Turgut Alp Edis

21702587

GE461 Introduction to Data Science – Project 4

Part A

In this part of the project, initially principal component analysis (PCA) is applied, and first two principal components are taken. The projection of these two components can be seen in Figure 1.

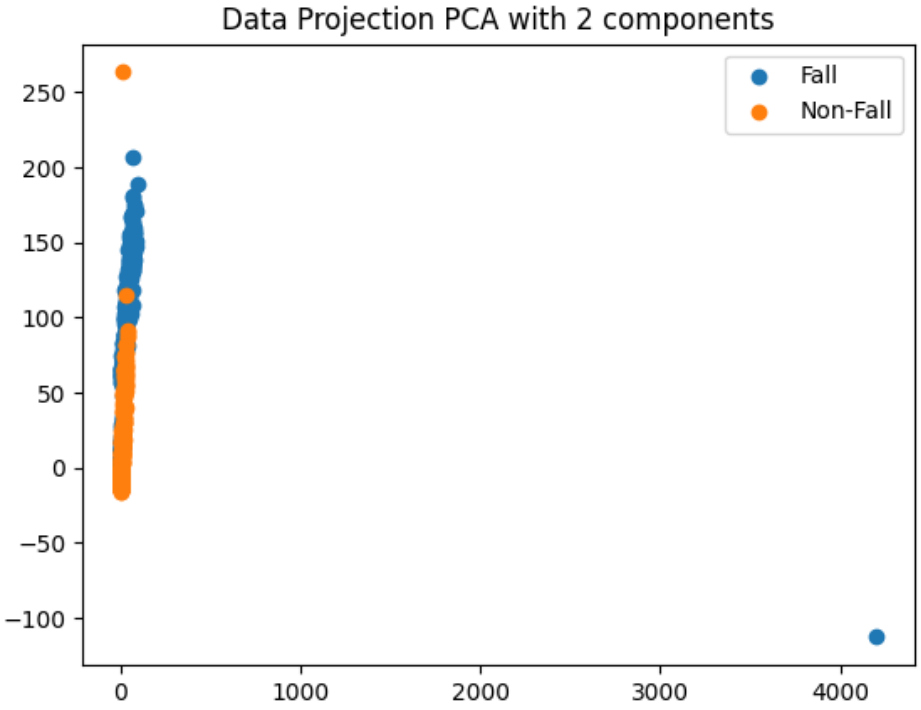

Figure 1. PCA Projection with 2 Principal Components

However, there are two values which make difficult to visualize whole PCA dataset. These values can be clearly seen in Figure 1 and can be interpreted that there are the minimum and maximum values of the components. Therefore, these values are removed from the dataset to display components more clearly as it can be seen in Figure 2.

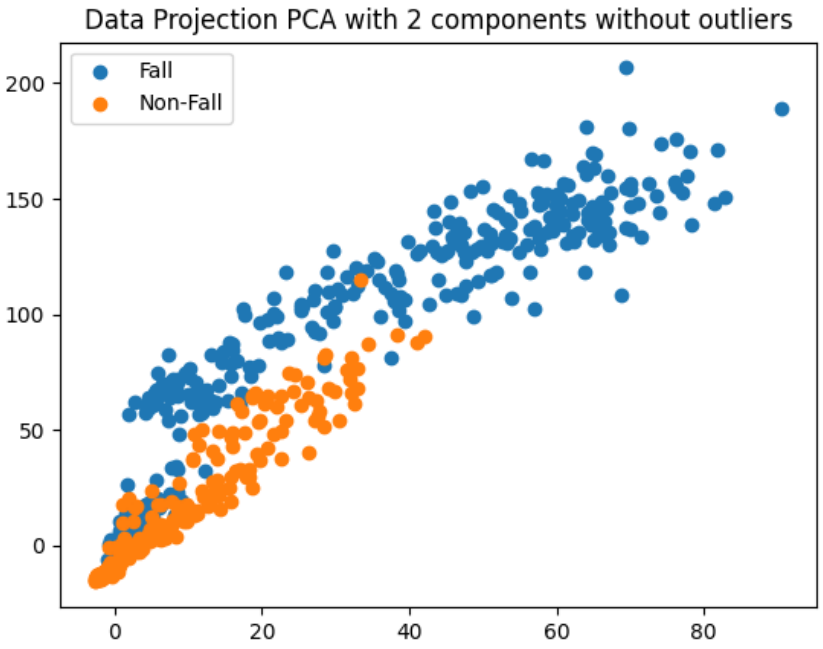

Figure 2. PCA Projection with 2 Principal Components Without Outliers

Since the data is not clear with outliers, the dataset without outliers is used for k-means clustering. Clusters can be observed in Figure 3 to 9.

