GE461 Introduction to Data Science

Project 4 Report

İrem Ecem Yelkanat
21702624
Section 1

Part A

The number of features of the input data that is given for the assignment, which is 306, is considerably large, therefore, it is hard to visualize the data. For this reason, principal component analysis (PCA) is applied on the given input data to reduce the feature dimension of the given input to 2 dimensions. Top two principal components are extracted by PCA, which can capture total of 83.83% of variance, where the first principal component can capture 75.31% of variance and the second principal component can capture 8.5% of variance. The projection of the given input to top two principal components can be seen in *Figure 1, Figure 2* and *Figure 3*.

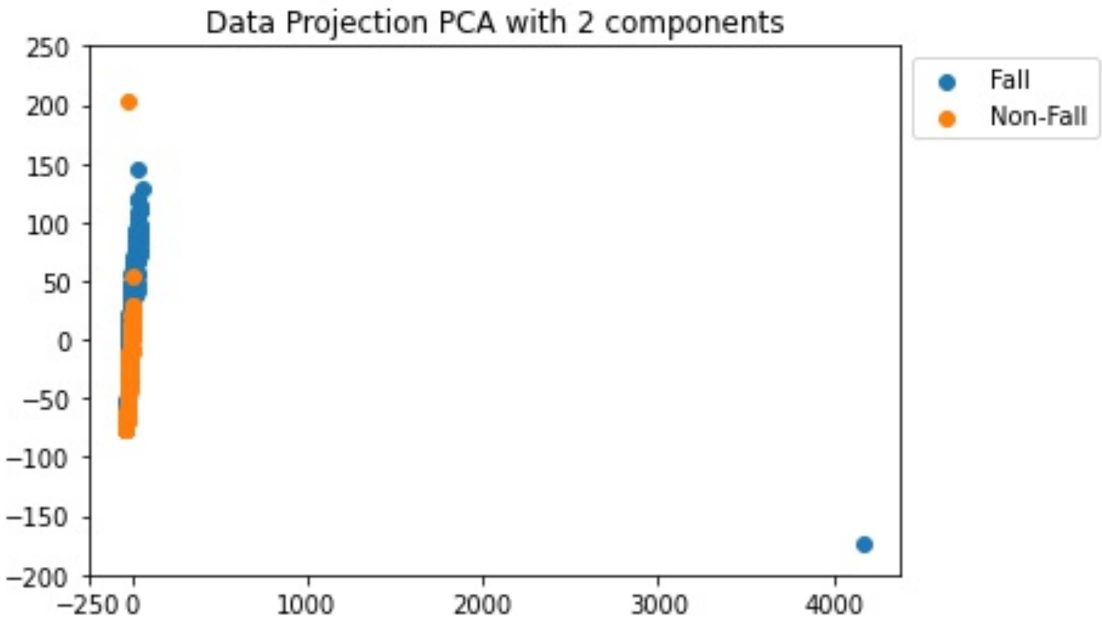

Figure 1: Projection of the Given Input to Top Two PC's

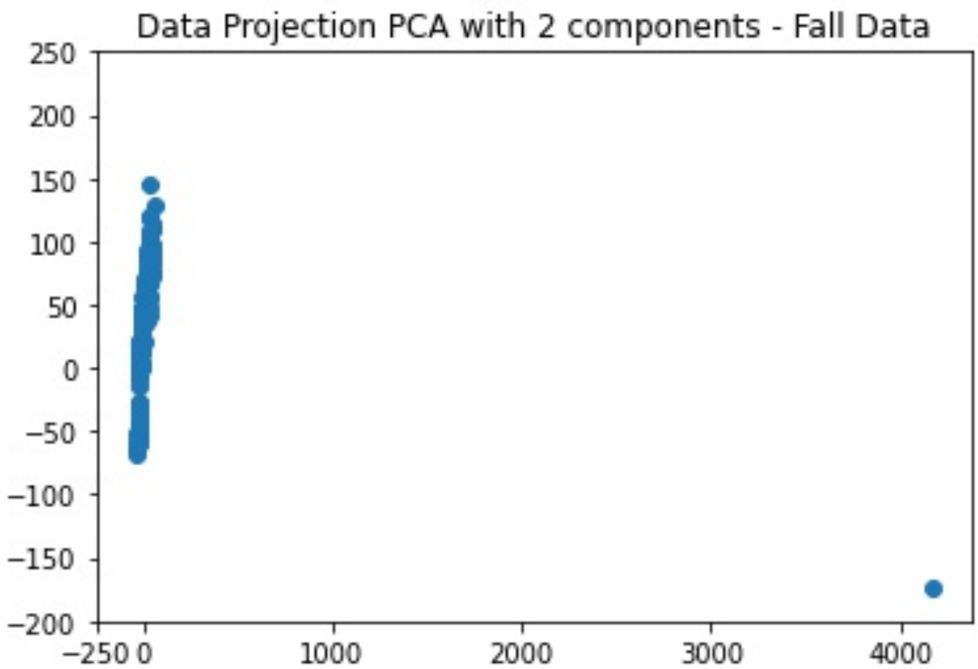

Figure 2: Projection of the Given Input to Top Two PC's - Fall Data

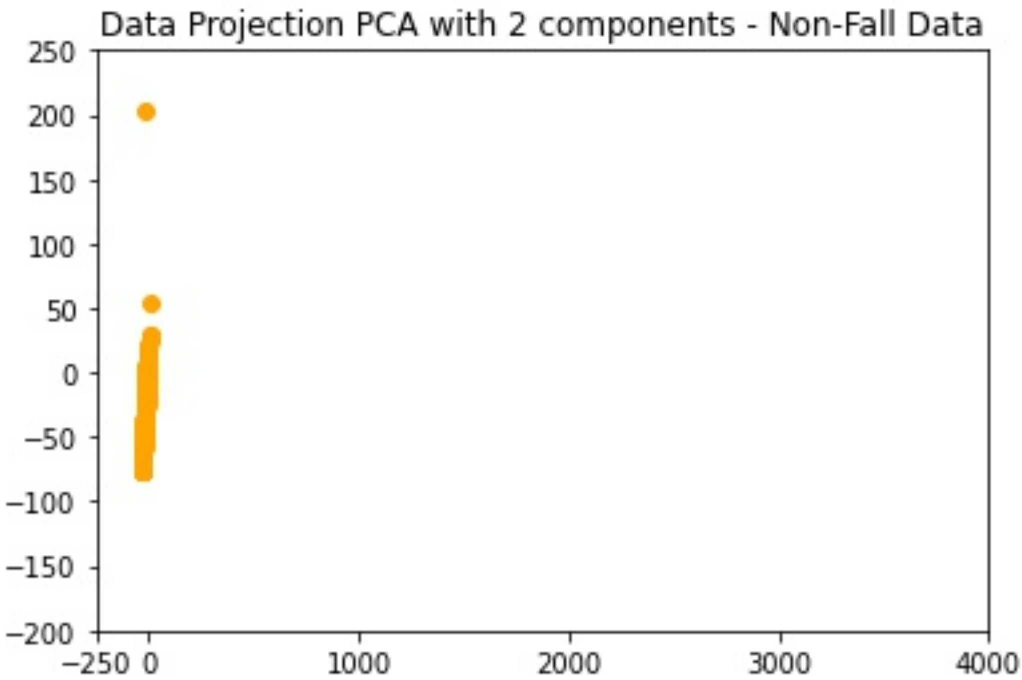

Figure 3: Projection of the Given Input to Top Two PC's - NonFall Data

With the new projected data, which is the projection of the input data to the top two principal components, k-means clustering algorithm is used to separate data into clusters. For the k values {2,3,4,5,6,7,8,9,10}, k-means algorithm is run and the results of separating the data are visualized in the figures *Figure 4* through *Figure 12*.

