IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Human Activity Recognition Using Accelerometer Data

K DurgaPrasad[1]
Student, Department of Information Technology
GMR Institute of Technology, Rajam, India

Abstract: Now-a-days, smart devices are playing a vital role to recognize the human activities and become a well-known field of research. Human activity recognition, is a motile procedure of recognizing different environments. Walking, standing, going upstairs, going downstairs, running is some of examples of such kind of actions. We use smart phone accelerometer to detect the human's activity. By using machine learning techniques on the publicly available datasets. Using trained models in this field to reach the fundamental goal of this model which is recognition and categorization of activity taking place. Random forest, support vector machine, logistic regression, decision tree, k-nearest neighbor has been used to analyses and compare the performance. Besides using these models, the dimension of the dataset is reduced through feature selection process. The results of these algorithms are compared and presented in the form of tables and graphs and identified to be the best algorithm based on accuracy results. This kind of action prediction model can be cast-off to provide astute info about lots of human beings just by making them comprise a smartphone with them.

Keywords: Human Activity Recognition, Machine Learning, Random Forest, Decision Tree, Support Vector Machine, Logistic Regression, K-Nearest Neighbor.

I. INTRODUCTION

In today's world, this has been proven that smart devices such as smartphones, smartwatches and so on have been extensively used. People carry mobile devices around for most part of the day. This enables us to track activity of the person carrying the smartphone using the data of the sensors of their smartphones. All of these smart devices are fixed with some sensor devices like cameras, gyroscopes, and accelerometers. By using accelerometer data on a smart phone, a person's activity can be stored automatically and sent to a server where it can be processed to recognize activity. HAR is one of the important technologies to monitor the dynamism of a person and this can be done with the support of Machine learning techniques

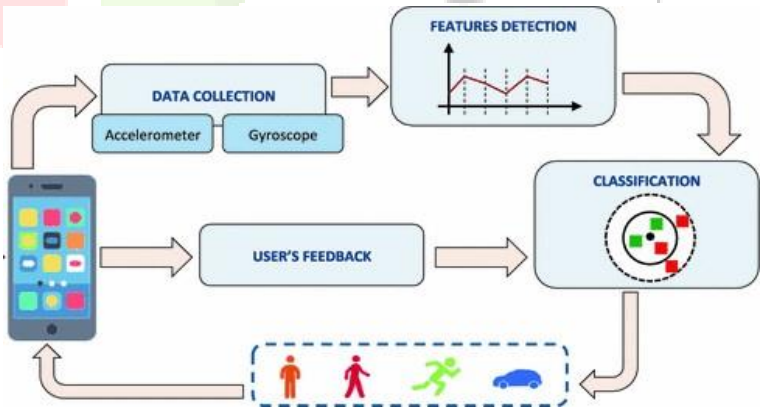

Fig-1: Overview of human activity recognition using data collected from smartphone

Fig-1 shows an overview of the whole process. The data from UCI machine learning repository will be used processed through machine learning to produce predictive classification models that will be used to classify the physical activities of the person. Main challenge in HAR is that human body activities are very intricate and its real-time recognition is very tough. To overcome this challenge machine learning is applied on sensors data like accelerometer data.

In any application of machine learning, feature selection is the most important step for creating any classifier. The features which are more informative, are selected before classification for generalized results. The features select for classification are mean, standard deviation, peak and resultant values. These features help classify data from the tri-axial accelerometer, x, y, and z values into classes of walking, standing, running, sitting etc. The detection of activities like walking and sitting can provide useful information a few people's health condition.

With the implementation of this system in day to day lives people can monitor and also be aware of activities happening at real time and surveillance over areas like security, health care, child monitoring can be managed in an effective way.

II. LITERATURE SURVEY:

In paper [1] Zameer Gulzar, A.Anny Leema, I.Malaserene proposes human activity analysis is a popular activity in the growing industry and applied different machine learning algorithms to measure the effectiveness of various machine learning classification algorithms. Author implemented machine learning algorithms in real world datasets so that their accuracy can be studied and effective conclusion are drawn. Author performed comparative study among the applied various techniques KNN, SVM, Random Forest, Neural Networks, Logistic Regression and Naïve Bayes. In them, Logistic Regression and Neural network gave good results when compared to others. Author proposed that these results can be used for making smart watches and similar devices which can track a user's activity and notify the person's daily activity log.

