Human Activity Recognition Project Report

1. Project Description

In this project, different machine learning algorithms are applied to classify human physical activities (sitting, standing, walking, running, going upstairs, going downstairs) carried out by humans based on accelerometer data. The aim is to determine the best classification method by comparing the accuracies and success measures of the algorithms.

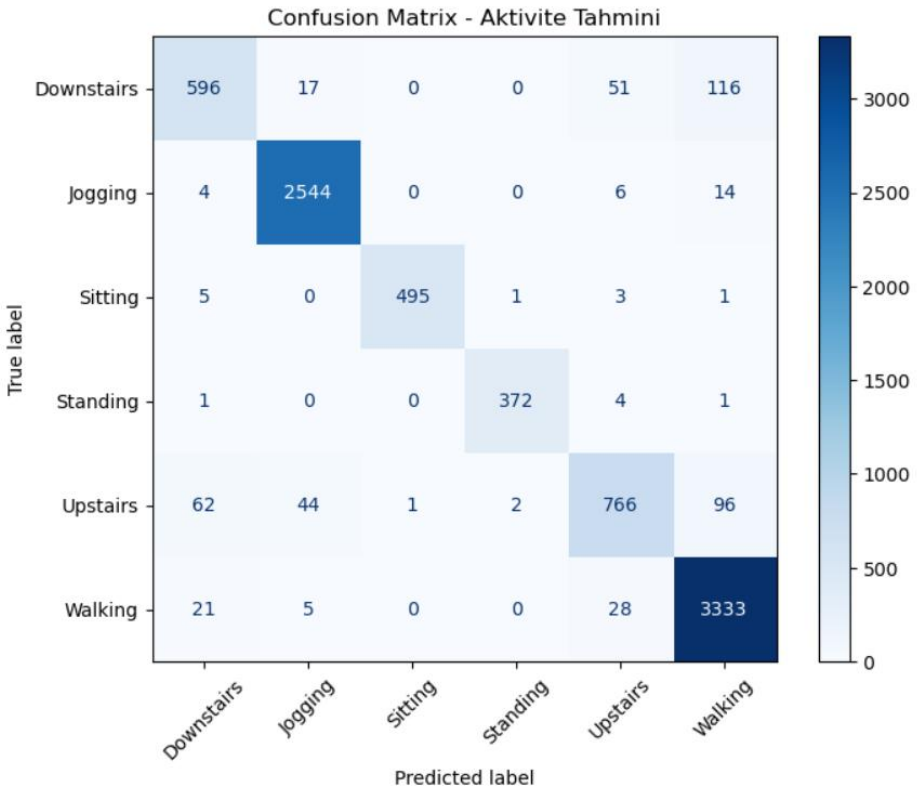
2. Dataset and Feature Extraction

The dataset includes time series accelerometer data in x, y, z axes. Statistical features (mean, standard deviation, min, max) were calculated from the data with the windowing method. 12 features were generated for each window and the most dominant activity during the window time was considered as the label.

3. Models Used

3.1 Random Forest

Random Forest is an ensemble learning technique in which numerous decision trees are combined to create a more robust and powerful classification model. 100 trees were utilized



