

Nome: ALEXSANDRO COSTA DA SILVA  
Pontuação Total da Avaliação: 6.72 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): {( 'Acertar a resposta aproximada de  $P(\text{"não urgente"}) = 70/100 = 0.70$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"atraso"}|\text{"não urgente"}) = 12/70 \approx 0.17$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"urgente"}) = 30/100 = 0.30$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"imediatamente"}|\text{"não urgente"}) = 5/70 \approx 0.07$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"atraso"}|\text{"urgente"}) = 8/30 \approx 0.27$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"problema"}|\text{"urgente"}) = 10/30 \approx 0.33$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"problema"}|\text{"não urgente"}) = 10/70 \approx 0.14$ , 0.32), ( 'Acertar a resposta aproximada de  $P(\text{"imediatamente"}|\text{"urgente"}) = 15/30 = 0.5$ , 0.32) }

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

1A) U (Urgente)  $\hat{=}$   $30/100 = 0,3$   
NU (NÃO urgente)  $\hat{=}$   $70/100 = 0,7$   
Probabilidades Condicionais:  
"imediatamente"  $= 20/100 = 0,2$   
"problema"  $= 20/100 = 0,2$   
"atraso"  $= 20/100 = 0,2$   
 $P(\text{"imediatamente"}|U) = 15/20 = 0,75$   
 $P(\text{"imediatamente"}|NU) = 5/20 = 0,25$   
 $P(\text{"problema"}|U) = 10/20 = 0,5$   
 $P(\text{"problema"}|NU) = 10/20 = 0,5$   
 $P(\text{"atraso"}|U) = 8/20 = 0,4$   
 $P(\text{"atraso"}|NU) = 15/20 = 0,6$

=====

**Correção proposta pela IA:**

=====

**Considerações:**

A resposta do estudante apresenta alguns erros e imprecisões em relação às probabilidades condicionais. Embora o estudante tenha calculado corretamente as probabilidades de uma mensagem ser "Urgente" (0,3) e "Não Urgente" (0,7), os cálculos das probabilidades condicionais para cada palavra apresentam erros.

=====Feedback:

O estudante errou ao calcular as probabilidades condicionais  $P(\text{"imediatamente"}|U)$ ,  $P(\text{"problema"}|U)$ ,  $P(\text{"atraso"}|U)$ ,  $P(\text{"imediatamente"}|NU)$ ,  $P(\text{"problema"}|NU)$  e  $P(\text{"atraso"}|NU)$ . As probabilidades condicionais devem ser calculadas com base no número total de mensagens "Urgentes" ou "Não Urgentes" que contêm a palavra específica, em relação ao número total de mensagens "Urgentes" ou "Não Urgentes". Por exemplo,  $P(\text{"imediatamente"}|U) = 15/30 = 0,5$  e  $P(\text{"imediatamente"}|NU) = 5/70 \approx 0,07$ . Além disso,  $P(\text{"problema"}|U) = 10/30 \approx 0,33$ ,  $P(\text{"problema"}|NU) = 10/70 \approx 0,14$ ,  $P(\text{"atraso"}|U) = 8/30 \approx 0,27$  e  $P(\text{"atraso"}|NU) = 12/70 \approx 0,17$ .

=====Pontuação:

1.92

=====

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

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

1B)  $P(U|\text{"imediatamente"}) = (0,3*0,75)/0,2 = 1,125$

$P(NU|\text{"imediatamente"}) = (0,3*0,25)/0,2 = 0,375$

$P(U|\text{"problema"}) = (0,3*0,5)/0,2 = 0,75$

$P(NU|\text{"problema"}) = (0,7*0,5)/0,2 = 1,75$

$P(U|\text{"imediatamente"} \text{ e } \text{"problema"}) = 0,3*1,125*0,75 = 0,25$

$P(NU|\text{"imediatamente"} \text{ e } \text{"problema"}) = 0,7*0,375*1,75 = 0,46$

A mensagem é classificada como "Não Urgente", pois a maior probabilidade foi a de  $P(NU|\text{"imediatamente"} \text{ e } \text{"problema"}) = 0,46$ .

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma tentativa de aplicar o teorema de Bayes para classificar a mensagem como "Urgente" ou "Não Urgente". No entanto, há erros significativos na aplicação do teorema e nos cálculos. O estudante não calculou corretamente as probabilidades condicionais nem aplicou corretamente o teorema de Bayes para as condições específicas dadas.

