HUMAN ACTIVITY RECOGNITION WITH SMARTPHONES USING MACHINE LEARNING MODELS: A COMPARATIVE STUDY

Riddhi Sharma

Delhi Technological University

ABSTRACT

Human Activity Recognition (HAR) is a significant field of study having applications in healthcare, sports, and human-computer interaction. This study investigates the categorization of activities based on sensor data from cellphones. The dataset is made up of accelerometer and gyroscope measurements taken by 30 subjects while conducting six different activities: walking, walking upstairs, walking downstairs, sitting, standing, and laying. The dataset was analyzed using a variety of machine learning techniques, including Support Vector Machines (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN). The data was pre-processed, features were retrieved, and models were trained and tested using cross-validation. Grid search was utilized to optimize the hyperparameters for each model. The findings show that machine learning is successful in HAR when employing smartphone sensor data.

INTRODUCTION

Human Activity Recognition (HAR) is a fundamental challenge in wearable and ubiquitous computing. It is the prediction and classification of human actions based on sensor data obtained from wearable devices, such as smartphones with inbuilt inertial sensors. These sensors include accelerometers and gyroscopes, which detect the device's direction and movement in three dimensions. The goal of this project is to create and test machine learning models that can properly categorize activities into one of six predetermined categories using sensor data obtained from cellphones. Individuals conduct these tasks on a regular basis, and they are critical for applications such as health monitoring, fitness tracking, and contextual computing.

Accurate HAR systems have major implications for a wide range of real-world applications, including senior health monitoring, fitness tracking, tailored coaching, and context-aware services. This research helps to improve user experience and the quality and efficiency of tailored services in wearable and ubiquitous computing settings by accurately predicting human behaviors using smartphone sensor data.

DATASET DESCRIPTION

The Human Activity Recognition database was built from the recordings of 30 study participants performing activities of daily living (ADL) while carrying a waist-mounted smartphone with embedded inertial sensors. The dataset consists of sensor signals (accelerometer and gyroscope) sampled at a constant rate of 50Hz. The dataset has been randomly partitioned into two sets:

- **Training Set:** 70% of the volunteers were selected for generating the training data.
- **Test Set:** 30% of the volunteers were selected for generating the test data.

Each participant performed six activities:

- WALKING
- WALKING_UPSTAIRS

- WALKING_DOWNSTAIRS
- SITTING
- STANDING
- LAYING

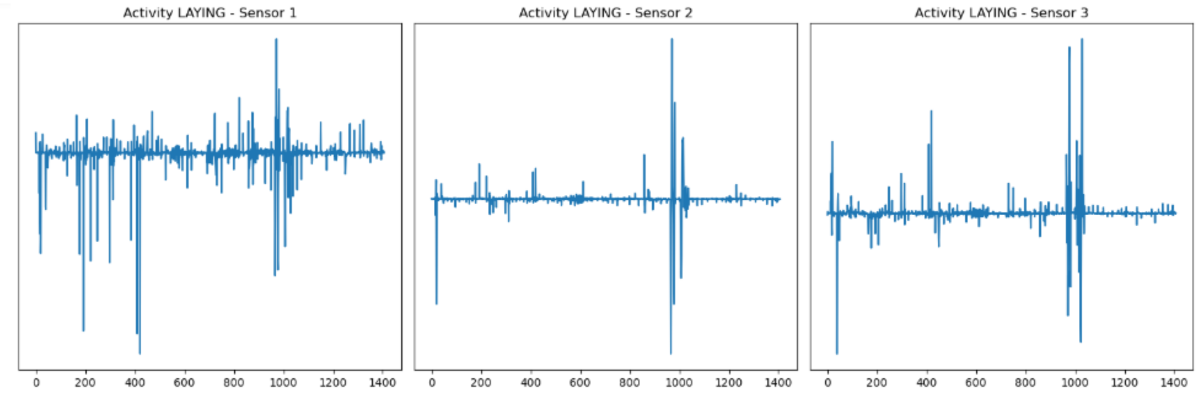
Sensor signals were pre-processed as follows:

