

Nome: Alisson Chaves Ferreira

Pontuação Total da Avaliação: 6.82 pontos

=====

Correção da Questão 1:

Pergunta: 1a) 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 "imediatamente": 15 (Mensagens Urgentes) e 5 (Mensagens Não Urgentes); Palavra Presente "problema": 10 (Mensagens Urgentes) e 10 (Mensagens Não Urgentes); Palavra Presente "atraso": 8 (Mensagens Urgentes) e 12 (Mensagens Não Urgentes). 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".  
Diretriz(es):  $\{('P("urgente")=30/100=0.30', 0.32), ('P("atraso")| "não urgente")=12/70 \sim 0.17', 0.32), ('P("imediatamente")| "urgente")=15/30=0.5', 0.32), ('P("imediatamente")| "não urgente")=5/70 \sim 0.07', 0.32), ('P("atraso")| "urgente")=8/30 \sim 0.27', 0.32), ('P("problema")| "urgente")=10/30 \sim 0.33', 0.32), ('P("não urgente")=70/100=0.70', 0.32), ('P("problema")| "não urgente")=10/70 \sim 0.14', 0.32)\}$

=====Resposta do estudante da questão 1:

1A) De acordo com a questão, e considerando as mensagens urgentes (U) e as não urgentes (N) assim como suas palavras: imediatamente (I), problema (P) e atraso (A), aqui está a seguinte solução:

Mensagem contendo imediatamente ser urgente e não urgente:

$$P(I|U) = 15/30 = 0.5$$

$$P(I|N) = 5/70 = 0.071428571$$

Mensagem contendo problema ser urgente e não urgente:

$$P(P|U) = 10/30 \approx 0.3333$$

$$P(P|N) = 10/70 = 0.142857142$$

Mensagem contendo atraso ser urgente e não urgente:

$$P(A|U) = 8/30 \approx 0.2666$$

$$P(A|N) = 12/70 = 0.171428571$$

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma boa compreensão sobre como calcular as probabilidades condicionais para cada palavra em relação às mensagens "Urgentes" e "Não Urgentes". No entanto, falta a resposta para as probabilidades de uma mensagem ser "Urgente" e "Não Urgente" com base no conjunto de dados fornecido, que são  $P("Urgente") = 30/100 = 0.30$  e  $P("Não Urgente") = 70/100 = 0.70$ .

=====Feedback:

Os cálculos condicionais para as palavras "imediatamente", "problema" e "atraso" estão corretos. Para "imediatamente",  $P("I"|U) = 15/30 = 0.5$  e  $P("I"|N) = 5/70 \approx 0.07$ . Para "problema",  $P("P"|U) = 10/30 \approx 0.33$  e  $P("P"|N) = 10/70 \approx 0.14$ . Para "atraso",  $P("A"|U) = 8/30 \approx 0.27$  e  $P("A"|N) = 12/70 \approx 0.17$ . Esses cálculos demonstram uma boa compreensão do conceito de probabilidade condicional.

=====Pontuação:

2.271219512195122

=====

Correção da Questão 2:

Pergunta: 1b) 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 "imediatamente": 15 (Mensagens Urgentes) e 5 (Mensagens Não Urgentes); Palavra Presente "problema": 10 (Mensagens Urgentes) e 10 (Mensagens Não Urgentes); Palavra Presente "atraso": 8 (Mensagens Urgentes) e 12 (Mensagens Não Urgentes). 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".

Diretriz(es):  $\{(P("Urgente"|"imediatamente"&"problema")=(P("urgente")*P("imediatamente"|"urgente")*P("problema"|"urgente"))/(P("problema")*P("imediatamente"))), \text{podendo simplificar para } P("Urgente"|"imediatamente"&"problema")=P("urgente")*P("imediatamente"|"urgente")*P("problema"|"urgente") \sim 0.30 * 0.5 * 0.33 \sim 0.05', 1.0), ('Classificar o e-mail como "URGENTE" baseado no cálculo do máximo argumento entre } P("Urgente"|"imediatamente"&"problema") \text{ e } P("Não Urgente"|"imediatamente"&"problema")', 0.5), (P("Não Urgente"|"imediatamente"&"problema")=(P("Não urgente")*P("imediatamente"|"Não urgente")*P("problema"|"Não urgente"))/(P("problema")*P("imediatamente"))), \text{podendo simplificar para } P("Não Urgente"|"imediatamente"&"problema")=P("Não urgente")*P("imediatamente"|"Não urgente")*P("problema"|"Não urgente") \sim 0.70 * 0.07 * 0.014 \sim 0.001', 1.0)\}$

=====Resposta do estudante da questão 2:

1B) Classifica-se a mensagem contendo "imediatamente" e "problema" como urgente  
 $P(U|"imediatamente"&"problema") = 0.30 \times 0.5 \times 0.3333 / 0.15 \times 0.1 = 0.03333$

Classifica-se a mensagem contendo "imediatamente" e "problema" como não urgente  
 $P(\bar{U}|\bar{A}|\bar{B}) = 0.70 \times 0.07114285 \times 0.142857142 / 0.5 \times 0.10 = 0.0231$

=====Correção proposta pela IA:

=====Considerações:

A resposta do estudante apresenta alguns erros e simplificações incorretas na aplicação do Teorema de Bayes para classificar a mensagem como "Urgente" ou "Não Urgente".

