Teste 1 - Versão 1
09/04/2024

Aprendizagem Automática
90 minutos

Nome: _____

ID: _____

Problema	Valores	Classificação
1	3	
2	2	
3	1.5	
4	1.5	
5	3	
6	5	
7	4.5	
Total	20.5	

Problema 1 (Pré-processamento, 3 valores)

1.) Para cada uma das afirmações seguintes indica se esta é verdadeira ou falsa. Circula a opção correta.

1.1) A técnica de imputação por média substitui os valores omissos pela média de toda a linha. (0.5 Valores)

a. Verdadeiro

b. Falso

1.2) O método KNN (K-Nearest Neighbors) pode ser usado para imputar valores omissos num conjunto de dados. (0.5 Valores)

a. Verdadeiro

b. Falso

1.3) Feature scaling é um passo de pré-processamento necessário para todos os algoritmos de aprendizagem automática. (0.5 Valores)

a. Verdadeiro

b. False

1.4) Feature Scaling pode ajudar a acelerar a convergência e treino de modelos de aprendizagem automática. (0.5 Valores)

a. Verdadeiro

b. Falso

1.5) É possível aplicar métodos de feature selection baseados em wrappers a conjuntos de dados que não possuem rótulos (labels). (0.5 Valores)

a. Verdadeiro

b. Falso

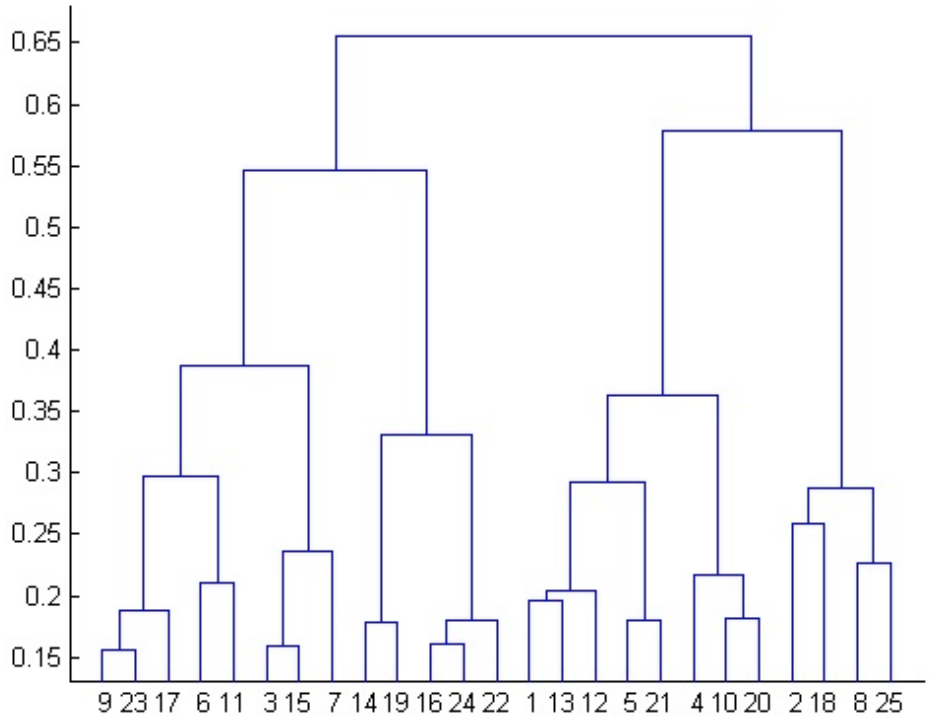
1.6) Os métodos de feature selection embebidos são frequentemente mais eficientes computacionalmente do que os métodos baseados em filtros. (0.5 Valores)

a. Verdadeiro

b. Falso

Problema 2 (Clustering, 2 Valores)

2.1) Dado o seguinte dendograma, qual seria o número de clusters mais apropriado? (1 Valor)



a. 2
b. 4
c. 6
d. 8
e. nenhuma das anteriores

2.2) Suponha que deseje agrupar 7 observações em 3 clusters usando o algoritmo de clustering K-Means. Após a primeira iteração, os clusters C1, C2, C3 têm as seguintes observações:

C1: {(2,2), (4,4), (6,6)} C2: {(0,4), (4,0)} C3: {(5,5), (9,9)}

Qual será a distância de Manhattan para a observação (9, 9) e o centróide do cluster C1 na segunda iteração? (1 Valor)

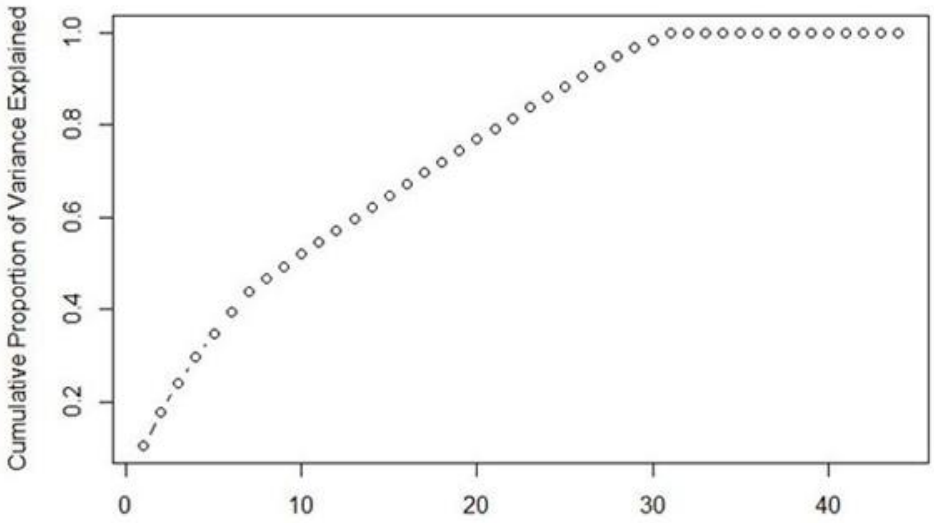
a. $5 \cdot \sqrt{2}$
b. $13 \cdot \sqrt{2}$
c. 10
d. nenhuma das anteriores

Problema 3 (Redução de Dimensionalidade, 1.5 Valores)

3.1) Qual das seguintes afirmações está correta para t-SNE e PCA? (0.75 Valores)

a. t-SNE é linear, enquanto PCA é não linear
b. t-SNE e PCA são ambos lineares
c. t-SNE e PCA são ambos não lineares
d. t-SNE é não linear, enquanto PCA é linear

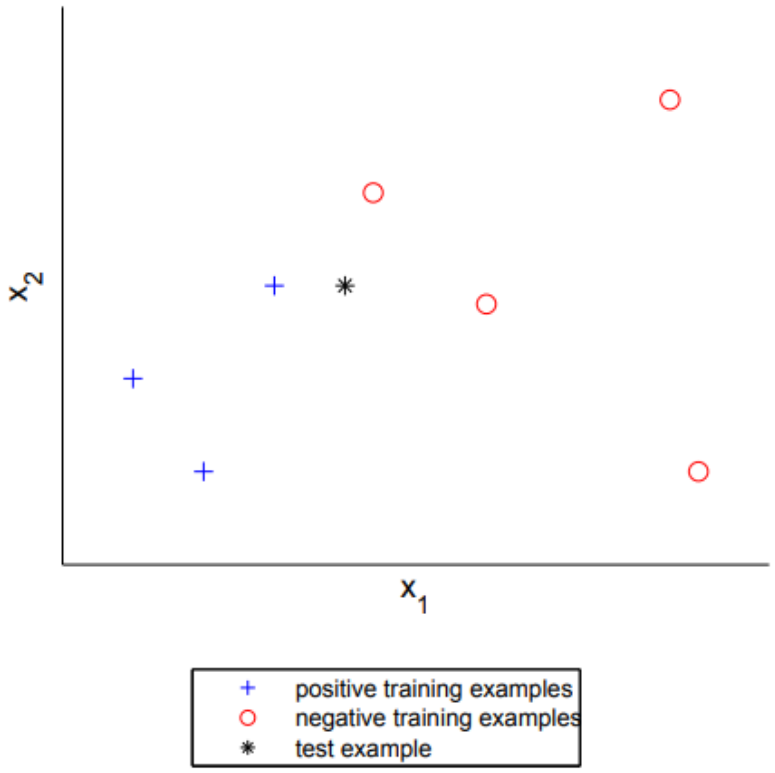
3.2) Tendo em consideração o gráfico seguinte, qual é o número ótimo de componentes principais? (0.75 Valores)



a. 7
b. 20
c. 30
d. 40

Problema 4 (KNN, 1.5 Valores)