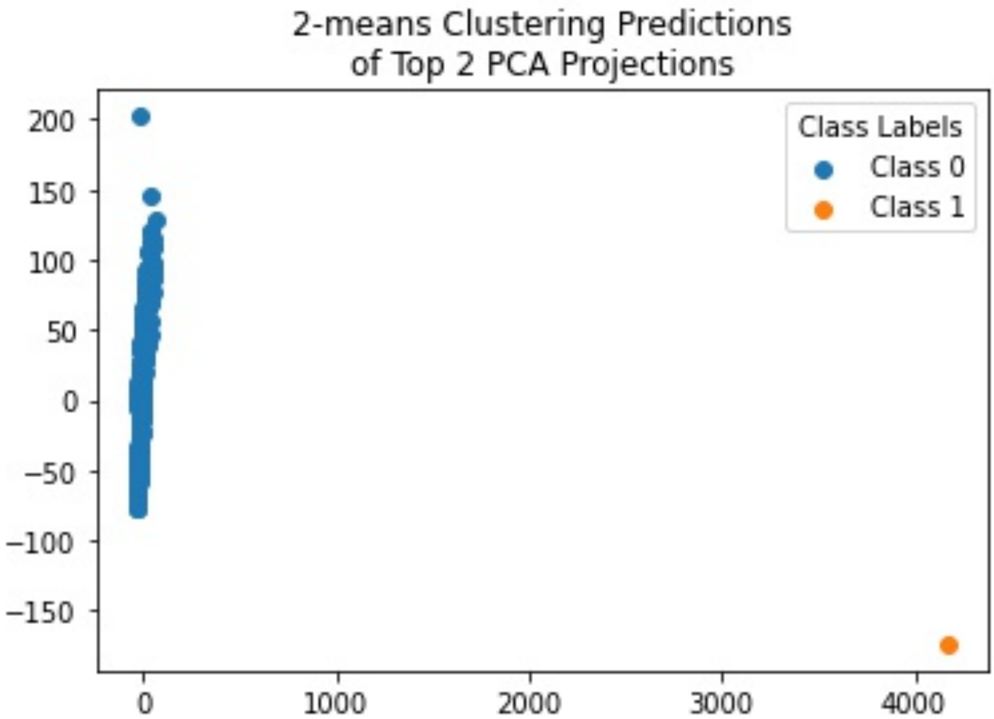

Figure 4: 2-means Clustering Predictions of Top 2 PCA Projections

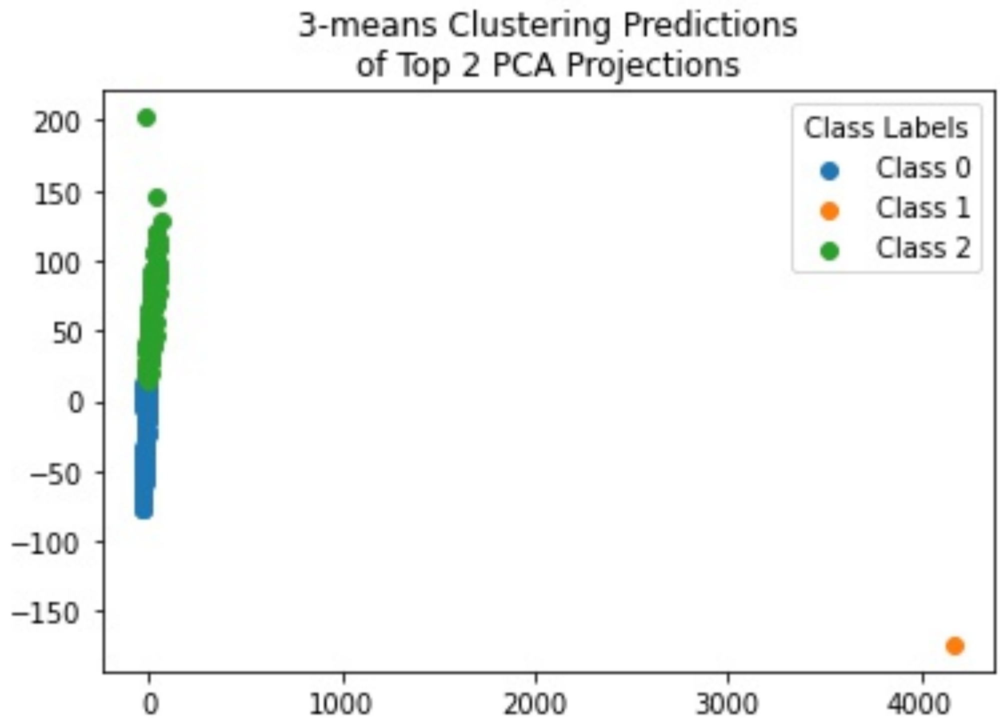

Figure 5: 3-means Clustering Predictions of Top 2 PCA Projections

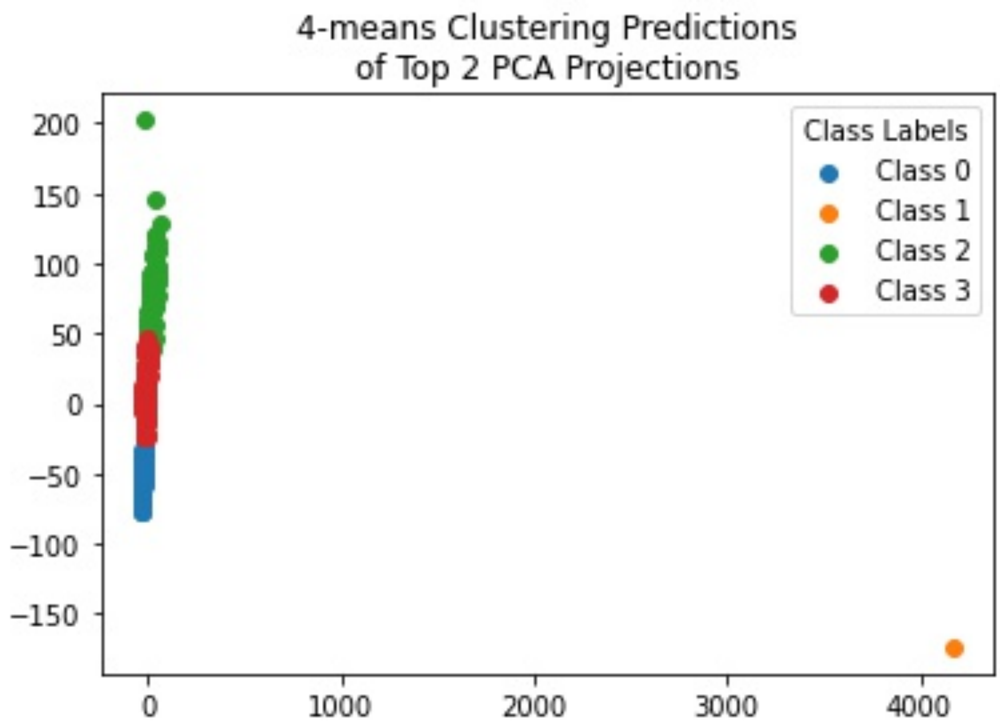

Figure 6: 4-means Clustering Predictions of Top 2 PCA Projections

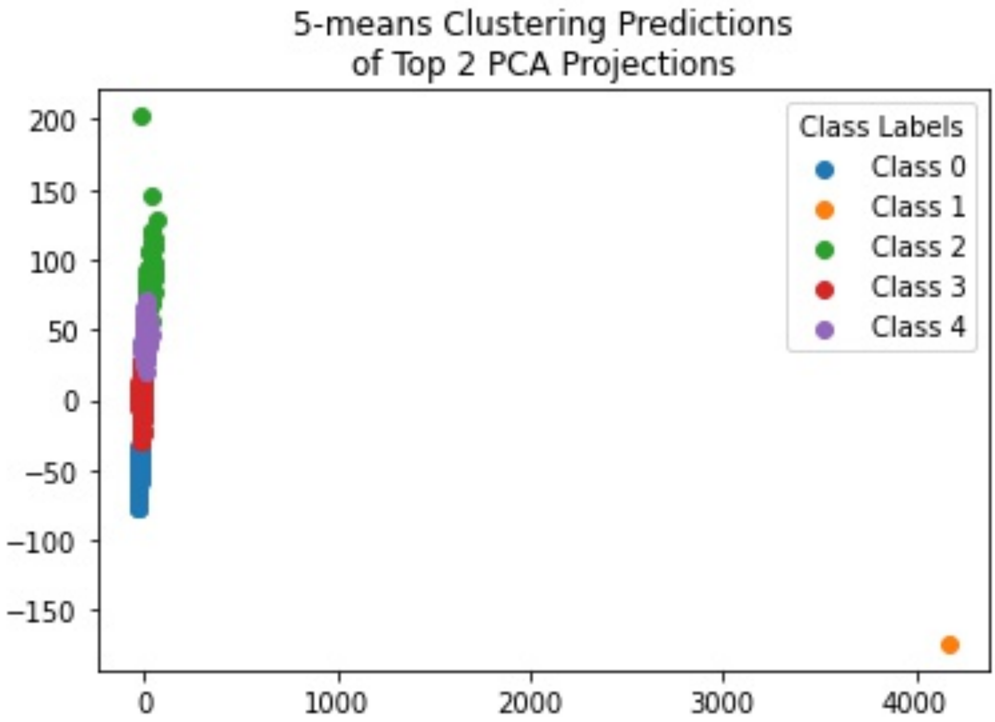

Figure 7: 5-means Clustering Predictions of Top 2 PCA Projections

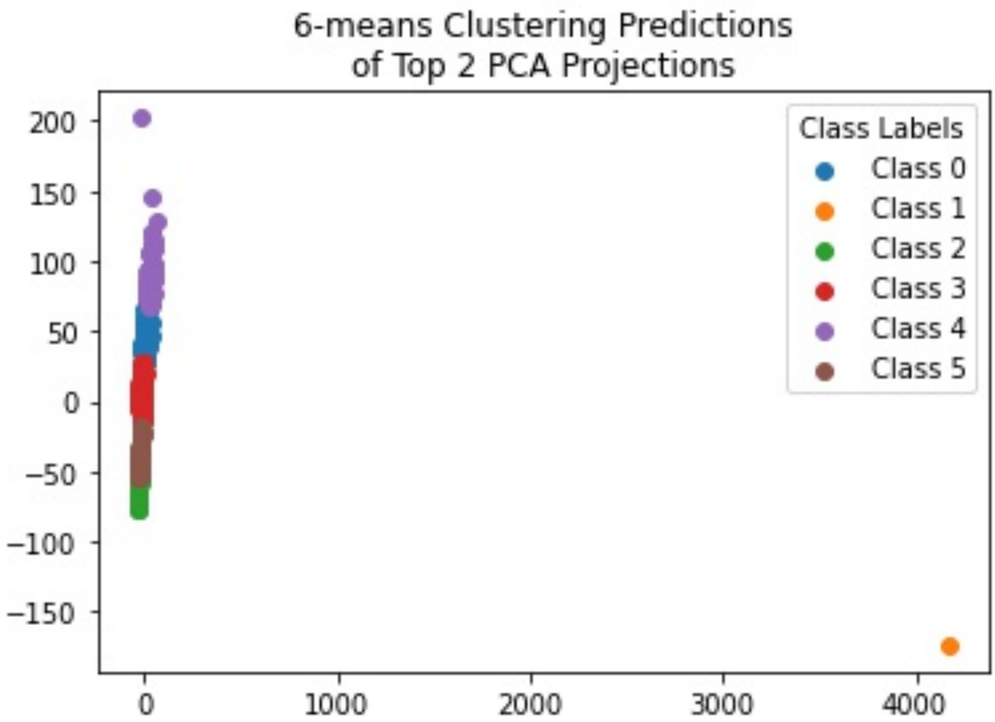

Figure 8: 6-means Clustering Predictions of Top 2 PCA Projections

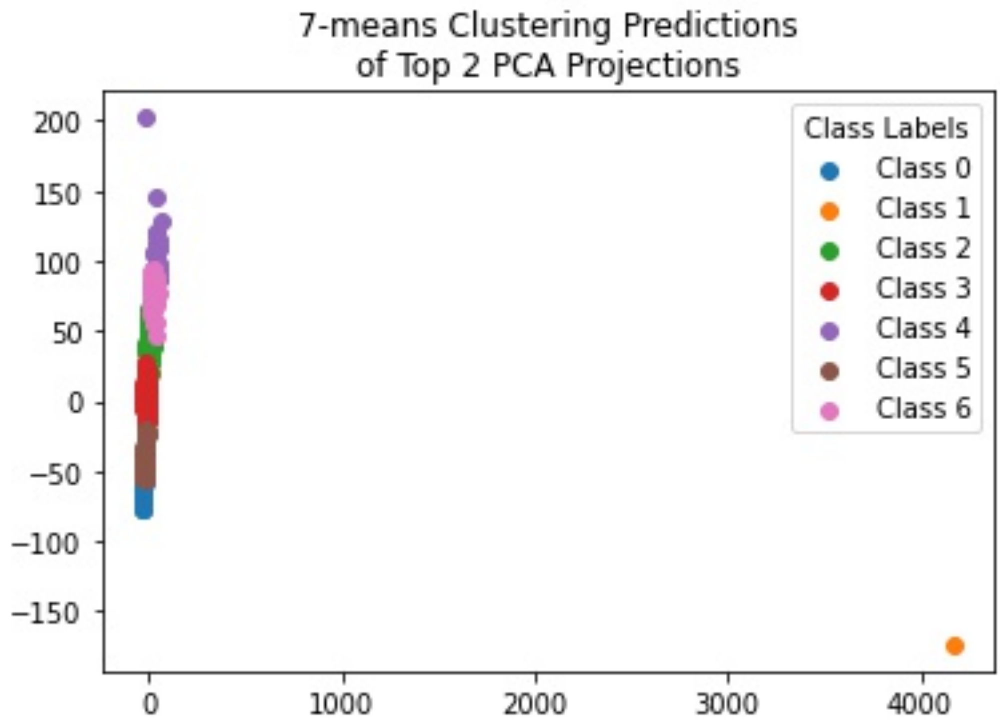

Figure 9: 7-means Clustering Predictions of Top 2 PCA Projections

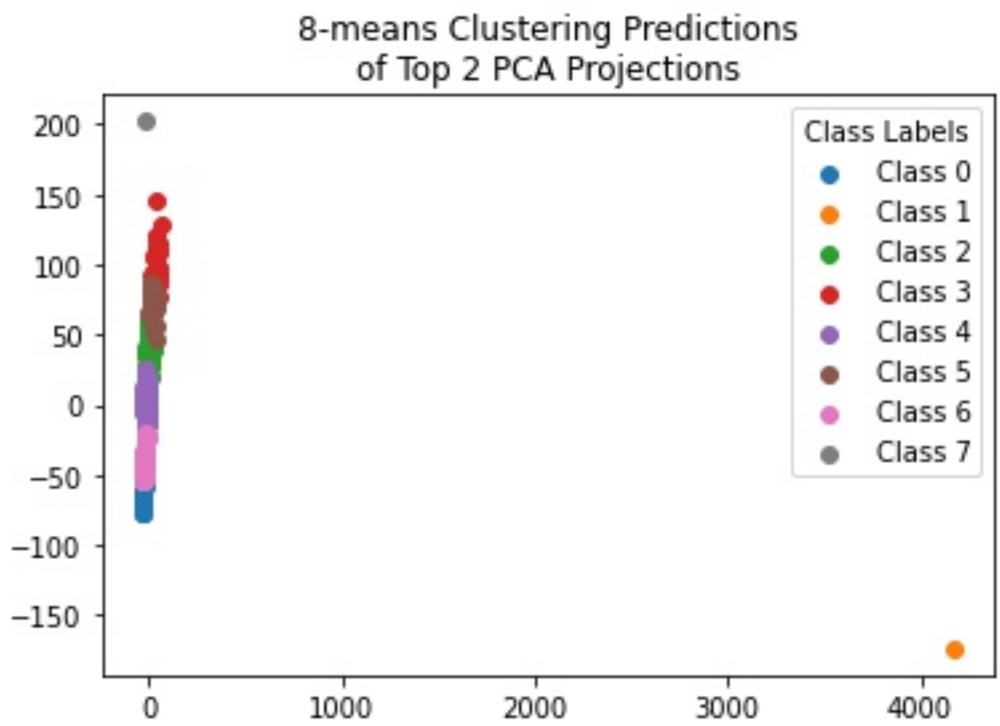

Figure 10: 8-means Clustering Predictions of Top 2 PCA Projections

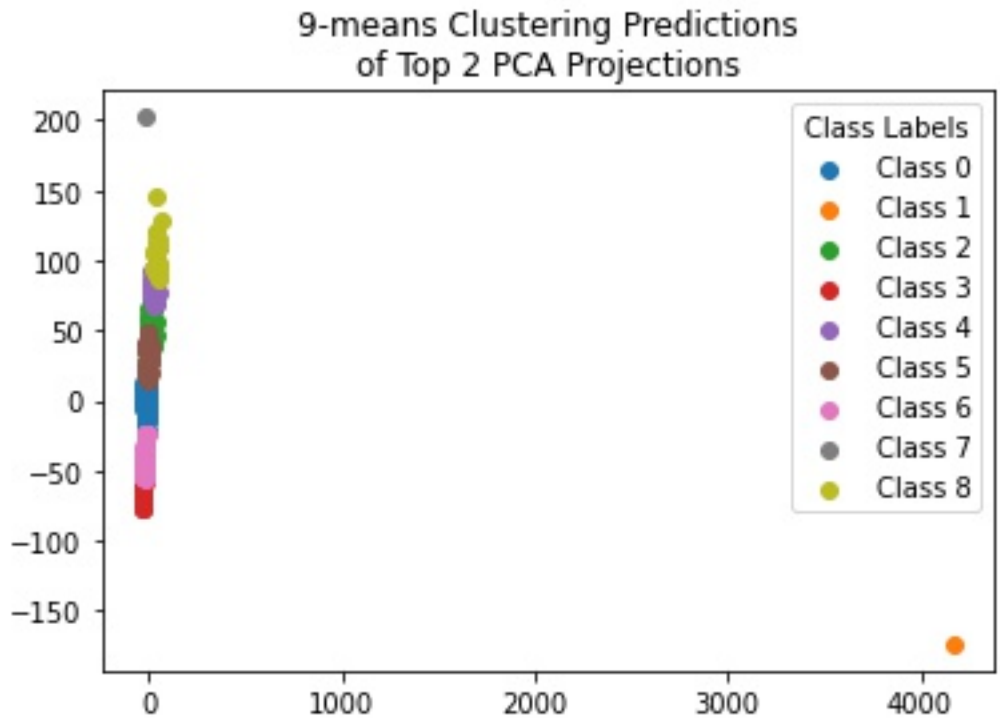

