

Knowledge Discovery in Databases II  
WS 2015/2016

Übungsblatt 10: Bagging and Boosting

**Aufgabe 10-1**      **Cohen's Kappa**

Gegeben seien die folgenden Konfusionsmatrizen zu den Zeitpunkten  $t = 1, 2, 3$ :

	$t = 1$			$t = 2$			$t = 3$	
	positiv	negativ		positiv	negativ		positiv	negativ
positiv	37	14	positiv	65	8	positiv	90	4
negativ	17	32	negativ	7	20	negativ	5	1

Berechnen Sie Accuracy und Cohen's Kappa, und vergleichen Sie die Ergebnisse.

**Aufgabe 10-2**      **Bagging and Boosting**

Bagging and Boosting are two of the most common ensemble techniques. In this exercise, we will implement two classifiers applying bagging and boosting using decision trees as base learner. The classifier and the base learner should use the sklearn framework.

- (a) Download and inspect the code template `python_ensemble.py`.
- (b) Implement a class `weighted_bootstrap` for generating a bootstrap given a certain distribution over the sample set.
- (c) Implement a bagging classifier using the `Bootstrap` class and `DecisionTree` as base learner.
- (d) Implement a boosting classifier using the `Bootstrap` class and `DecisionTree` as base learner.
- (e) Test both classifiers and the base classifier on the `ionsphere` data set.