Information Mining - winter semester 2019

Exercise sheet 11

Exercise 1: Combination of multiple models

(a) In the lecture methods to combine multiple models are shown. Explain briefly the underlying idea behind the following principles: (i) Bagging, (ii) Boosting and (iii) stacking.

(b) Use *RapidMiner* to classify the example data[1] with the stacking-method. Evaluate the results with a 10-fold cross-validation.

Use `Naive Bayes` and `k-NN` as learning methods for the models. As the learning method for the stacking use `Decision Tree`.

Exercise 2: k-nn in RapidMiner

In the lecture we have seen that k-nn can be used for instance based classification. An instance is classified by its k neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor. In this exercise use the weather data (knnWeatherDataTraining.arff[2]) and apply k-nn clustering to classify the following instance. Note, you should consider normalisation before applying k-nn (page 131-132 in chapter 4).

outlook=sunny, temperature=60, humidity=76, windy=TRUE, play=?

[1] `http://www.is.inf.uni-due.de/courses/im_ws19/uebung/data_a23.csv`

[2] `http://www.is.inf.uni-due.de/courses/im_ws19/uebung/knnWeatherDataTraining.arff`