Exercise 1: Performance Measures

(a) Given the following confusion matrices

$$M_1 = \begin{pmatrix} 0 & 10 \\ 0 & 990 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 10 & 0 \\ 10 & 980 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 10 & 10 \\ 0 & 980 \end{pmatrix},$$

each of which corresponds to a classifier. Compute the accuracy, F_1 score, G measure/mean, BAC and MCC of each classifier.

(b) What are the population counterparts of the class-specific variants in the multiclass setting of true positive rate, positive predictive value and true negative rate?

Exercise 2: PR-curve

Draw the PR-curve for the following data set:

Truth	**Score**
Pos	0.95
Pos	0.86
Pos	0.69
Neg	0.65
Pos	0.59
Neg	0.52
Pos	0.51
Neg	0.39
Neg	0.28
Neg	0.18
Pos	0.15
Neg	0.06

Exercise 3: MetaCost

Implement the MetaCost algorithm and use it with some classifier of your choice on an imbalanced data set of your choice, where the cost-matrix is given by the cost-sensitive heuristic we saw in the lecture. Compare the confusion matrices of the underlying classifier and the MetaCost classifier as well as their total costs.