Figure 11: 9-means Clustering Predictions of Top 2 PCA Projections

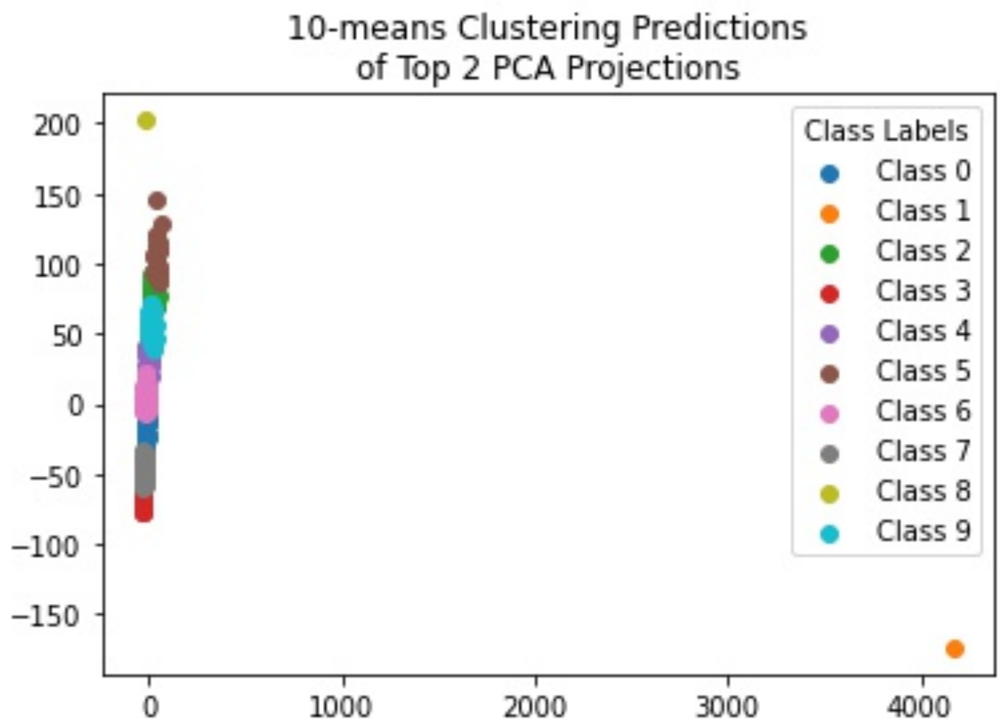

Figure 12: 10-means Clustering Predictions of Top 2 PCA Projections

As it can be seen in figures *Figure 1* through *Figure 12,* since there are two outlier samples in the projected data, the clusters obtained cannot be displayed clearly. For this reason, these two outliers are removed from the projected data and the result can be seen in the *Figure 13*.

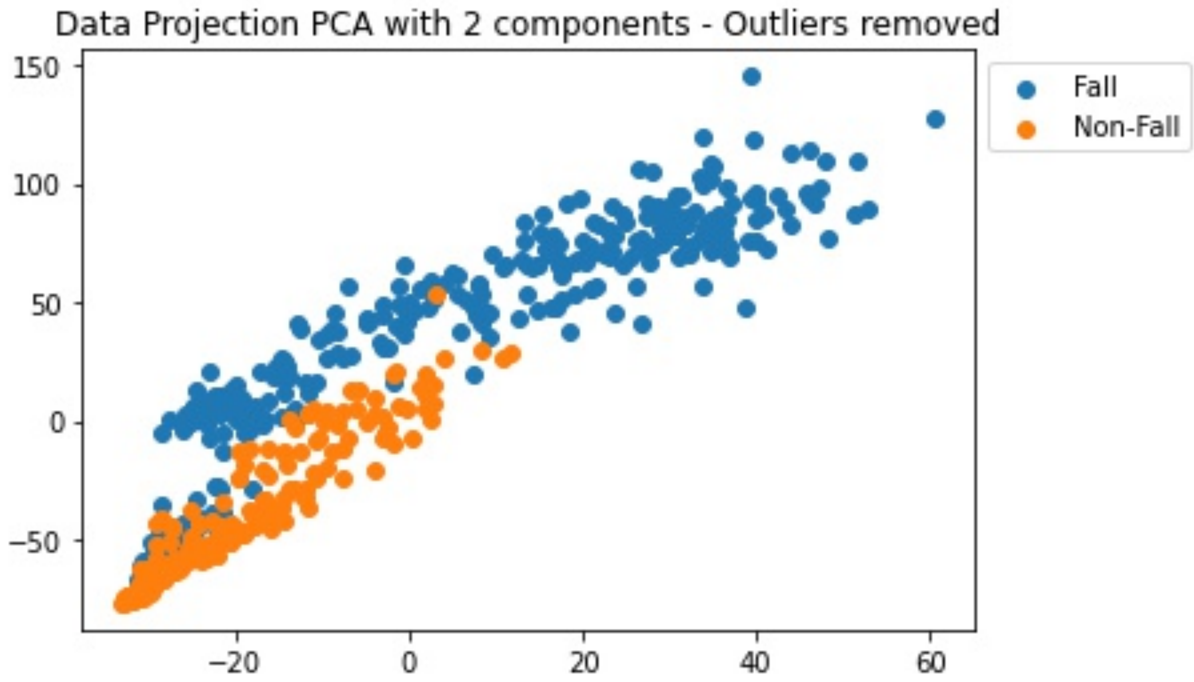

Figure 13: Projection of the Given Input to Top Two PC's - Outliers Removed

With the projected dataset without the outliers, k-means clustering algorithm is used for the k values {2,3,4,5,6,7,8,9,10}. The results of separating the data is visualized in the figures *Figure 14* through Figure

