

LAB 14

Due: Friday 12/08/2023 @ 11:59pm EST

The purpose of labs is to give you some hands on experience programming the things we've talked about in lecture. This lab will focus on learning (i.e. "fitting") a Linear Regression model

Task 0: Setup

Included with this file is a new file called `requirements.txt`. This file is rather special in Python: it is the convention used to communicate dependencies that your code depends on. For instance, if you open this file, you will see entries for `numpy` and `scikit-learn`: two python packages we will need in order to run the code in this lab. You can download and install these python packages through `pip`, which is python's package manager. While there are many ways of invoking `pip`, I like to do the following:

```
python3 -m pip install -r requirements.txt
```

(I let `python3` figure out which `pip` is attached to it rather than the more common usage: `pip install -r requirements.txt`)

Task 1: class LinearRegression (100 points)

In the file `linearRegression.py`, you will find a class called `LinearRegression`. Remember, linear regression learns a *hyperplane* (i.e. a line in higher dimensions) using the equation:

$$\vec{\theta}^* = [\mathbf{X}, \vec{1}]^\dagger \vec{y}_{gt}$$

where \mathbf{A}^\dagger is the *pseudoinverse* of matrix \mathbf{A} , \vec{y}_{gt} is the vector of ground truth values, and $[\mathbf{X}, \vec{1}]$ is concatenating a column of 1s to feature matrix \mathbf{X} .

When it is time to make predictions, we use the following equation:

$$\hat{\vec{y}} = [\mathbf{X}, \vec{1}] \vec{\theta}^*$$

. Your job is complete the following methods:

- the `fit` method: This method is where you are handed training data \mathbf{X} and \vec{y}_{gt} and need to compute $\vec{\theta}^*$
- the `predict` method: This method is where you are handed feature matrix \mathbf{X} and are asked to make predictions on that data using the parameters $\vec{\theta}^*$ that you computed from the `fit`.

Feel free to run the `LinearRegression.py` file, there is some basic testing at the bottom of the file that will be run if you run the file. I would recommend adding some of your own testing (what shape is what, etc) to increase your confidence in the correctness of your solution before you submit to the autograder. The autograder will take a few minutes to run!

Task 2: Submit Your Lab

To complete your lab, please **only turn in the `LinearRegression.py` file** on Gradescope. You shouldn't have to worry about zipping it up or anything, just drag and drop it in.