=====Feedback:

Para calcular  $P("Urgente"|"imediatamente"&"problema")$ , o estudante deve considerar a probabilidade a priori de uma mensagem ser urgente, que é  $30/100 = 0.3$ , e as probabilidades condicionais de encontrar as palavras "imediatamente" e "problema" em mensagens urgentes, que são  $15/30 = 0.5$  para "imediatamente" e  $10/30 = 0.3333$  para "problema". Já para  $P("Não Urgente"|"imediatamente"&"problema")$ , as probabilidades condicionais são  $5/70 = 0.0714$  para "imediatamente" e  $10/70 = 0.1429$  para "problema", com uma probabilidade a priori de 0.7 para mensagens não urgentes. Além disso, o denominador do Teorema de Bayes deve levar em consideração a probabilidade de encontrar essas palavras em qualquer tipo de mensagem.

=====Pontuação:

1.1015873015873017

=====Correção da Questão 3:

Pergunta: 2a) Árvores de decisão são modelos amplamente utilizados para classificação e regressão. 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.

Diretriz(es):  $\{( 'Demonstrar o cálculo de entropia final como sendo } H("depois da partição")=P("amostras irem para conjunto 1")*H("conjunto 1 após partição")+P("amostras irem para conjunto 2")*H("conjunto 2 após partição")', 1.0), ('Explicar que o conceito de Information Gain (IG) como sendo } IG=H("antes da partição")-H("depois da partição") \text{ é essencial para decidir qual atributo/partição escolher a cada nível da árvore}', 1.0), ('Demonstrar o cálculo de entropia inicial como sendo } H("antes da partição")=-p("classe1")*\log_2(P("classe1"))-...-p("classeN")*\log_2(P("classeN"))', 1.0)\}$

=====Resposta do estudante da questão 3:

2A) O ganho de informação, ou entropia funciona trabalhando e obtendo os menores graus de diferença entre os atributos de uma árvore de decisão. Ele trabalha estimando o grau de cada atributo.

Vamos supor que temos uma tabela com o atributo vento, esse atributo pode ser "Sim" ou "Não". Dependemos desse atributo para calcular a entropia. 5 são "Sim" e 7 são "Não" para jogar basquete, totalizando 12 atributos.

$H(S) = -(5/12 \times \log_2(5/12) + 7/12 \times \log_2(7/12))$

$H(S) = -(0.416666 \times 1.1609 + 0.3333 \times 0.1666)$

$H(S) = 0.72593372$

Aqui conseguimos medir a entropia da influencia do vento na decisão de jogar ou não basquete naquele dia.

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma compreensão básica do conceito de entropia e sua aplicação em árvores de decisão, mas carece de profundidade e precisão em relação às diretrizes fornecidas. O estudante calculou a entropia inicial corretamente, considerando a probabilidade de cada atributo, mas não prosseguiu com o cálculo do ganho de informação (Information Gain, IG) e não aplicou isso para decidir qual atributo escolher a cada nível da árvore de decisão.

=====Feedback:

Um ponto forte da resposta é a tentativa de calcular a entropia da influência do vento na decisão de jogar ou não basquete, mostrando que o estudante tem uma compreensão básica da fórmula de entropia. No entanto, para melhorar, o estudante precisa entender e aplicar corretamente a fórmula de ganho de informação ( $IG = H(\text{"antes da partição"}) - H(\text{"depois da partição"})$ ) e ilustrar como isso é usado para escolher o melhor atributo em cada nó da árvore de decisão. Além disso, o estudante deve demonstrar como calcular a entropia após a partição, considerando a probabilidade de as amostras irem para cada conjunto após a divisão e o cálculo da entropia para cada um desses conjuntos.

=====Pontuação:

1.5

=====

Correção da Questão 4:

Pergunta: 2b) Árvores de decisão são modelos amplamente utilizados para classificação e regressão. 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?

Diretriz(es): {'Apontar a limitação da quantidade de níveis da árvore (profundidade máxima atingida) como um possível critério de parada.', 1.25), ('Apontar qualquer outra limitação plausível para que uma árvore de decisão não continue a realizar partições com base em um número de amostras presentes em um dado nó.', 1.25), ('Apontar a limitação de número insuficiente de amostras no nó para realizar nova partição como um possível critério de parada.', 1.25), ('Apontar a limitação de entropia das classes (labels) atingir zero no nó como um possível critério de parada.', 1.25)}

=====Resposta do estudante da questão 4:

2B) Para garantir um algoritmo eficiente em um modelo de decisão, posso usar dois critérios de parada.

Um deles é a profundidade da árvore que irei definir até que nível essa busca será feita, isso aumentará a eficiência impedindo que ela chegue a níveis de busca desnecessários e aumente o processamento.

Um outro atributo útil para garantir que tenhamos um critério de parada eficiente é a aplicação do monitoramento da homogeneidade, ou pureza das classes dentro do conjunto. Se em algum momento estes atributos assumem um valor muito próximo, não há motivo para que se continue e gaste poder de processamento à toa, assim será logo finalizado.

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante aborda dois critérios de parada para a construção de uma árvore de decisão: a profundidade da árvore e a homogeneidade ou pureza das classes dentro do conjunto. Embora esses critérios sejam relevantes, a resposta pode ser mais precisa e abrangente em relação às diretrizes fornecidas.

=====Feedback:

Um ponto forte da resposta é a menção à profundidade da árvore como um critério de parada, o que está alinhado com a diretriz sobre limitar a quantidade de níveis da árvore. No entanto, para melhorar e atingir a nota máxima, seria útil mencionar explicitamente outras limitações plausíveis, como o número insuficiente de amostras no nó para realizar nova partição ou a limitação de entropia das classes atingir zero no nó. Além disso, uma explicação mais detalhada sobre como esses critérios contribuem para a eficiência do modelo seria benéfica.

=====Pontuação:  
1.9461206896551724  
=====