Figure 3. K=2 Clustering Without Outliers

Figure 4. K=3 Clustering Without Outliers

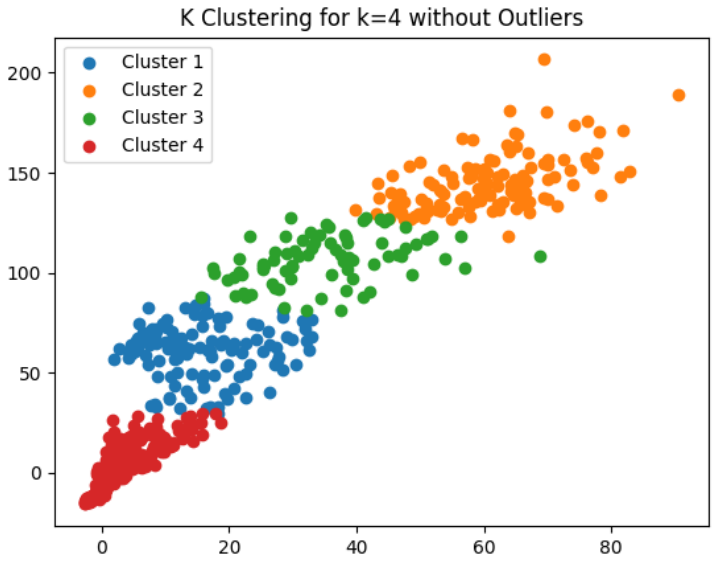

Figure 5. K=4 Clustering Without Outliers

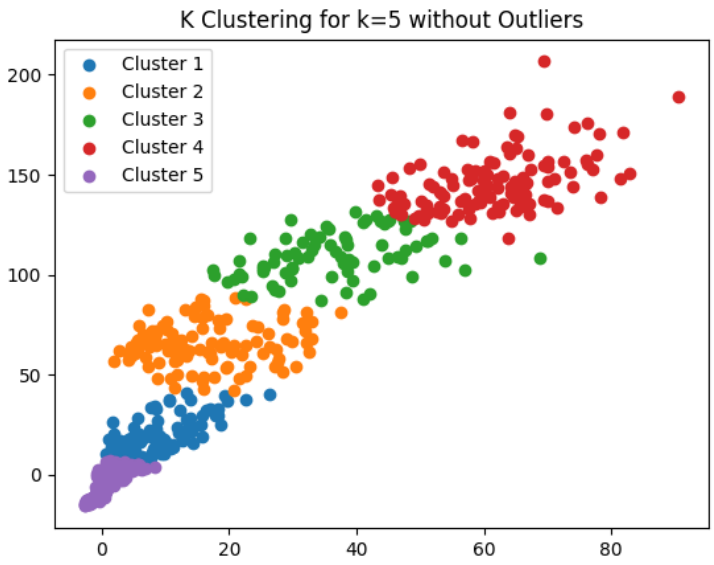

Figure 6. K=5 Clustering Without Outliers

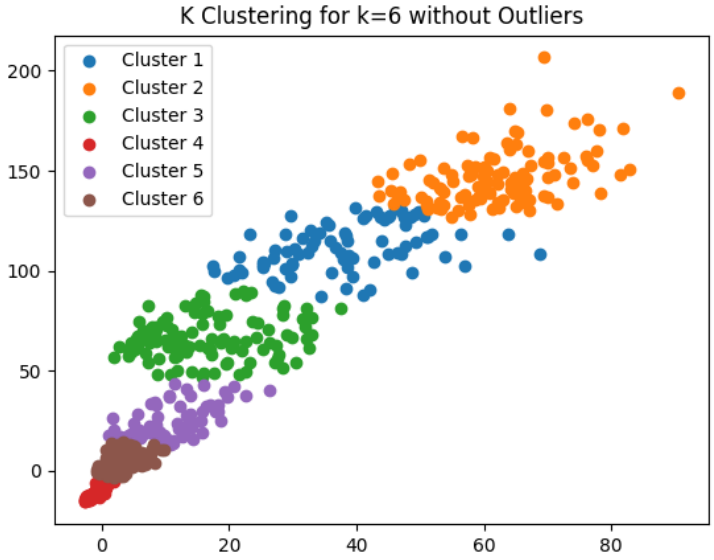

Figure 7. K=6 Clustering Without Outliers

Figure 8. K=7 Clustering Without Outliers

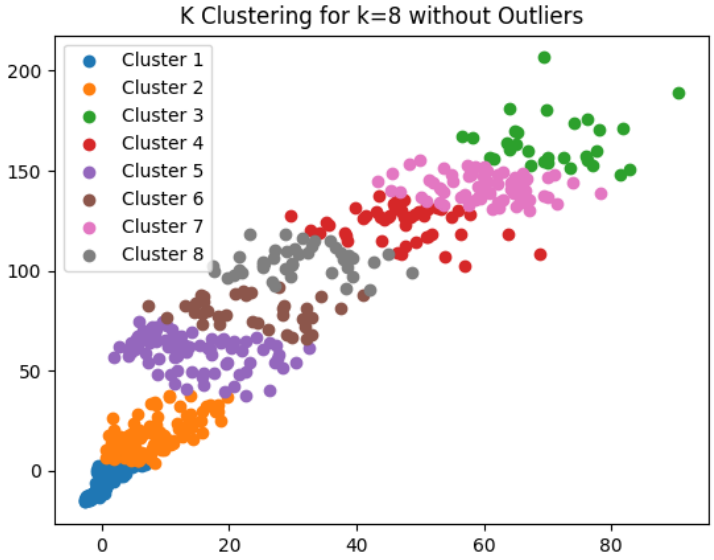

Figure 9. K=8 Clustering Without Outliers

As it can be seen from Figure 3 to 9, after K=4, overlapping of the points is started. It is enough to display 7 figures that includes K=2 to K=8 since after K=4, the data points start to overlap with each other.

When K=2, consistency is so low that I could not calculate the percentage because the shapes of the clusters and PCA are not fit. However, when the calculation is done according to the values, then, 51 of 566 datapoints are misclassified. Therefore, the consistency is 91 % which is quite large.

As a result, K means clustering can be used for Fall Detection.

Part B

From the sklearn's SVM class, regularization parameter, kernel type and degree are used to find the best model on the dataset. Different combinations of these values are used.

Used Parameters:

Regularization Parameter: [1e-3, 1e-2, 1e-1, 1e0, 1e1]

Degree: [1,2,3,4,5]

Kernel: [Linear, Poly, Rbf, Sigmoid]

The result of each combination is in Table 1. Their results are in descending order in order to see the best SVM parameters. According to the table, the best parameters are the top 10 of this table. Though, it is not 100 % accurate, its accuracy is big enough to ignore the error which is 98.8 %.

