

Nome: CARLOS SANTIAGO NASCIMENTO FERREIRA

Pontuação Total da Avaliação: 9.40 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  $P(\text{"problema"}|\text{"urgente"})=10/30\sim0.33$ ', 0.32}, ('Acertar a resposta aproximada de  $P(\text{"não urgente"})=70/100=0.70$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"atraso"}|\text{"urgente"})=8/30\sim0.27$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"problema"}|\text{"não urgente"})=10/70\sim0.14$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"urgente"})=30/100=0.30$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"imediatamente"}|\text{"urgente"})=15/30=0.5$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"atraso"}|\text{"não urgente"})=12/70\sim0.17$ ', 0.32), ('Acertar a resposta aproximada de  $P(\text{"imediatamente"}|\text{"não urgente"})=5/70\sim0.07$ ', 0.32)}

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

1A) Para calcular a probabilidade de uma mensagem ser urgente ou não urgente, segue aqui as probabilidades:

$P(\text{Urgente}) = 30/100$  (que é o total de mensagens urgentes sobre o nosso total de mensagens especificado na questão)

$P(\text{Não-Urgente}) = 70/100$  (que é o total de mensagens não urgentes sobre o nosso total de mensagens especificado na questão)

Após isso, temos as probabilidades condicionais:

$P(\text{Imediatamente} | \text{Urgente}) = 15/30$  (Total de mensagens que contem imediatamente sobre todas as palavras que são urgentes)

$P(\text{Imediatamente} | \text{Não-Urgente}) = 5/70$  (Total de mensagens que contem imediatamente sobre todas as palavras que não são urgentes)

$P(\text{Problema} | \text{Urgente}) = 10/30$

$P(\text{Problema} | \text{Não-Urgente}) = 10/70$

$P(\text{Atraso} | \text{Urgente}) = 8/30$

$P(\text{Atraso} | \text{Não-Urgente}) = 12/70$

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante está correta. Ele calculou corretamente as probabilidades de uma mensagem ser "Urgente" e "Não Urgente", que são  $30/100 = 0,30$  e  $70/100 = 0,70$ , respectivamente. Além disso, ele calculou corretamente as probabilidades condicionais para cada palavra em relação às mensagens "Urgentes" e "Não Urgentes". Por exemplo,  $P(\text{Imediatamente} | \text{Urgente}) = 15/30 = 0,5$ ,  $P(\text{Imediatamente} | \text{Não-Urgente}) = 5/70 = 0,07$ ,  $P(\text{Problema} | \text{Urgente}) = 10/30 = 0,33$ ,  $P(\text{Problema} | \text{Não-Urgente}) = 10/70 = 0,14$ ,  $P(\text{Atraso} | \text{Urgente}) = 8/30 = 0,27$  e  $P(\text{Atraso} | \text{Não-Urgente}) = 12/70 = 0,17$ .

=====Feedback:

O estudante demonstrou um bom entendimento do conceito de probabilidade condicional e aplicou corretamente as fórmulas para calcular as probabilidades. No entanto, é importante notar que a resposta não fornece nenhuma interpretação ou discussão adicionais sobre os resultados. Para melhorar, o estudante poderia discutir como essas probabilidades condicionais podem ser usadas para classificar novas mensagens como "Urgentes" ou "Não Urgentes". Além disso, seria útil se o estudante fornecesse exemplos concretos de como essas probabilidades podem ser aplicadas em um cenário real.

=====Pontuação:

2.4

=====

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): {( '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), ( '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), ( '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) }

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

1B) Para Urgente:

$$P(\text{Urgente} | \text{Imediatamente, Problema}) = (P(\text{Urgente}) * P(\text{Imediatamente} | \text{Urgente}) * P(\text{Problema} | \text{Urgente})) / P(\text{Problema}) * P(\text{Imediatamente})$$

Vamos calcular as probabilidades da seguinte maneira:

$$P(\text{Problema}) = 20/100 \text{ (Soma da quantidade mensagens que contém a palavra "problema" sobre o total de mensagens especificado na questão)}$$

$$P(\text{Imediatamente}) = 20/100 \text{ (Soma da quantidade mensagens que contém a palavra "imediatamente" sobre o total de mensagens especificado na questão)}$$

$$P(\text{Urgente}) = 30/100 \text{ (que é o total de mensagens urgentes sobre o nosso total de mensagens especificado na questão)}$$

$$P(\text{Não-Urgente}) = 70/100 \text{ (que é o total de mensagens não urgentes sobre o nosso total de mensagens especificado na questão)}$$

$$P(\text{Imediatamente} | \text{Urgente}) = 15/30 \text{ (Total de mensagens que contem imediatamente sobre todas as palavras que são urgentes)}$$

$$P(\text{Imediatamente} | \text{Não-Urgente}) = 5/70 \text{ (Total de mensagens que contem imediatamente sobre todas as palavras que não são urgentes)}$$

$$P(\text{Problema} | \text{Urgente}) = 10/30$$

$$P(\text{Problema} | \text{Não-Urgente}) = 10/70$$

Simplificando o denominador e simplificando os valores fracionários, ficaria assim:

$$P(\text{Urgente} | \text{Imediatamente, Problema}) = (0.3 * 0.5 * 0.33) = 0.0495$$

