k-Nearest-Neighbours

Lab Session 1
Machine Learning: Pattern Recognition
Master Artificial Intelligence

Camiel Verschoor	Steven Laan
StudentID: 10017321	StudentID: 6036031
UvAnetID: 6229298	UvAnetID: 6036031
Verschoor@uva.nl	S.Laan@uva.nl

September 11, 2012

1 Data Visualization

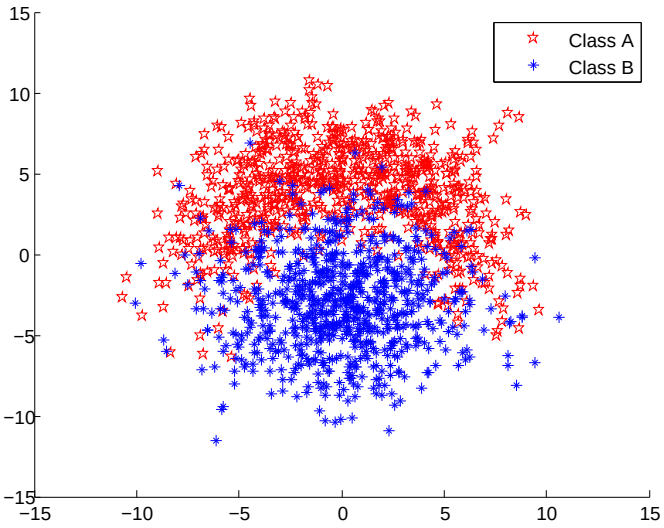

Figure 1: Data visualization.

In figure 1 the data in the training set, which consists of the two classes, A (red) and B (blue).

2 k-Nearest Neighbours

A kNN classifier, with $k = 1$, is trained on the trainings data. The performance is evaluated on the test data. This resulted in the following confusion matrix:

	True	**False**
Positive	206	44
Negative	36	214

From the confusion matrix the error rate is computed by the following formula:

$$1 - accuracy = 1 - \frac{tp + fp}{tp + tn + fp + fn} \tag{1}$$

The error rate and other statistical measures of the classifier are presented in the table below:

Accuracy	84.0%
Precision	82.4%
Recall	85.1%
F-measure	83.7%
Error rate	16.0%

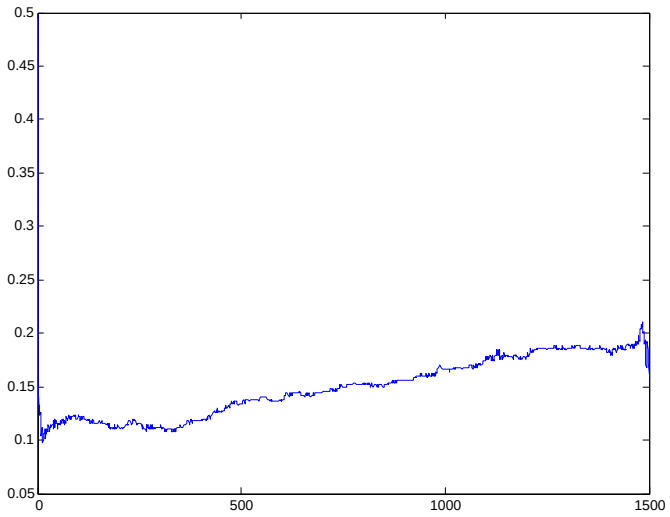

Figure 2: Graph of the error rate with various K

3 Cross Validation