

Nome: AlvaroMaiaChaves

Pontuação Total da Avaliação: 4.66 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): $\{(P("problema"|"não urgente")=10/70 \sim 0.14', 0.32), (P("imediatamente"|"urgente")=15/30=0.5', 0.32), (P("atraso"|"não urgente")=12/70 \sim 0.17', 0.32), (P("urgente")=30/100=0.30', 0.32), (P("não urgente")=70/100=0.70', 0.32), (P("problema"|"urgente")=10/30 \sim 0.33', 0.32), (P("atraso"|"urgente")=8/30 \sim 0.27', 0.32), (P("imediatamente"|"não urgente")=5/70 \sim 0.07', 0.32)\}$

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

1a) $p(U) = 30/100 = 0,3$

$p(\bar{U}) = 70/100 = 0,7$

$P(I|U) = 15/0,3 = 50\%$

$P(I|\bar{U}) = 5/0,7 = 7,14\%$

$P(P|U) = 10/0,3 = 33,3\%$

$P(P|\bar{U}) = 10/0,7 = 14,2\%$

$P(A|U) = 8/0,3 = 26,6\%$

$P(A|\bar{U}) = 12/0,7 = 17,14\%$

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma boa compreensão do problema, calculando corretamente as probabilidades iniciais de uma mensagem ser "Urgente" ($p(U) = 30/100 = 0,3$) e "Não Urgente" ($p(\bar{U}) = 70/100 = 0,7$). No entanto, ao calcular as probabilidades condicionais para cada palavra, o estudante cometeu erros brutos de cálculo.

=====Feedback:

Para as probabilidades condicionais, o estudante não aplicou corretamente as fórmulas de probabilidade condicional. Por exemplo, $P(I|U) = 15/30 = 0,5$, $P(I|\bar{U}) = 5/70 = 0,071$ (aproximadamente), $P(P|U) = 10/30 = 0,33$, $P(P|\bar{U}) = 10/70 = 0,143$ (aproximadamente), $P(A|U) = 8/30 = 0,267$ (aproximadamente) e $P(A|\bar{U}) = 12/70 = 0,171$ (aproximadamente). O estudante deve revisar a aplicação das fórmulas de probabilidade condicional e considerar as frequências absolutas das palavras nas mensagens "Urgentes" e "Não Urgentes".

=====Pontuação:

2.155820895522388

=====

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): $\{(P("Urgente"|"imediatamente" \& "problema")=(P("urgente")*P("imediatamente"|"urgente")*P("problema"|"urgente"))/(P("problema")*P("imediatamente"|"urgente"))), \text{podendo simplificar para } P("Urgente"|"imediatamente" \& "problema")=P("urgente")*P("imediatamente"|"urgente")*P("problema"|"ur$

nte") $\sim 0.30 * 0.5 * 0.33 \sim 0.05$, 1.0), ('Classificar o e-mail como "URGENTE" baseado no cálculo do máximo argumento entre $P("Urgente"|"imediatamente"&"problema")$ e $P("Não Urgente"|"imediatamente"&"problema")$ ', 0.5), ($P("Não Urgente"|"imediatamente"&"problema") = (P("Não urgente") * P("imediatamente"|"Não urgente") * P("problema"|"Não urgente")) / (P("problema") * P("imediatamente"))$), 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) $P(U|I \& P) = [P(U) * P(I|U)] + [P(U) * P(P|U)] = 0,3 * 50 + 0,3 * 33,3 = 24,99\%$

$P(\bar{U}|\bar{I} \& P) = [P(\bar{U}) * P(\bar{I}|\bar{U})] + [P(U) * P(P|\bar{U})] = 0,7 * 7,14 + 0,7 * 14,2 = 14,93\%$

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta erros significativos na aplicação do teorema de Bayes para calcular as probabilidades de uma mensagem ser "Urgente" ou "Não Urgente" com base nas palavras presentes. Primeiramente, o estudante não aplicou corretamente a fórmula do teorema de Bayes, que requer a divisão pela probabilidade total de ocorrerem as palavras "imediatamente" e "problema" em qualquer mensagem, seja ela urgente ou não. Além disso, os cálculos para as probabilidades condicionais apresentados estão incorretos, não refletindo a aplicação correta do teorema de Bayes.

