

**Curso:** Mestrado em Engenharia Biomédica  
**U.C.:** Aprendizagem e Extração do Conhecimento

Ficha de Exercícios 05	
Docente:	Hugo Peixoto   António Abelha
Tema:	RapidMiner – Classificação
Ano Letivo:	2022-2023 – 1º Semestre
Duração da aula:	2 horas

## 1. Regressão Linear

Importe o dataset heart-failure.csv e prepare o processo de Data Mining da forma que considerar mais correta.

- Crie um subprocesso, denominado “DataPrep”. Neste subprocesso deverá incluir todos os passos de Preparação dos Dados para o aplicar o modelo.

[1] Justifique a utilização de todos os operadores de preparação de dados que incluiu no seu modelo.

[2] Inclua os operadores necessários para gerar o modelo e a respetiva avaliação. Utilize 3 dos algoritmos experimentados na aula e utilizando as seguintes medidas de avaliação, Accuracy, Precision, Recall e Specificity, documente os resultados obtidos.

$$\text{Accuracy} = (TP+TN)/(TP+FP+FN+TN)$$

$$\text{Precision} = TP/(TP+FP)$$

$$\text{Recall} = TP/(TP+FN)$$

$$\text{Specificity} = TN / (TN + FP)$$

[3] Haverá alguma intervenção no modelo que possa melhorar os resultados obtidos. Avalie os novos modelos, recorrendo às métricas acima apresentadas para validar as alterações abaixo propostas:

- Experimente perceber, por exemplo através das árvores de decisão, quais os atributos mais importantes. No operador “*Select Attributes*” faça refletir essa importância.
- Experimente Normalizar o atributo *platelets*, recorrendo ao operador “*Normalize*” – “*range transformation*”;