

# Inteligência Artificial

Guilherme Henrique de Souza Nakahata

Universidade Estadual do Paraná - Unespar

25 de Julho de 2024

- **Erro de Classificação**

$$err(f) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(y_i \neq f(x_i))$$

$$err(f) = \frac{\text{exemplos\_incorretamente\_classificados}}{\text{total\_exemplos\_classificados}}$$

- **Acurácia de Classificação**

$$acc(f) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(y_i = f(x_i)) = 1 - err(f)$$

$$acc(f) = \frac{\text{exemplos\_corretamente\_classificados}}{\text{total\_exemplos\_classificados}}$$

- Problema da acurácia e da taxa de erro?

- Problema da acurácia e da taxa de erro?
- Distribuição de classes;
- Classes desbalanceadas;
- A - 99 %, B - 0,75%, C - 0,25%;
- 99 % acurácia.

- Fornece um limiar abaixo do qual o erro do modelo deve ficar;
- $\text{erro} - \text{maj}(T) = 1 - \text{argmax}_{distribuição}(c_i) \quad i = 1, \dots, k;$
- Exemplo:

- **Número de exemplos:**
- **Número de classes:**
- **Distribuição de classes:**

Cabeça $X_1$	Peso $X_2$	Sorriso $X_3$	Classe $Y$
Redonda	8	Não	Feliz
Triangular	11	Sim	Neutro
Redonda	6	Sim	Feliz
Quadrada	9	Não	Triste
Quadrada	10	Sim	Neutro
Triangular	6	Não	Triste
Redonda	5	Sim	Feliz
Redonda	15	Não	Triste
Quadrada	12	Sim	Neutro
Redonda	7	Não	Feliz
Triangular	13	Sim	Neutro
Quadrada	10	Sim	Neutro

# Exemplo

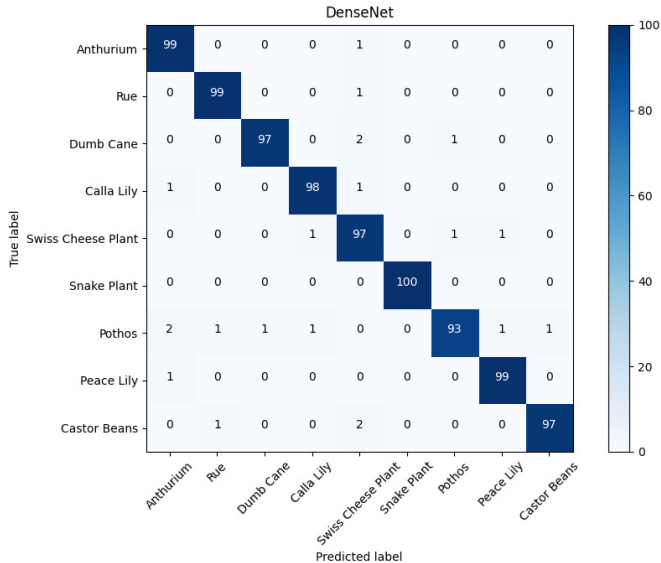
- **Número de exemplos:**  
 $n = 12$
- **Número de classes:**  $k = 3$ 
  - $C1 = \text{feliz}$
  - $C2 = \text{neutro}$
  - $C3 = \text{triste}$
- **Distribuição de classes:**
  - $\text{distr}(C1) = \frac{4}{12} = 33.3\%$
  - $\text{distr}(C2) = \frac{5}{12} = 41.7\%$
  - $\text{distr}(C3) = \frac{3}{12} = 25\%$
- A classe  $C2$  (neutro) é **majoritária**  $\rightarrow \frac{5}{12}$  (41.7%)
- A classe  $C3$  (triste) é **minoritária**  $\rightarrow \frac{3}{12}$  (25%)
- Erro majoritário:  
 $1 - \frac{5}{12} = \frac{7}{12} = 58.3\%$