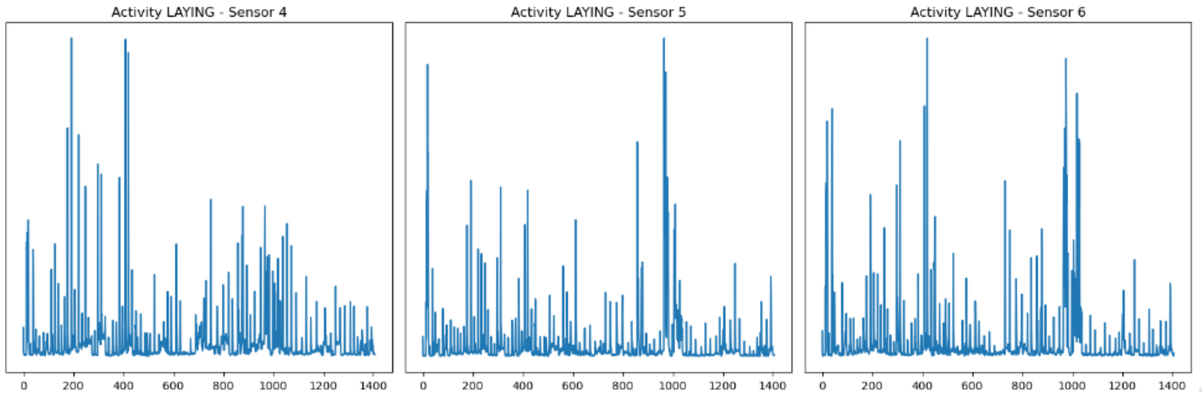
- Noise filters were applied to sensor signals.
- The data was segmented using fixed-width sliding windows of 2.56 seconds each with 50% overlap (128 readings per window).
- The accelerometer input was split into body acceleration and gravity using a Butterworth low-pass filter with a cutoff frequency of 0.3 Hz.
- Each window yielded a 561-feature vector containing variables from both the temporal and frequency domains.

The tests involved 30 individuals aged 19 to 48 years. Each participant completed six different tasks while carrying a Samsung Galaxy S II smartphone around their waist. The smartphone's inbuilt accelerometer and gyroscope recorded the following data:

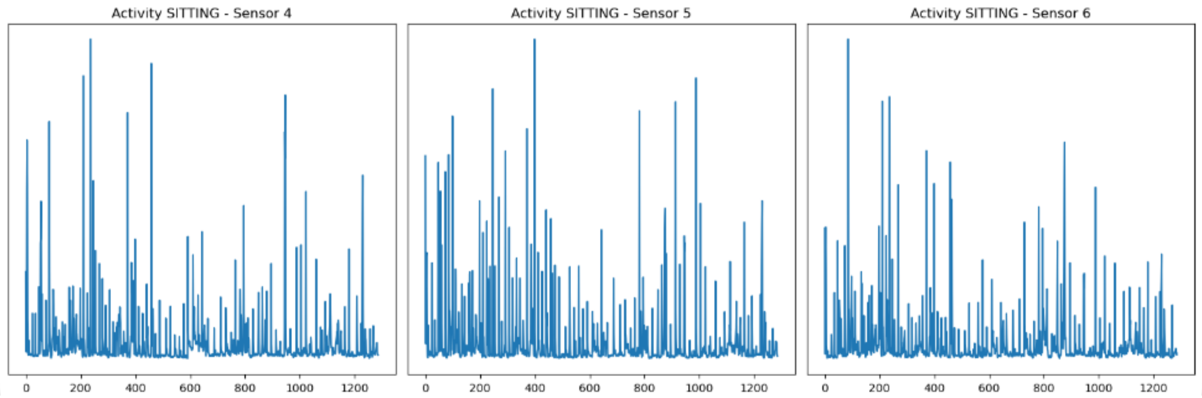
- The accelerometer measures triaxial linear acceleration (total acceleration).
- The gyroscope measures triaxial angular velocity.
- The trials were video recorded so that the data could be manually labeled.
- Every record in the collection contains:
    - The accelerometer produces triaxial acceleration (total acceleration), as does the predicted body acceleration.
    - The gyroscope measures triaxial angular velocity.
    - A 561-feature vector including temporal and frequency domain variables.
    - The activity label indicates one of the six activities.
    - An identification for the subject who conducted the experiment.

