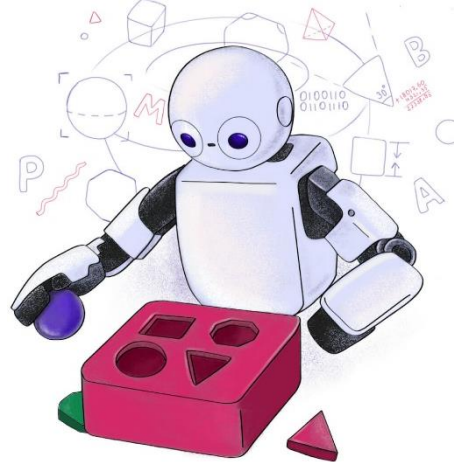


T320 - Introdução ao Aprendizado de Máquina II: *Classificação (Parte V)*



Recapitulando

- Anteriormente, vimos como lidar com problemas de classificação que envolvem mais de duas classes, também chamados de problemas de ***classificação multi-classes*** através das abordagens:
 - Um-Contra-Resto
 - Um-Contra-Um
 - Regressão Softmax
- Nesta aula, veremos as ***métricas*** mais utilizadas para medir o ***desempenho de classificadores***.

Métricas para avaliação de classificadores

- As métricas para avaliação do desempenho de classificadores que estudaremos são:
 - Taxa de erro e acurácia
 - Matriz de confusão
 - Várias métricas podem ser extraídas da matriz.
 - Pontuação-F (*F-score*)
 - Curva Característica Operacional do Receptor (do inglês, *Receiver Operating Characteristic* - ROC)

Métricas para avaliação de classificadores

Taxa de erro e acurácia

- A **taxa de erro**, é 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**, $y(i)$ é o valor esperado e $\hat{y}(\mathbf{x}(i))$ é a saída do classificador. Observe que $p_e(\hat{y}(\mathbf{x})) \in [0, 1]$.

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

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

Métricas para avaliação de classificadores

Matriz de Confusão

- O nome, **matriz de confusão**, deriva do fato de que ela torna fácil verificar se o classificador está se **confundindo** (ou seja, **rotulando incorretamente** os exemplos).
- A **matriz de confusão** contabiliza o número de classificações corretas e incorretas para cada uma das Q classes existentes.

- A **matriz de confusão**, $C \in \mathbb{R}^{Q \times Q}$, é definida como

Quantidade de exemplos realmente 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 exemplos da classe 1 foram corretamente atribuídos à classe 1.
- C_{12} indica quantos exemplos da classe 2 foram atribuídos à classe 1.

- A diagonal de C fornece o número de classificações corretas.
- A q -ésima **linha** indica o total de exemplos que foram classificados como pertencentes a q -ésima classe.
- A q -ésima **coluna** indica o total de exemplos realmente pertencentes à q -ésima classe.
- A informação apresentada na matriz permite verificar quais classes o **classificador** tem maior dificuldade em classificar.

Métricas para 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.
- **IMPORTANTE:** 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 para 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.

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

TP	FP
FN	TN

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

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

TP	FP
FN	TN

- **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 para avaliação de classificadores

Matriz de Confusão

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

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

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

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

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

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

- **Especificidade:** também conhecida como ***taxa de verdadeiros negativos***. É a proporção de exemplos da classe negativa corretamente classificados.

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

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

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, obtém-se os valores das métricas para cada uma das Q classes.
- Veja o exemplo abaixo para $Q = 3$, ou seja, C_1 , C_2 e C_3 .

Classe C_1 é a positiva.

Classes Estimadas	$+(C_1)$	Verdadeiro Positivo (TP)	Falso Positivo (FP)	Falso Positivo (FP)
	$-(C_2)$	Falso Negativo (FN)	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)
	$-(C_3)$	Falso Negativo (FN)	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)
		$+(C_1)$	$-(C_2)$	$-(C_3)$
Classes Verdadeiras				

Classe C_2 é a positiva.

Classes Estimadas	$-(C_1)$	Verdadeiro Negativo (TN)	Falso Negativo (FN)	Verdadeiro Negativo (TN)
	$+(C_2)$	Falso Positivo (FP)	Verdadeiro Positivo (TP)	Falso Positivo (FP)
	$-(C_3)$	Verdadeiro Negativo (TN)	Falso Negativo (FN)	Verdadeiro Negativo (TN)
		$-(C_1)$	$+(C_2)$	$-(C_3)$
Classes Verdadeiras				

Classe C_3 é a positiva.

Classes Estimadas	$-(C_1)$	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)	Falso Negativo (FN)
	$-(C_2)$	Verdadeiro Negativo (TN)	Verdadeiro Negativo (TN)	Falso Negativo (FN)
	$+(C_3)$	Falso Positivo (FP)	Falso Positivo (FP)	Verdadeiro Positivo (TP)
		$-(C_1)$	$-(C_2)$	$+(C_3)$
Classes Verdadeiras				

Observações importantes quanto à matriz de confusão

- **Precisão** diz o quão exato é o modelo em relação a ***todos os exemplos classificados como positivos***, ou seja, quantos deles são realmente positivos.

