Esad İsmail Tök
21801679
GE461-01

GE 461: Introduction to Data Science

Project W13: Telehealth- Fall Detection

PART-A

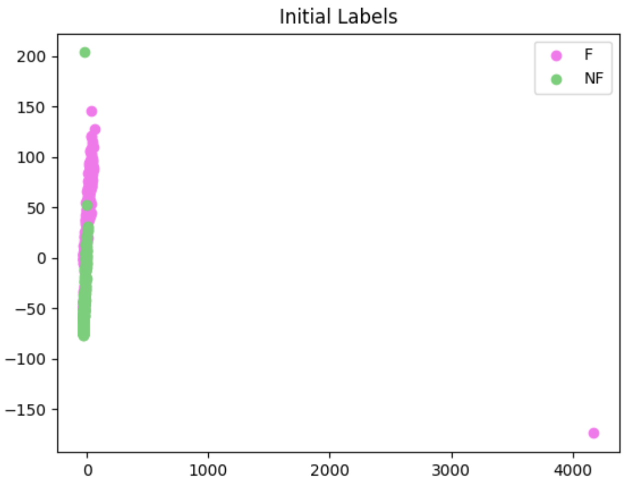


Figure-1

The dataset contains two outliers, one with the highest value on the first principal component (x-axis) and the other with the highest value on the second principal component (y-axis). Aside from these outliers, the data exhibits a reasonable level of separability. However, the presence of these outliers can potentially pose challenges when applying the K-means algorithm.

The first principal component accounts for 75.3072% of the total variance in the data, while the second principal component explains 8.51159% of the variance.

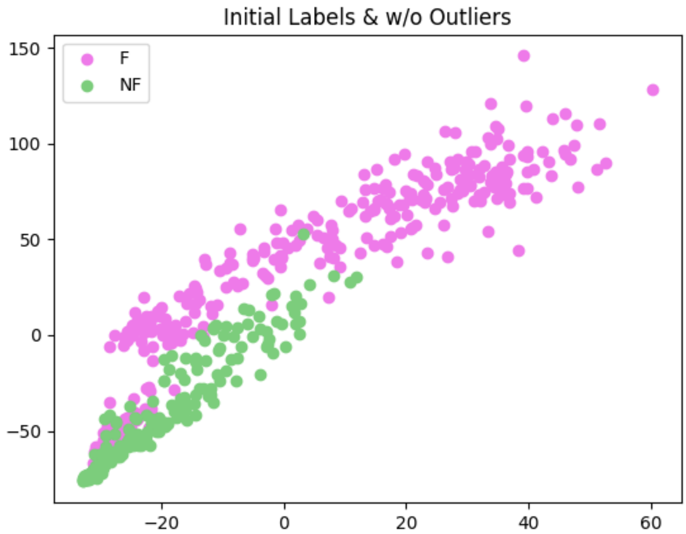


Figure-2

Figure-2 demonstrates the effectiveness of PCA in addressing the distortion caused by the two outliers when focusing on the left section of the initial plot.