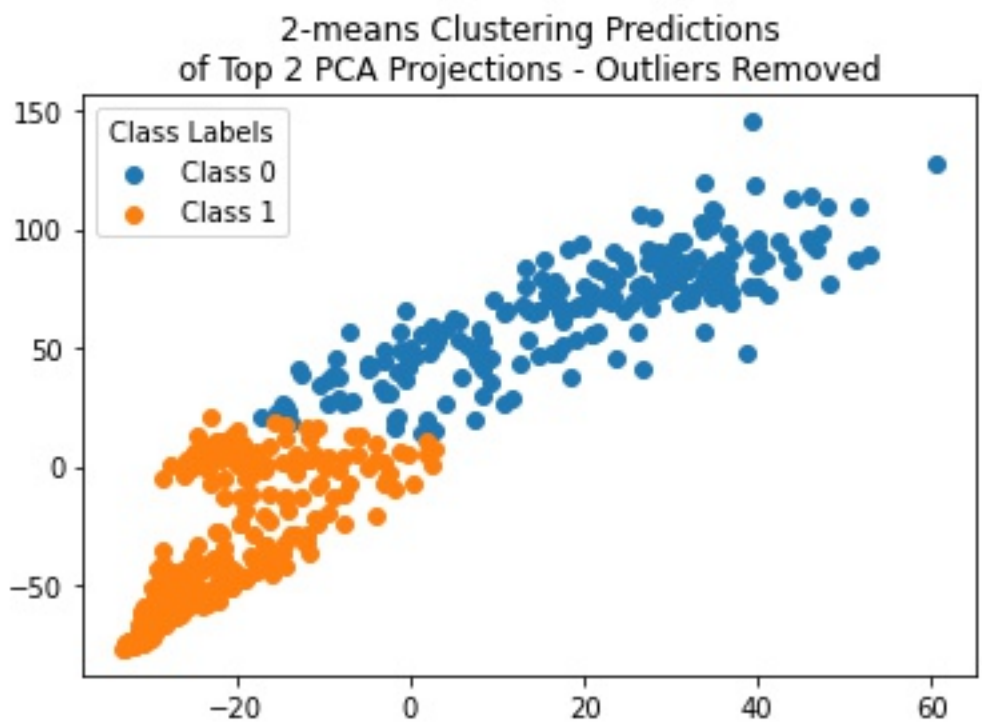

Figure 14: 2-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

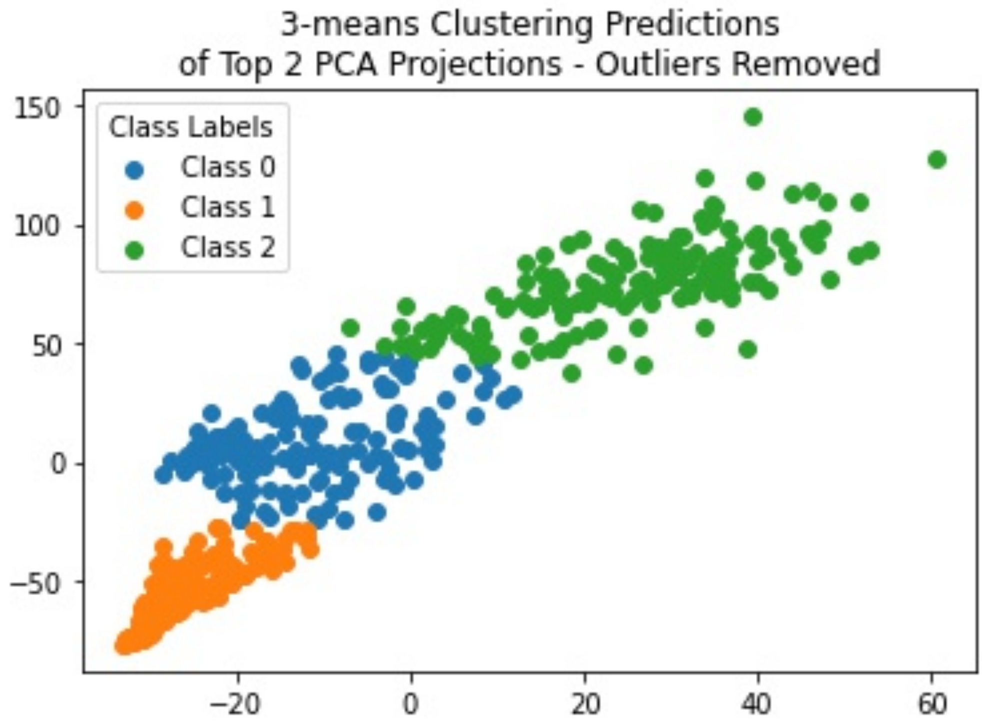

Figure 15: 3-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

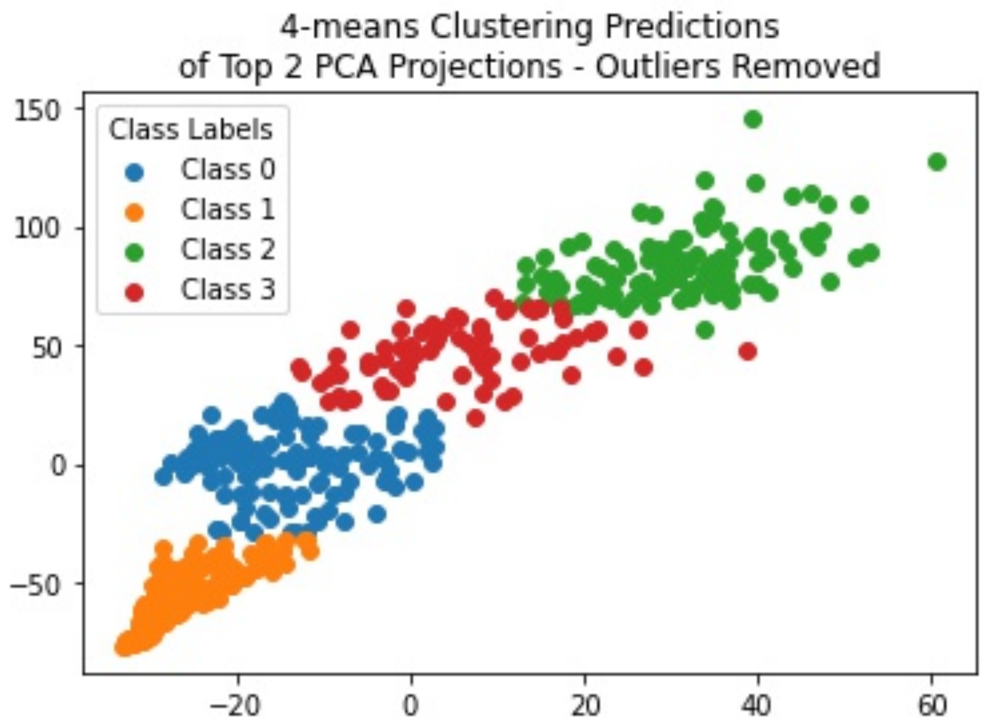

Figure 16: 4-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

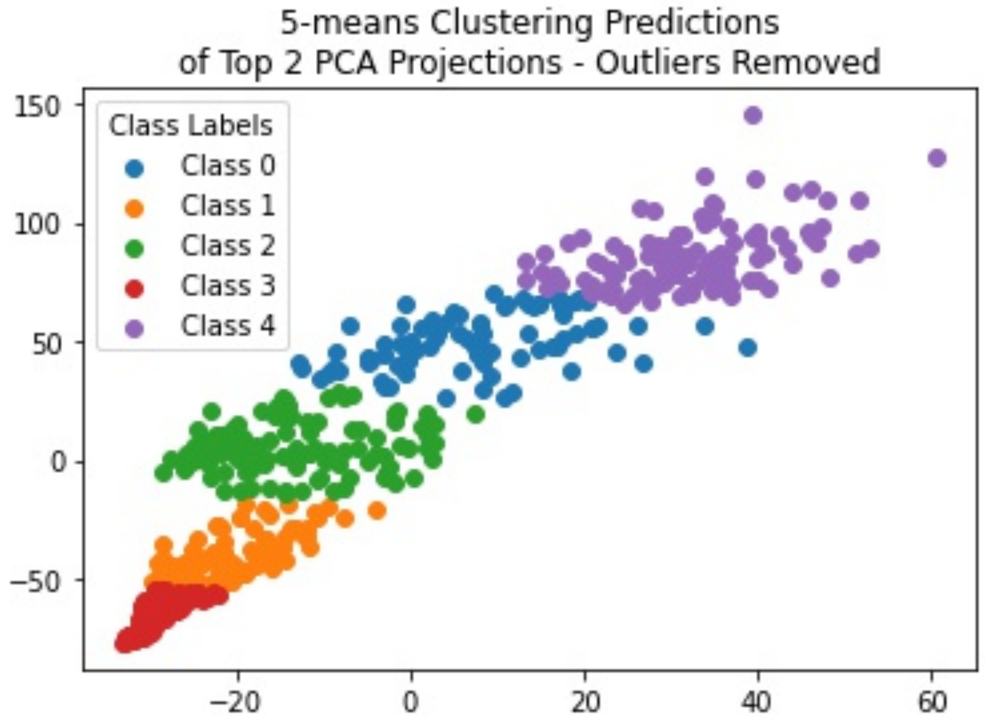

Figure 17: 5-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

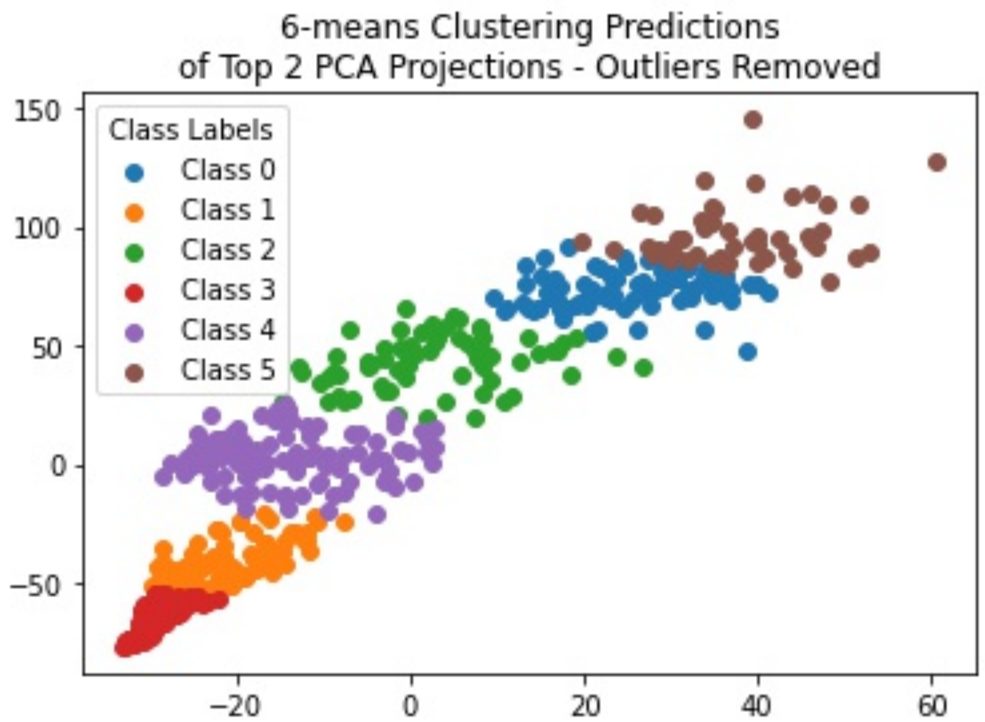

Figure 18: 6-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

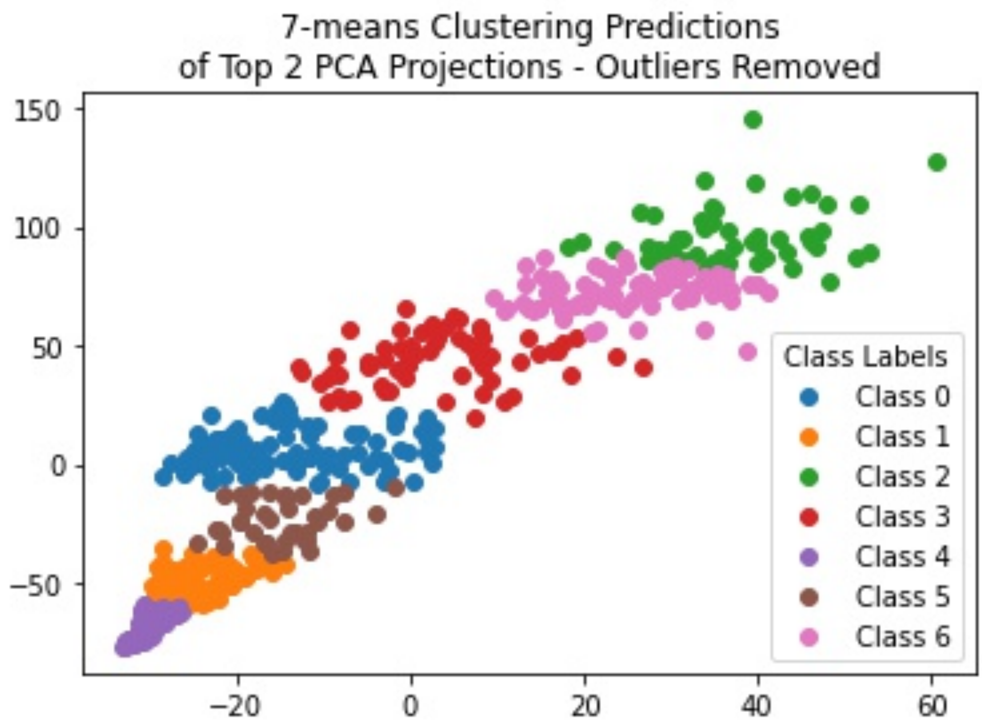

Figure 19: 7-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

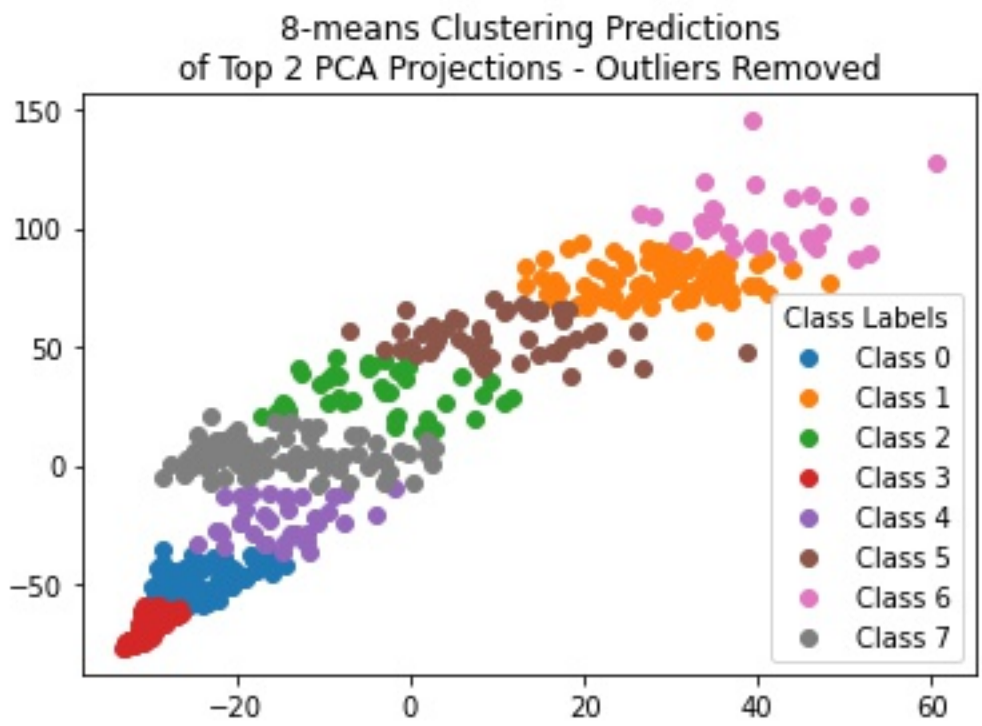

Figure 20: 8-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

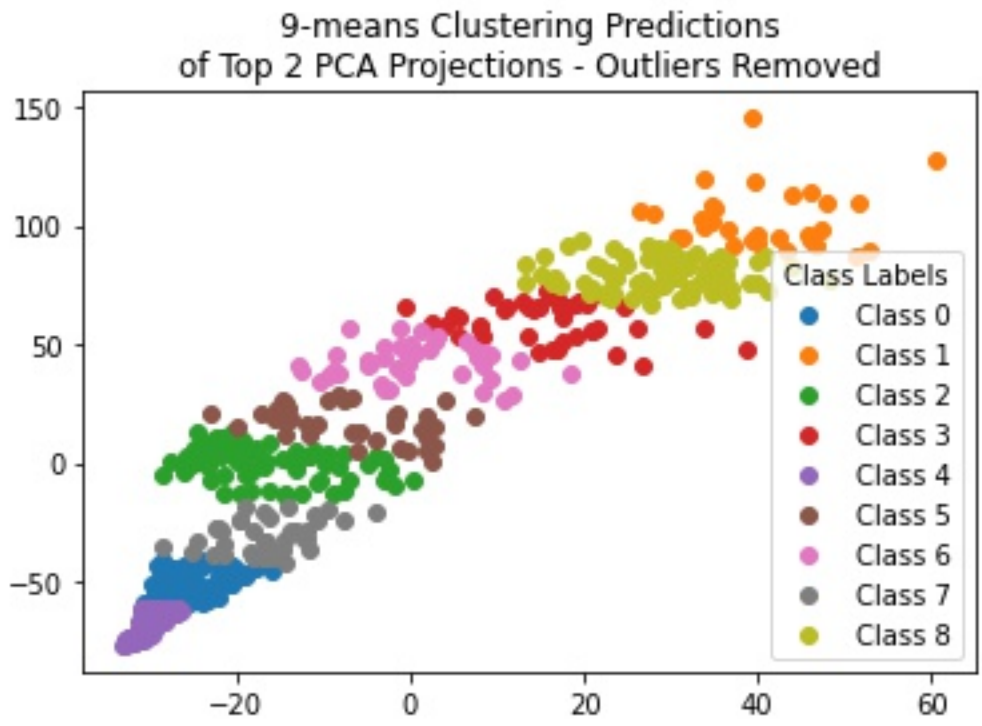

Figure 21: 9-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

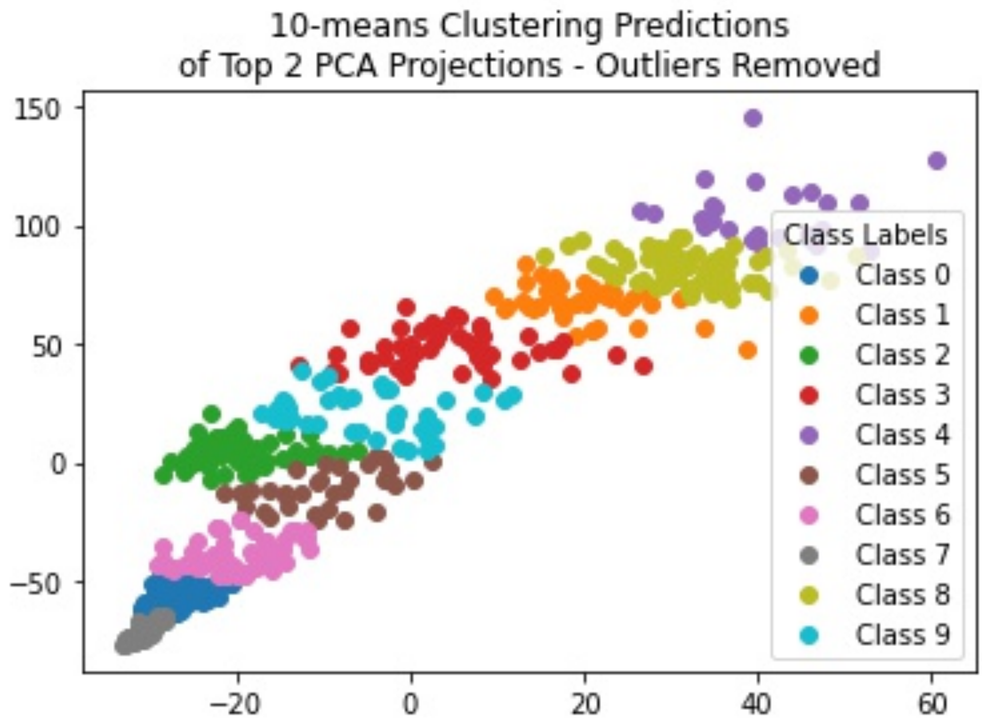

Figure 22: 10-means Clustering Predictions of Top 2 PCA Projections - Outliers Removed

From the first part of the illustrations, in figures *Figure 4* through *Figure 12*, it can be seen that there are two outliers in the projected data, especially the sample near to the bottom right corner of the plot. These outliers affect the k-means clustering algorithm considerably. When k is 2, because of the outlier point at bottom right, the algorithm separates the data inefficiently as outlier point considered as one class and the rest of all data samples considered as another class. When k is greater than or equal to 5, it can be seen that the clusters overlap with each other, and the separation into clusters is not clear. The optimal number of clusters can be considered as 3 or 4. For k is 3, class 1 and class 2 will correspond to fall samples, and class 0 will correspond to non-fall samples. For k is 4, class 1 and class 2 will correspond to fall samples, and class 0 and class 2 will correspond to non-fall samples.

