

CD - 2024-2 Avaliação 2b

Nome e Sobrenome: _____

Disciplina: Ciéncia dos Dados Semestre: 2024.2

Professor: Dr. Diego Pinheiro

Observações: As respostas deverão ser escritas à CANETA. NÃO SERÃO ACEITAS QUESTÕES APENAS COM RESPOSTAS);

Informações preliminares

As questões a seguir são relacionadas a um modelo de classificação para, a partir da quantidade de admissões em um hospital (x_1), predizer se os pacientes vão falecer ($y = 1$) ou não ($y = 0$) conforme os dados a seguir.

x_1	0	1	2	3	4	5	6	7	8	9	10	11
y	0	0	0	0	1	1	0	1	1	1	1	1

Questão 1 (0,5 pontos)

A regressão logística utiliza a **função logit**, que transforma as probabilidades $Pr(y = 1|x_1)$ na escala **log-odds**, permitindo modelar a relação entre as variáveis preditoras e a variável de resposta binária de forma linear.

$$\text{logit}(x_1) = \ln \left[\frac{Pr(y=1|x_1)}{1 - Pr(y=1|x_1)} \right] = w_0 + w_1 x_1$$

Considere um modelo de regressão logística onde estamos analisando o **odds de falecimento** em função da quantidade de admissões (x_1). Sabemos que o **Odds Ratio (OR)** é calculado como a razão entre os **Odds** para $x_1 + 1$ e os **Odds** para x_1 . Ou seja, $OddsRatio(x_1) = \frac{Odds(x_1+1)}{Odds(x_1)}$. Com base na equação do modelo de regressão logística fornecido, forneça a fórmula geral para calcular o **Odds Ratio (OR)**.

$$OddsRatio(x_1) =$$

Questão 2 (0,5 pontos)

Considerando um modelo inicial $\mathbf{w}^{[0]} = \begin{bmatrix} -1.85 \\ 0.39 \end{bmatrix}$, qual a probabilidade $Pr(y = 1|x_1)$ de falecimento (**com duas casas decimais**) para os pacientes ?

$Pr(y = 1 x_1 = 0)$	
$Pr(y = 1 x_1 = 1)$	
$Pr(y = 1 x_1 = 2)$	
$Pr(y = 1 x_1 = 3)$	
$Pr(y = 1 x_1 = 4)$	
$Pr(y = 1 x_1 = 5)$	
$Pr(y = 1 x_1 = 6)$	
$Pr(y = 1 x_1 = 7)$	
$Pr(y = 1 x_1 = 8)$	
$Pr(y = 1 x_1 = 9)$	
$Pr(y = 1 x_1 = 10)$	
$Pr(y = 1 x_1 = 11)$	

Questão 3 (0,5 pontos)

Qual o **F₁ score** do modelo $\mathbf{w}^{[0]}$ (**com duas casas decimais**) considerando $Pr(y = 1|x_1) \geq 0.5$?

F₁ score	
----------------------------	--

Questão 4 (0,5 pontos)

A partir do modelo $\mathbf{w}^{[0]} = \begin{bmatrix} -1.85 \\ 0.39 \end{bmatrix}$ e considerando $\lambda = 0.9$, $\left(\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}^T} \Big|_{\mathbf{w}=\mathbf{w}^{[0]}} \right) = \begin{bmatrix} -0.38 \\ -6.50 \end{bmatrix}$, $\left(\frac{\partial^2 E(\mathbf{w})}{\partial \mathbf{w} \partial \mathbf{w}^T} \Big|_{\mathbf{w}=\mathbf{w}^{[0]}} \right)^{-1} = \begin{bmatrix} 1.90 & -0.28 \\ -0.28 & 0.06 \end{bmatrix}$, qual seria um melhor modelo $\mathbf{w}^{[1]}$?

w^[1]	
------------------------	--

--	--

Questão 5 (0,5 pontos)

Considerando o modelo $\mathbf{w}^{[1]}$ atualizado anteriormente, qual a probabilidade $Pr(y = 1|x_1)$ de falecimento (**com duas casas decimais**) para os pacientes ?

$Pr(y = 1 x_1 = 0)$
$Pr(y = 1 x_1 = 1)$
$Pr(y = 1 x_1 = 2)$
$Pr(y = 1 x_1 = 3)$
$Pr(y = 1 x_1 = 4)$
$Pr(y = 1 x_1 = 5)$
$Pr(y = 1 x_1 = 6)$
$Pr(y = 1 x_1 = 7)$
$Pr(y = 1 x_1 = 8)$
$Pr(y = 1 x_1 = 9)$
$Pr(y = 1 x_1 = 10)$
$Pr(y = 1 x_1 = 11)$

Questão 6 (0,5 pontos)

Qual o **F₁score** do modelo $\mathbf{w}^{[1]}$ (**com duas casas decimais**) $Pr(y = 1|x_1) \geq 0.5$?

F₁score	
---------------------------	--