

Questão 1:

- 1) Em primeiro lugar, vale ressaltar que existem diversos problemas que são resolvidos por aprendizado de máquina, como por exemplo: classificação, regressão, agrupamento e regras de associação. O primeiro tipo, sendo o mais utilizado, tem como objetivo principal prever/descrever uma classe, sendo um atributo nominal. Já o segundo tipo se assemelha bastante do primeiro, sendo sua principal diferença o tipo de atributo, que nesse caso é numérico. O terceiro tipo tem como finalidade agrupar as instâncias de acordo com atributos de classificação, assim criando grupos de instancias similares. Por fim, o último tipo de problema busca associações entre as instâncias.
- 2) O aprendizado de máquina engloba várias abordagens. No aprendizado supervisionado, modelos aprendem a partir de exemplos com entradas e rótulos. O aprendizado não supervisionado descobre padrões em dados sem rótulos, com técnicas como K-means e mapas auto-organizáveis. O semi-supervisionado mistura dados rotulados e não rotulados. No aprendizado por reforço, agentes buscam maximizar recompensas em interações. O deep learning usa redes neurais profundas para aprender representações complexas. Cada abordagem tem usos específicos e vantagens, aplicadas de acordo com a natureza dos dados e objetivos.
- 3) O processo de KDD (Knowledge Discovery in Databases) é uma abordagem estruturada para extrair conhecimento de conjuntos de dados. Ele engloba etapas como seleção de dados relevantes, pré-processamento para limpeza e organização, transformação dos dados para melhor análise, aplicação de algoritmos de mineração para identificar padrões, avaliação e interpretação dos resultados, visualização para comunicação eficaz e, se apropriado, implementação das descobertas. Esse processo é iterativo e visa transformar dados em insights acionáveis, melhorando a tomada de decisões e compreensão.

Questão 2

Questão 2

①

$$E = \sum -p_i \log_2 p_i$$

CONC

$$E = -\frac{1}{2} \cdot \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} = 1$$

TEMPO

0-10 - 6
10-30 - 2
30-60 - 2
>60 - 2

$$E\left(\frac{1}{2}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{6} \log_2 \frac{1}{6} - \frac{1}{6} \log_2 \frac{1}{6} - \frac{1}{6} \log_2 \frac{1}{6} = 1,103$$

TIPO

FRANCES - 2
TAILANDES - 4
HAMBURGER - 4
ITALIANO - 2

RES

SIM - 5
NAO - 2

$$E\left(\frac{5}{12}, \frac{7}{12}\right) = -\frac{5}{12} \log_2 \frac{5}{12} - \frac{7}{12} \log_2 \frac{7}{12} = 0,1294$$

$$E\left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right) = -\frac{1}{6} \log_2 \frac{1}{6} - \frac{1}{3} \log_2 \frac{1}{3} - \frac{1}{3} \log_2 \frac{1}{3} - \frac{1}{6} \log_2 \frac{1}{6} = 0,861$$

CHUVA

SIM = 5
NAO = 7 = 0,1294

FOME

SIM = 7
NAO = 5 = 0,1294

BEIXASAB

SIM = 9
NAO = 5 = 0,1294

BAR

SIM = 6
NAO = 6 = 1

ALTERNATIVO

SIM = 6
NAO = 6 = 1

PRECO

R = 7
RA = 2
RAR = 3

$$E\left(\frac{7}{12}, \frac{2}{12}, \frac{3}{12}\right) = -\frac{7}{12} \log_2 \frac{7}{12} - \frac{2}{12} \log_2 \frac{2}{12} - \frac{3}{12} \log_2 \frac{3}{12} =$$

1,138

CLIENTE

VENHUM - 2
ALGUNS - 4
CHEIO - 6

$$E\left(\frac{2}{12}, \frac{4}{12}, \frac{6}{12}\right) = 1,145$$

RAIZ

2

COM:

$$\text{ALTERNATIVO} = E\left(\frac{5}{6}, \frac{1}{6}\right) = 0.63$$

$$\text{BAR} = E\left(\frac{3}{6}, \frac{3}{6}\right) = 1$$

$$\text{SEX/BAR} = E\left(\frac{5}{6}, \frac{1}{6}\right) = 0.63$$

$$\text{FONE} = E\left(\frac{4}{6}, \frac{2}{6}\right) = 0.91 \quad \leftarrow \text{SEGUNDA 10 WNEC}$$

$$\text{PREÇO} = E\left(\frac{4}{6}, \frac{2}{6}\right) = 0.91$$

$$\text{CHUVA} = E\left(\frac{2}{6}, \frac{4}{6}\right) = 0.91$$

$$\text{RES} = E\left(\frac{2}{6}, \frac{4}{6}\right) = 0.91$$

$$\text{TIPO} = E\left(\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{2}{6}\right) = 1.18$$

$$\text{EMPO} = E\left(\frac{2}{6}, \frac{2}{6}, \frac{2}{6}\right) = 1.58$$

A