In paper [2] Harpreet kaur lohia, Simran kaur Dari, Shravanti kalwal, Roshini Singh used the data available at the UCI machine learning repository which has been modified for his research. Author used supervised machine learning algorithms to produce predictive classification models that will be used to classify physical activities of the person into six categories namely sitting, standing, laying, walking, upstairs, downstairs. From the analysis conducted in this paper out of different algorithms used the accurate result that they have got is from random after applying bagging. Also, the confusion matrix analysis concluded that the misclassification error in the Random Forest model was zero than other models. It can be obviously said that the random forest model can be efficiently applied when used for activity recognition using the data provided by the accelerometer and gyroscope of a smartphone.

In paper [3] Mst. Alema Khatun, Mohammad Abu Yousuf introduced three algorithms CNN, SVM and Random Forest respectively and compared the advantages and disadvantages of these algorithms. In this paper, Author used accelerometer and gyroscope to find out the subject's acceleration and angular velocity. Author wants to show CNN gives best results when compared to other algorithms. Machine learning works in a hand-crafted way, which relies on domain knowledge or human experience. They found that the proposed solutions were better than others and CNN has great use in HAR.CNN giving good performance.

In paper [4] Rimsha Khan, Muhammad Abbas, Rubia Anjum, Fatima Waheed worked related to human activity recognition using accelerometer data is analysed and presented in the paper. Recognizing activity in real time is great challenge which can be efficiently only through using machine learning techniques. In this paper some important machine learning techniques have been analysed and their results compared using MATLAB classification Learner and Weka tool. In MATLAB classification Learner the best model is Bagged Tree with 96.2% accuracy results is identified and it also gives most accurate results compared to others.

In paper [5] Ashim Saha, Tulika Sharma, Harshika Batra, Anupreksha Jain, Vanbal Pal have designed a smartphone-based recognition system in this study, through which they are able to recognize overall six activities (standing, sitting, walking, walking upstairs, walking downstairs). The activity data were trained and tested with four supervised machine learning algorithms: Logistic Regression, K-Nearest Neighbor, Support Vector Machine and Random Forest classifier separately. The best classification rate found in our experiment was 95.995% attained by the Logistic Regression algorithm. This classification attained by the Logistic Regression is robust in orientation and the position of smartphone.

III. DATA COLLECTION

The dataset used in this paper from UCI machine learning repository. [1] The original dataset was derived by carrying experiments on 30 different individual volunteers within the age of 19-40 years. In the experiment the volunteer was made to perform six activities namely sitting, standing, walking, walking downstairs, walking upstairs. The dataset contains 563 columns and 7325 rows. During the experiment the volunteer was wearing a smartphone on the waist. Using the sensors which are embedded in smartphone, Accelerometer and Gyroscope on the give smartphone 3-axial linear acceleration and 3-axial angular velocity were recorded at a constant of 50Hz.

IV. METHODOLOGY

In this paper, [4] the problem mainly depends on classification modelling. So, machine learning predictive models will decide the activity of a person, after predictive models are formulated using the training data. Classification should be done for six categories namely sitting, standing, laying, walking, walking upstairs, walking downstairs. By using supervised machine learning algorithms Logistic regression, Random Forest, Support Vector Machine, Decision Tree, k-Nearest Neighbor (KNN) train the data and find the activity of the person.

Logistic Regression:

Logistic Regression is a powerful supervised machine learning algorithm used for binary classification problems.[1] In this model, the probabilities give the possible outcomes of a single test are modelled with a logistic function. Logistic regression essentially uses a logistic function to reduce outliers and classify the data.

$$Logistic\ function = \frac{1}{1+e^{-x}}$$

Multinomial logistic regression can model setups where there are more than two possible discrete outcomes. When a model learns the training data too meticulously, it fails to fit new data or predict unseen observations reliably. This condition is called overfitting. Logistic function will be used to reduce the overfitting.

