Assignment 3

Lecturer: Prof. Moshe Sipper, TA: Shachar Schnapp

Spring 2022

Submission guidelines. **Please read and follow carefully**:

- Submit the assignment in groups of 2 or 3.
- Submit via Moodle.
- The submission should include two separate files:
  1. A Jupyter notebook file that includes all of the experiments you've done,
  2. A Python file that includes your final model.
- For questions, use the exercise forum, or if they are not of public interest, send them via the course requests system.

Kaggle competition

In this excise we conduct a Kaggle-style competition. In such a competition we take one dataset and divide it into three parts: `train`, `dev` and `test`. You get two datasets: `train` and `dev`. Given these, your goal is to produce the most accurate regressor, use **mean absolute error** as a accuracy function. After that, you submit your model and we'll train it using the `train` dataset (same train dataset that you have), and test it on the `test` that is hidden from you. The group that achieves the best accuracy wins— the top three groups will receive a bonus for the assignment:

- First: 30 points
- Second: 20 points
- Third: 10 points

Part 1. Your answers for the this part should be included in Jupyter file: `experiment.ipynb`. You need to justify by experiments each of the steps that lead to your model, specifically focusing on:

- Show the preliminary data analysis that you did.
- Show all the different models that you tried, and explain how you chose the best one.
- Show the preprocessing that you ran on the dataset and how it affects the model's accuracy.

- Describe the hyperparameter search and how each hyperparameter affects the model's accuracy.

Part 2. Your answers for this part should be included in Python file: `model.py`.

- Use our model template `model.py` (in the included ass3.zip), the template includes two functions: `fit` and `predict`. For your convenience we created a naïve implementation. Change the template in accordance with your own model, based on experiments that you did in the previous part.
- Don't forget to set the model with the base hyperparameters, based on the experiment that you did in the previous part.
- You're allowed to add any function that you want (including a preprocessing function), but you can't add any files (you need to submit only the Python file `model.py`).
- Before you submit the code, run the test file and make sure that it runs properly (also included in ass3.zip).

Good luck!