Regularization Parameter	Kernel Type	Degree	Validation Accuracy
0.001	linear		0.988235
0.01	linear		0.988235
0.1	linear		0.988235
1	linear		0.988235
1	poly	[1, 2, 3, 4, 5]	0.988235
1	rbf		0.988235
10	linear		0.988235
10	poly	[1, 2, 3, 4, 5]	0.988235
10	poly	[1, 2, 3, 4, 5]	0.988235
10	rbf		0.988235
1	sigmoid		0.964706
10	sigmoid		0.952941
1	poly	[1, 2, 3, 4, 5]	0.929412
10	poly	[1, 2, 3, 4, 5]	0.929412
0.1	rbf		0.917647
0.1	poly	[1, 2, 3, 4, 5]	0.905882
1	poly	[1, 2, 3, 4, 5]	0.905882
0.1	sigmoid		0.894118
10	poly	[1, 2, 3, 4, 5]	0.894118
0.1	poly	[1, 2, 3, 4, 5]	0.870588
10	poly	[1, 2, 3, 4, 5]	0.870588
0.1	poly	[1, 2, 3, 4, 5]	0.847059
1	poly	[1, 2, 3, 4, 5]	0.776471
1	poly	[1, 2, 3, 4, 5]	0.647059
0.001	poly	[1, 2, 3, 4, 5]	0.588235
0.001	poly	[1, 2, 3, 4, 5]	0.588235
0.001	poly	[1, 2, 3, 4, 5]	0.588235
0.001	poly	[1, 2, 3, 4, 5]	0.588235
0.001	poly	[1, 2, 3, 4, 5]	0.588235
0.001	rbf		0.588235
0.001	sigmoid		0.588235
0.01	poly	[1, 2, 3, 4, 5]	0.588235
0.01	poly	[1, 2, 3, 4, 5]	0.588235
0.01	poly	[1, 2, 3, 4, 5]	0.588235
0.01	poly	[1, 2, 3, 4, 5]	0.588235
0.01	poly	[1, 2, 3, 4, 5]	0.588235
0.01	rbf		0.588235
0.01	sigmoid		0.588235
0.1	poly	[1, 2, 3, 4, 5]	0.588235
0.1	poly	[1, 2, 3, 4, 5]	0.588235

Table 1. Results of SVM Classifier

From sklearn's another library MLPClassifier, the parameters are quite different than SVM but the alpha value of MLP is similar as regularization value of SVM.

Used Parameters:

Alpha: [1e-3, 1e-2, 1e-1, 1e0, 1e1]

Hidden layer: [(8, 8), (16, 16), (32, 32), (64, 64)]

Activation Function: [logistic, tanh, relu]

Learning rates: [1e-3, 1e-2, 1e-1]

The result of each combination is in Table 2. Their results are in descending order in order to see the best MLP parameters. According to the table, the best parameters are the first 89 combinations of the parameters in the Table 2. The amount of correct combinations are quite large compared to SVM, so it can be interpreted that MLP classifier is the best parameter for classification of data.

Hidden Layer Size	Activation Function	Regularization Size (Alpha)	Learning Rate	Validation Accuracy
(8, 8)	logistic	0.001	0.001	1
(8, 8)	logistic	0.001	0.01	1
(8, 8)	logistic	0.01	0.001	1
(8, 8)	logistic	0.01	0.01	1
(8, 8)	logistic	0.01	0.1	1
(8, 8)	logistic	1	0.01	1
(8, 8)	logistic	10	0.01	1
(8, 8)	tanh	0.001	0.001	1
(8, 8)	tanh	0.001	0.01	1
(8, 8)	tanh	0.01	0.001	1
(8, 8)	tanh	0.01	0.01	1
(8, 8)	tanh	0.1	0.001	1
(8, 8)	tanh	0.1	0.01	1
(8, 8)	tanh	1	0.001	1
(8, 8)	tanh	10	0.01	1
(8, 8)	relu	0.001	0.001	1
(8, 8)	relu	0.001	0.01	1
(8, 8)	relu	0.001	0.1	1
(8, 8)	relu	0.01	0.001	1
(8, 8)	relu	0.01	0.01	1
(8, 8)	relu	0.01	0.1	1