Şekil 1: Random Forest – Sınıf Bazlı Precision / Recall / F1-Score

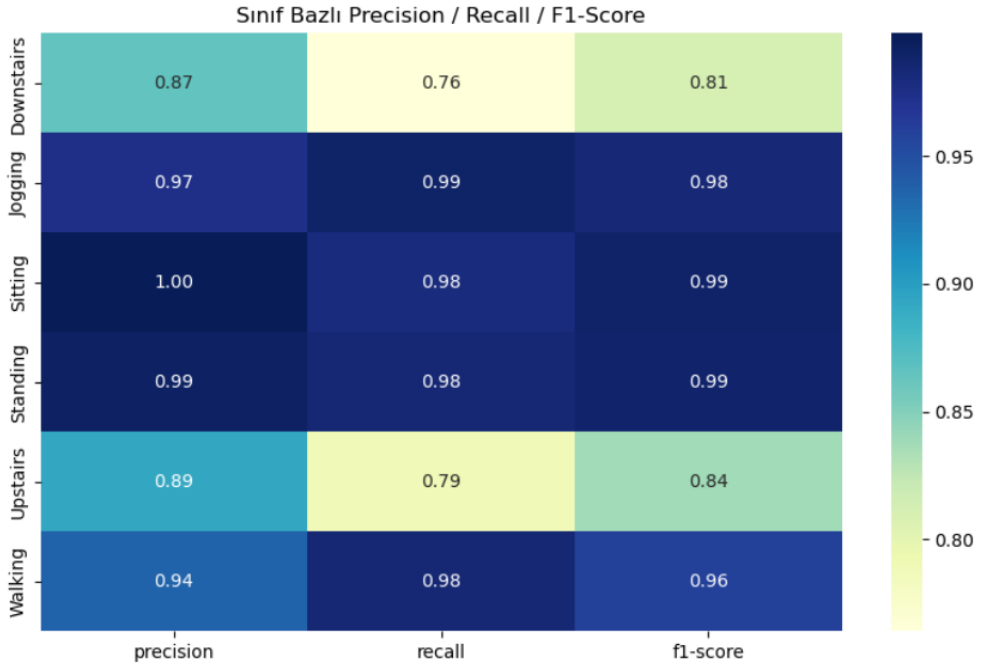
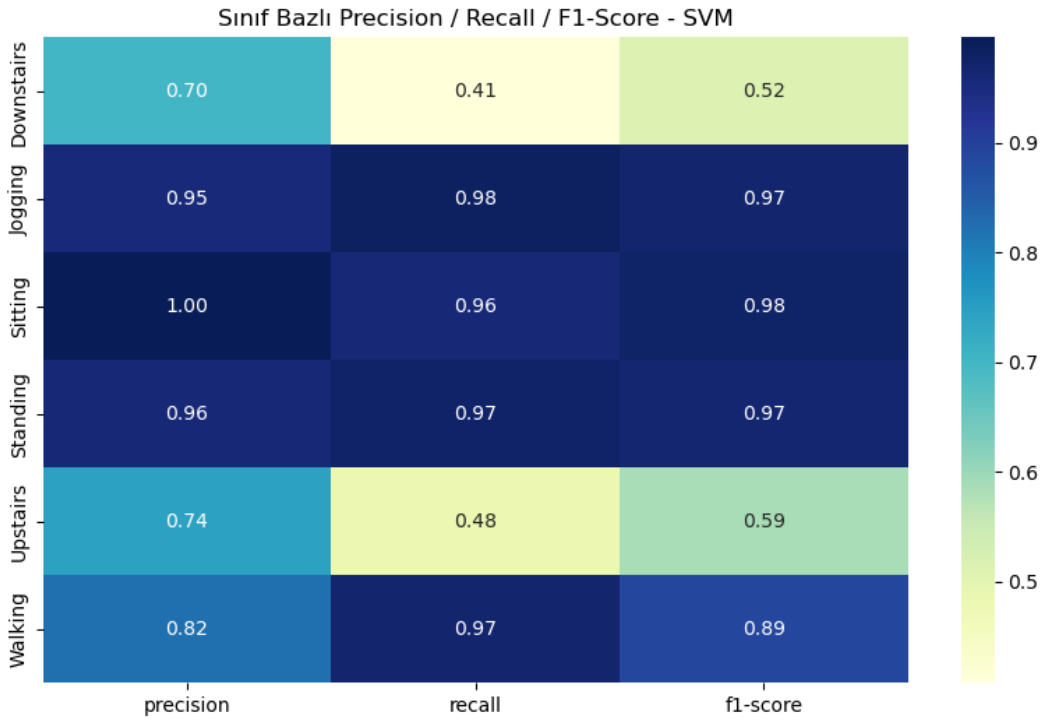