Agora, realizando o cálculo para não-urgente:

$$P(\text{Não-Urgente} | \text{Imediatamente, Problema}) = (P(\text{Não-Urgente}) * P(\text{Imediatamente} | \text{Não-Urgente}) * P(\text{Problema} | \text{Não-Urgente})) / P(\text{Problema}) * P(\text{Imediatamente})$$

Simplificando o denominador novamente, e substituindo os valores, ficaria da seguinte maneira:

$$P(\text{Não-Urgente} | \text{Imediatamente, Problema}) = 0.7 * 0.07 * 0.14 = 0.00686$$

Ou seja, utilizando o teorema de Bayes podemos classificar a mensagem como Urgente, já que ao comparar as duas probabilidades, vemos que:

$$P(\text{Urgente} | \text{Imediatamente, Problema}) = 0.0495$$

$$P(\text{Não-Urgente} | \text{Imediatamente, Problema}) = 0.00686$$

$$P(\text{Urgente} | \text{Imediatamente, Problema}) > P(\text{Não-Urgente} | \text{Imediatamente, Problema})$$

$$0.0495 > 0.00686$$

Então, a mensagem seria classificada como urgente.

=====

Correção proposta pela IA:

=====

#### =====Considerações:

A resposta do estudante demonstra uma compreensão básica do teorema de Bayes e sua aplicação na classificação de mensagens como "Urgente" ou "Não Urgente". No entanto, há room para melhoria na clareza e precisão dos cálculos. O estudante calculou corretamente as probabilidades  $P(\text{Urgente}|\text{Imediatamente, Problema})$  e  $P(\text{Não-Urgente}|\text{Imediatamente, Problema})$ , mas poderia ter fornecido mais detalhes sobre o processo de simplificação.

#### =====Feedback:

Pontos fortes da resposta incluem a identificação correta das probabilidades necessárias para aplicar o teorema de Bayes e a comparação das probabilidades para classificar a mensagem. No entanto, para melhorar, o estudante poderia ter explicitado melhor o cálculo do denominador comum para ambas as probabilidades, garantindo que todas as partes do teorema de Bayes estejam claramente representadas. Além disso, uma discussão mais detalhada sobre a interpretação dos resultados e porque a mensagem foi classificada como "Urgente" com base nas probabilidades calculadas seria benéfica.

#### =====Pontuação:

2.3

#### =====

#### 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) Para visualizarmos isso,

Hipoteticamente, digamos que tenhamos um problema: Uma tabela de informações contendo os jogos de vôlei que um grupo jogou ao decorrer da semana, e com isso eles queriam saber se é possível prever pelas partições dessa tabelas se haverá jogo ou não no dia.

Digamos que a nossa entropia inicial, a nossa primeira ramificação seja jogo ou não joga. Se em 14 jogos no total, 7 jogamos e 7 não jogamos, vamos calcular a entropia:

A fórmula:

$$H(x) = -p_1 \log_2(p_1) - p_2 \log_2(p_2)$$

Onde os "p" se referem a probabilidade. Para ficar mais claro, vamos utilizar a fórmula com a seguinte unidade joga.

$$P(\text{Joga}=\text{sim}) = 7/14 \text{ (Quantidade de vezes jogadas sobre o total de jogos, ser o nosso } p_1)$$

$$P(\text{Joga}=\text{não}) = 7/14 \text{ (Quantidade de vezes jogadas sobre o total de jogos, ser o nosso } p_2)$$

$$H(\text{Joga}) = -0.5 \log_2(0.5) - 0.5 \log_2(0.5) = 0.84$$

Agora, vamos selecionar uma partição (uma ou um conjunto de colunas da nossa tabela) para compor o cálculo de entropia e ganho

Digamos que iremos utilizar a coluna de Temperatura com  $90^\circ\text{F}$  para a nossa partição, então separaremos para o caso de Temperatura  $> 90^\circ\text{F}$  e outro para Temperatura  $\leq 90^\circ\text{F}$ .