When k is 2, the degree of consistency of separating the classes into two clusters is found to be 55%, which is low, as stated before, this is the result of outlier point considered as one class and the rest of all data samples considered as another class. In other words, the algorithm places all of the samples except the putter into the same cluster. This behavior is slightly better than random assignment, which is poor.

From the second part of the illustrations, in figures *Figure 14* through *Figure 22*, it can be seen that the two outliers in the projected data are removed and the data can be displayed more clearly. When k-means clustering algorithm is run with k value 2, the data can be separated into clusters more clearly compared to the separation result when data contain outliers. The optimal number of clusters can be considered as 3 or 4, since the separation in the clusters is similar to ground truth labels. When k is greater than or equal to 5, the resulting separation does not provide additional information since there are two ground truth values in the original dataset.

When k is 2, the degree of consistency of separating the classes into two clusters is found to be 78%, which is higher than the first part, because there is no outliers in the data that affect the result of separating data into clusters.

Excluding the outliers in second part and third part of the illustrations, 2-means clustering algorithm, which is an unsupervised learning algorithm, performs considerably well on the data. Given the reduced dimension is 2 out of 306, it is promising that fall detection is possible with the given data with more features. Since k-means algorithm does not consider the shape of the data, in our case they are stacked as is can be seen in *Figure 13* and they are not similar to perfect circular shape, supervised learning algorithms may learn these patterns as well and supervised learning algorithms perform well on fall detection.