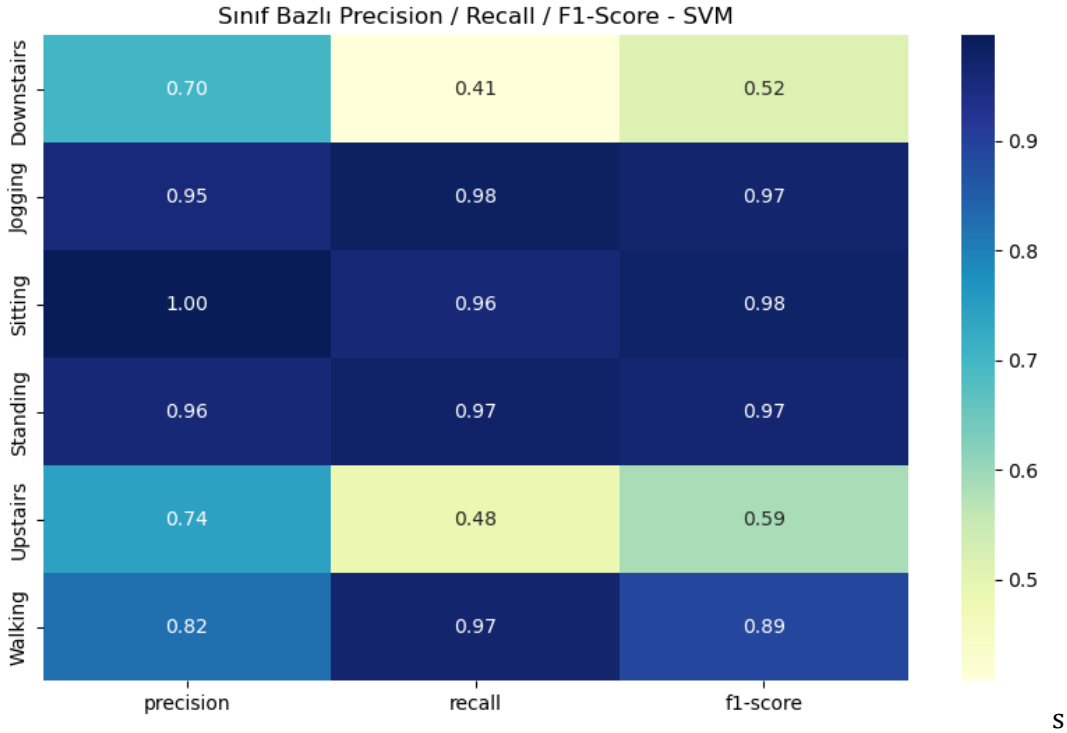
Figure 2: Random Forest Confusion Matrix

Random Forest model exhibited more than 94% accuracy with significant performance particularly in activities such as ambulation, sitting, and running. There was relatively decreased performance metrics for both stair climbing and descending tasks; however, overall results were excellent.

3.2 Decision Tree The Decision Tree is an interpretable and comprehensible machine learning model in which the classification of data is achieved by branching based on some specified criteria. Despite the huge benefits of this model, particularly in understandability, it may demonstrate limited performance in capturing sophisticated patterns.patterns.



