1. Each row is an example with 201 attributes. The first 200 attributes are features, and the last one is the label.
2. Originally, the dataset includes 16 classes, and I removed 4 classes with less than 200 labeled examples. So we have 12 classes for our classification.
3. Both the “.txt” and “.npy” files for the data are uploaded
4. Train file is used for training including the cross validation, and test file is used for testing the accuracy
5. For each class, I randomly selected 100 examples for training and 100 examples for testing. So in both the training and testing data files, we have 100\*12 examples.