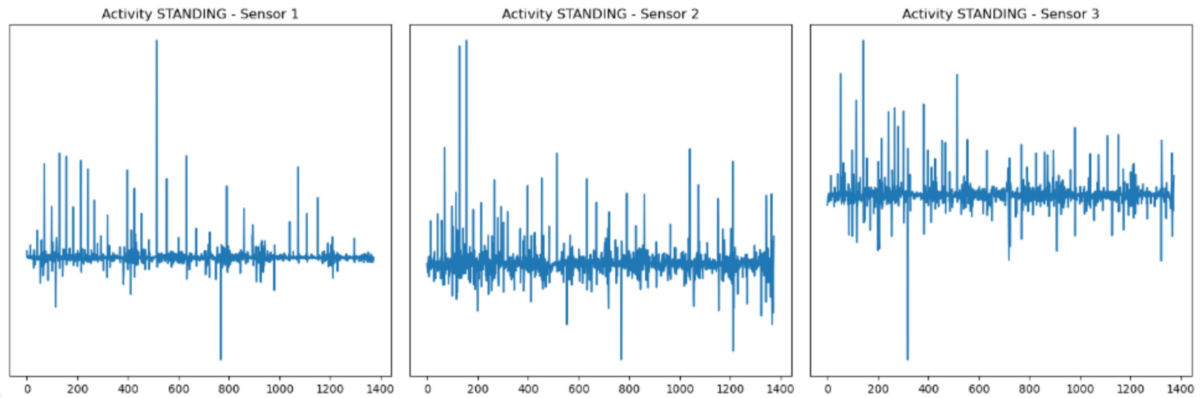
DATASET VISUALISATION

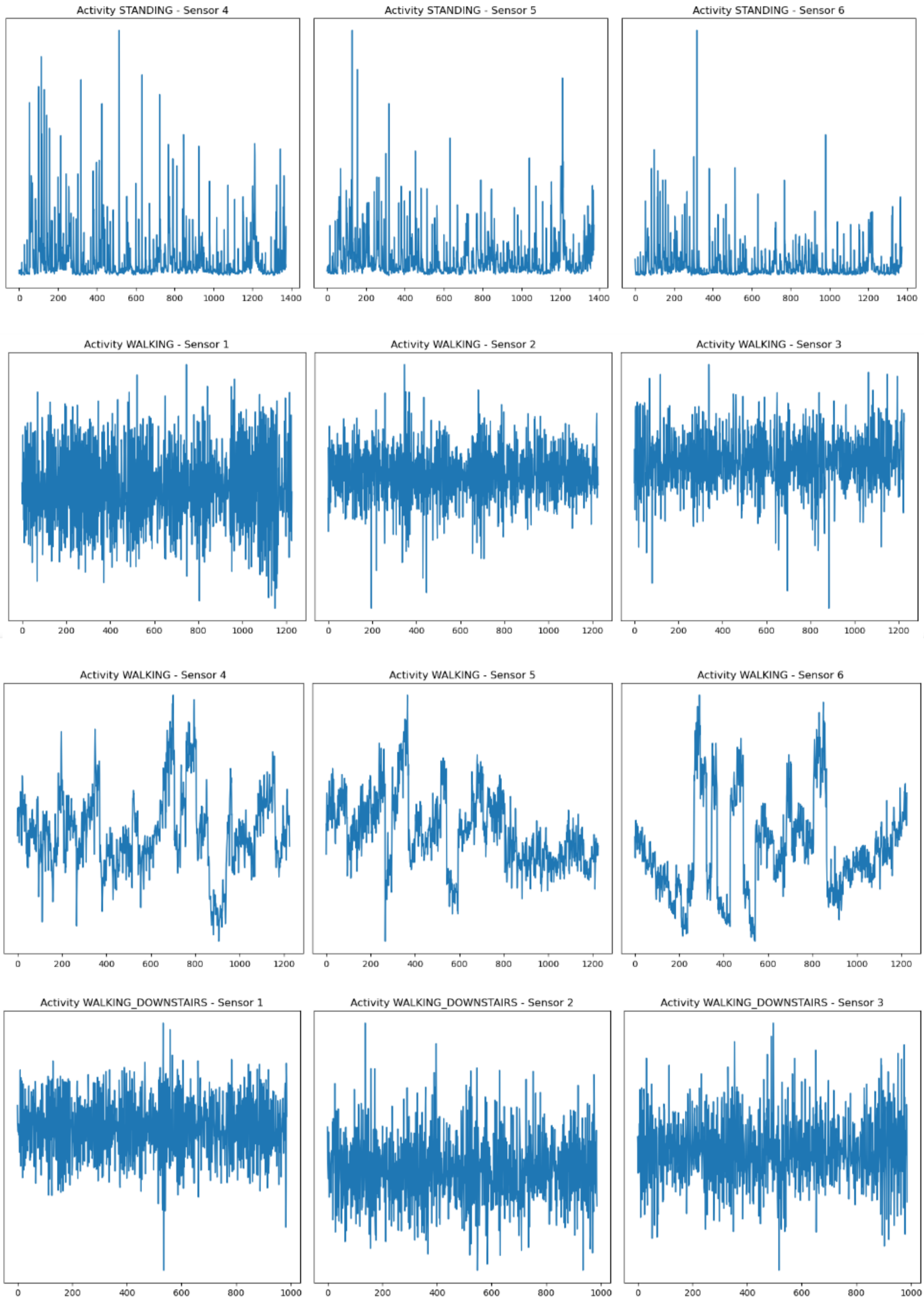


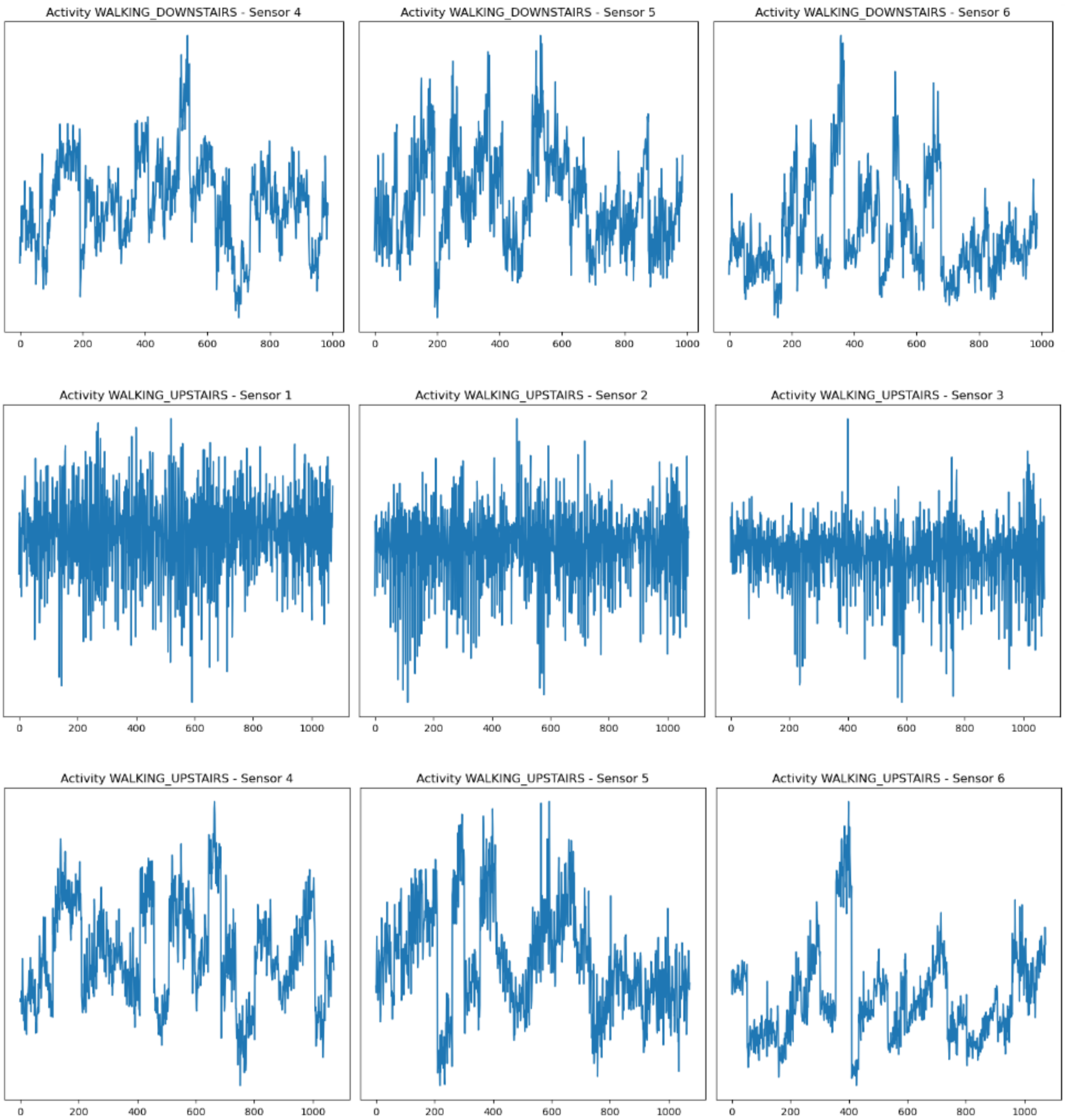

Activity LAYING - Sensor 4
Activity LAYING - Sensor 5
Activity LAYING - Sensor 6
0 200 400 600 800 1000 1200 1400