Şekil 3: Decision Tree – Sınıf Bazlı Precision / Recall / F1-Score



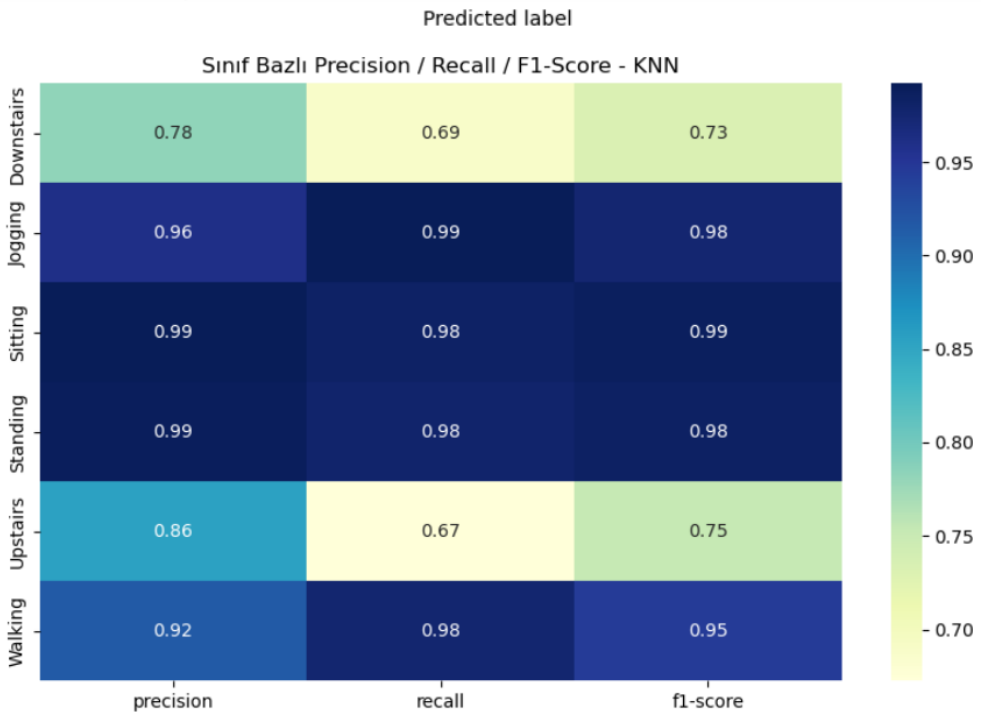
s

Figure 4: Decision Tree – Confusion Matrix

The Decision Tree model worked effectively in sitting, standing and running activities (f1-score: 97–99%). There was low success in Downstairs and Upstairs activities. This suggests that the model has difficulty distinguishing between complex transition activities. If we examine the confusion matrix, it can be seen that these classes are confused with one another.

3.3 K-Nearest Neighbors (KNN)

The K-Nearest Neighbors model makes predictions by finding the nearest neighbors in the training data in classification problems. The model is very effective, especially when the dataset is reasonably clean and the level of similarity is high. However, its performance is slow when used on large datasets with many instances.



