

Universidade Kimpa Vita
Departamento do Ensino e Investigação da Engenharia Informática
Disciplina de Inteligência Artificial II

I- Teórica – 6 valores

- 1) Como explicar o conceito de "Overfitting"?
- 2) Quais Como explicaria o termo "Viés"?
- 3) Temos 5 pessoas com idades x iguais a 5, 12, 22, 25, 30, os y 's correspondentes são 0, 1, 0, 1, 1. Se $w_0 = -6.3$ e $w_1 = 0.45$, obtenha a probabilidade de gerar os y 's acima com o modelo logístico. Me forneça o resultado dessas duas pessoas com 11 e 55 de idade

Função:

$$p(x) = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x))} .$$

II- Prática – 14 valores

Problema

Trabalha com uma base de dados que mapeia o preço de smartphones em 4 faixas possíveis de acordo com as características do aparelho. O trabalho está dividido em 2 partes:

Parte 1: Faça uma Clusterização dos dados descobrindo qual o número ideal de grupos. Compare sua resposta com os valores possíveis de price_range.

- *Sua análise levou a mais ou menos clusters que o esperado?*
- *Baseado nos valores das amostras pertencentes a cada um dos clusters formados, o que eles significam?*

Parte 2: Construa um modelo de classificação de smartphones. Não precisa utilizar todas as features. Faça a seleção que julgar pertinente e justifique suas escolhas.

- *Quais métricas escolherá para avaliar o modelo?*
- *Quais foram os resultados obtidos e o que eles significam?*

Importante: Precisa ler a base, limpá-la, pensar em quais dados serão incorporados nos modelos, padronizar dados, etc. Quanto as configurações do modelo deixo a sua preferência.

Chave da Prova

I- Teórica – 6 Valores

1) Como explicar o conceito de "Overfitting"?

R/ Quando um modelo é treinado demais com um conjunto específico de dados, ele pode se tornar muito bom em reconhecer e prever padrões nesses dados específicos, mas não consegue generalizar bem para novos dados.

2) Quais Como explicaria o termo "Viés"?

R/ O viés se refere à diferença sistemática entre as previsões de um modelo e os valores reais. O viés ocorre quando um modelo é muito simples e não consegue capturar a complexidade dos dados. Nesse caso, o modelo subestima as previsões e tem um viés alto.

3) Temos 5 pessoas com idades x iguais a 5, 12, 22, 25, 30, os y 's correspondentes são 0, 1, 0, 1, 1. Se $w_0 = -6.3$ e $w_1 = 0.45$, obtenha a probabilidade de gerar os y 's acima com o modelo de regressão logística. Me forneça o resultado dessas duas pessoas com 11 e 45 de idade

R/ A probabilidade de gerar y para uma pessoa com 11 anos é de 0.11 e para uma pessoa com 45 anos é de 0.99.

Função:

$$p(x) = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x))} .$$

II- Prática – 14 Valores

Problema

Trabalha com uma base de dados que mapeia o preço de smartphones em 4 faixas possíveis de acordo com as características do aparelho. O trabalho está dividido em 2 partes: Parte 1: Faça uma Clusterização dos dados descobrindo qual o número ideal de grupos. Compare sua resposta com os valores possíveis de `price_range`.

- Sua análise levou a mais ou menos clusters que o esperado?
- Baseado nos valores das amostras pertencentes a cada um dos clusters formados, o que eles significam?

Parte 2: Construa um modelo de classificação de smartphones. Não precisa utilizar todas as features. Faça a seleção que julgar pertinente e justifique suas escolhas.

- Quais métricas escolherá para avaliar o modelo?
- Quais foram os resultados obtidos e o que eles significam?

Importante: Precisa ler a base, limpá-la, pensar em quais dados serão incorporados nos modelos, padronizar dados, etc. Quanto as configurações do modelo deixo a sua preferência.

Resposta

