



Curso: Mestrado Integrado em Informática – Engenharia do Conhecimento

U.C.: Descoberta de Conhecimento

Folha de Exercícios FE07	
Docente	Hugo Peixoto
Tema:	Árvores de Decisão
Turma:	PL
Ano Letivo:	2018-19 – 1º Semestre
Duração da aula:	2 horas

1. Decision Trees

- 1) Conduzir um estudo para encontrar os dados dos passageiros do Titanic. Pesquisar por “Titanic passenger list”.
- 2) Abrir o Excel e seleccionar um conjunto robusto de dados. Deve conter atributos como, sexo, idade, classe, e se possível país de origem. O atributo ID do passageiro e nome não devem ser usados para o processo de decisão.
- 3) Adicionar ao dataset uma coluna que identifique:
 - a. sobrevivente
 - b. falecido
 - c. desconhecidoDenomine a coluna de “Survival_Result”.
- 4) Importar para o RapidMiner não esquecer de colocar o operador Set Role nos atributos que forem desnecessários.
- 5) Duplicar o dataset no excel, mas retirar a coluna “Survival_Result”.
- 6) Adicionar pessoas aleatórias ao data set como os colegas ao lado. Decidindo, por exemplo, em que classe embarcaram.
- 7) Correr o modelo usando diferentes configurações nas características da do algoritmo de decisão.
- 8) Observar os resultados e documentar as evidências.