(8, 8)	relu	0.1	0.001	1
(8, 8)	relu	1	0.001	1
(8, 8)	relu	1	0.01	1
(8, 8)	relu	1	0.1	1
(16, 16)	logistic	0.001	0.001	1
(16, 16)	logistic	0.001	0.01	1
(16, 16)	logistic	0.001	0.1	1
(16, 16)	logistic	0.01	0.001	1
(16, 16)	logistic	0.01	0.01	1
(16, 16)	logistic	1	0.01	1
(16, 16)	logistic	10	0.01	1
(16, 16)	tanh	0.001	0.001	1
(16, 16)	tanh	0.001	0.01	1
(16, 16)	tanh	0.01	0.001	1
(16, 16)	tanh	0.01	0.01	1
(16, 16)	tanh	0.1	0.001	1
(16, 16)	tanh	0.1	0.01	1
(16, 16)	tanh	1	0.001	1
(16, 16)	relu	0.001	0.001	1
(16, 16)	relu	0.001	0.01	1
(16, 16)	relu	0.001	0.1	1
(16, 16)	relu	0.01	0.001	1
(16, 16)	relu	0.01	0.01	1
(16, 16)	relu	0.1	0.001	1
(16, 16)	relu	0.1	0.01	1
(16, 16)	relu	0.1	0.1	1
(16, 16)	relu	1	0.001	1
(16, 16)	relu	10	0.01	1
(32, 32)	logistic	0.001	0.001	1
(32, 32)	logistic	0.001	0.01	1
(32, 32)	logistic	0.001	0.1	1
(32, 32)	logistic	0.01	0.001	1
(32, 32)	logistic	0.01	0.01	1
(32, 32)	logistic	0.1	0.001	1
(32, 32)	logistic	0.1	0.01	1
(32, 32)	logistic	1	0.01	1
(32, 32)	tanh	0.001	0.001	1
(32, 32)	tanh	0.001	0.01	1
(32, 32)	tanh	0.01	0.001	1
(32, 32)	tanh	0.01	0.01	1
(32, 32)	tanh	0.1	0.001	1

(32, 32)	tanh	0.1	0.01	1
(32, 32)	relu	0.001	0.001	1
(32, 32)	relu	0.001	0.01	1
(32, 32)	relu	0.01	0.001	1
(32, 32)	relu	0.01	0.01	1
(32, 32)	relu	0.1	0.001	1
(32, 32)	relu	0.1	0.01	1
(32, 32)	relu	1	0.01	1
(32, 32)	relu	10	0.1	1
(64, 64)	logistic	0.001	0.001	1
(64, 64)	logistic	0.001	0.01	1
(64, 64)	logistic	0.01	0.001	1
(64, 64)	logistic	0.01	0.01	1
(64, 64)	logistic	0.01	0.1	1
(64, 64)	logistic	0.1	0.001	1
(64, 64)	logistic	1	0.1	1
(64, 64)	tanh	0.001	0.001	1
(64, 64)	tanh	0.001	0.01	1
(64, 64)	tanh	0.01	0.001	1
(64, 64)	tanh	0.01	0.01	1
(64, 64)	tanh	0.1	0.001	1
(64, 64)	tanh	0.1	0.01	1
(64, 64)	relu	0.001	0.001	1
(64, 64)	relu	0.001	0.01	1
(64, 64)	relu	0.01	0.001	1
(64, 64)	relu	0.1	0.001	1
(64, 64)	relu	0.1	0.01	1
(8, 8)	logistic	0.001	0.1	0.988235
(8, 8)	logistic	0.1	0.001	0.988235
(8, 8)	logistic	0.1	0.01	0.988235
(8, 8)	logistic	1	0.001	0.988235
(8, 8)	logistic	10	0.001	0.988235
(8, 8)	tanh	1	0.01	0.988235
(8, 8)	tanh	10	0.001	0.988235
(8, 8)	relu	0.1	0.01	0.988235
(8, 8)	relu	0.1	0.1	0.988235
(8, 8)	relu	10	0.001	0.988235
(8, 8)	relu	10	0.01	0.988235
(16, 16)	logistic	0.1	0.001	0.988235
(16, 16)	logistic	0.1	0.1	0.988235
(16, 16)	logistic	1	0.001	0.988235

