Lab 8

Due: Friday 11/21/2025 @ 11:59pm EST

The purpose of labs is to give you some hands on experience programming the things we've talked about in lecture (and some we will talk about in lecture). The purpose of this lab is for you to learn how to program some basic supervised-learning models.

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 pythn 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`. 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 *pseudo-inverse* of $\mathbf{A}$, $\vec{y}_{gt}$ 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 (on feature matrix $\mathbf{X}$), we use the following equation:

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

Your job is to complete the following methods:

- `LinearRegression.fit`. This method is where you are handed training data $\mathbf{X}$ and $\vec{y}_{gt}$ and you need to compute $\vec{\theta}^*$

- `LinearRegression.predict`. This method is where you are handed feature matrix $\mathbf{X}$ (it may not be the same feature matrix as the training data), and are asked to produce predictions on that data using $\vec{\theta}^*$ that you computed from a prior call of the `fit` method.

Feel free to run the `LinearRegression.py` file, there is some basic testing at the bottom of the file that will be executed. I would recommend adding some testing of your own (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

Please drag and drop your complete `LinearRegression.py` file on gradescope!