=====Feedback:

Para atingir a nota máxima, o estudante precisa aplicar corretamente o teorema de Bayes. Isso significa calcular as probabilidades condicionais de cada palavra dentro de mensagens urgentes e não urgentes e então usar essas probabilidades para encontrar a probabilidade de uma mensagem ser urgente dado que contém as palavras "imediatamente" e "problema". Por exemplo,

$P("Urgente"|"imediatamente"&"problema") = (P("Urgente") * P("imediatamente"|"Urgente") * P("problema"|"Urgente")) / (P("imediatamente") * P("problema"))$, considerando $P("imediatamente")$ e $P("problema")$ calculadas com base na ocorrência dessas palavras em todas as mensagens. Além disso, é crucial garantir que as probabilidades condicionais sejam calculadas corretamente com base nos dados fornecidos.

=====Pontuação:

0.24806201550387597

=====

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 inicial como sendo $H("antes da partição") = -p("classe1") * \log_2(P("classe1")) - \dots - p("classeN") * \log_2(P("classeN"))$ ', 1.0), ('Explicar que o conceito de Information Gain (IG) como sendo $IG = H("antes da partição") - H("depois da partição")$ é essencial para decidir qual atributo/partição escolher a cada nível da árvore', 1.0), ('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)}

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

2a) O ganho de informa  o se da quando se usa o calculo de entropia ao criar um novo n   ou seja, ser   escolhido de acordo com o calculo a vari  vel que naquele momento mostra um ganho de informa  o, como exemplo simples considere,

Uma tabela de disponibilidade de dias para se jogar v  lei onde nas colunas contem os dados:

-tempo: nublado, chuvoso, limpo, nublado, limpo, limpo, nublado, chuvoso

-joga: joga, n  oJoga, joga, joga, n  oJoga, n  oJoga, joga, joga

Agora temos que calcular a entropia principal ou seja da classe joga ou n  o joga:

$h(joga) = -p(joga) * \log_2(p(joga)) - \hat{\blacksquare} p(joga) * \log_2(p(joga))$

$h(n  oJoga) = -p(n  oJoga) * \log_2(p(n  oJoga)) - \hat{\blacksquare} p(n  oJoga) * \log_2(p(n  oJoga))$

contando que temos uma propor  o para quem joga = 5/3 e para quem n  oJoga = 3/5

Substituindo,

$$h(joga) = -5/3 * \log_2(5/3) - 5/3 * \log_2(5/3)$$

$$h(joga) = -1,666 * 0,736 - 1,66 * 0,736$$

$$h(joga) = -2,452$$

$$h(n\tilde{A}oJoga) = -3/5 * \log_2(3/5) - 3/5 * \log_2(3/5)$$

$$h(n\tilde{A}oJoga) = -0,6 * 0,736 - 0,6 * 0,736$$

$$h(n\tilde{A}oJoga) = 0,8832$$

$$h(joga|n\tilde{A}oJoga) = -h(joga) - (n\tilde{A}oJoga)$$

$$h(joga|n\tilde{A}oJoga) = 2,452 - 0,8832$$

Assim a entropia de jogar fica,

$$h(joga|n\tilde{A}oJoga) = 1,568$$

Agora para ver qual atributo mostra um ganho de informa  o, nesse caso tempo:

$$h(nublado|joga) = -p(nublado|joga) * \log_2(p(nublado|joga))$$

$$-p(nublado|n\tilde{A}oJoga) * \log_2(p(nublado|n\tilde{A}oJoga))$$

$$h(nublado|joga) = -3/5 * \log_2(3/5) - 0/3 * \log_2(0/3)$$

$$h(nublado|joga) = -0,6 * 0,736$$

$$h(nublado|joga) = 0,4416$$

$$h(chuvoso|joga) = -p(chuvoso|joga) * \log_2(p(chuvoso|joga))$$

$$-p(chuvoso|n\tilde{A}oJoga) * \log_2(p(chuvoso|n\tilde{A}oJoga))$$

$$h(chuvoso|joga) = -1/5 * \log_2(1/5) - 1/3 * \log_2(1/3)$$

$$h(chuvoso|joga) = -0,2 * 2,32 - 0,33 * 1,58$$

$$h(chuvoso|joga) = 0,464 + 0,52$$

$$h(chuvoso|joga) = 0,984$$

$$h(limpo|joga) = -p(limpo|joga) * \log_2(p(limpo|joga)) - p(limpo|n\tilde{A}oJoga) * \log_2(p(limpo|n\tilde{A}oJoga))$$

