



INSTITUTO FEDERAL

Ceará

Campus Aracati

AVALIAÇÃO AP1 N2 – 2024.2

DISCIPLINA DE INTELIGÊNCIA ARTIFICIAL

Naive Bayes e Árvore de Decisão

Nome:

Data:

Diretriz: responda essas questões utilizando um editor de texto no computador e, ao término, envie para valter.costa@ifce.edu.br.

- 1) Uma empresa está desenvolvendo um sistema para classificar mensagens recebidas como "Urgente" ou "Não Urgente" com base nas palavras presentes na mensagem. Foi analisado um conjunto de 100 mensagens, e os dados a seguir foram coletados:

Mensagens Urgentes: 30

Mensagens Não Urgentes: 70

Palavra Presente	Mensagens Urgentes	Mensagens Não Urgentes
"imediatamente"	15	5
"problema"	10	10
"atraso"	8	12

- a) (2,5 pontos) Calcule as probabilidades de uma mensagem ser "Urgente" e "Não Urgente" com base no conjunto de dados fornecido. Adicionalmente, determine as probabilidades condicionais para cada palavra ("imediatamente", "problema" e "atraso") em relação às mensagens "Urgentes" e "Não Urgentes".
- b) (2,5 pontos) Suponha que uma nova mensagem contenha as palavras "imediatamente" e "problema". Calcule a probabilidade de ser uma mensagem "Urgente" e de ser "Não Urgente" utilizando o teorema de Bayes e classifique a mensagem como "Urgente" ou "Não Urgente".

Considere:

$$P(A|B, \dots, Z) = \frac{P(B|A)P(\dots|A)P(Z|A)P(A)}{P(B)P(\dots)P(Z)}$$

- 2) Árvores de decisão são modelos amplamente utilizados para classificação e regressão.
- a) (2,5 pontos) Explique como o conceito de ganho de informação é utilizado na construção de uma árvore de decisão. Detalhe mostrando o uso do cálculo de entropia e ganho de informação em um problema hipotético.
- b) (2,5 pontos) Suponha que você está desenvolvendo um sistema de suporte à decisão, aponte dois critérios de parada que você utilizaria na construção da árvore de decisão para garantir um modelo eficiente?