

1ª Avaliação

Questão 1: (0,5 pt)

Não é um exemplo típico de aplicação de Mineração de Dados:

- a) Prever se clientes de uma empresa de telefonia podem migrar para o concorrente.
- ☒ b) Checar se o cálculo da folha de pagamento está correto. *e*
- c) Detectar tentativa de invasão de rede.
- d) Tentar reduzir a evasão escolar.

Questão 2: (0,5 pt)

Analisamos dados coletados de algas marinhas desconhecidas, buscando categorizá-las segundo atributos em comum. Que tipo de tarefa de mineração de dados é esta?

- a) Classificação.
- ☒ b) Agrupamento. *e*
- c) Regressão.
- d) Outros.

Questão 3: (1 pt)

Sobre o caso apresentado na pergunta anterior, podemos dizer que este tipo de tarefa de mineração de dados é:

- a) Supervisionada.
- ☒ b) Não supervisionada. *e*

Questão 4: (1 pt)

Prever a altura de bebês a partir de sua altura aos dois anos de idade. Que tipo de tarefa de mineração de dados é esta?

- a) Classificação.
- b) Agrupamento.
- c) Regras de Associação.
- ☒ d) Regressão. *e*

Questão 5: (1 pt)

Seu cliente quer potencializar as vendas em seu e-commerce e lhe contrata para tentar ofertar novos produtos aos clientes à medida que eles comprem. Que tipo de tarefa de mineração de dados é esta?

- a) Classificação.
- b) Agrupamento.
- ☒ c) Regras de Associação. *e*
- d) Regressão.

Questão 6: (1 pt)

O analista dividiu dados em dois grupos de forma aleatória, usou o primeiro grupo para que o algoritmo conhecesse os dados, e o segundo grupo para ver como o algoritmo se sai prevendo a fraude. Que técnica de avaliação foi utilizada?

- ☒ a) Hold out.
- ☐ b) Matriz de Confusão.
- ☐ c) Validação Cruzada.
- ☐ d) Cálculo de lembrança.

Questão 7: (1 pt)

O modelo foi construído e teve uma performance excelente no laboratório, mas sua performance não foi tão boa quando colocado em produção. O que provavelmente ocorreu?

- ☒ a) Superajuste.
- ☐ b) Generalização.
- ☐ c) Matriz de confusão calculada errada.
- ☐ d) Lembrança muito baixa.

Questão 8: (1 pt)

Um modelo criado em cima de uma relação com 60 atributos não tem um bom desempenho. O analista volta a base de dados original e coleta mais 30 atributos, porém o desempenho do modelo não melhora. Qual é a provável causa da falha em se criar um bom modelo?

- ☐ a) Superajuste.
- ☐ b) Generalização.
- ☐ c) Maldição da dimensionalidade.
- ☒ d) Classe Rara.

Questão 9: (1 pt)

Sobre o ambiente para mineração de dados Weka, pergunta-se: Quais seus quatro aplicativos e qual a função de cada um?

Questão 10: (1 pt)

Caracterize de forma sucinta: 1- Mineração de Dados, 2-Aprendizagem supervisionada, 3-Aprendizagem não-supervisionada.

Questão 11: (1 pt)

De forma sucinta, faça a distinção entre os classificadores dos tipos: Árvore de decisão, Naive Bayes e KNN. (max 20 linhas)