

# BCC740 - Inteligência Artificial

## Prova 3

1. Para o modelo linear múltiplo (modelo linear com múltiplas variáveis de entrada), considere o seguinte conjunto de dados:

$x_1$	$x_2$	$y$
0	0	2
1	0	4
0	1	3
1	1	7

Monte a matriz  $X$  (com coluna de 1's) e o vetor  $\mathbf{y}$ .

2. Por que é comum adicionar uma coluna de uns na matriz  $X$ ?
3. Com os dados da questão anterior, calcule as matrizes  $X^\top X$  e  $X^\top \mathbf{y}$ .
4. Inverta a matriz  $X^\top X$  pelo método que achar mais adequado.
5. Obtenha os pesos  $\hat{\mathbf{w}}$  utilizando a fórmula dos mínimos quadrados:

$$\hat{\mathbf{w}} = (X^\top X)^{-1} X^\top \mathbf{y}.$$

6. Considere a função sigmoide:

$$\sigma(z) = \frac{1}{1 + e^{-z}}.$$

Considere também, o conjunto de dados abaixo:

$x_1$	$x_2$	$y$
0	1	1
1	0	0

Considere pesos  $w_0 = 0.5$ ,  $w_1 = 1$ ,  $w_2 = -1$ . Calcule  $z$  e  $\hat{y} = \sigma(z)$  para cada linha.

7. Considere a função de verossimilhança abaixo:

$$L(\mathbf{w}) = \prod_{i=1}^n \hat{y}_i^{y_i} (1 - \hat{y}_i)^{1-y_i}$$

- (a) Por que ela não é comumente utilizada em problemas de classificação? Qual é a alternativa?
- (b) Escreva a expressão da função de perda de entropia cruzada (- log verossimilhança) para dois exemplos rotulados  $y_i \in \{0, 1\}$  e previsões  $\hat{y}_i = \sigma(z_i)$ ,  $i = 1, 2$ .

8. Considere o conjunto de dados abaixo:

$x_1$	$x_2$	$y$
0	1	1
1	0	0

Considere o conjunto de pesos iniciais  $w_0 = 1, w_1 = 1, w_2 = 1$ . Qual será o valor dos pesos após a primeira iteração do algoritmo de descida do gradiente com  $\eta = 1$ ? Veja a formula das derivadas parciais da *loss* abaixo.

$$\frac{\partial L}{\partial w_j} = \sum_{i=1}^n (\hat{y}_i - y_i) x_{ij}$$