ECS 330: EECS LABORATORY-II

PROJECT REPORT

Project Id: 17 - **SHIRSHAKK PURKAYASTHA**
Faculty Advisor: Dr. Kundan Kandhaway 18247

Github Repository: *https://github.com/Shirshakk-P/Smartphone-based-Recognition-of-Human-activities*

"Smartphone-based Recognition of Human Activities"

OBJECTIVE:

The objective of this project was to develop ML models with a broader aim to classify the different human activities with the help of data accumulated by smartphone sensors.

DATASET:

http://archive.ics.uci.edu/ml/datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions

DATASET INFORMATION:

Dataset Characteristics:

1) Multi-Variate Data 2) Time Series Data

The experiments were carried out with a group of 30 volunteers within an age bracket of 19-48 years.
They performed a protocol of activities composed of six basic activities:

>>3 Static Postures (standing,sitting,lying).
>>3 Dynamic Activities (walking,walking downstairs,walking upstairs).

The experiment also included postural transitions that occurred between the static postures:

stand-to-sit	***sit-to-stand***
sit-to-lie	***lie-to-sit***
stand-to-lie	***lie-to-stand.***

All the participants were wearing a smartphone (Samsung Galaxy S II) on the waist during the experiment execution.
Data captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz using the embedded accelerometer and gyroscope of the device. The experiments were video-recorded to label the data manually.

The obtained dataset was randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.

The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window).
The sensor acceleration signal, which has gravitational and body motion components, was separated using a Butterworth low-pass filter into body acceleration and gravity.
The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of 561 features was obtained by calculating variables from the time and frequency domain.

Attribute Information:

The dataset is divided in two parts-

1. **Inertial Sensor Data:**
- Raw triaxial signals from the accelerometer and gyroscope of all the trials with with participants.
- The labels of all the performed activities.

2. **Records of activity windows:**
- A 561-feature vector with time and frequency domain variables.
- Its associated activity label.
- An identifier of the subject who carried out the experiment.

Activity Labels:

1. WALKING	2. WALKING_UPSTAIRS
3. WALKING_DOWNSTAIRS	4. SITTING
5. STANDING	6. LAYING
7. STAND_TO_SIT	8. SIT_TO_STAND
9. SIT_TO_LIE	10. LIE_TO_SIT
11. STAND_TO_LIE	12. LIE_TO_STAND

- Features are normalized and bounded within [-1,1].
- The units used for the accelerations (total and body) are 'g'(acceleration due to gravity-> 9.80665 m/sec2).
- The gyroscope units are rad/sec.

LEARNING MODELS:

This project implemented three Machine Learning Models to classify the Human Activity Recognition Dataset, viz:

1. Logistic Regression (LR)
2. K-Nearest Neighbours (KNN)
3. Support Vector Machines (SVM)

COMPARISON METRIC:

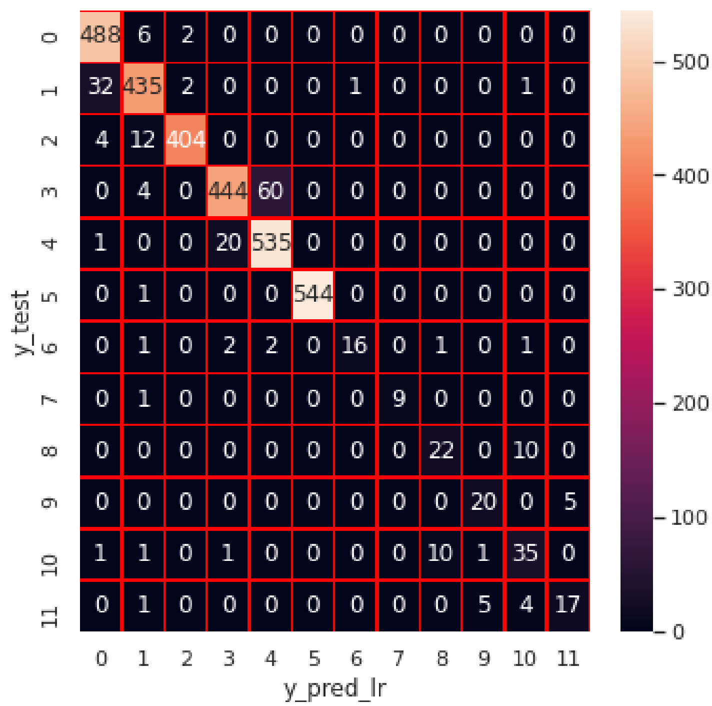
The f1-score is chosen as the comparison metric for the three trained models.

