

DAIS2021: Assignment #3

Decision Trees, Perceptron

Kyra Ahrens, Tom Weber, Lennart Clasmeier
{ahrens,tomweber}@informatik.uni-hamburg.de,
lennart.clasmeier@studium.uni-hamburg.de

University of Hamburg — May 2, 2021
Due: 16.05.2021 8 p.m.

Introduction

In this assignment, we want to get familiar with decision trees and perceptrons. We will have a look at how decision trees are created and how learning in a perceptron works.

The assignment has 100 points in total, distributed over the tasks. You need to get at least 50 points to pass this assignment.

1 Decision Trees I: 10 Points

In order to familiarise ourselves with the application of decision trees to simple problems, hands-on experience with calculating Gini indices is the core of this task. It is modelled as part of a moodle quiz. Please solve questions 1-10 of the quiz for this task. You will find it on the moodle page of the course or by clicking this link:

<https://lernen.min.uni-hamburg.de/mod/quiz/view.php?id=55838>.

2 Decision Trees II: 40 points

Next, we will have a look at creating decision trees for more complex problems via the Python library *scikit-learn* and a different criterion, entropy. Please do all the tasks specified (cf. section 2.3) in the attached Jupyter notebook *DecisionTrees.ipynb*.

3 Perceptron I: 10 points

Now, we want to gain some insight in neural processing algorithms. The perceptron is a simple one-layer neural network which can learn different functions. We are interested in learning Boolean functions. The task is modelled as part of a moodle quiz. Please solve questions 11-20 of the quiz for this task. You will find it on the moodle page of the course or by clicking this link:

<https://lernen.min.uni-hamburg.de/mod/quiz/view.php?id=55838>.

4 Perceptron II: 40 points

In this task, we will have a look at how we can learn Boolean functions algorithmically. Please do all the tasks specified in the attached Jupyter notebook *Perceptron.ipynb*.

5 Next Assignment

Now that you got some understanding of decision trees and perceptrons, we are going more practical in the next tutorial, i.e. we will run numerous experiments with popular datasets. Critical to all experiments is a thorough evaluation and comparison of different methods and hyperparameters. Therefore, refresh your mind on the following topics:

- Classifier training:
 - Data normalization
 - Sampling train and test data, cross-validation
- Classifier evaluation:
 - Training vs. test error
 - Evaluation metrics, confusion matrix