

## Homework 3

*Hand Out: April.29**Due: May.18*

- Feel free to talk to other students in the class when doing the homework. You should, however, write down your solution yourself. You also must indicate on each homework with whom you collaborated and cite any other sources you use including Internet sites.
- You will write your solution in LaTeX and submit the **pdf file through Gradescope**. You also need to submit the **zipped LaTeX files to CCLE**. We will grade your homework based on the final version of the pdf file submitted to Gradescope. We will not grade the zipped Latex files on CCLE. However, failure to submitting your LaTeX files to CCLE will incur 2 points penalty out of 100 points.
- The homework (both pdf and zipped Latex source files) is due at **1:59 PM before the class**.

The following questions are from *Understanding Machine Learning: From Theory to Algorithms* by Shai Shalev-Shwartz and Shai Ben-David. It can be found here <http://www.cs.huji.ac.il/shais/UnderstandingMachineLearning/> by courtesy of the authors.

1. (10 Points) Exercise **9.1**
2. (10 Points) Exercise **9.3**
3. (10 Points) Exercise **9.5**
4. (10 Points) Exercise **18.1**
5. (10 Points) Exercise **18.2**
6. (10 Points) Exercise **19.1**

In the following two problems, you are asked to implement Perceptron algorithm and Linear Regression algorithm in Python. To set up the environment, install the following software & packages:

- python (<https://www.python.org/downloads/>)
- numpy (<http://www.numpy.org/>)
- scipy (<http://www.scipy.org/>)
- matplotlib (<http://matplotlib.org/>)
- scikit-learn (<http://scikit-learn.org/stable/>)

Note that, you are allowed to use other programming languages for your implementation. If so, you may need to create an csv data loader yourself and read the data from `./data/*.csv`. Make sure to submit your source code to CCLE. Your answer will NOT be graded if we didn't see your submission.

7. (15 Points) Implement the Perceptron algorithm on Iris Dataset for the task of binary classification. The skeleton code is given in `iris.py`.

- Run the skeleton code, report the testing error obtained by the Perceptron model implemented in scikit-learn.
  - Replace skeleton code line 45-47 with your implementation.
  - Report your error on the testing set. Report the number of iterations it took the algorithm to converge on the training set.
8. (15 Points) Implement the linear regression algorithm on Boston House Dataset for the task of regression. The skeleton code is given in `boston.py`.
- Run the skeleton code, report the mean squared testing error obtained by the LinearRegression model implemented in scikit-learn.
  - Replace skeleton code line 47-49 with your implementation.
  - Report your mean squared error on the testing set.