Şekil 5: KNN – Sınıf Bazlı Precision / Recall / F1-Score

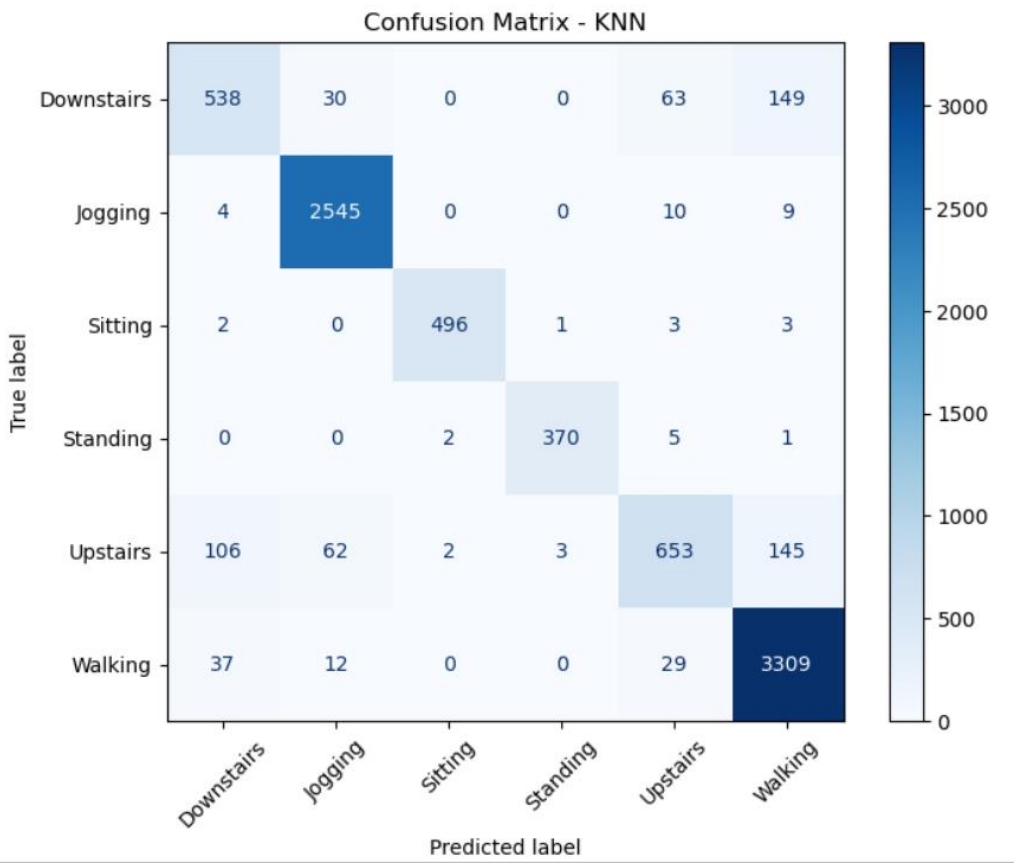
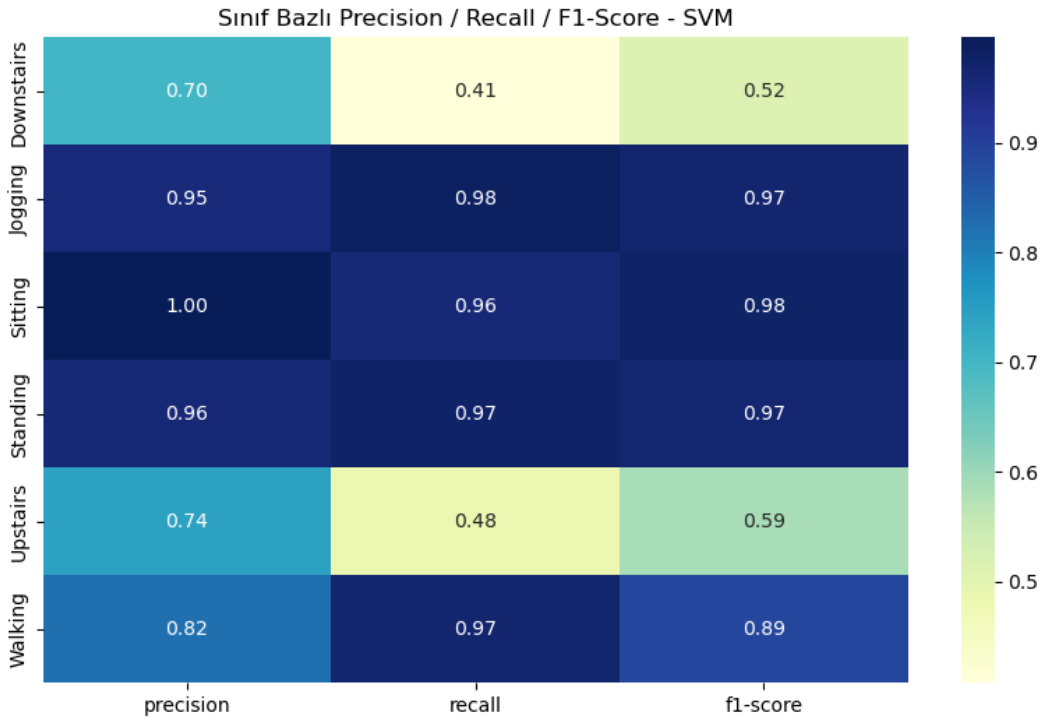


Figure 6: K-Nearest Neighbors – Confusion Matrix

KNN model has achieved very high success in sitting, standing, walking and running activities. f1-score values are between 95% and 99%. Success is low in climbing and going down stairs (Downstairs, Upstairs) activities. Especially in the Upstairs class, the recall value dropped to 67%. This shows that the model is unable to distinguish transition movements.