Cabeça $X_1$	Peso $X_2$	Sorriso $X_3$	Classe $Y$
Redonda	8	Não	Feliz
Triangular	11	Sim	Neutro
Redonda	6	Sim	Feliz
Quadrada	9	Não	Triste
Quadrada	10	Sim	Neutro
Triangular	6	Não	Triste
Redonda	5	Sim	Feliz
Redonda	15	Não	Triste
Quadrada	12	Sim	Neutro
Redonda	7	Não	Feliz
Triangular	13	Sim	Neutro
Quadrada	10	Sim	Neutro

- Métricas tradicionais;
- Consideram a distribuição de classes;
- **Precisão;**
- **Revocação;**
- Obtidas a partir da **matriz de confusão**;
- Mapeia os erros e acertos de um classificador:
  - Valores esperados;
  - Valores previstos;
- Elementos da diagonal principal correspondem aos acertos;
- Valores fora os erros;
- Matriz de confusão ideal?



# Matriz de confusão



<b>Exemplo</b>	$X_1$	$X_2$	$X_3$	$Y$	$h(X)$
$z_1$	$\alpha$	$s$	2	+	+
$z_2$	$\alpha$	$s$	1	-	+
$z_3$	$b$	$n$	1	+	-
$z_4$	$b$	$s$	2	-	-
$z_5$	$c$	$n$	2	+	+

		<b>Predita</b>		<b>Total</b>
		+	-	
<b>Verdadeira</b>	+			
	-			
<b>Total</b>				

<b>Exemplo</b>	$X_1$	$X_2$	$X_3$	$Y$	$h(X)$
$z_1$	$\alpha$	$s$	2	+	+
$z_2$	$\alpha$	$s$	1	-	+
$z_3$	$b$	$n$	1	+	-
$z_4$	$b$	$s$	2	-	-
$z_5$	$c$	$n$	2	+	+

		<b>Predita</b>		<b>Total</b>
		+	-	
<b>Verdadeira</b>	+	1	2	3
	-	1	1	2
<b>Total</b>		2	3	5

- VP (Verdadeiros Positivos) = Exemplos positivos classificados como positivos;
- FP (Falsos Positivos) = Exemplos negativos classificados como positivos Verdadeira;
- FN (Falsos Negativos) = Exemplos positivos classificados como negativos;
- VN (Verdadeiros Negativos) = Exemplos negativos classificados como negativos;

$$err(h) = \frac{FP + FN}{VP + FP + FN + VN}$$

$$acc(h) = \frac{VP + VN}{VP + FP + FN + VN}$$

- Precisão;
- Revocação;
- F-score;

# Métricas para classificação

- Precisão;
  - "Entre todas as instâncias que foram previstas como positivas, quantas realmente são positivas?"

$$\text{Precisão}(C_i) = \frac{\text{Corretamente\_Reconhecidos}_{C_i}}{\text{Total\_Reconhecidos}_{C_i}}$$

- Revocação;
  - "Entre todas as instâncias que realmente são positivas, quantas foram corretamente identificadas pelo modelo?"

$$\text{Revocação}(C_i) = \frac{\text{Corretamente\_Reconhecidos}_{C_i}}{\text{Total\_exemplos}_{C_i}}$$

- F-score;
  - Média harmônica da precisão e da revocação, proporcionando um balanço entre as duas métricas;

$$\text{Medida - F}(C_i) = \frac{2 \times \text{revocação}_{C_i} \times \text{precisão}_{C_i}}{\text{revocação}_{C_i} + \text{precisao}_{C_i}}$$

# Métricas para classificação

		Pedita		Total
		+	-	
Verdadeira	+	1	2	3
	-	1	1	2
Total		2	3	5

		Pedita	
		+	-
Verdadeira	+	VP	FN
	-	FP	VN

# Métricas para classificação

		Preditá	
		+	-
Verdadeira	+	VP	FN
	-	FP	VN

- $\text{Precisão}(+) = \frac{VP}{VP+FP}$
- $\text{Precisão}(-) = \frac{VN}{VN+FN}$
- $\text{Revocação}(+) = \frac{VP}{VP+FN}$
- $\text{Revocação}(-) = \frac{VN}{VN+FP}$