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

TP	FP
FN	TN

- 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 (verdadeiro positivo)***, um ***falso positivo*** significa que um ***ham (verdadeiro negativo)*** foi classificado como ***spam***. O usuário de email pode perder emails importantes se a **precisão** for baixa.

- **Recall** calcula quantos exemplos realmente positivos o classificador captura em relação a todos exemplos positivos.

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

TP	FP
FN	TN

- O **recall** é uma boa medida para determinar a qualidade de um classificador quando houver um alto custo associado a ***falsos negativos***.

- Por exemplo, na classificação de doenças, se um paciente doente (***verdadeiro positivo***) 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

$$\text{Precisão} = \frac{TP}{TP + FP} \quad \begin{array}{|c|c|} \hline TP & FP \\ \hline FN & TN \\ \hline \end{array} \quad \text{Recall} = \frac{TP}{TP + FN}$$

- 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 as duas.

Métricas para avaliação de classificadores

Pontuação-F

- As métricas de **precisão** e **recall** são analisadas conjuntamente através de uma métrica que combina ambas métricas, chamada de **pontuação-F** (ou **F-score**).
- Ela realiza uma **média harmônica ponderada** dada pela equação F_m 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**:

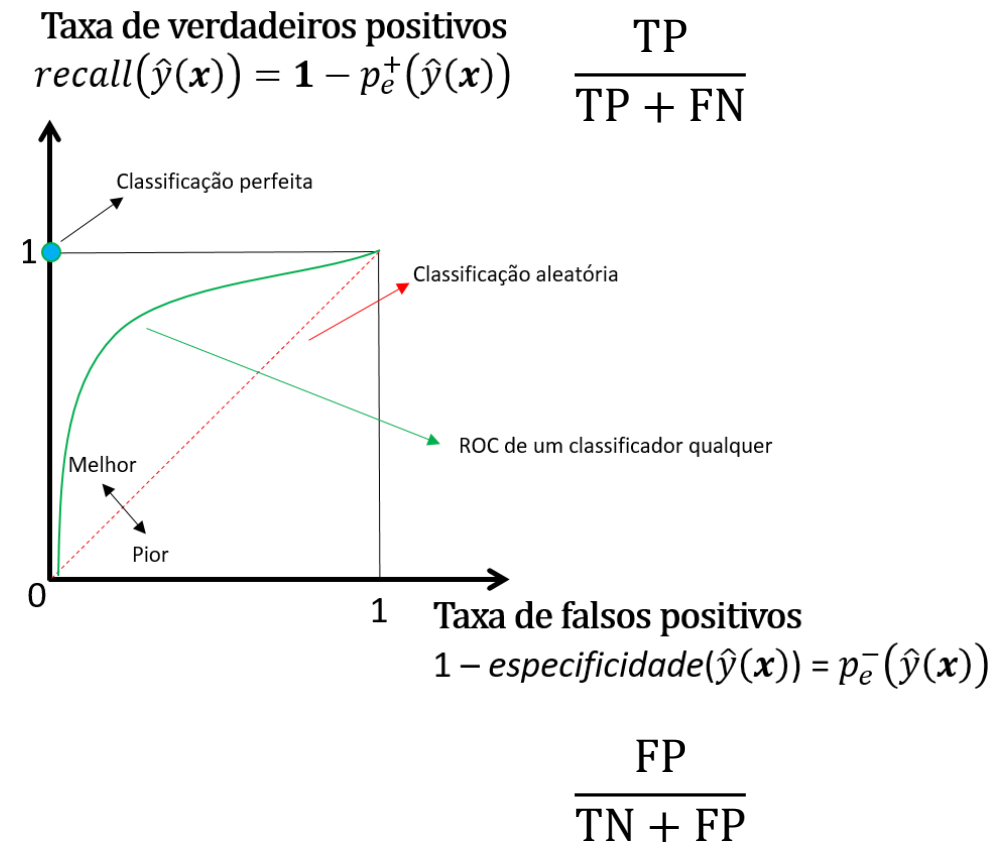
$$F_1 = 2 \frac{\text{recall}(\hat{y}(x)) \times \text{precisão}(\hat{y}(x))}{\text{recall}(\hat{y}(x)) + \text{precisão}(\hat{y}(x))} = \frac{TP}{TP + \frac{FN + FP}{2}}.$$

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

Métricas para avaliação de classificadores

Curva Característica de Operação do Receptor (Curva ROC)