Esad İsmail Tök
21801679
GE461-01

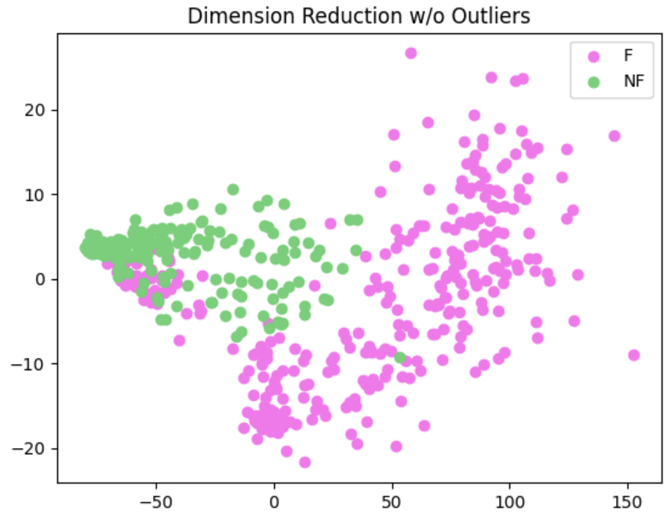


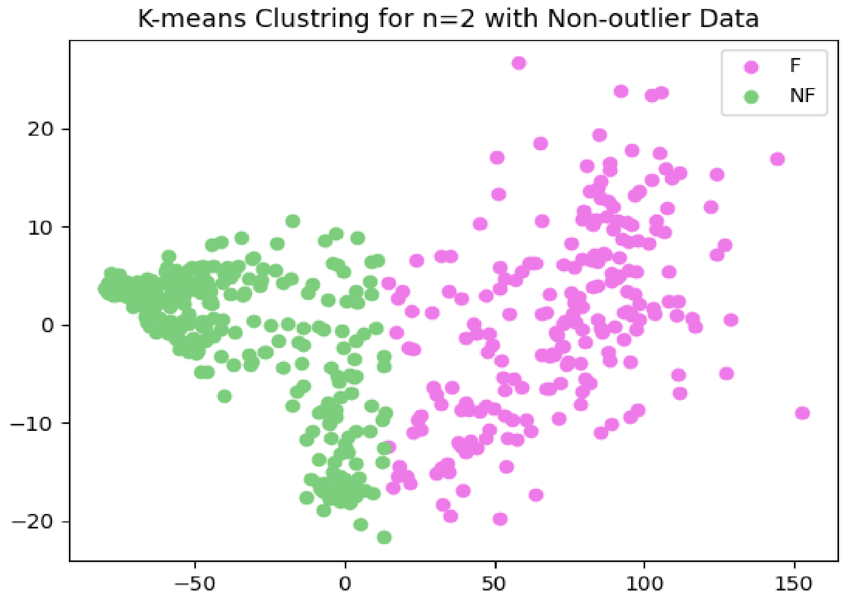
Figure-3

Figure-3 illustrates the outcome of performing PCA on the dataset without the outliers, reducing the samples from 566 to 564. A visual inspection suggests that this approach yields a more favorable representation for unsupervised learning, as it produces improved results in the absence of outliers.

Figure-4

As anticipated, the performance of the k-means algorithm is hindered when there exist outliers with significantly larger values compared to the rest of the data. In this case, the pink cluster, which is expected to represent the "F" class, has been manually labeled as such. However, the results are unsatisfactory in the opposite direction as well.

Esad İsmail Tök
21801679
GE461-01



When applying the K-means algorithm to the dimension-reduced data without outliers, significantly improved results are obtained compared to the previous approach.

Figure-5

During multiple iterations of the K-means algorithm with increasing values of k, the results are generally reasonable, except for the outlier in the bottom right. These results are particularly useful when searching for sub-clusters in the data. However, in our specific case where there are only two classes (F and NF), increasing k does not provide any meaningful insights about the data.