# Métricas para classificação

		Predita		Total
		+	-	
Verdadeira	+	1	2	3
	-	1	1	2
Total		2	3	5

		Predita	
		+	-
Verdadeira	+	VP	FN
	-	FP	VN

- $\text{Precisão}(+) = \frac{VP}{VP+FP} = \frac{1}{2} = 0,5$
- $\text{Precisão}(-) = \frac{VN}{VN+FN} = \frac{1}{3} = 0,33$
- $\text{Revocação}(+) = \frac{VP}{VP+FN} = \frac{1}{3} = 0,33$
- $\text{Revocação}(-) = \frac{VN}{VN+FP} = \frac{1}{2} = 0,5$

# Métricas para classificação

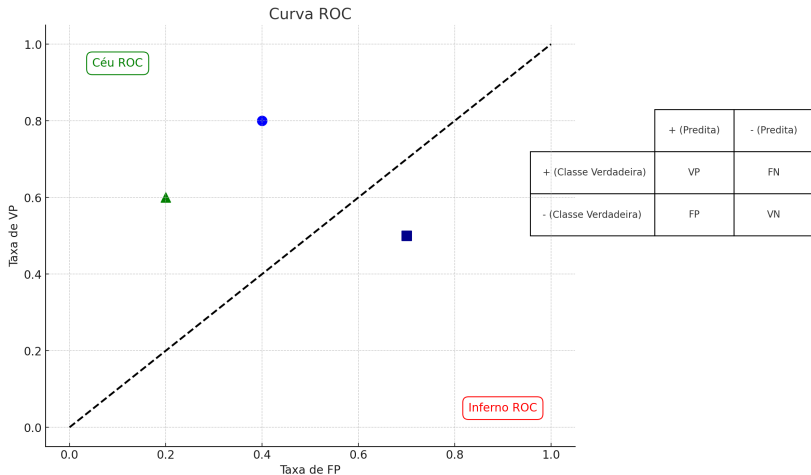
		Predita		Total
		+	-	
Verdadeira	+	1	2	3
	-	1	1	2
Total		2	3	5

		Predita	
		+	-
Verdadeira	+	VP	FN
	-	FP	VN

- $\text{Precisão}(+) = \frac{VP}{VP+FP} = \frac{1}{2} = 0,5$
- $\text{Precisão}(-) = \frac{VN}{VN+FN} = \frac{1}{3} = 0,33$
- $\text{Revocação}(+) = \frac{VP}{VP+FN} = \frac{1}{3} = 0,33$
- $\text{Revocação}(-) = \frac{VN}{VN+FP} = \frac{1}{2} = 0,5$
- $F1(+) = \frac{2 \times 0,33 \times 0,5}{0,33 + 0,5} = 0,40$
- $F1(-) = \frac{2 \times 0,5 \times 0,33}{0,5 + 0,33} = 0,40$

- Análise ROC;
- Forma de avaliar classificadores;
- **Problemas binários;**
- Curva ROC (Receiver Operating CHaracteristics);
- Espaço ROC;
- Gráfico bidimensional;
- Taxa de VP é traçada no eixo Y;
- Taxa de FP é traçada no eixo X;
- Ponto (0,1) Representa 100 % de acerto;
- Ponto (1,0) Representa 100 % de erro.

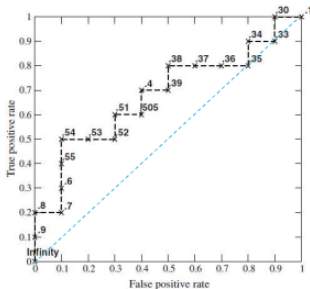
# Métricas para classificação



- Classificadores binários com saídas contínuas ou probabilísticas;
- Saída discretizadas;
- Limiar;
- Curva no espaço ROC;
- **Curva ROC**;
- Mostra o comportamento do classificador;
- Conforme o limiar varia de  $-\infty$  para  $+\infty$ .