Part B

In this part of the question, the data is splitter into three sets which are training, validation and test sets, with the proportions of 70%, 15% and 15% respectively. For the supervised learning stage, support-vector-machine classifier (SVM) and multi-layer perceptron classifier (MLP) is used. For both models, different number of hyper parameters is used to tune the parameters on the validation set by training the models with training set. With the best hyper parameters, classification accuracy of models on test set is calculated. In addition the two models, SVM and MLP are compared and evaluated.

SVM

For the SVM model, regularization parameter, kernel type, degree and kernel coefficient are used as hyper parameters to find the best model on the validation data. Different combinations of the following set of hyper parameters are used.

- c_values (regularization parameter): [1e-3, 1e-2, 1e-1, 1e0, 1e1]
- kernel_types: [“linear", "poly", "rbf", "sigmoid"]
- degrees: [1, 2, 3, 4, 5]
- gamma_values (kernel coefficient): ["scale", "auto"]

SVM model is trained with above hyper parameters, and the validation scores for each different combination of these parameters are recorded in SVM_data.xlsx. The resulting table consisting of the results can be found in Appendix A.

In order to decide on the best combination of hyper parameters, the results are sorted in descending order with respect to their resulting validation accuracies. 9 out of 75 different combinations of hyper parameters resulted in validation accuracy of 100%, and from those combinations the following combination is chosen as its regularization parameter is lower than other’s which I considered it to be better because punishing the learning rate too much with L2 regularization may result in poorer results on test set.

- c_value (regularization parameter): 1e-3
- kernel_type: "poly"
- degree: 4
- gamma_value (kernel coefficient): "auto"

The SVM model with the hyper parameters above resulted in 100% accuracy on the test set.

MLP

For the MLP model, hidden layer size, activation function, solver, alpha and learning rate are used as hyper parameters to find the best model on the validation data. Different combinations of the following set of hyper parameters are used.

- hidden_layer_sizes: [(8,8), (16,16), (32,32), (64,64)]
- activation_functions: ["logistic", "tanh", "relu"]
- solvers: ["sgd", "adam"]
- alphas: [1e-3, 1e-2, 1e-1, 1e0, 1e1]
- learning_rates: [1e-3, 1e-2, 1e-1]

MLP model is trained with above hyper parameters, and the validation scores for each different combination of these parameters are recorded in MLP_data.xlsx. The resulting table consisting of the results can be found in Appendix B.

In order to decide on the best combination of hyper parameters, the results are sorted in descending order with respect to their resulting validation accuracies. 25 out of 360 different combinations of hyper parameters resulted in validation accuracy of 100%, and from those combinations the following combination is chosen as its hidden layer size is larger, which means that it can capture more complex patterns in the data, and its regularization parameter is lower than other's which I considered it to be better because punishing the learning rate too much with L2 regularization may result in poorer results on test set.

- hidden_layer_size: (64,64)
- activation_function: "relu"
- solver: "sgd"
- alpha: 1e-3
- learning_rate: 1e-3

The MLP model with the hyper parameters above resulted in 100% accuracy on the test set.

Interpretation and Comparison

The success of fall detection based o wearable sensors is can be considered excellent, as both of the models resulted in test accuracy of 100%, for the given train-validation-test sets.

For the given data and problem, MLP and SVM performed well in terms of accuracy. Comparing the number of combinations used to train the model and the number of combinations that the validation accuracy is 100%, which are 0.12 for SVM and 0.07 for MLP, it can be said that SVM model performs better in terms of the accuracy proportions. In addition, when experimenting with the model's performance with different set of hyper parameters on test set, it was seen that SVM models performed generally better than MLP models, provided that the combinations of hyper parameters that result in the same validation accuracy in both models are compared. Furthermore, when training the models, SVM models are trained faster than MLP models.

For SVM and MLP models, small set of hyper parameters are used to train the models due to time complexity. For different combinations of hyper parameters and different split of train-validation-test datasets may result in different outcomes in terms of the performance and accuracy of the models. In general case, the SVM models with kernel functions which is not linear performs better than MLP models in terms of accuracy, in both low dimensional and high dimensional data sets. In addition, the complexity of SVM is increased when the dimension of the data increases while the complexity of SVM models is not dependent on the dimension of the data. SVM models are more efficient than MLP models as the perform better on unseen data.

Appendix A

Regularization Parameter	Kernel Type	Degree	rnel Coefficie	Validation Accuracy
0,001	poly	4	auto	1
0,01	poly	3	auto	1
0,01	poly	4	auto	1
0,1	poly	3	auto	1
0,1	poly	4	auto	1
1	poly	3	auto	1
1	poly	4	auto	1
10	poly	3	auto	1
10	poly	4	auto	1
0,001	poly	3	auto	0,988235294117647
0,001	poly	5	auto	0,988235294117647
0,01	linear			0,988235294117647
0,01	poly	2	auto	0,988235294117647
0,01	poly	5	auto	0,988235294117647
0,1	linear			0,988235294117647
0,1	poly	2	auto	0,988235294117647
0,1	poly	5	auto	0,988235294117647
1	linear			0,988235294117647
1	poly	1	auto	0,988235294117647
1	poly	2	auto	0,988235294117647
1	poly	5	auto	0,988235294117647
10	linear			0,988235294117647
10	poly	1	auto	0,988235294117647
10	poly	2	auto	0,988235294117647
10	poly	5	auto	0,988235294117647
0,001	linear			0,976470588235294
0,001	poly	2	auto	0,976470588235294
0,01	poly	1	auto	0,976470588235294
0,1	poly	1	auto	0,976470588235294
1	poly	1	scale	0,976470588235294
1	rbf		scale	0,976470588235294
10	poly	1	scale	0,976470588235294
10	poly	2	scale	0,976470588235294
10	rbf		scale	0,976470588235294
1	sigmoid		scale	0,964705882352941
10	sigmoid		scale	0,941176470588235
0,1	rbf		scale	0,917647058823529
10	poly	3	scale	0,917647058823529
0,001	poly	1	auto	0,905882352941176
1	rbf		auto	0,905882352941176
10	rbf		auto	0,905882352941176

1	poly	2	scale	0,894117647058824
0,1	poly	2	scale	0,882352941176471
1	poly	3	scale	0,882352941176471
10	poly	4	scale	0,882352941176471
0,1	sigmoid		scale	0,870588235294118
0,1	poly	1	scale	0,858823529411765
0,1	rbf		auto	0,823529411764706
10	poly	5	scale	0,8
1	poly	4	scale	0,670588235294118
1	poly	5	scale	0,623529411764706
0,001	poly	1	scale	0,588235294117647
0,001	poly	2	scale	0,588235294117647
0,001	poly	3	scale	0,588235294117647
0,001	poly	4	scale	0,588235294117647
0,001	poly	5	scale	0,588235294117647
0,001	rbf		scale	0,588235294117647
0,001	rbf		auto	0,588235294117647
0,001	sigmoid		scale	0,588235294117647
0,001	sigmoid		auto	0,588235294117647
0,01	poly	1	scale	0,588235294117647
0,01	poly	2	scale	0,588235294117647
0,01	poly	3	scale	0,588235294117647
0,01	poly	4	scale	0,588235294117647
0,01	poly	5	scale	0,588235294117647
0,01	rbf		scale	0,588235294117647
0,01	rbf		auto	0,588235294117647
0,01	sigmoid		scale	0,588235294117647
0,01	sigmoid		auto	0,588235294117647
0,1	poly	3	scale	0,588235294117647
0,1	poly	4	scale	0,588235294117647
0,1	poly	5	scale	0,588235294117647
0,1	sigmoid		auto	0,588235294117647
1	sigmoid		auto	0,588235294117647
10	sigmoid		auto	0,141176470588235

Table 1: Results for Different Hyper Parameter Combinations for SVM Model

Appendix B

Hidden Layer Size	Activation Function	Solver	Alpha	Learning Rate	Validation Accuracy
(8, 8)	relu	sgd	0,001	0,001	1
(8, 8)	relu	sgd	0,001	0,01	1
(8, 8)	relu	sgd	0,01	0,001	1
(8, 8)	relu	sgd	0,01	0,01	1
(8, 8)	relu	sgd	0,1	0,001	1
(8, 8)	relu	sgd	0,1	0,01	1
(8, 8)	relu	sgd	1	0,001	1
(8, 8)	relu	adam	0,001	0,001	1
(8, 8)	relu	adam	0,001	0,01	1
(8, 8)	relu	adam	0,01	0,001	1
(8, 8)	relu	adam	0,01	0,01	1
(8, 8)	relu	adam	0,1	0,001	1
(8, 8)	relu	adam	0,1	0,01	1
(8, 8)	relu	adam	1	0,001	1
(8, 8)	relu	adam	1	0,01	1
(16, 16)	relu	adam	0,001	0,001	1
(16, 16)	relu	adam	0,01	0,001	1
(16, 16)	relu	adam	0,1	0,001	1
(64, 64)	relu	sgd	0,001	0,001	1
(64, 64)	relu	sgd	0,01	0,001	1
(64, 64)	relu	sgd	0,1	0,001	1
(64, 64)	relu	sgd	1	0,001	1
(64, 64)	relu	adam	0,001	0,001	1
(64, 64)	relu	adam	0,01	0,001	1
(64, 64)	relu	adam	0,1	0,001	1
(8, 8)	logistic	sgd	0,001	0,001	0,988235294117647
(8, 8)	logistic	sgd	0,001	0,01	0,988235294117647
(8, 8)	logistic	sgd	0,001	0,1	0,988235294117647
(8, 8)	logistic	sgd	0,01	0,001	0,988235294117647
(8, 8)	logistic	sgd	0,01	0,01	0,988235294117647
(8, 8)	logistic	sgd	0,01	0,1	0,988235294117647
(8, 8)	logistic	sgd	0,1	0,001	0,988235294117647
(8, 8)	logistic	sgd	0,1	0,01	0,988235294117647
(8, 8)	logistic	sgd	0,1	0,1	0,988235294117647
(8, 8)	logistic	sgd	1	0,001	0,988235294117647
(8, 8)	logistic	sgd	1	0,01	0,988235294117647
(8, 8)	logistic	sgd	1	0,1	0,988235294117647
(8, 8)	logistic	sgd	10	0,01	0,988235294117647
(8, 8)	logistic	adam	0,001	0,001	0,988235294117647
(8, 8)	logistic	adam	0,001	0,01	0,988235294117647
(8, 8)	logistic	adam	0,01	0,001	0,988235294117647

