



MEIC – ALAMEDA

2017/18

Sistemas de Apoio à Decisão

Lab 4 – KNN

Goals:

- KNN
 - Accuracy and Confusion Matrix
 - Training strategies
1. Load the iris data. Train a classifier using the KNN algorithm. Keep all parameters with their default values Use percentage split with 70%. Keep class as the class attribute.
 - a. What is the accuracy achieved?
 - b. And the number FP and FN for Iris-virginica?
 - c. And for Iris-setosa?
 - d. Compare the results achieved through Cross-validation with 10 folds.
 2. Load the glass data. Train a classifier using the KNN algorithm. Keep all parameters with their default values, but the number of neighbors (KNN). Keep type as the class attribute
 - a. What is the accuracy with 1 neighbor?
 - b. And with 5, 10, 15, 50 and 100 neighbors?
 - c. How does the accuracy change?
 - d. Is any of the models in overfitting?

R packages

- caret
- e1071

Technique	Weka	R
kNN	<u>weka.classifiers.lazy.IBk</u>	<code>caret.train(method='knn')</code> <code>caret.knn3</code>