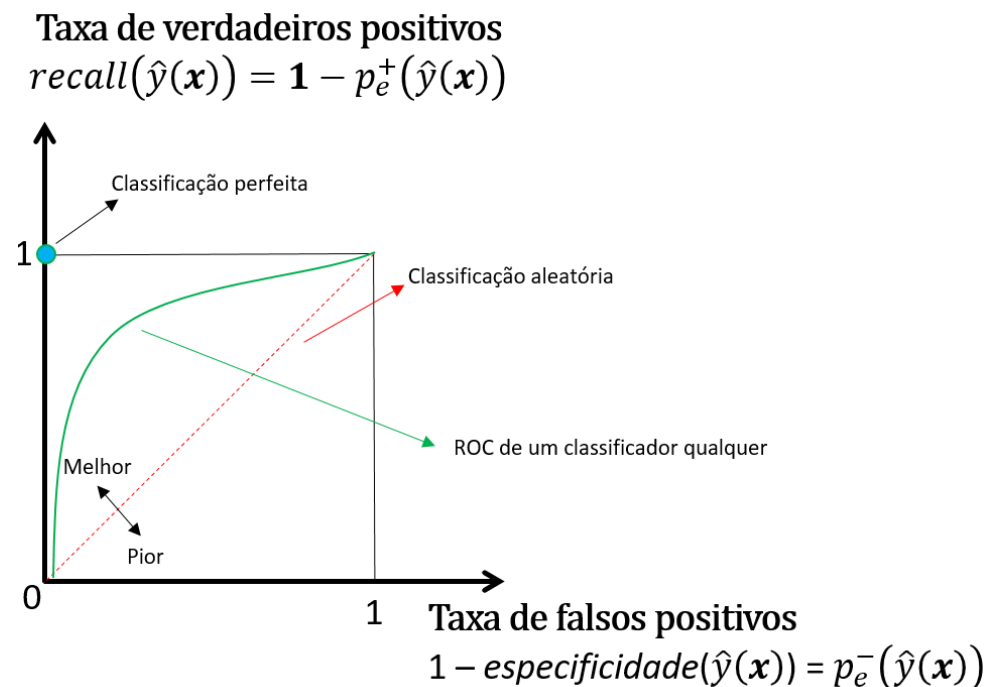
$$F_1 = 2 \frac{\text{recall}(\hat{y}(x)) \times \text{precisão}(\hat{y}(x))}{\text{recall}(\hat{y}(x)) + \text{precisão}(\hat{y}(x))} = \frac{TP}{TP + \frac{FN + FP}{2}}.$$

- Valores de F_1 próximos de 1 indicam que o **classificador** obteve bons resultados tanto de **precisão** quanto de **recall**.

Métricas de avaliação de classificadores

Curva Característica Operacional do Receptor (ROC)

- É um gráfico, conforme mostrado ao lado, que ilustra a performance de um **classificador binário** conforme seu **limiar de discriminação** é variado.
- A curva é criada plotando-se o **recall** em função da **taxa de falsos positivos** para vários valores de **limiar de discriminação**.
- Quanto mais à esquerda e para cima estiver a **curva ROC** de um **classificador**, melhor será o seu desempenho.
- A linha diagonal em vermelho, está associada a um **classificador puramente aleatório**. Um bom **classificador** fica o mais longe possível dessa linha (em direção ao canto superior esquerdo).
- Um **classificador perfeito** teria um ponto no canto superior esquerdo da curva ROC, representando 100% de **recall** (ou seja, sem falsos negativos) e 100% de **especificidade** (ou seja, sem falsos positivos).



Métricas de avaliação de classificadores

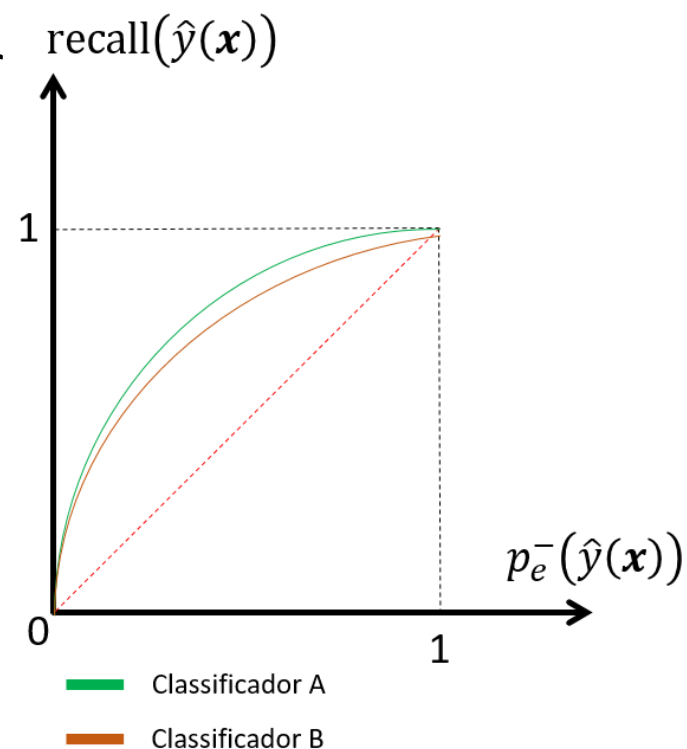
Curva Característica Operacional do Receptor (ROC)

- A forma usual de se comparar **classificadores** consiste em criar uma **curva ROC** para cada um.
- Em geral, **classificadores** produzem uma saída real (i.e, uma probabilidade) para cada exemplo de entrada.
- Normalmente, estas saídas são, então, discretizadas para que se tenha a decisão final: por exemplo, se $h_a(x(i))$ ultrapassa um determinado **limiar, T** , ela é mapeada no valor 1 (classe positiva, C_2); caso contrário, ela é mapeada no valor 0 (classe negativa, C_1).
- Sendo assim, ao plotar a **taxa de verdadeiro positivo** (ou **recall**) versus a **taxa de falso positivo** para diferentes valores de **limiar, T** , obtemos a **curva ROC** associada a um **classificador**.

Métricas de avaliação de classificadores

Curva Característica Operacional do Receptor (ROC)

- Por exemplo, considere as **curvas ROC** na figura ao lado. Para decidir qual o melhor **classificador**, podemos tomar como base a **área sob a curva (ASC) ROC**.
- A **ASC** é outra métrica da qualidade de um classificador, um número entre 0 e 1. Quanto maior a **ASC**, melhor será o classificador.
- Neste exemplo, o classificador A tem melhor desempenho, pois tem **área sob a curva ROC** maior do que a do classificador B.
- **Vantagens da curva ROC**
 - Possibilita a análise de diferentes métricas de desempenho independente do **limiar** escolhido.
 - Auxilia o estudo de diferentes **limiares** para lidar com problemas de desbalanceamento nos dados (i.e., nos quais as classes possuem tamanhos discrepantes).
- **Desvantagens**
 - Apropriada para problemas de **classificação binária**.
 - No caso multi-classes, devemos utilizar as estratégias **um-contra-o-resto** ou **um-contra-um** e plotar várias **curvas ROC**.



Tarefas

- **Quiz:** “*T320 - Quiz - Classificação (Parte V)*” que se encontra no MS Teams.
- **Exercício Prático:** [Laboratório #5](#).
 - Pode ser baixado do MS Teams ou do GitHub.
 - Pode ser respondido através do link acima (na nuvem) ou localmente.
 - [Instruções para resolução e entrega dos laboratórios](#).
 - **Atividades podem ser feitas em grupo, mas as entregas devem ser individuais.**

Obrigado!