

MiniTeste-3

(para avaliação individual)

Para ser submetido até as 23h59 do dia 6/1/2021.

1. Considere o dataset “unbalanced”. Compare o desempenho dos algoritmos k-means e EM quando o número de partições é determinado pelo algoritmo *EM*. Considerar o número de partições (clusters) obtidas. Deve “desligar” o atributo classe para tornar mais real a avaliação. Compare estes resultados com os obtidos pelo algoritmo *DBSCAN*.
2. Considere o dataset de disciplinas (student_courses.bas, descarregar do blackboard). Apresente para este dataset exemplos de regras redundantes, produtivas e não produtivas e regras significativas. Comente os exemplos obtidos/usados
3. Considere o dataset adult (descarregar do blackboard). Usando o sistema CAREN e os seus vários métodos para Subgroup Mining apresente e comente regras que denunciam discriminação de género e idade em termos de rendimentos anuais. Notar: o atributo classe deste dataset Adult caracteriza os rendimentos anuais dos indivíduos registados i.e. $\leq 50K$ indica rendimento anual menor que 50 000 USD (e o seu oposto).