Decision Tree:

Decision Tree [4] is a supervised learning technique that can be used for both classification and regression problems, but frequently it is chosen for solving classification problems. It is called Decision tree, because similar to a tree, it starts with a root node, which

expands on further branches and constructs a tree-like structure.[7] By calculating how much accuracy each split will cost us, using a function. The split that costs least are chosen, this algorithm is recursive in nature as the groups formed can be sub-divided using same strategy. Due to this process, this algorithm is also known as the Greedy Algorithm, as we have an more desire of letting down the cost. This makes the root node as finest predictor/classifier.

Random Forest:

Random Forest [3] is a widespread machine learning algorithm that belongs to the supervised learning technique. [1] Classification and Regression problems in ML be done by using it. It is based on the concept of ensemble learning, which is a process of joining multiple classifiers to solve a intricate problem and to increase the performance of the model.[7] Random Forest is a classifier that has a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of dataset. The larger number of trees in the forest leads to higher accuracy and avoids the problem of overfitting.

K-Nearest Neighbor:

The k-nearest neighbors (KNN) [4] are a simple, easy to implement supervised machine learning algorithm that can be used to solve both classification and regression problems. K-NN algorithm assumes the similar between the new data and available cases and put the new case into the category that is most similar to the available categories.[8] K-NN algorithm stores all the existing data and classifies a new data point based on the comparison. This means when new data appears then it can be simply classified into a well matching set category by using k-NN algorithm.

Support Vector Machine:

Support Vector Machine (SVM) [2] is one of the best Supervised learning algorithms which is used for classification as well as Regression problems. However, mainly it is used for classification problems in ML. The goal of the SVM algorithm is to create the best line that can segregate n-dimensional space into classes. This best line is called a hyper plane. This hyper plane will be used to put the new data point in the correct category. SVM takes the extreme points/vectors that support in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is called as Support Vector Machine. As it having positive and negative hyperplane makes algorithm more accurate.

V. RESULT & DISCUSSIONS

The table below represents the resultant data, which is subsequently used for training and testing the predictive models. After preparing the final dataset, apply the machine learning techniques to get the accuracy.

Activities	Number of Examples	Percentages
Laying	537	18.23
Sitting	491	16.67
Standing	532	18.06
Walking	496	16.84
Walking Downstairs	420	14.26
Walking Upstairs	470	15.96
Total	2946	100

Table-1: Activities

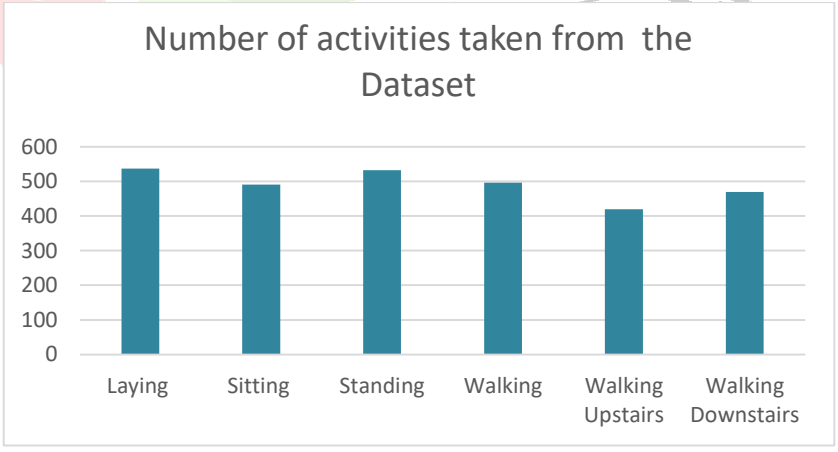

Fig-1: Number of activities taken from the dataset

After analyzing the data, train the model using machine learning techniques and get the accuracy for each model as follows in the table below.

