

MTAT.03.227 Machine Learning

Practice session 10

November 25-27, 2019

Exercise 1. Boosting of decision stumps

Answer the questions below for the dataset D consisting of the following two-dimensional data points and binary labels:

$$\begin{array}{lll} x_1 = (1, 2), & x_2 = (0, 0), & x_3 = (-1, -1), \\ y_1 = +1, & y_2 = -1, & y_3 = +1 \end{array}$$

A decision tree with a single decision node is known as the *decision stump*. Let us define a learning algorithm \mathcal{A} as a decision stump learner which selects its decision node as the one with minimal error. If several such exist then it chooses the one with the most positive predictions.

(a) Apply the AdaBoosting algorithm on the above data D with $T = 3$ iterations and the decision stump learner \mathcal{A} . For this determine all the obtained instance weights w_{ti} and model weights α_t .

(b) Determine the final output score $H(x)$ of each of the instances. Are all instances classified correctly when thresholding $H(x)$ at 0? If not, how many iterations T does it take to reach correct classifications?

Exercise 2. Bagging of decision stumps

Consider the bagging algorithm applied on the above data D , ensemble size $T = 3$, and the decision stump learner \mathcal{A} as defined in Exercise 1.

(a) List all possible prediction vectors $(\hat{y}_1, \hat{y}_2, \hat{y}_3)$ that individual trees in the bagging ensemble might provide as predictions on the training data.

(b) *Note that this subtask involves combinatorics that we don't assume you know for Test 3.* What is the probability that the majority vote of the obtained ensemble gives correct predictions to all instances? Hint: for this first calculate the probability of each of the outcomes in subtask (a).