Trabalho 1 – Aprendizado Automático

A partir da base de dados enviada e do estudo sobre pré-processamentos, apresentar os resultados obtidos, utilizando o modelo Decision Tree (DT), e discutir seus resultados dentro dos processos para Data Mining.

1) A partir da base de dados:

a. Descreva a base de dados descrendo os atributos, numéricos e categóricos, classificando segundo a escala (nominal ou razão) e a cardinalidade (discreta,

contínua, binária).

b. Descreva cada um dos atributos segundo frequência, mínimo e máximo valor,

dia desvios padrão, conforme o caso.

c. Avalie os resultados dos processos abaixo, caso sejam utilizados na base de

dados, após o processo de classificação com DT ter sido utilizado.

i. limpeza de dados (outlier, missing)

ii. normalização/transformação

iii. discretização

2) Faça uma análise de seleção de variáveis utilizando os métodos que julgar necessários, considerando o tipo de dado. Abaixo uma tabela é indicada para que possa usar como inspiração avaliação de importância composta dos métodos que considerar.

Dica: utilize uma tabela métodos x ranking com atributos selecionados

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Método de Seleção 1 | Método de Seleção 1 | Método de Seleção 1 | Método de Seleção 1 |  | | Atributo 1 |  |  |  |  |  | | Atributo 2 |  |  |  |  |  | | ... |  |  |  |  |  | | Atributo n |  |  |  |  |  | | | | |
|  | | | |
|  | | | |
|  | | | |