$$h(limpo|joga) = -1/5 * \log_2(1/4) - 2/3 * \log_2(2/3)$$

$$h(limpo|joga) = -0,2 * 2 - 0,66 * 0,58$$

$$h(limpo|joga) = 0,4 + 1,24$$

$$h(limpo|joga) = 1,64$$

Agora comparamos com a entropia de jogar

$$h(joga) = 1,568 - h(nublado|joga) +$$

$$1,568 - h(chuvoso|joga)$$

$$1,568 - h(limpo|joga)$$

$$h(joga) = 1,568 - 0,4416 + 1,568 - 0,984 + 1,568 - 1,64$$

$$h(joga) = 1,126 + 0,584 - 0,072$$

No caso a coluna tempo tem uma influencia naquela na escolha de jogar ou n  o

$$h(joga) = 1,638$$

Corre  o proposta pela IA:

Considera  es:

A resposta do estudante apresenta uma boa tentativa de explicar o conceito de ganho de informa  o na constru  o de uma  rvore de decis  o, mas carece de clareza e precis  o em alguns pontos. O estudante tentou calcular a entropia inicial e a entropia ap  s a parti  o, mas os c lculos apresentados cont m erros e falta de detalhes.

Feedback:

O estudante precisa melhorar a explica  o do conceito de Information Gain (IG) e como ele   utilizado para decidir qual atributo ou parti  o escolher a cada n vel da  rvore. Al m disso, os c lculos de entropia precisam ser revistos e apresentados de forma clara e precisa. O estudante deve demonstrar o c lculo de entropia inicial usando a f rmula $H(\text{"antes da parti  o"}) = -p(\text{"classe1"}) * \log_2(P(\text{"classe1"})) - \dots - p(\text{"classeN"}) * \log_2(P(\text{"classeN"}))$, e explicar como o Information Gain   calculado como $IG = H(\text{"antes da parti  o"}) - H(\text{"depois da parti  o"})$.

Pontua  o:

1.3532608695652173

Corre  o da Quest  o 4:

Pergunta: 2b)  rvores de decis  o s o modelos amplamente utilizados para classifica  o e regress  o. Suponha que voc  est  desenvolvendo um sistema de suporte   decis  o, aponte dois crit rios de parada voc  utilizaria na constru  o da  rvore de decis  o para garantir um modelo eficiente?

Diretriz(es): {'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 da quantidade de níveis da árvore (profundidade máxima atingida) 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), ('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)}

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

2b) O tamanho máximo da árvore para que não exceda e comece a gerar dados errôneos, e limitar o número de atributos por n^3 para que se alcance um tamanho uniforme em cada n^3 .

=====

Correção proposta pela IA:

=====

=====Considerações:

A resposta do estudante apresenta uma visão básica sobre os critérios de parada para a construção de árvores de decisão, mas não aborda completamente as diretrizes fornecidas. O estudante menciona o tamanho máximo da árvore e a limitação do número de atributos por nó, o que está relacionado à profundidade da árvore, mas não explora outros critérios importantes como a limitação devido à entropia das classes ou a insuficiência de amostras em um nó.

=====Feedback:

Um ponto forte da resposta é a menção ao tamanho máximo da árvore, o que está alinhado com a diretriz sobre a limitação da quantidade de níveis da árvore. No entanto, para atingir a nota máxima, seria necessário incluir mais critérios de parada como a limitação do número de amostras no nó ou a redução da entropia das classes a zero. Além disso, a resposta poderia ser mais clara e específica sobre como esses critérios contribuem para a eficiência do modelo.

=====Pontuação:

0.9077669902912622

=====