S.no	Model	Accuracy %
1	Support Vector Machine	97.6
2	K Neighbors classifier	97.5
3	Decision Tree Classifier	94.3
4	Random Forest Classifier	98
5	Gaussian NB	72.8
6	Ridge Classifier	97.9
7	Logistic Regression	99

Table 2: - Different ML Algorithm with their accuracy

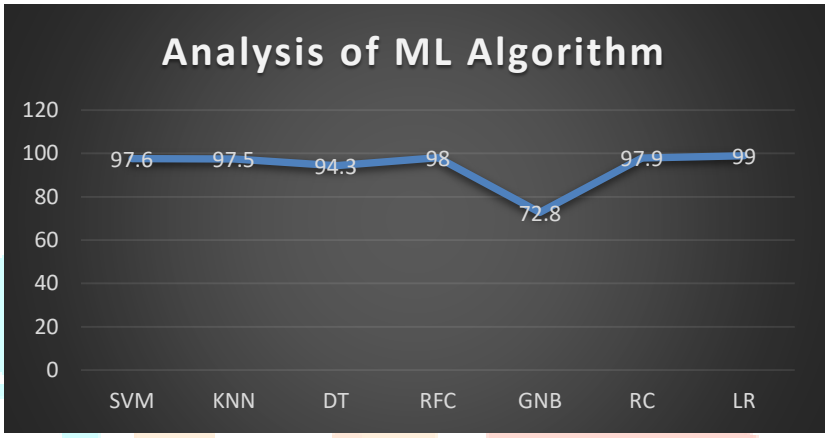

Fig 2: - Analysis of ML Algorithm

VI. CONCLUSION

From the analysis conducted in this paper, different machine learning algorithm are applied on the data. Comparative study performed among the applied various techniques KNN, SVM, Random Forest, Decision Tree, Gaussian NB, Ridge Classifier, Logistic Regression. After analysing their approaches for Human Activity Recognition using accelerometer data. Logistic Regression gives good result when compared to others. So Logistic Regression model can be effectively utilized when used for activity recognition using the data provided by accelerometer and gyroscope of a smartphone.

REFERENCES:

[1] Zameer Gulzar, A.Anny Leema, I.Malaserene.”Human Activity Analysis using Machine Learning Classification Techniques”. ISSN:22783075, Volume9, Issue2, December,2019, ijitee

[2] Mst.Alema Khatun, Mohammad Abu Yousuf,” Human Activity Recognition Using Smartphone Sensor Based on Selective Classifiers”, 2020 2nd International Conference on Sustainable Technology of Industry 4.0(STI),19-20 December, IEEE Xplore.

[3] Rimsha Khan, Muhammad Abbas, Rubia Anjum, Fatima Waheed,”Evaluating Machine Learning Techniques on Human Activity Using Accelerometer Data”,2020 IEEE Xplore

[4] Harpreet kaur lohia, Simran kaur Dari, Shravanti Kalwal, Roshini Singh, ”Human Activity Recognition using Machine Learning Techniques”,JETIR June 2022,Volume 9,Issue 6, ISSN:2349-5162.

[5] Ashim Saha, Tulika Sharma, Harshika Batra, Anupreksha Jain, Vanbal Pal, “Human Action Recognition Using Smartphone Sensors”, 2020 International Conference on Computational Performance Evaluation (ComPE), North-Eastern Hill University, Shillong, Meghalaya, India. July 2–4, 2020 IEEE

[6] Swati Patil, Harshal Lonkar, Ashitosh Supekar, Pranav Khandebharad, Mayur Limbore, “Human Activity Recognition Using Machine Learning”, International Journal of Engineering Research and Technology(IJERT),ISSN:2278-0181,Vol.10 Issue01,April-2021

[7] Prof Jayashree M Kudari, Abhilash Banerjee,”Recognition of Human Activities using Machine Learning Algorithms”,International Research Journal Of Engineering and Technology(IRJET),volume:07 Issue:09,sep 2020

[8]Ahatsham Hayat, Morgado Dias, Bikram Pratim Bhuyan, Ravi Tomar,”Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches”,mdpi.com/journal/information,doi:10.3390/info13060275,information 2022

[9]Prof.Hansa Shingrakhia,”Human Activity Recognition using deep learning”,2018 JETIR May 2018,Volume 5, Issue 5,ISSN-2349-5162

[10]Pinky Paul, Thomas George,” An Effective Approach for Human Activity Recognition on Smartphone”, march 2015 IEEE