(8, 8)	logistic	adam	0,01	0,01	0,988235294117647
(8, 8)	logistic	adam	0,1	0,001	0,988235294117647
(8, 8)	logistic	adam	1	0,001	0,988235294117647
(8, 8)	logistic	adam	1	0,01	0,988235294117647
(8, 8)	logistic	adam	10	0,001	0,988235294117647
(8, 8)	logistic	adam	10	0,01	0,988235294117647
(8, 8)	tanh	sgd	0,001	0,01	0,988235294117647
(8, 8)	tanh	sgd	0,001	0,1	0,988235294117647
(8, 8)	tanh	sgd	0,01	0,01	0,988235294117647
(8, 8)	tanh	sgd	0,1	0,01	0,988235294117647
(8, 8)	tanh	sgd	1	0,01	0,988235294117647
(8, 8)	tanh	sgd	10	0,001	0,988235294117647
(8, 8)	tanh	sgd	10	0,01	0,988235294117647
(8, 8)	tanh	adam	0,001	0,001	0,988235294117647
(8, 8)	tanh	adam	0,001	0,01	0,988235294117647
(8, 8)	tanh	adam	0,01	0,001	0,988235294117647
(8, 8)	tanh	adam	0,01	0,01	0,988235294117647
(8, 8)	tanh	adam	0,1	0,001	0,988235294117647
(8, 8)	tanh	adam	0,1	0,01	0,988235294117647
(8, 8)	tanh	adam	1	0,001	0,988235294117647
(8, 8)	tanh	adam	10	0,001	0,988235294117647
(8, 8)	relu	sgd	1	0,01	0,988235294117647
(8, 8)	relu	sgd	10	0,001	0,988235294117647
(8, 8)	relu	sgd	10	0,01	0,988235294117647
(8, 8)	relu	adam	10	0,001	0,988235294117647
(8, 8)	relu	adam	10	0,01	0,988235294117647
(16, 16)	logistic	sgd	0,001	0,001	0,988235294117647
(16, 16)	logistic	sgd	0,001	0,01	0,988235294117647
(16, 16)	logistic	sgd	0,001	0,1	0,988235294117647
(16, 16)	logistic	sgd	0,01	0,001	0,988235294117647
(16, 16)	logistic	sgd	0,01	0,01	0,988235294117647
(16, 16)	logistic	sgd	0,01	0,1	0,988235294117647
(16, 16)	logistic	sgd	0,1	0,001	0,988235294117647
(16, 16)	logistic	sgd	0,1	0,01	0,988235294117647
(16, 16)	logistic	sgd	0,1	0,1	0,988235294117647
(16, 16)	logistic	sgd	1	0,001	0,988235294117647
(16, 16)	logistic	sgd	1	0,01	0,988235294117647
(16, 16)	logistic	sgd	1	0,1	0,988235294117647
(16, 16)	logistic	sgd	10	0,001	0,988235294117647
(16, 16)	logistic	sgd	10	0,01	0,988235294117647
(16, 16)	logistic	adam	0,001	0,001	0,988235294117647
(16, 16)	logistic	adam	0,001	0,01	0,988235294117647
(16, 16)	logistic	adam	0,01	0,001	0,988235294117647
(16, 16)	logistic	adam	0,01	0,01	0,988235294117647

(16, 16)	logistic	adam	0,1	0,001	0,988235294117647
(16, 16)	logistic	adam	0,1	0,01	0,988235294117647
(16, 16)	logistic	adam	1	0,001	0,988235294117647
(16, 16)	logistic	adam	10	0,001	0,988235294117647
(16, 16)	logistic	adam	10	0,1	0,988235294117647
(16, 16)	tanh	sgd	0,001	0,01	0,988235294117647
(16, 16)	tanh	sgd	0,001	0,1	0,988235294117647
(16, 16)	tanh	sgd	0,01	0,01	0,988235294117647
(16, 16)	tanh	sgd	0,01	0,1	0,988235294117647
(16, 16)	tanh	sgd	0,1	0,01	0,988235294117647
(16, 16)	tanh	sgd	0,1	0,1	0,988235294117647
(16, 16)	tanh	sgd	1	0,01	0,988235294117647
(16, 16)	tanh	sgd	10	0,001	0,988235294117647
(16, 16)	tanh	sgd	10	0,01	0,988235294117647
(16, 16)	tanh	adam	0,001	0,001	0,988235294117647
(16, 16)	tanh	adam	0,001	0,01	0,988235294117647
(16, 16)	tanh	adam	0,01	0,001	0,988235294117647
(16, 16)	tanh	adam	0,01	0,01	0,988235294117647
(16, 16)	tanh	adam	0,1	0,001	0,988235294117647
(16, 16)	tanh	adam	0,1	0,01	0,988235294117647
(16, 16)	tanh	adam	1	0,001	0,988235294117647
(16, 16)	tanh	adam	1	0,01	0,988235294117647
(16, 16)	tanh	adam	10	0,001	0,988235294117647
(16, 16)	tanh	adam	10	0,01	0,988235294117647
(16, 16)	relu	sgd	0,001	0,001	0,988235294117647
(16, 16)	relu	sgd	0,001	0,01	0,988235294117647
(16, 16)	relu	sgd	0,01	0,001	0,988235294117647
(16, 16)	relu	sgd	0,01	0,01	0,988235294117647
(16, 16)	relu	sgd	0,1	0,001	0,988235294117647
(16, 16)	relu	sgd	0,1	0,01	0,988235294117647
(16, 16)	relu	sgd	1	0,01	0,988235294117647
(16, 16)	relu	sgd	10	0,001	0,988235294117647
(16, 16)	relu	sgd	10	0,01	0,988235294117647
(16, 16)	relu	adam	0,001	0,01	0,988235294117647
(16, 16)	relu	adam	0,01	0,1	0,988235294117647
(16, 16)	relu	adam	0,1	0,01	0,988235294117647
(16, 16)	relu	adam	1	0,001	0,988235294117647
(16, 16)	relu	adam	1	0,01	0,988235294117647
(16, 16)	relu	adam	10	0,001	0,988235294117647
(16, 16)	relu	adam	10	0,01	0,988235294117647
(32, 32)	logistic	sgd	0,001	0,001	0,988235294117647
(32, 32)	logistic	sgd	0,001	0,01	0,988235294117647
(32, 32)	logistic	sgd	0,001	0,1	0,988235294117647
(32, 32)	logistic	sgd	0,01	0,001	0,988235294117647

(32, 32)	logistic	sgd	0,01	0,01	0,988235294117647
(32, 32)	logistic	sgd	0,01	0,1	0,988235294117647
(32, 32)	logistic	sgd	0,1	0,001	0,988235294117647
(32, 32)	logistic	sgd	0,1	0,01	0,988235294117647
(32, 32)	logistic	sgd	0,1	0,1	0,988235294117647
(32, 32)	logistic	sgd	1	0,001	0,988235294117647
(32, 32)	logistic	sgd	1	0,01	0,988235294117647
(32, 32)	logistic	sgd	1	0,1	0,988235294117647
(32, 32)	logistic	sgd	10	0,001	0,988235294117647
(32, 32)	logistic	sgd	10	0,01	0,988235294117647
(32, 32)	logistic	sgd	10	0,1	0,988235294117647
(32, 32)	logistic	adam	0,001	0,001	0,988235294117647
(32, 32)	logistic	adam	0,001	0,01	0,988235294117647
(32, 32)	logistic	adam	0,001	0,1	0,988235294117647
(32, 32)	logistic	adam	0,01	0,001	0,988235294117647
(32, 32)	logistic	adam	0,01	0,01	0,988235294117647
(32, 32)	logistic	adam	0,01	0,1	0,988235294117647
(32, 32)	logistic	adam	0,1	0,001	0,988235294117647
(32, 32)	logistic	adam	0,1	0,01	0,988235294117647
(32, 32)	logistic	adam	1	0,001	0,988235294117647
(32, 32)	logistic	adam	10	0,001	0,988235294117647
(32, 32)	tanh	sgd	0,001	0,001	0,988235294117647
(32, 32)	tanh	sgd	0,001	0,01	0,988235294117647
(32, 32)	tanh	sgd	0,01	0,001	0,988235294117647
(32, 32)	tanh	sgd	0,01	0,01	0,988235294117647
(32, 32)	tanh	sgd	0,01	0,1	0,988235294117647
(32, 32)	tanh	sgd	0,1	0,001	0,988235294117647
(32, 32)	tanh	sgd	0,1	0,01	0,988235294117647
(32, 32)	tanh	sgd	0,1	0,1	0,988235294117647
(32, 32)	tanh	sgd	1	0,001	0,988235294117647
(32, 32)	tanh	sgd	1	0,01	0,988235294117647
(32, 32)	tanh	sgd	10	0,001	0,988235294117647
(32, 32)	tanh	sgd	10	0,01	0,988235294117647
(32, 32)	tanh	adam	0,001	0,001	0,988235294117647
(32, 32)	tanh	adam	0,001	0,01	0,988235294117647
(32, 32)	tanh	adam	0,01	0,001	0,988235294117647
(32, 32)	tanh	adam	0,01	0,01	0,988235294117647
(32, 32)	tanh	adam	0,1	0,001	0,988235294117647
(32, 32)	tanh	adam	0,1	0,01	0,988235294117647
(32, 32)	tanh	adam	1	0,001	0,988235294117647
(32, 32)	tanh	adam	1	0,01	0,988235294117647
(32, 32)	tanh	adam	10	0,001	0,988235294117647
(32, 32)	tanh	adam	10	0,01	0,988235294117647
(32, 32)	relu	sgd	0,001	0,001	0,988235294117647

(32, 32)	relu	sgd	0,001	0,01	0,988235294117647
(32, 32)	relu	sgd	0,01	0,001	0,988235294117647
(32, 32)	relu	sgd	0,01	0,01	0,988235294117647
(32, 32)	relu	sgd	0,1	0,001	0,988235294117647
(32, 32)	relu	sgd	0,1	0,01	0,988235294117647
(32, 32)	relu	sgd	1	0,001	0,988235294117647
(32, 32)	relu	sgd	1	0,01	0,988235294117647
(32, 32)	relu	sgd	10	0,001	0,988235294117647
(32, 32)	relu	sgd	10	0,01	0,988235294117647
(32, 32)	relu	adam	0,001	0,01	0,988235294117647
(32, 32)	relu	adam	0,01	0,001	0,988235294117647
(32, 32)	relu	adam	0,01	0,01	0,988235294117647
(32, 32)	relu	adam	0,1	0,001	0,988235294117647
(32, 32)	relu	adam	0,1	0,01	0,988235294117647
(32, 32)	relu	adam	0,1	0,1	0,988235294117647
(32, 32)	relu	adam	1	0,001	0,988235294117647
(32, 32)	relu	adam	1	0,01	0,988235294117647
(32, 32)	relu	adam	10	0,001	0,988235294117647
(32, 32)	relu	adam	10	0,01	0,988235294117647
(64, 64)	logistic	sgd	0,001	0,001	0,988235294117647
(64, 64)	logistic	sgd	0,001	0,01	0,988235294117647
(64, 64)	logistic	sgd	0,001	0,1	0,988235294117647
(64, 64)	logistic	sgd	0,01	0,001	0,988235294117647
(64, 64)	logistic	sgd	0,01	0,01	0,988235294117647
(64, 64)	logistic	sgd	0,01	0,1	0,988235294117647
(64, 64)	logistic	sgd	0,1	0,001	0,988235294117647
(64, 64)	logistic	sgd	0,1	0,01	0,988235294117647
(64, 64)	logistic	sgd	0,1	0,1	0,988235294117647
(64, 64)	logistic	sgd	1	0,001	0,988235294117647
(64, 64)	logistic	sgd	1	0,01	0,988235294117647
(64, 64)	logistic	sgd	1	0,1	0,988235294117647
(64, 64)	logistic	sgd	10	0,001	0,988235294117647
(64, 64)	logistic	sgd	10	0,01	0,988235294117647
(64, 64)	logistic	adam	0,001	0,001	0,988235294117647
(64, 64)	logistic	adam	0,001	0,01	0,988235294117647
(64, 64)	logistic	adam	0,01	0,001	0,988235294117647
(64, 64)	logistic	adam	0,01	0,01	0,988235294117647
(64, 64)	logistic	adam	0,1	0,001	0,988235294117647
(64, 64)	logistic	adam	0,1	0,1	0,988235294117647
(64, 64)	logistic	adam	1	0,001	0,988235294117647
(64, 64)	logistic	adam	10	0,001	0,988235294117647
(64, 64)	tanh	sgd	0,001	0,01	0,988235294117647
(64, 64)	tanh	sgd	0,01	0,01	0,988235294117647
(64, 64)	tanh	sgd	0,01	0,1	0,988235294117647

