



Universidade Federal de Santa Maria – Centro de Tecnologia

Peso: 4

Disciplina: Mineração de Dados (ELC1098)

Professor: Dr. Joaquim Assunção

Atenção: Leia atentamente as questões e escreva de forma clara e legível.

Nome: _____

Curso: _____ Data: ____/____/____

- 1) Na limpeza de um conjunto de dados ruidoso e cheio de valores faltantes. Cite 3 abordagens para complementar os valores faltantes (1.5). Para cada uma escreva um possível problema de usar essa abordagem. (1.5)

- 2) Indique a alternativa que contém uma etapa que não é associada ao processo de KDD (2).
- a) Seleção e transformação.
 - b) Pré-Processamento.
 - c) Mineração de dados.
 - d) Limpeza e integração.
 - e) Previsão (*forecasting*).
 - f) Interpretação (*evaluation*).

- 3) Brevemente, descreva a relação entre KDD, mineração de dados, e análise de dados. (2)

4) Analise o seguinte conjunto e responda:

4.1) Caso nosso projeto de mineração tenha como objetivo encontrar correlação entre produtos e lucro. Quais informações podemos descartar? (1)

4.2) Considerando esse objetivo (4.1), devo alterar (adicionar ou editar) algo nesse conjunto? Por que? (1)

ID	Produto	Descricao	Valor compra	Valor venda	Unidades vendidas
1	Monitor LG 29p	Monitor LG 29p Entradas 1,2,3. 29p. Full HD IPS 29UM68	250	290	29
...	...		...	...	...
999	Monitor Sansung 30p	Monitor Sansung 30p Entr. 1,2,3,4. 30p. 4K	500	600	45

5) Em geral, dado um *dataset* X, qual a diferença entre a extração de um subconjunto, e a redução de dimensionalidade de X? (1)