ML MODEL 1: **LOGISTIC REGRESSION**

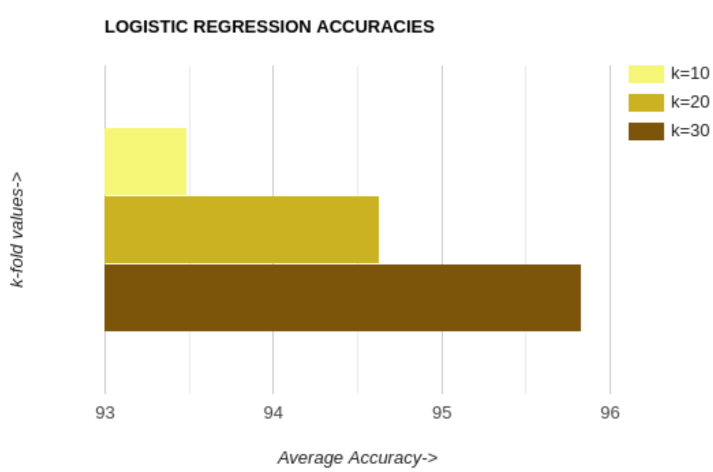
We trained the dataset on a *Logistic Regression* model with k-fold cross-validation values of (10,20,30). The performance of the LR Model on the corresponding k-fold values is given by:

k-fold Value	AVERAGE ACCURACY	LR SCORE
k = 10	93.4862	93.8962
k = 20	94.6312	93.8962
k = 30	95.8291	93.8962

CONFUSION MATRIX:



GRAPHICAL COMPARISON:



OBSERVATION:

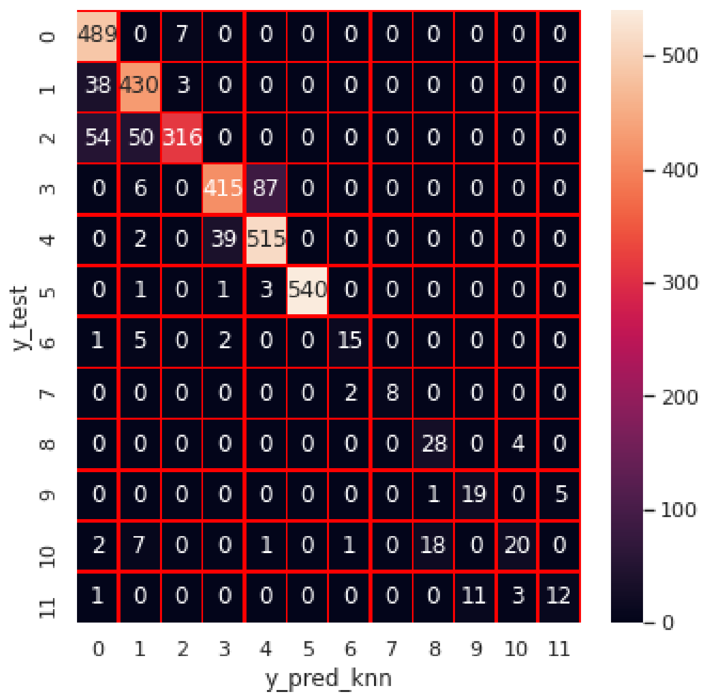
A LR Model with k-fold value of 30 gives the maximum average accuracy.