(64, 64)	tanh	sgd	0,1	0,01	0,988235294117647
(64, 64)	tanh	sgd	0,1	0,1	0,988235294117647
(64, 64)	tanh	sgd	1	0,001	0,988235294117647
(64, 64)	tanh	sgd	1	0,01	0,988235294117647
(64, 64)	tanh	sgd	1	0,1	0,988235294117647
(64, 64)	tanh	sgd	10	0,001	0,988235294117647
(64, 64)	tanh	sgd	10	0,01	0,988235294117647
(64, 64)	tanh	adam	0,001	0,001	0,988235294117647
(64, 64)	tanh	adam	0,001	0,01	0,988235294117647
(64, 64)	tanh	adam	0,01	0,001	0,988235294117647
(64, 64)	tanh	adam	0,01	0,01	0,988235294117647
(64, 64)	tanh	adam	0,1	0,01	0,988235294117647
(64, 64)	tanh	adam	1	0,001	0,988235294117647
(64, 64)	tanh	adam	1	0,01	0,988235294117647
(64, 64)	tanh	adam	10	0,001	0,988235294117647
(64, 64)	tanh	adam	10	0,01	0,988235294117647
(64, 64)	relu	sgd	0,001	0,01	0,988235294117647
(64, 64)	relu	sgd	0,01	0,01	0,988235294117647
(64, 64)	relu	sgd	0,1	0,01	0,988235294117647
(64, 64)	relu	sgd	1	0,01	0,988235294117647
(64, 64)	relu	sgd	10	0,001	0,988235294117647
(64, 64)	relu	sgd	10	0,01	0,988235294117647
(64, 64)	relu	adam	0,001	0,01	0,988235294117647
(64, 64)	relu	adam	0,01	0,01	0,988235294117647
(64, 64)	relu	adam	0,1	0,01	0,988235294117647
(64, 64)	relu	adam	1	0,001	0,988235294117647
(64, 64)	relu	adam	1	0,01	0,988235294117647
(64, 64)	relu	adam	10	0,001	0,988235294117647
(64, 64)	relu	adam	10	0,01	0,988235294117647
(8, 8)	logistic	sgd	10	0,001	0,976470588235294
(8, 8)	logistic	adam	0,1	0,01	0,976470588235294
(8, 8)	tanh	sgd	0,001	0,001	0,976470588235294
(8, 8)	tanh	sgd	0,01	0,001	0,976470588235294
(8, 8)	tanh	sgd	0,1	0,001	0,976470588235294
(8, 8)	tanh	sgd	0,1	0,1	0,976470588235294
(8, 8)	tanh	sgd	1	0,001	0,976470588235294
(8, 8)	tanh	adam	10	0,01	0,976470588235294
(8, 8)	relu	adam	1	0,1	0,976470588235294
(8, 8)	relu	adam	10	0,1	0,976470588235294
(16, 16)	logistic	sgd	10	0,1	0,976470588235294
(16, 16)	logistic	adam	0,1	0,1	0,976470588235294
(16, 16)	logistic	adam	1	0,01	0,976470588235294
(16, 16)	logistic	adam	10	0,01	0,976470588235294
(16, 16)	tanh	sgd	0,001	0,001	0,976470588235294

(16, 16)	tanh	sgd	0,01	0,001	0,976470588235294
(16, 16)	tanh	sgd	0,1	0,001	0,976470588235294
(16, 16)	tanh	sgd	1	0,001	0,976470588235294
(16, 16)	relu	sgd	1	0,001	0,976470588235294
(16, 16)	relu	adam	0,001	0,1	0,976470588235294
(16, 16)	relu	adam	0,01	0,01	0,976470588235294
(32, 32)	logistic	adam	1	0,01	0,976470588235294
(32, 32)	logistic	adam	1	0,1	0,976470588235294
(32, 32)	tanh	sgd	10	0,1	0,976470588235294
(64, 64)	logistic	sgd	10	0,1	0,976470588235294
(64, 64)	logistic	adam	0,1	0,01	0,976470588235294
(64, 64)	logistic	adam	1	0,01	0,976470588235294
(64, 64)	tanh	sgd	0,001	0,001	0,976470588235294
(64, 64)	tanh	sgd	0,001	0,1	0,976470588235294
(64, 64)	tanh	sgd	0,01	0,001	0,976470588235294
(64, 64)	tanh	sgd	0,1	0,001	0,976470588235294
(64, 64)	tanh	sgd	10	0,1	0,976470588235294
(64, 64)	tanh	adam	0,1	0,001	0,976470588235294
(8, 8)	logistic	adam	0,1	0,1	0,964705882352941
(8, 8)	relu	adam	0,001	0,1	0,964705882352941
(8, 8)	relu	adam	0,1	0,1	0,964705882352941
(16, 16)	relu	adam	10	0,1	0,964705882352941
(32, 32)	logistic	adam	0,1	0,1	0,964705882352941
(32, 32)	tanh	sgd	1	0,1	0,964705882352941
(32, 32)	relu	adam	0,001	0,001	0,964705882352941
(64, 64)	logistic	adam	10	0,01	0,964705882352941
(64, 64)	relu	adam	10	0,1	0,964705882352941
(8, 8)	tanh	adam	1	0,01	0,952941176470588
(8, 8)	relu	adam	0,01	0,1	0,952941176470588
(16, 16)	logistic	adam	0,001	0,1	0,952941176470588
(16, 16)	logistic	adam	1	0,1	0,952941176470588
(32, 32)	logistic	adam	10	0,1	0,952941176470588
(32, 32)	tanh	sgd	0,001	0,1	0,952941176470588
(64, 64)	logistic	adam	1	0,1	0,952941176470588
(8, 8)	logistic	adam	0,001	0,1	0,941176470588235
(8, 8)	tanh	sgd	0,01	0,1	0,941176470588235
(8, 8)	tanh	sgd	10	0,1	0,941176470588235
(8, 8)	tanh	adam	0,001	0,1	0,941176470588235
(16, 16)	logistic	adam	0,01	0,1	0,941176470588235
(16, 16)	tanh	sgd	1	0,1	0,941176470588235
(32, 32)	logistic	adam	10	0,01	0,941176470588235
(32, 32)	tanh	adam	0,01	0,1	0,941176470588235
(64, 64)	logistic	adam	0,001	0,1	0,941176470588235
(8, 8)	tanh	sgd	1	0,1	0,929411764705882

(8, 8)	tanh	adam	0,01	0,1	0,929411764705882
(8, 8)	tanh	adam	10	0,1	0,929411764705882
(16, 16)	tanh	adam	10	0,1	0,929411764705882
(16, 16)	relu	adam	1	0,1	0,929411764705882
(32, 32)	tanh	adam	0,001	0,1	0,929411764705882
(32, 32)	tanh	adam	0,1	0,1	0,929411764705882
(64, 64)	tanh	adam	0,001	0,1	0,929411764705882
(64, 64)	tanh	adam	0,1	0,1	0,929411764705882
(64, 64)	tanh	adam	10	0,1	0,929411764705882
(8, 8)	logistic	sgd	10	0,1	0,917647058823529
(16, 16)	tanh	sgd	10	0,1	0,917647058823529
(16, 16)	tanh	adam	0,01	0,1	0,917647058823529
(16, 16)	tanh	adam	1	0,1	0,917647058823529
(64, 64)	logistic	adam	0,01	0,1	0,917647058823529
(64, 64)	tanh	adam	0,01	0,1	0,917647058823529
(8, 8)	logistic	adam	0,01	0,1	0,905882352941176
(8, 8)	logistic	adam	1	0,1	0,905882352941176
(8, 8)	tanh	adam	0,1	0,1	0,905882352941176
(32, 32)	tanh	adam	1	0,1	0,905882352941176
(64, 64)	logistic	adam	10	0,1	0,905882352941176
(64, 64)	relu	adam	1	0,1	0,905882352941176
(8, 8)	logistic	adam	10	0,1	0,894117647058824
(16, 16)	relu	adam	0,1	0,1	0,882352941176471
(32, 32)	tanh	adam	10	0,1	0,882352941176471
(32, 32)	relu	adam	10	0,1	0,858823529411765
(64, 64)	relu	adam	0,1	0,1	0,858823529411765
(8, 8)	tanh	adam	1	0,1	0,847058823529412
(16, 16)	tanh	adam	0,001	0,1	0,835294117647059
(64, 64)	tanh	adam	1	0,1	0,8
(16, 16)	tanh	adam	0,1	0,1	0,741176470588235
(64, 64)	relu	adam	0,001	0,1	0,682352941176471
(64, 64)	relu	adam	0,01	0,1	0,6
(8, 8)	relu	sgd	0,001	0,1	0,588235294117647
(8, 8)	relu	sgd	0,01	0,1	0,588235294117647
(8, 8)	relu	sgd	0,1	0,1	0,588235294117647
(8, 8)	relu	sgd	1	0,1	0,588235294117647
(8, 8)	relu	sgd	10	0,1	0,588235294117647
(16, 16)	relu	sgd	0,001	0,1	0,588235294117647
(16, 16)	relu	sgd	0,01	0,1	0,588235294117647
(16, 16)	relu	sgd	0,1	0,1	0,588235294117647
(16, 16)	relu	sgd	1	0,1	0,588235294117647
(16, 16)	relu	sgd	10	0,1	0,588235294117647
(32, 32)	relu	sgd	0,001	0,1	0,588235294117647
(32, 32)	relu	sgd	0,01	0,1	0,588235294117647

(32, 32)	relu	sgd	0,1	0,1	0,588235294117647
(32, 32)	relu	sgd	1	0,1	0,588235294117647
(32, 32)	relu	sgd	10	0,1	0,588235294117647
(32, 32)	relu	adam	0,001	0,1	0,588235294117647
(32, 32)	relu	adam	0,01	0,1	0,588235294117647
(32, 32)	relu	adam	1	0,1	0,588235294117647
(64, 64)	relu	sgd	0,001	0,1	0,588235294117647
(64, 64)	relu	sgd	0,01	0,1	0,588235294117647
(64, 64)	relu	sgd	0,1	0,1	0,588235294117647
(64, 64)	relu	sgd	1	0,1	0,411764705882353
(64, 64)	relu	sgd	10	0,1	0,411764705882353

Table 2: Results for Different Hyper Parameter Combinations for MLP Model