This assignment does not count toward the final grade.

Project 0

Due No Due Date **Points** 0

Note: This is project that aims to provide a general idea about how a project looks like throughout this course and will only be a PASS/FAIL project.

We have collected a dataset of a person grasping various objects. For each object, the person has performed multiple grasps. For each grasp, we have recorded joint angles (via a Cyberglove) and forearm EMG data (via a Myo armband).

In this project, we will only look at joint angle data. Using clustering, you must decide if the joint angles in this dataset cluster together, and if these clusters indeed correspond to the various objects being grasped.

Step 1

Check out Project 0 from the SVN server:

```
svn co svn+ssh://YOUR_UNI_HERE@roamopenserver.me.columbia.edu/home/YOUR_UNI_HERE/assignments/project0
```

Install the virtual environment for the project:

```
cd project0/grasp_clustering
pipenv install --dev
```

Step 2

You must fill in a python file called *cluster_grasps.py*. Your scripts should work as follows:

- It implements a class called GraspClustering that at least contains two methods: train(self) and predict(self, test_data). More details of these method will be described in the following bullet points.
- train(self) takes no inputs and returns nothing. It should load the training data, run a clustering algorithm on the training data, and store the fitted model somewhere where you can access it later from the predict(...) function. Most likely, you will store this model using member variables of the GraspClustering class.
- predict(self, test_data) takes an argument test_data as input and should return the labels of these data. The test_data will be in shape (batch_size, data_dimension) and the returned value should be in shape (batch_size, 1).

The data file for training, *object_grasping_30sec_no_labels.csv*, is located in "data" directory inside the package. We have provided a standardized way to load these data into memory for training. To load the data, simply run the following scripts:

```
import os

from load_data import load_data
```

```
training_data_path = os.environ['TRAIN_DATA_PATH']  # retrieve environmental variable from .env file
training_data = load_data(training_data_path)  # load the data as a numpy array format
```

Step 3 - Autograding

Once you believe your script works as expected, run the autograder. All the grading scripts will be located in the grading directory. To run grading, please run this command under the grasp_clustering folder:

```
pipenv run python grading/grade_assignment.py
```

If you see PASS - congratulations! you have passed this teaser. You are ready for the real project, when the time comes.

If you see FAIL - the performance of your clustering algorithm should be improved.

What the Autograder does

Our Autograder does the following:

- Import the class GraspClustering that is implemented by you.
- Run the train method with the data that is loaded with the provided data loader.
- Finally, it predicts the labels of test data and compare them with ground truth to calculate a score. If the score is > 0.72, you will pass.

Submitting

To submit, commit your assignment to SVN. **Make sure to double-check that the version you submitted is indeed the one that works.** If you inadvertently introduce a typo or some other problem that makes your submission crash, your score might be severely affected.

There is no partial credit for each item. Note you are NOT allowed to use Neural Network for this assignment. You will get a chance to explore neural network methods in the next project.

The deadline for this project is February 14th. Be sure to review the **grading and collaboration policy** for the course.

Good luck!

Data files

We are providing you with two data files for this project:

- *object_grasping_30sec_no_labels.csv* contains training data. Note that is has no labels. It is formatted as follows:
  - The first row contains column headers. Each following row represents a grasp
  - The first 8 columns contain EMG data (which we are ignoring for now).
  - The next 15 columns contain joint angle (glove) data.
- object_grasping_10sec.csv contains test data. This file is not used directly by your code (only by our Autograder) so its format is not described here.