4.) Na figura seguinte estão representados dados de treino e um único ponto de teste para a tarefa de classificação dados dois atributos contínuos X1 e X2. Para cada valor de k, circule a previsão do classificador k-nearest neighbor para o ponto de teste representado (*).



4.1) Previsão para k=1: (0.5 Valores)

a. positive (+)

b. negative (o)

4.2) Previsão para k=3: (0.5 Valores)

a. positive (+)

b. negative (o)

4.3) Previsão para k=5: (0.5 Valores)

a. positive (+)

b. negative (o)

Problema 5 (Regressão Linear/Logística, 3 Valores)

5) Considera a tabela abaixo. Supõe que queremos prever a variavel y sabendo as features x1, x2 usando um modelo de regressão linear $\hat{y} = \beta 1x1 + \beta 2\ x2$.

x1	x2	y
1	1	2
3	-1	-6
-1	2	7

5.1) Quais os valores dos coeficientes $\beta = (\beta 1, \beta 2)$ que minimizam a soma dos erros ao quadrado $(\sum(y-\hat{y})^2)$? (1.5 Valores)

a. (2, -1)

b. (3, -1)

c. (-1, 2)

d. (-1, 3)

e. nenhuma das anteriores

5.2) Suponha que lhe é atribuída a seguinte tarefa de classificação: prever o target $Y \in \{0, 1\}$ dados dois atributos de valores reais $X1 \in R$ e $X2 \in R$. Após treinar um modelo de regressão logística obtivemos os seguintes coeficientes: $\beta 0 = -0.5$, $\beta 1 = 1$ e $\beta 2 = 2$. Usando este modelo, qual seria a sua previsão para o exemplo: x1=0, x2=2. (1.5 Valores)

a. 0.97

b. 3.5

c. 0.03

d. nenhuma das anteriores

Problema 6 (Arvores de Decisão, 5 Valores)

6.) Suponha que lhe sejam dados seis pontos de treino (listados na tabela a seguir) para um problema de classificação com dois atributos binários X1, X2 e três classes $Y \in \{1, 2, 3\}$. Vamos usar uma árvore de decisão baseada no ganho de informação.