Vamos dizer que de acordo com nossa tabela, nos dias em que jogaram e a temperatura estava acima de  $90^\circ\text{F}$  foram 3

$$P(\text{Temperatura} > 90 | \text{Joga} = \text{Sim}) = 3/7 \text{ (Total de dias que a temperatura estava acima de } 90 \text{ sobre a quantidade de dias que teve jogo)}$$

$$P(\text{Temperatura} > 90 | \text{Joga} = \text{Não}) = 4/7$$

Vamos dizer que de acordo com nossa tabela, nos dias em que jogaram e a temperatura estava acima de  $90^\circ\text{F}$  foram 5

$$P(\text{Temperatura} \leq 90 | \text{Joga} = \text{Sim}) = 4/7 \text{ (Total de dias que a temperatura estava acima de } 90 \text{ sobre a quantidade de dias que teve jogo)}$$

$$P(\text{Temperatura} \leq 90 | \text{Joga} = \text{Não}) = 3/7$$

$$H(\text{Temperatura} > 90) = -0.42 \log_2(0.42) - 0.57 \log_2(0.57)$$

$$H(\text{Temperatura} \leq 90) = -0.57 * \log_2(0.57) - 0.57 * \log_2(0.42)$$

Após o computador realizar os cálculos, faremos o cálculo de Entropia para o conjunto S:

$$H(S|\text{Temperatura}) = 7/14 \text{ (Total de Joga sim dividido pelo total de itens do conjunto)} *$$

$$H(\text{Temperatura} > 90) * 7/14 \text{ (Total de Joga não dividido pelo total de itens do conjunto)} *$$

$$H(\text{Temperatura} \leq 90) = -0.57 * \log_2(0.57) - 0.57 * \log_2(0.42)$$

Após isso, precisamos calcular o ganho, que vai se dar pela subtração da entropia inicial, menos o resultado do  $H(S|\text{Temperatura})$ .

$$G(S|\text{Temperatura}) = 0.84 - H(S|\text{Temperatura})$$

Esse cálculo é importante para ramificar a nossa Árvore e a quantidade do ganho vai ser referente ao quanto essa partição é importante para a resolução do nosso problema.

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma boa compreensão do conceito de ganho de informação e entropia na construção de árvores de decisão. Ele corretamente explica a fórmula da entropia inicial e calcula a entropia para o problema hipotético. Além disso, o estudante também apresenta a ideia de calcular o ganho de informação depois de realizar uma partição nos dados.

=====Feedback:

No entanto, o estudante não explicita claramente a fórmula do ganho de informação como sendo  $IG = H(\text{"antes da partição"}) - H(\text{"depois da partição"})$ , embora tenha mencionado a ideia de subtrair a entropia inicial da entropia depois da partição. Além disso, os cálculos da entropia final apresentados parecem confusos e não seguem claramente a fórmula  $H(\text{"depois da partição"}) = P(\text{"amostras irem para conjunto 1"}) * H(\text{"conjunto 1 após partição"}) + P(\text{"amostras irem para conjunto 2"}) * H(\text{"conjunto 2 após partição"})$ . Para atingir a nota máxima, o estudante deveria explicitar essas fórmulas e realizar os cálculos de forma clara e organizada.

=====Pontuação:

2.2

=====

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

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

2B) Citarei vários critérios interessantes para serem usados de parada para Árvore de decisão.

Quantidade de Filhos em um Nó: Podemos utilizar isso principalmente como ponto de parada, porque a quantidade de itens em um nó, além de gerar um gasto computacional alto, pode gerar uma quantidade de partições que podem prejudicar o nosso problema. Uma versão alternativa a essa seria também parar em quantidade x de nós-veis em uma Árvore, os motivos são os mesmos. Pureza dos Nós: Se em minha Árvore, a entropia dos itens do nó está muito próxima, é um critério de parada porque cada item está muito similar, então as divisões dentro da Árvore vão ficando redundante.

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante menciona dois critérios de parada válidos na construção de árvores de decisão, que são a quantidade de filhos em um nó e a pureza dos nós. Esses critérios estão de acordo com as práticas comuns de construção de árvores de decisão para evitar sobre-ajuste e garantir um modelo eficiente.

=====Feedback:

O estudante apresentou uma resposta satisfatória, pois mencionou dois critérios relevantes. No

entanto, seria útil se o estudante fornecesse mais detalhes sobre como esses critérios são implementados em árvores de decisão e como eles se relacionam com a otimização do modelo. Além disso, exemplos práticos ou ilustrações poderiam ajudar a clarificar a aplicação desses critérios.

=====Pontuação:

2.5

=====