Activity SITTING - Sensor 1
Activity SITTING - Sensor 2
Activity SITTING - Sensor 3
0 200 400 600 800 1000 1200


Activity SITTING - Sensor 4
Activity SITTING - Sensor 5
Activity SITTING - Sensor 6
0 200 400 600 800 1000 1200


Activity STANDING - Sensor 1
Activity STANDING - Sensor 2
Activity STANDING - Sensor 3
0 200 400 600 800 1000 1200 1400


Activity STANDING - Sensor 4
Activity STANDING - Sensor 5
Activity STANDING - Sensor 6
0 200 400 600 800 1000 1200 1400
Activity WALKING - Sensor 1
Activity WALKING - Sensor 2
Activity WALKING - Sensor 3
0 200 400 600 800 1000 1200
Activity WALKING - Sensor 4
Activity WALKING - Sensor 5
Activity WALKING - Sensor 6
0 200 400 600 800 1000 1200
Activity WALKING_DOWNSTAIRS - Sensor 1
Activity WALKING_DOWNSTAIRS - Sensor 2
Activity WALKING_DOWNSTAIRS - Sensor 3
0 200 400 600 800 1000



METHODOLOGY:

Data Preprocessing

- Data preprocessing is an important step in preparing raw sensor data to train machine learning models. This section describes the procedures used to clean, extract features, segment, and encode data for successful analysis and model training.

    1. **Noise Filter:**
    Raw sensor data frequently contains noise and artifacts, which can degrade the accuracy of activity detection algorithms. To mitigate this, noise filtering techniques were used to clean up the sensor readings. These strategies try to

smooth out noise-induced changes while keeping the underlying patterns that describe human activity

2. **Standardization:**
   Standardization is a preprocessing technique that converts data to a mean of zero and a standard deviation of one. This procedure guarantees that all input characteristics have a comparable scale, avoiding specific features from dominating the learning algorithm due to their greater range.

3. **Label Encoding:**
   Label encoding is the process of converting categorical activity labels into numerical values that can be used for modeling. In this project, the following activities were encoded into numerical values:

   - Walking: 1
   - Walking Upstairs: 2
   - Walking Downstairs: 3
   - Sitting: 4
   - Standing: 5
   - Laying: 6

   Label encoding enables the machine learning models to understand and process the activity labels during training and evaluation.

4. **Segmentation:**
   Segmentation divides continuous sensor data into fixed-width sliding windows with a predetermined overlap. This method enables the model to incorporate temporal relationships in the data and ensures that each segment has a succession of sensor readings describing the actions done. For this project, fixed-width sliding windows with a duration of 2.56 seconds (128 readings per window) and a 50% overlap were employed.

Machine Learning Models

1. **Support Vector Machine (SVM)**

   - ➢ SVM was used because of its capacity to handle high-dimensional data and nonlinear correlations.
   - ➢ For best performance, hyperparameters (C, kernel) were tuned using GridSearchCV.

2. **Random Forest classifier**

   - ➢ Random Forest was used for its ensemble learning approach, which effectively manages noise while avoiding overfitting.
   - ➢ GridSearchCV was used to tune hyperparameters (n_estimators, max_depth, min_samples_split, min_samples_leaf).

3. K-Nearest Neighbors (KNN)

- KNN was used since it is simple and effective for classification jobs.
- GridSearchCV was used to tune hyperparameters such as n_neighbors, weights, and metric.

RESULTS AND EVALUATION

1. Before hyperparameter tuning

Algorithms	Accuracy	Precision	Recall	F1 Score
Support Vector Machine	0.9594	0.9602	0.9594	0.9592
Random Forest	0.9309	0.9321	0.9309	0.9308
K-Nearest Neighbor	0.8762	0.8821	0.8762	0.8753

2. After hyperparameter tuning

Algorithms	Accuracy	Precision	Recall	F1 Score
Support Vector Machine	0.9601	0.9610	0.9601	0.9599
Random Forest	0.9325	0.9300	0.9265	0.9300
K-Nearest Neighbor	0.9035	0.9095	0.9035	0.9022

Support Vector machine performs the best with an accuracy of 95.94% and 96.01% before and after hyperparameter tuning respectively and F1 Score of 0.9592 and 0.9599 respectively.

REFERENCES

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. *21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN* 2013. Bruges, Belgium 24-26 April 2013.