(16, 16)	logistic	10	0.001	0.988235
(16, 16)	tanh	10	0.001	0.988235
(16, 16)	tanh	10	0.01	0.988235
(16, 16)	relu	1	0.01	0.988235
(16, 16)	relu	10	0.001	0.988235
(32, 32)	logistic	1	0.001	0.988235
(32, 32)	logistic	10	0.001	0.988235
(32, 32)	tanh	1	0.001	0.988235
(32, 32)	tanh	1	0.01	0.988235
(32, 32)	tanh	10	0.001	0.988235
(32, 32)	tanh	10	0.01	0.988235
(32, 32)	relu	1	0.001	0.988235
(32, 32)	relu	1	0.1	0.988235
(32, 32)	relu	10	0.001	0.988235
(32, 32)	relu	10	0.01	0.988235
(64, 64)	logistic	1	0.001	0.988235
(64, 64)	logistic	10	0.001	0.988235
(64, 64)	logistic	10	0.01	0.988235
(64, 64)	tanh	1	0.001	0.988235
(64, 64)	tanh	10	0.001	0.988235
(64, 64)	tanh	10	0.01	0.988235
(64, 64)	relu	0.001	0.1	0.988235
(64, 64)	relu	0.01	0.01	0.988235
(64, 64)	relu	1	0.001	0.988235
(64, 64)	relu	10	0.001	0.988235
(64, 64)	relu	10	0.01	0.988235
(8, 8)	logistic	0.1	0.1	0.976471
(8, 8)	logistic	1	0.1	0.976471
(8, 8)	relu	10	0.1	0.976471
(16, 16)	logistic	0.1	0.01	0.976471
(16, 16)	tanh	0.01	0.1	0.976471
(16, 16)	tanh	1	0.01	0.976471
(32, 32)	logistic	0.01	0.1	0.976471
(32, 32)	logistic	1	0.1	0.976471
(32, 32)	relu	0.001	0.1	0.976471
(64, 64)	logistic	0.001	0.1	0.976471
(64, 64)	logistic	0.1	0.1	0.976471
(64, 64)	logistic	1	0.01	0.976471
(64, 64)	relu	1	0.1	0.976471
(8, 8)	tanh	0.001	0.1	0.964706
(16, 16)	tanh	0.001	0.1	0.964706

(32, 32)	logistic	10	0.01	0.964706
(64, 64)	tanh	1	0.01	0.964706
(16, 16)	logistic	1	0.1	0.952941
(16, 16)	tanh	0.1	0.1	0.952941
(32, 32)	tanh	0.1	0.1	0.952941
(64, 64)	tanh	0.001	0.1	0.952941
(64, 64)	tanh	0.01	0.1	0.952941
(64, 64)	tanh	1	0.1	0.952941
(16, 16)	logistic	0.01	0.1	0.941176
(16, 16)	tanh	1	0.1	0.941176
(32, 32)	logistic	10	0.1	0.941176
(32, 32)	tanh	0.01	0.1	0.941176
(64, 64)	relu	0.01	0.1	0.941176
(64, 64)	relu	10	0.1	0.941176
(8, 8)	tanh	0.01	0.1	0.929412
(16, 16)	logistic	10	0.1	0.929412
(16, 16)	tanh	10	0.1	0.917647
(16, 16)	relu	10	0.1	0.917647
(32, 32)	tanh	10	0.1	0.917647
(64, 64)	relu	1	0.01	0.917647
(8, 8)	tanh	0.1	0.1	0.905882
(32, 32)	logistic	0.1	0.1	0.905882
(64, 64)	logistic	10	0.1	0.905882
(8, 8)	logistic	10	0.1	0.894118
(64, 64)	tanh	0.1	0.1	0.894118
(8, 8)	tanh	10	0.1	0.882353
(32, 32)	tanh	1	0.1	0.882353
(64, 64)	tanh	10	0.1	0.858824
(32, 32)	tanh	0.001	0.1	0.847059
(8, 8)	tanh	1	0.1	0.823529
(64, 64)	relu	0.1	0.1	0.823529
(64, 64)	logistic	0.1	0.01	0.8
(16, 16)	relu	1	0.1	0.635294
(32, 32)	relu	0.01	0.1	0.635294
(16, 16)	relu	0.01	0.1	0.611765
(32, 32)	relu	0.1	0.1	0.588235

Table 2. Results of MLP Classifier

As a result, the data points of the result of top 2 PC of PCA of fall detection dataset is quite confusing due to the outliers. When the outliers removed each point can be seen much more clearly. When the k-means classifier is applied, after K=4, the points are overlapping. When K=2, this classifier can be used for fall detection since the error percentage is low. However, SVM and MLP classifiers are better than k means since their error percentage is very low which can be observed in table 1 and 2. When it is required to choose one of these top 2 classifiers (SVM and MLP), choosing MLP can be more accurate since its error percentage is 0 for 89 different combinations.