- Duas classes (Class):  $p$  e  $n$ ;
- Conjunto de teste com 20 exemplos;
  - 10 positivos ( $p$ ) e 10 negativos ( $n$ );
- Score probabilidade de  $p$  gerada por um classificador;
- Cada ponto no gráfico ROC é rotulado pelo limiar de score que o produz;
  - A curva ROC é a linha gerada a partir desses pontos;
  - Taxa de VP =  $\frac{VP}{P}$ ;
  - Taxa de FP =  $\frac{FP}{N}$ ;
- Curva é baseada nas taxas de VP e FP;
- Não dependem da distribuição de classes;
- Area sobe a curva ROC (**AUC-ROC**).

# Métricas para classificação

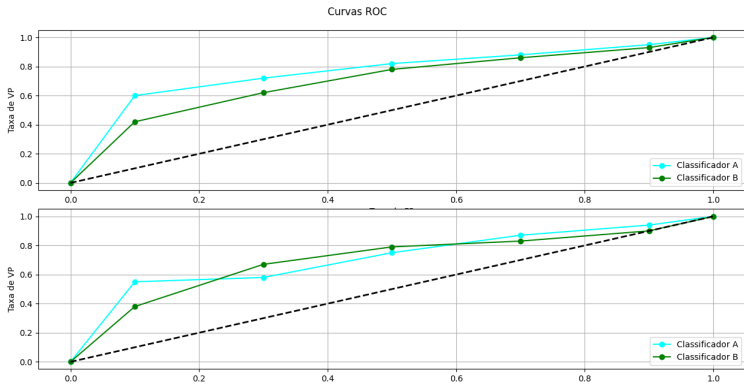


Inst#	Class	Score	Inst#	Class	Score
1	p	.9	11	p	.4
2	p	.8	12	n	.39
3	n	.7	13	p	.38
4	p	.6	14	n	.37
5	p	.55	15	n	.36
6	p	.54	16	n	.35
7	n	.53	17	p	.34
8	n	.52	18	n	.33
9	p	.51	19	p	.30
10	n	.505	20	n	.1

- Podem comparar desempenho de mais classificadores;
- Quanto mais próxima do canto superior esquerdo melhor;
- Intersecção entre curvas?

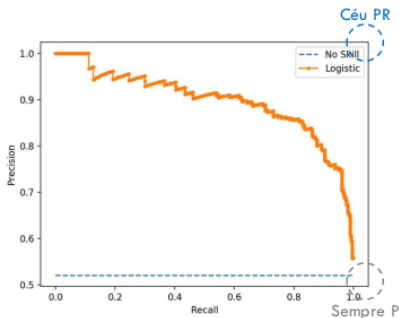


# Métricas para classificação



# Métricas para classificação

- Curva PR;
- Gráfico bidimensional;
- Precisão traçada no eixo Y;
- Revocação traçada no eixo X;



- Avaliação em termos de erro e acertos;
- **Regressão?**

- Avaliação em termos de erro e acertos;
- **Regressão?**
- Tamanho do erro;
- Previsão;
- Saída contínua.

# Métricas para Regressão

**Erro quadrático médio** (*mean-squared error*) é a métrica mais comum;

- **Raiz do erro quadrático médio** (*root mean-squared error*);
- Tende a maximizar o efeito de valores muito errados;

$$\text{mse}(h) = \frac{1}{n} \sum_{i=1}^n (y_i - h(x_i))^2$$

$$\text{rmse}(h) = \sqrt{\text{mse}(h)}$$

**Erro absoluto médio** (*mean absolute error*)

- Trata todos os erros igualmente de acordo com a sua magnitude;

$$\text{mae}(h) = \frac{1}{n} \sum_{i=1}^n |y_i - h(x_i)|$$

## Exemplo de Cálculo do MSE e RMSE:

- Suponha que temos as previsões  $h(x) = [2.5, 0.0, 2.1, 1.4]$
- E os valores reais  $y = [3.0, -0.5, 2.0, 1.0]$