X1	X2	Y
1	1	1
1	1	1
1	1	2
1	0	3
0	0	2
0	0	3

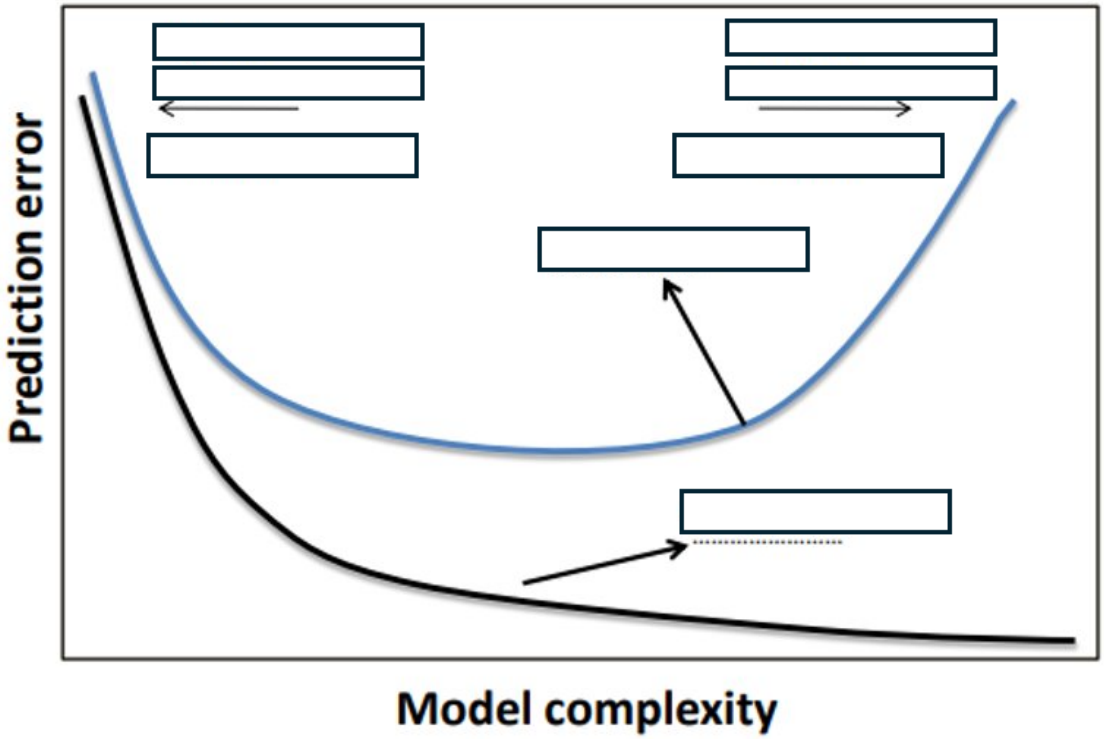
6.1) Desenha os splits possíveis (para X1 e X2). (1.5 Valores)

6.2) Calcula o ganho de informação (information gain) para cada um dos splits. (2 Valores)

6.3) Qual dos splits escolherias? Desenha a árvore de decisão resultante desse split apenas. Certifica-te que legendas todos os ramos e nós/folhas. Como classificarias um exemplo com X1=0 e X2=1? (1.5 Valores)

Problema 7 (Aprendizagem Automática Geral, 4.5 Valores)

7.1) A imagem seguinte mostra as curvas de erro de treino e validação para um modelo com complexidade crescente.



Legenda a imagem com as seguintes expressões: “training error”, “validation error”, “low variance”, “high variance”, “low bias”, “high bias”, “overfitting” e “underfitting”. (1.5 Valores)

7.2) Para cada uma das descrições listadas abaixo, circula se o desenho experimental é adequado ou problemático. Se considerar que é problemático, indique brevemente os problemas com a abordagem.

7.2.1) Uma equipa de trabalho afirma ter alcançado um grande sucesso após obter uma precisão de classificação (accuracy) de 98% numa tarefa de classificação binária onde uma classe é muito rara. Os seus dados consistiam em 100 exemplos positivos e 10 000 exemplos negativos. (0.75 Valores)

a. Adequado

b. Problemático

7.2.2) Uma equipa de trabalho realizou um processo de feature selection e reduziu o número de features do seu dataset para um subset menor. Em seguida, dividiram os dados nos datasets de treino e teste. Treinaram o seu modelo nos dados de treino e reportaram o melhor erro de teste que obtiveram. (0.75 Valores)

a. Adequado

b. Problemático

7.3) Foram estudados vários métodos para controlar o overfitting para diversos classificadores. Abaixo, encontram-se listados vários classificadores e ações que podem afetar o seu bias e variância. Indique (circulando) como o bias e a variância mudam em resposta à ação:

7.3.1) Reduzir o número de folhas numa arvore de decisão: (0.5 Valores)

Bias	**Variância**
Diminuir	Diminuir
Aumentar	Aumentar
Permanecer inalterado	Permanecer inalterado

7.3.2) Aumentar o k num classificador k-nearest neighbor: (0.5 Valores)

Bias	**Variância**
Diminuir	Diminuir
Aumentar	Aumentar
Permanecer inalterado	Permanecer inalterado

7.3.3) Usar regularização num modelo de regressão linear/logística: (0.5 Valores)

Bias	**Variância**
Diminuir	Diminuir
Aumentar	Aumentar
Permanecer inalterado	Permanecer inalterado