



Universidade do Minho

Departamento de Informática

Mestrado [Integrado] em Engenharia Informática

Perfil de Sistemas Inteligentes

Aprendizagem e Extração de Conhecimento

4º/1º Ano, 1º Semestre

Ano letivo 2018/2019

Ficha Prática nº 2

5 de novembro de 2018

**Tema** Prediction Model Selection

**Enunciado** Nesta ficha pretende-se que o aluno adquira conhecimentos associados à avaliação e selecção de modelos de previsão de machine learning em linguagem Python através do uso das funcionalidades existentes na biblioteca sklearn.

**Tarefas** Utilize a resolução da Ficha Prática nº 1 para a resolução desta Ficha Prática.

As tarefas a completar são as seguintes:

Ex. 1) Através das funcionalidades disponíveis da biblioteca sklearn, averigue a performance dos quatro modelos de previsão treinados (Logistic Regression, Gaussian Naive Bayes, K-Nearest Neighbors, Support Vector Machines) utilizando os seguintes métodos:

1. Train/Test Split (80% -20%)
2. 10-Fold Cross Validation

Ex. 2) De acordo com os resultados do ex. 1, seleccione o modelo que apresenta melhor performance de previsão, justificando a sua resposta.

Ex. 3) Na literatura científica, uma das técnicas mais utilizadas para analisar a performance de modelos de previsão tem sido o K-Fold Cross Validation, apresentando várias variantes. Investigue o conjunto de variantes existentes na literatura, listando o conjunto de técnicas e definindo vantagens e desvantagens de cada uma.