=====Feedback:

Para chegar à solução correta, o estudante deveria primeiro calcular as probabilidades condicionais corretas, como  $P(\text{"imediatamente"}|\text{"Urgente"})$ ,  $P(\text{"problema"}|\text{"Urgente"})$ ,  $P(\text{"imediatamente"}|\text{"Não Urgente"})$  e  $P(\text{"problema"}|\text{"Não Urgente"})$ , e então aplicar o teorema de Bayes considerando as probabilidades anteriores de ser "Urgente" ou "Não Urgente". Além disso, a simplificação e aplicação correta do teorema de Bayes para as condições dadas ("imediatamente" e "problema") são cruciais. É importante notar que a resposta do estudante não permite uma classificação precisa com base nos cálculos apresentados devido a erros nos passos iniciais de cálculo.

=====Pontuação:

0.5

=====

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

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

2A) Ganho de informação é o resultado da diferença da entropia inicial calculada para a

entropia de alguma variável para melhorar o resultado da classificação. O exemplo hipotético seria para dizer se uma turma de pessoa pode jogar no dia, dando as variáveis como clima, vento, umidade e outras. Começando com o cálculo da entropia inicial, que seria contar todos os casos, separar os que de jogar são verdadeiros e os que são falsos com a seguinte fórmula  $H(\text{Joga}) = -P(\text{verdadeiro})\log_2(P(\text{verdadeiro})) - P(\text{falso})\log_2(P(\text{falso}))$ . Depois disso poderia ser analisado os casos em que estava ventando separar quando os casos verdadeiros e falsos e calcular a entropia de cada um para quando o Jogar for verdadeiro e falso com as seguintes fórmulas  $H(\text{Joga}|\text{VentoVerdadeiro}) = -P(\text{JogaVerdadeiro})\log_2(P(\text{JogaVerdadeiro})) - P(\text{JogaFalso})\log_2(P(\text{JogaFalso}))$  e  $H(\text{Joga}|\text{VentoFalso}) = -P(\text{JogaVerdadeiro})\log_2(P(\text{JogaVerdadeiro})) - P(\text{JogaFalso})\log_2(P(\text{JogaFalso}))$ , depois de ter os dois resultados aplica-los na seguinte fórmula  $H(\text{Vento}) = P(\text{VentoVerdadeiro}) * H(\text{Joga}|\text{VentoVerdadeiro}) + P(\text{VentoFalso}) * H(\text{Joga}|\text{VentoFalso})$  e com o resultado disso podemos calcular o ganho de informação usando  $H(\text{Joga}) - H(\text{Vento})$ .

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante demonstra um entendimento básico do conceito de ganho de informação e sua aplicação na construção de árvores de decisão. Ele corretamente identificou que o ganho de informação é calculado como a diferença entre a entropia inicial e a entropia após a partição. Além disso, o estudante forneceu um exemplo hipotético para ilustrar o uso do cálculo de entropia e ganho de informação.

=====Feedback:

No entanto, a resposta poderia ser melhorada com mais detalhes e precisão na explicação do cálculo da entropia e do ganho de informação. Por exemplo, seria útil explicar que o cálculo da entropia inicial é feito como  $H(\text{"antes da partição"}) = -p(\text{"classe1"})\log_2(P(\text{"classe1"})) - \dots - p(\text{"classeN"})\log_2(P(\text{"classeN"}))$ . Além disso, o estudante poderia ter explicitado que o ganho de informação é calculado como  $IG = H(\text{"antes da partição"}) - H(\text{"depois da partição"})$ , onde  $H(\text{"depois da partição"})$  é calculado como  $P(\text{"amostras irem para conjunto 1"}) * H(\text{"conjunto 1 após partição"}) + P(\text{"amostras irem para conjunto 2"}) * H(\text{"conjunto 2 após partição"})$ .

=====Pontuação:

1.8

=====

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): {'Mencionar e/ou explicar dois ou mais critérios de parada válidos na construção de árvores de decisão.', 2.5}, {'Não mencionar e/ou explicar qualquer critério de parada na construção de árvores de decisão.', 0.0}, {'Mencionar e/ou explicar dois apenas um critério de parada válido na construção de árvores de decisão.', 1.25}}

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

2B) Os critérios de paradas usados seriam o de profundidade máxima e o de quantidade limite por folhas, usando um desses dois modelos é garantido que o sistema não passe do limite esperado e pare quando estiver em uma das condições sem perder a pureza.

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante menciona dois critérios de parada válidos na construção de árvores de decisão, que são a profundidade máxima e a quantidade limite por folhas. Isso está alinhado com a rubrica que solicita a menção e/ou explicação de dois ou mais critérios de parada válidos.

=====Feedback:

Um feedback positivo é que o estudante conseguiu identificar dois critérios importantes para evitar overfitting e garantir que o modelo não se torne demasiado complexo. No entanto, uma melhoria necessária seria a explicação mais detalhada de como esses critérios funcionam e por que são importantes. Além disso, seria benéfico mencionar outros critérios de parada possíveis,

como o tamanho mínimo de amostra por nó ou o critério de pureza.

=====Pontuação:

2.5

=====