

Universidade Federal da Grande Dourados
Faculdade de Ciências Exatas e Tecnologia
Bacharelado em Engenharia de Computação – Tópicos Avançados em Aprendizagem de Máquina

Atividade em grupo (até 4 pessoas) – apenas 1 aluno fará o upload no Moodle com o nome de todos os envolvidos.

1. Considere a base de dados Funcionarios.csv. Faça o pré-processamento desta base, considerando as tarefas:
 - a. Redução de Dimensionalidade (vertical e horizontal, se necessário)
 - b. Limpeza dos Dados (correção de inconsistências, duplicidade de registros, dados incompletos ou ausentes)
 - c. Transformação (codificação categórica para numérica, normalização)

Entregar:

- Com os dados ainda não normalizados (e já atendidas os Itens a e b), no próprio .csv mostre o resultado para alguma análise estatística descritiva para descrever esta base de dados, tal como a média e a moda.

2. Usando alguma linguagem de programação, automatize o processo de normalização dos dados.

Entregar:

- No mesmo .csv, a base de dados pré-processada (atendidos os Itens a,b e c do Exercício 1)

3. No mesmo .csv, entregar:

- Descrição das ações que foram realizadas. Por exemplo:
 - Algum atributo foi desconsiderado? Qual e por quê?
 - Houve dado ausente? Como resolveu esse problema?
 - Qual técnica de normalização foi utilizada?
- Qual tarefa de aprendizagem de máquina poderia ser aplicada à esta base de dados e para quê?