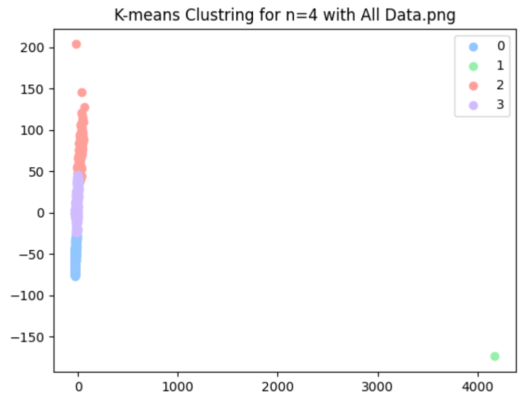


Figure-6

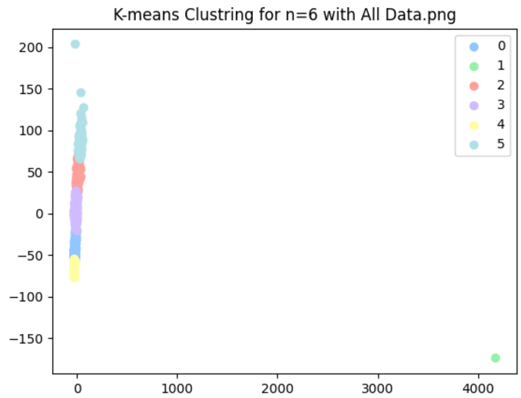


Figure-7

Esad İsmail Tök
21801679
GE461-01

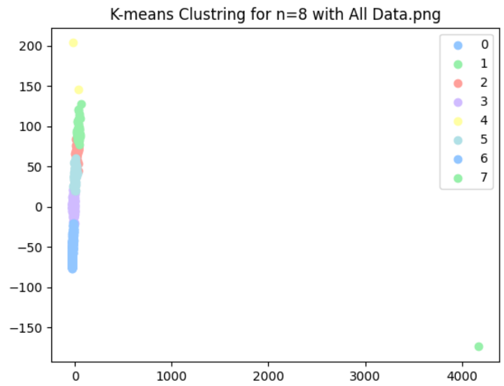


Figure-8

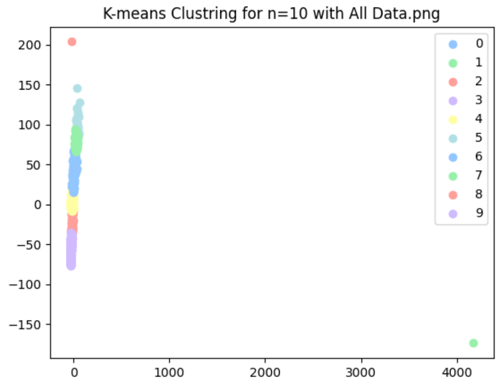


Figure-9

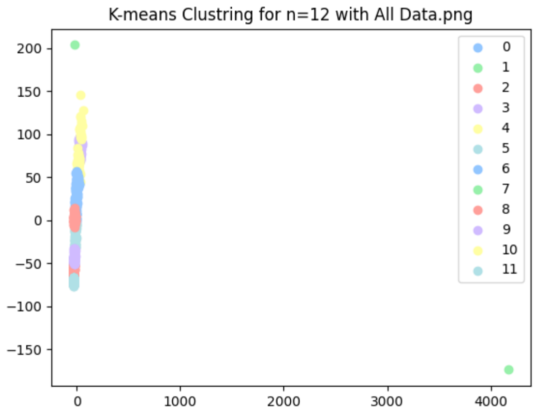


Figure-10

In conclusion, through exploration and analysis, it can be inferred that fall detection is feasible using the provided data. The insights gained from unsupervised learning techniques provide sufficient understanding and information about the data to support fall detection.

Esad İsmail Tök
21801679
GE461-01

PART-B

Results for SVM:

Hyperparameters: C = 1.0, degree = 4, max_iter = 100, shrinking = True

SVM Validation Accuracy: 100.00000%

SVM Test Accuracy: 95.74468%

Hyperparameters: C = 1.0, degree = 4, max_iter = 1000, shrinking = True

SVM Validation Accuracy: 100.00000%

SVM Test Accuracy: 97.87234%

Hyperparameters: C = 1.0, degree = 4, max_iter = 1000, shrinking = False

SVM Validation Accuracy: 97.36842%

SVM Test Accuracy: 100.00000%

Hyperparameters: C = 1.0, degree = 8, max_iter = 1000, shrinking = False

SVM Validation Accuracy: 100.00000%

SVM Test Accuracy: 95.74468%

Hyperparameters: C = 1.0, degree = 8, max_iter = 10000, shrinking = False

SVM Validation Accuracy: 97.36842%

SVM Test Accuracy: 97.87234%

Hyperparameters: C = 1.0, degree = 12, max_iter = 10000, shrinking = True, kernel = 'linear'

SVM Validation Accuracy: 100.00000%

SVM Test Accuracy: 97.87234%

Esad İsmail Tök
21801679
GE461-01

Hyperparameters: C = 1.0, degree = 12, max_iter = 10000, shrinking = False, kernel = 'linear'

SVM Validation Accuracy: 97.36842%

SVM Test Accuracy: 100.00000%

Results for MLP:

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(20, 2), random_state=42, learning_rate_init = 0.0001, max_iter = 200

MLP Validation Accuracy: 52.63158%

MLP Test Accuracy: 57.44681%

Hyperparameters: solver='adam', alpha=1e-5, hidden_layer_sizes=(20, 2), random_state=42, learning_rate_init = 0.0001, max_iter = 200

MLP Validation Accuracy: 73.68421%

MLP Test Accuracy: 82.97872%

Hyperparameters: solver='adam', alpha=1e-5, hidden_layer_sizes=(10, 2), random_state=42, learning_rate_init = 0.0001, max_iter = 200

MLP Validation Accuracy: 89.47368%

MLP Test Accuracy: 85.10638%

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(10, 2), random_state=42, learning_rate_init = 0.0001, max_iter = 200

MLP Validation Accuracy: 97.36842%

MLP Test Accuracy: 97.87234%

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=42, learning_rate_init = 0.0001, max_iter = 200

MLP Validation Accuracy: 100.00000%

MLP Test Accuracy: 97.87234%

Esad İsmail Tök
21801679
GE461-01

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=42, learning_rate_init = 0.001, max_iter = 200

MLP Validation Accuracy: 97.36842%

MLP Test Accuracy: 100.00000%

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=42, learning_rate_init = 0.01, max_iter = 200

MLP Validation Accuracy: 97.36842%

MLP Test Accuracy: 97.87234%

Hyperparameters: solver='adam', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=42, learning_rate_init = 0.1, max_iter = 200

MLP Validation Accuracy: 94.73684%

MLP Test Accuracy: 93.61702%

Hyperparameters: solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(5, 2), random_state=42, learning_rate_init = 0.1, max_iter = 200

MLP Validation Accuracy: 100.00000%

MLP Test Accuracy: 100.00000%

Ultimately, both algorithms yield excellent results, with MLP even achieving 100% accuracy. Both algorithms produce satisfactory outcomes, indicating the success of fall detection using wearable sensors.