

PERGUNTA 1

1. Considere as seguintes matrizes de confusão representando a avaliação de dois classificadores de sentimentos C1 e C2, respectivamente:

C1

| Previsto | Verdadeiro | |
|----------|------------|----------|
| | Positivo | Negativo |
| Positivo | 40 | 30 |
| Negativo | 10 | 20 |

C2

| Previsto | Verdadeiro | |
|----------|------------|----------|
| | Positivo | Negativo |
| Positivo | 25 | 25 |
| Negativo | 25 | 25 |

A partir das avaliações, qual classificador (C1 ou C2) você escolheria?
Selecione a alternativa correta:

Trata-se de um problema desbalanceado, portanto seria melhor utilizar o classificador C1.

Trata-se de um problema desbalanceado, portanto seria melhor utilizar o classificador C2.

Como trata-se de um problema balanceado, podemos escolher tanto C1 quanto C2.

Como trata-se de um problema balanceado, escolhemos C1 simplesmente olhando para a acurácia.

Como trata-se de um problema balanceado, escolhemos C2 simplesmente olhando para a acurácia.

1,68 pontos

PERGUNTA 2

1. A classificação de sentimentos a partir de textos é uma das principais aplicações de análise de sentimentos. Considerando o conjunto de dados a seguir, com informações de 8 (oito) textos, qual atributo preditivo (*feature*, característica) você escolheria para determinar, de forma mais precisa, o sentimento? Se julgar necessário, pode utilizar a equação de Bayes para tomar a decisão.

| ID | #AVALIAÇÕES | CATEGORIA | TEM PAIS VIVOS | SENTIMENTO |
|----|-------------|-----------|----------------|------------|
| 1 | MÉDIA | ESPORTE | SIM | POSITIVO |

| | | | | |
|---|-------|----------|-----|----------|
| 2 | BAIXA | POLICIAL | NÃO | NEGATIVO |
| 3 | ALTA | MODA | NÃO | POSITIVO |
| 4 | MÉDIA | ESPORTE | NÃO | POSITIVO |
| 5 | MÉDIA | POLICIAL | SIM | NEGATIVO |
| 6 | ALTA | POLICIAL | NÃO | NEGATIVO |
| 7 | BAIXA | ESPORTE | SIM | POSITIVO |
| 8 | BAIXA | POLICIAL | SIM | NEGATIVO |

SENTIMENTO.

ID.

CATEGORIA.

Não é possível utilizar somente um atributo.

#AVALIAÇÕES.

1,68 pontos

PERGUNTA 3

- Em uma ontologia, as relações semânticas representam um tipo de relação entre as palavras e os significados. A relação de parte entre duas palavras, ou seja, uma relação “é parte de” é o tipo de relação denominada:
 holonímia.
 meronímia.
 antonímia.
 hiponímia.
 sinonímia.

1,66 pontos

PERGUNTA 4

- Existem alguns métodos que conseguem calcular a similaridade semântica de dois elementos utilizando taxonomia, como o *Path Similarity*, que pode ser entendido como o menor caminho entre dois conceitos.

Considerando a taxonomia acima, qual é a similaridade entre as palavras “Burro” e “Animal” utilizando *Path Similarity*?

3
0,5
0,33
1

1,66 pontos

PERGUNTA 5

1. Considere um minerador de emoções, no qual, em um documento analisado, a avaliação “Bom atendimento” aparece 5 vezes, “Agilidade” 3 vezes, e “Qualidade” 1 vez.

As taxas de frequência (TF) normalizadas para “Bom atendimento”, “Agilidade” e “Qualidade”, respectivamente, são:

1; 0,6 e 0,2.
0,55; 0,33 e 0,11.
0,5; 0,5 e 0.
0,6; 1 e 0,2.
0,5; 0,3 e 0,1.

1,66 pontos

PERGUNTA 6

1. O objetivo da similaridade textual consiste em verificar o quão “próximos” são dois fragmentos de texto a partir _____ (lacuna 1) e de _____ (lacuna 2), em que o primeiro se refere à similaridade _____ (lacuna 3), e a segunda à similaridade _____ (lacuna 4).

Selecione a alternativa que representa os conteúdos adequados para preencher as lacunas do texto acima:

do conhecimento, sua ontologia, semântica, estrutural.
do significado, sua estrutura, léxica, semântica.
do significado, sua estrutura, semântica, léxica.
da estrutura, seu significado, semântica, léxica.
da ontologia, sua estrutura, semântica, estrutural.