ML MODEL 2: **k-NEAREST NEIGHBOURS**

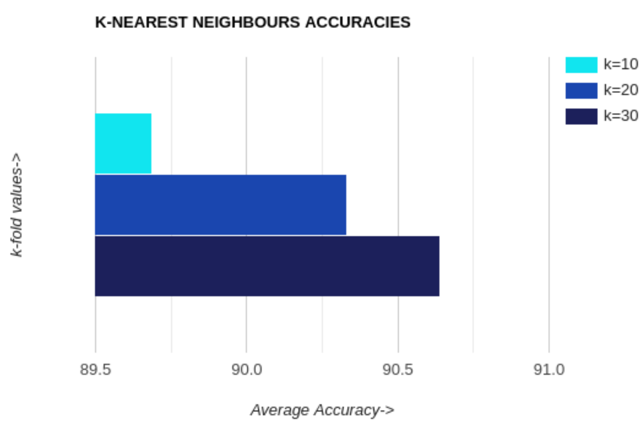
We trained the dataset on a ***k-Nearest Neighbours*** model with k-fold cross-validation values of (10,20,30). Also cluster-sizes are chosen as (10,20,30). The performance of the kNN Model on the corresponding k-fold values is given by:

k-fold Value	Cluster Size	AVERAGE ACCURACY	kNN SCORE
k = 10	10	89.2717	88.7729
k = 20	20	90.3313	88.7729
k = 30	30	90.6408	88.7729

CONFUSION MATRIX:



GRAPHICAL COMPARISON:



OBSERVATION:

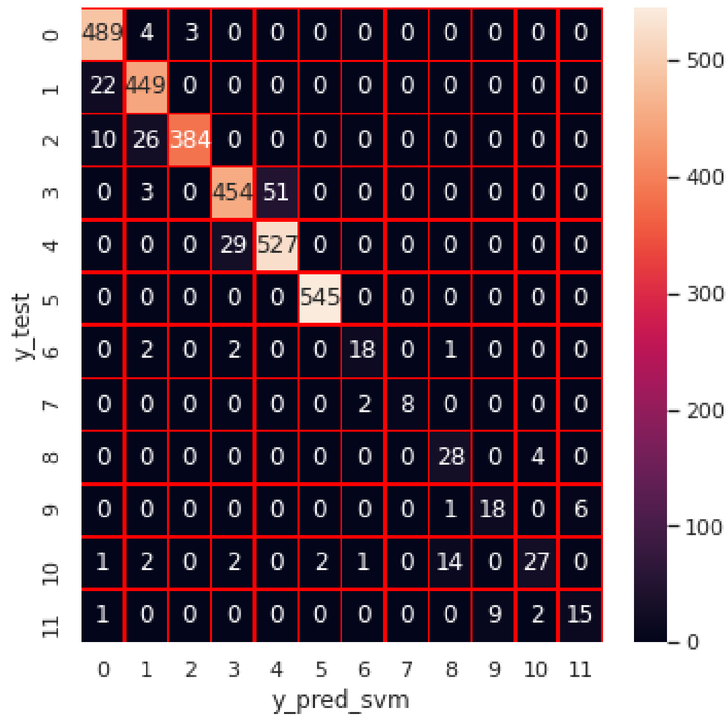
A Cluster size of 30 with a k-fold value of 30 gives the maximum average accuracy.

ML MODEL 3: SUPPORT VECTOR MACHINES

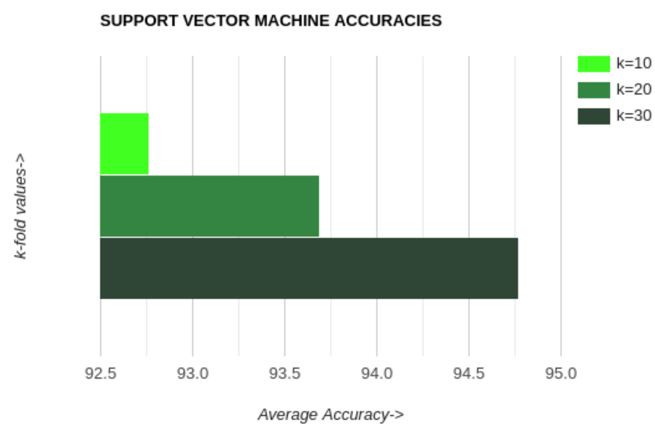
We trained the dataset on a *Support Vector Machine* model with k-fold cross-validation values of (10,20,30). The performance of the SVM Model on the corresponding k-fold values is given by:

k-fold Value	AVERAGE ACCURACY	SVM SCORE
k = 10	92.7650	93.6748
k = 20	93.6909	93.6748
k = 30	94.7735	93.6748

CONFUSION MATRIX:



GRAPHICAL COMPARISON:



OBSERVATION:

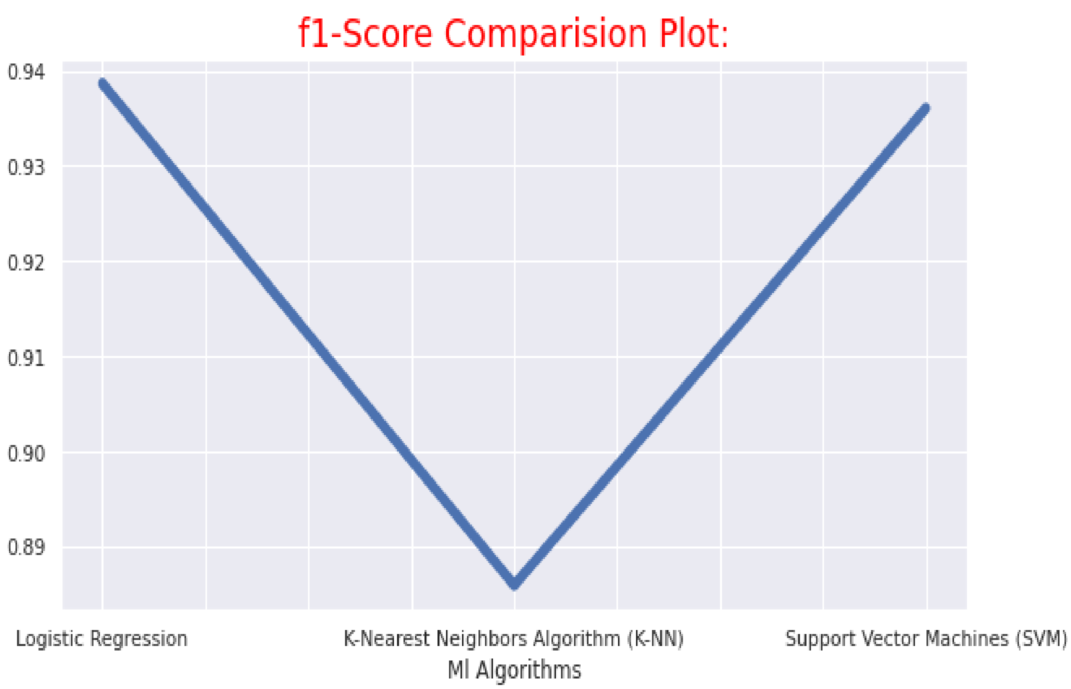
A SVM Model with k-fold value of 30 gives the maximum average accuracy.

f1-SCORES:

The corresponding f1-Scores of the three trained models is given by:

Serial No.	ML MODEL	f1-SCORE:
1.	LOGISTIC REGRESSION	0.938662
2.	k-NEAREST NEIGHBOURS	0.885910
3.	SUPPORT VECTOR MACHINES	0.936055

f1-SCORE PLOT:



CONCLUSIONS:

The highest classification accuracy is achieved by the LOGISTIC REGRESSION Model (k-fold value of 30) and the lowest by the k-NEAREST NEIGHBOURS Model (k-fold value of 10) in our predictions.

REFERENCES:

1. Human Activity Recognition on Smartphones With Awareness of Basic Activities and Postural Transitions.(Jorge-Luis Reyes-Ortiz, Luca Oneto, Alessandro Ghio, Albert SamÃ¡, Davide Anguita and Xavier Parra).
Artificial Neural Networks and Machine Learning â€" ICANN 2014. Lecture Notes in Computer Science. Springer. 2014.

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

3. Energy Efficient Smartphone-Based Activity Recognition using Fixed-Point Arithmetic.(Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, Jorge L. Reyes-Ortiz).
Journal of Universal Computer Science. Special Issue in Ambient Assisted Living: Home Care. Volume 19, Issue 9. May 2013

4. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine.(Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz).
4th International Workshop of Ambient Assited Living, IWAAL 2012, Vitoria-Gasteiz, Spain, December 3-5, 2012.
Proceedings. Lecture Notes in Computer Science 2012, pp 216-223.

5.Human Activity and Motion Disorder Recognition: Towards Smarter Interactive Cognitive Environments. (Jorge Luis Reyes-Ortiz, Alessandro Ghio, Xavier Parra-Llanas, Davide Anguita, Joan Cabestany, Andreu CatalÃ).
21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.

*****