$$\text{mse}(h) = \frac{1}{4} [(3.0 - 2.5)^2 + (-0.5 - 0.0)^2 + (2.0 - 2.1)^2 + (1.0 - 1.4)^2]$$

$$\text{mse}(h) = \frac{1}{4} [0.25 + 0.25 + 0.01 + 0.16] = \frac{1}{4} \times 0.67 = 0.1675$$

$$\text{rmse}(h) = \sqrt{0.1675} \approx 0.409$$

## Exemplo de Cálculo do MAE:

$$\text{mae}(h) = \frac{1}{4} [|3.0 - 2.5| + |-0.5 - 0.0| + |2.0 - 2.1| + |1.0 - 1.4|]$$

$$\text{mae}(h) = \frac{1}{4} [0.5 + 0.5 + 0.1 + 0.4] = \frac{1}{4} \times 1.5 = 0.375$$

- **Erro quadrático relativo;**
- **Raiz do Erro Quadrático Relativo** (Root Relative Squared Error)
- Pondera o erro de acordo com a sua **previsibilidade**:  
considera a distribuição dos valores em torno da média.

$$rse(h) = \frac{\sum_{i=1}^n (y_i - h(x_i))^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$rrse(h) = \sqrt{rse(h)}$$

## Exemplo de Erro Quadrático Relativo

**Valores reais:**  $y = [3, 5, 2, 7]$

**Predições:**  $h(x) = [2.5, 4.5, 3, 6.5]$

**Média dos valores reais:**  $\bar{y} = \frac{3+5+2+7}{4} = 4.25$

**Cálculo do  $rse(h)$ :**

$$\begin{aligned} rse(h) &= \frac{(3 - 2.5)^2 + (5 - 4.5)^2 + (2 - 3)^2 + (7 - 6.5)^2}{(3 - 4.25)^2 + (5 - 4.25)^2 + (2 - 4.25)^2 + (7 - 4.25)^2} \\ &= \frac{0.25 + 0.25 + 1 + 0.25}{1.5625 + 0.5625 + 5.0625 + 7.5625} = 0.065 \end{aligned}$$

**Cálculo do  $rrse(h)$ :**

$$rrse(h) = \sqrt{0.065} \approx 0.255$$



- **Erro Absoluto Relativo** (Relative Absolute Error);
- Mesma ponderação de valores de acordo com a distribuição em torno da média.

$$rae(h) = \frac{\sum_{i=1}^n |y_i - h(x_i)|}{\sum_{i=1}^n |y_i - \bar{y}|}$$

# Exemplo de Erro Absoluto Relativo

**Valores reais:**  $y = [3, 5, 2, 7]$

**Predições:**  $h(x) = [2.5, 4.5, 3, 6.5]$

**Média dos valores reais:**  $\bar{y} = 4.25$

**Cálculo do  $rae(h)$ :**

$$\begin{aligned} rae(h) &= \frac{|3 - 2.5| + |5 - 4.5| + |2 - 3| + |7 - 6.5|}{|3 - 4.25| + |5 - 4.25| + |2 - 4.25| + |7 - 4.25|} \\ &= \frac{0.5 + 0.5 + 1 + 0.5}{1.25 + 0.75 + 2.25 + 2.75} = 0.163 \end{aligned}$$

# Métricas para Regressão

- Independente da métrica;
- Quanto menor o erro **melhor**;
- O melhor modelo de regressão tende a ser melhor;
- Independentemente da métrica utilizada.

- Normalmente comparamos os desempenhos médios dos modelos;
- Diferença pequena;
- Testes estatísticos (testes de hipóteses);
- Verificar se a diferença nas médias é significativa;
  - Teste-t pareado;
  - Teste de Wilcoxon;
  - Teste de Friedman;

# Obrigado! Dúvidas?

Guilherme Henrique de Souza Nakahata

[guilhermenakahata@gmail.com](mailto:guilhermenakahata@gmail.com)

<https://github.com/GuilhermeNakahata/UNESPAR-2024>