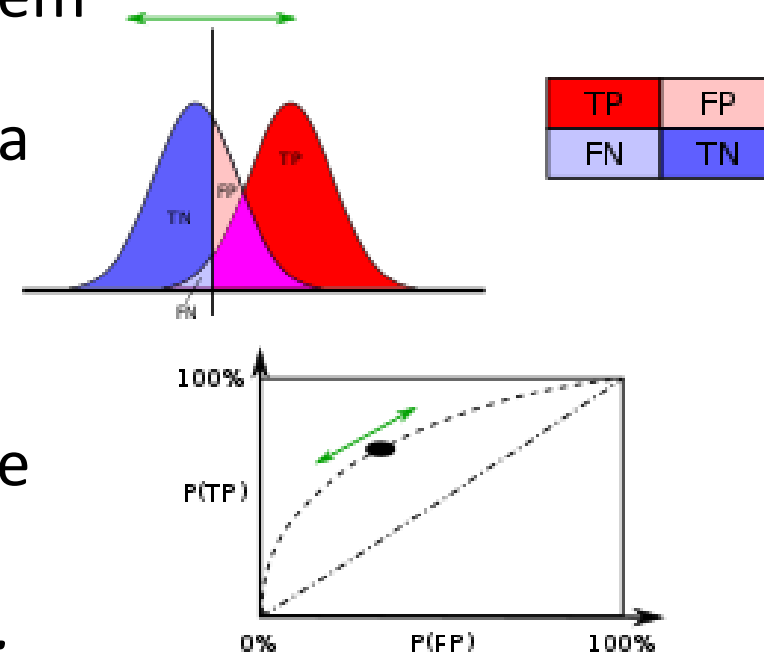
- Gráfico que mostra a performance de um **classificador binário** conforme seu **limiar de discriminação** é variado.
- A curva é criada plotando-se o **recall** em função da **taxa de falsos positivos** para vários valores de **limiar de discriminação**.
- Quanto mais à esquerda e para cima estiver a **curva ROC** de um **classificador**, melhor será o seu desempenho.
- A linha 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 para avaliação de classificadores

Curva Característica de Operação do Receptor (Curva ROC)

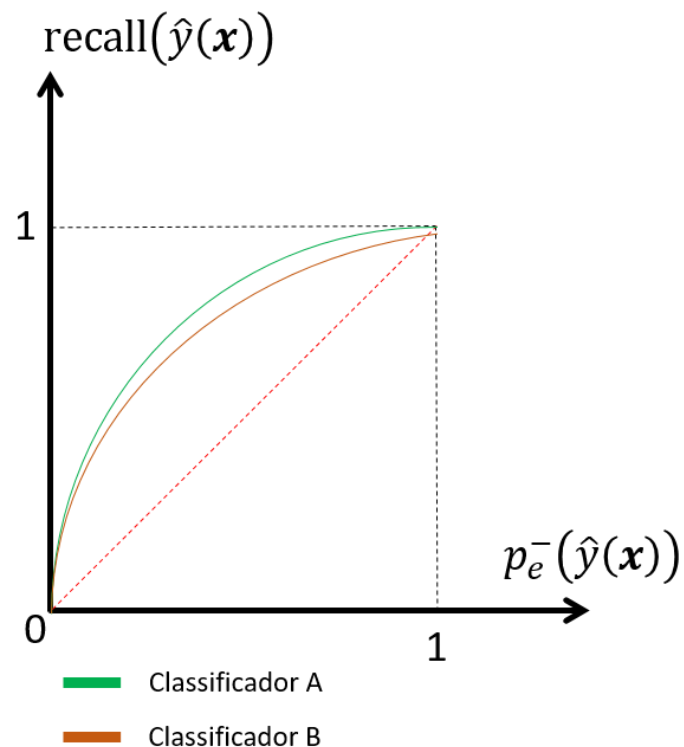
- A forma usual de se comparar **classificadores** consiste em criar uma **curva ROC** para cada um deles.
- Em geral, **classificadores** apresentam em sua saída uma probabilidade para cada exemplo de entrada.
- Normalmente, estas probabilidades são, então, discretizadas para que se tenha a decisão final: por exemplo, se o valor de $h_a(x(i))$ ultrapassa um determinado **limiar**, T , ele é mapeado no valor 1 (classe positiva, C_2); caso contrário, ele é mapeado no valor 0 (classe negativa, C_1).
- Sendo assim, ao plotarmos 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 para avaliação de classificadores

Curva Característica de Operação do Receptor (Curva ROC)



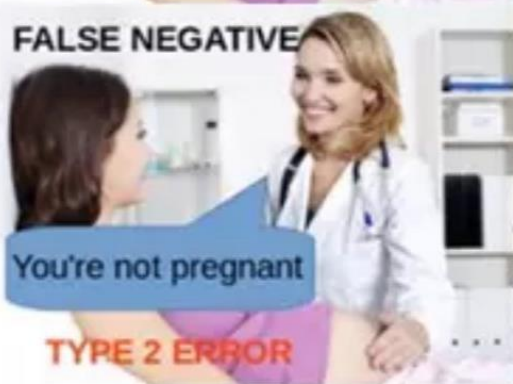
- Por exemplo, considere as **curvas ROC** na figura ao lado. Para decidir qual o melhor **classificador**, podemos tomar como base a **área sob a curva (ASC) ROC**.
- **ASC** é outra métrica da qualidade de um classificador. É um número entre 0 e 1. Quanto maior a **ASC**, melhor será o classificador.
- 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!

		Actual Values	
		1	0
Predicted Values	1	<p>TRUE POSITIVE</p> 	<p>FALSE POSITIVE</p>  <p>TYPE 1 ERROR</p>
	0	<p>FALSE NEGATIVE</p>  <p>TYPE 2 ERROR</p>	<p>TRUE NEGATIVE</p> 