3.4 Support Vector Machine (SVM) The Support Vector Machine (SVM) is an effective classification algorithm that proves to be efficient, particularly in high-dimensional data. It aims to find the hyperplane that maximally separates data points while generally optimizing the margin. This model achieves considerable success in



Şekil 7: SVM – Sınıf Bazlı Precision / Recall / F1-Score

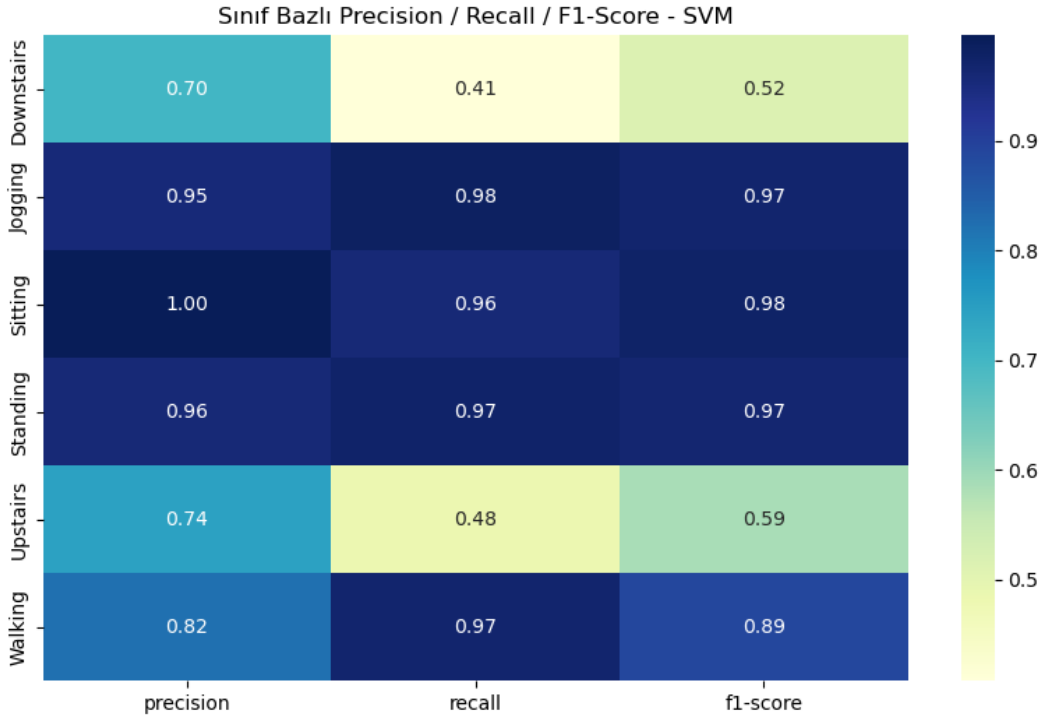


Figure 8: SVM – Confusion Matrix

The SVM model had good f1-scores (96–98%) for running, sitting and standing activities. But the success in the Downstairs and especially Upstairs activities was drastically reduced. The f1-score for Upstairs was only 59% and for Downstairs 52%. This situation shows that the model is struggling to distinguish transition movements.

4. Comparison of Model Performance

The table below will compare the accuracy, precision, recall, and F1-score mean of all the models. This will enable the selection of the best model.

5. Conclusion In the present study, human activities were systematically classified by comparing various machine learning models. Surprisingly, the Random Forest and K-Nearest Neighbors (KNN) models performed exceptionally well, particularly with distinct classes. Nevertheless, there was compromised performance in all the models regarding transitional movements, i.e., climbing stairs. Improved model performance in the future could be achieved by employing other windowing techniques and data augmentation strategies.

Muhammet İhsan Çakır 030722072

Emirhan Koç 030721088

Yahya Çınar 030721056