Smartphone Sensor Based Human Activity Recognition using Deep Learning Models

Ms.S.Roobini, Ms.J.Fenila Naomi

Abstract: The significant intent is to generate the model for anticipating the activities of a human that ensures the aversion of human life. Activity Recognition (AR) is monitoring the liveliness of a person by using smart phone. Smart phones are used in a wider manner and it becomes one of the ways to identify the human’s environmental changes by using the sensors in smart mobiles. Smart phones are equipped in detecting sensors like compass sensor, gyroscope, GPS sensor and accelerometer. The contraption is demonstrated to examine the state of an individual. Human Activity Recognition (HAR) framework collects the raw data from sensors and observes the human movement using different deep learning approach. Deep learning models are proposed to identify motions of humans with plausible high accuracy by using sensed data. HAR Dataset from UCI dataset storehouse is utilized. The performance of a framework is analyzed using Convolutional Neural Network with Long-Short Term Memory [ConvLSTM] and Recurrent Neural Network with Long-Short Term Memory [RNNLSTM] using only the raw data. The act of the model is analyzed in terms of exactness and efficiency. The designed activity recognition model can be manipulated in medical domain for predicting any disease by monitoring human actions.

Index Terms: ConvLSTM, Deep Learning, Human Activity Recognition, RNNLSTM.

I. INTRODUCTION

Smartphones have become a most useful tool in our daily life for communication with advanced technology provided intelligent assistance to the user in their everyday activities. The portable working framework with computing ability and interconnectivity, application programming interfaces (API) for executing outsiders’ tools and applications, mobile phones have highlights such as cameras, GPS, web browsers so on., and implanted sensors such as accelerometers, gyroscope and Magnetometer which permits the improvement of applications in view of client’s specific area, movement and context. To develop resourceful smart phone application, it is imperative to utilize context recognition and situational attention of the gadget’s client. Activity Recognition is one such platform for these devices which can be dealt by the implicit sensors and it is being used in various areas like business, medicinal services, security, transportation and so forth. Different kinds of sensors incorporate wearable sensors which can identify the movement and Bluetooth sensor which empowers the exchange of information from one gadget to another utilizing the information correspondence channels.

II. RELATED WORK

Min et al [14] designs two models, in which one uses only acceleration sensor data and other uses location information in addition to acceleration sensor data. Before feature extraction, the acceleration sensor data is divided into time segments which is said to be temporal segmentation. In order to handle streaming of data, sliding window technique is used.
The component vector of the time window is utilized as the contribution to the classifier. Utilizing area data, a progressively explicit and itemized area based classifier can be connected. The location information is not gathered from the framework and it is derived from different kinds of activities. In the Artificial Neural Network [ANN] classifier, the Xavier and ReLU are commonly utilized to decrease learning time in the field of machine learning. The learning rate is set to 0.01, and the Adam streamlining agent is utilized since it is known to accomplish great outcomes quick. Finally, the model with location information shows an accuracy of 95% and the model without location information does 90% accuracy.

Qingzhong et al [18] proposes a method for activity recognition in two steps. In step 1, accelerometer provides unrefined sensor data which recognizes the actions and the course of mobile phones indicating the gyroscope senses rotational movements which are large to detect by people. In step 2, feature extraction method is preformed for the sensed data. Machine learning models worked for action recognition, at that point the profound learning model dependent on convolutional neural system. The results are demonstrated to indicate consistency, it expected that the scope of activities recognition can be extended with the utilization of gathered information with multi-class dataset. The experimental results demonstrate that LibSVM performs better in all arrangement than FLD in terms of precision. From the results, it tends to be reasoned that accelerometer sensor (A-sensor) reading contribute more compared to G-sensor reading (Gyroscope sensor) but can increase the accuracy of the detection by utilizing both AI and profound learning calculations.

Erhan et al [4] proposes several supervised machine learning algorithms such as Decision trees, Support Vector Machines, K-nearest neighbours (KNN) and ensemble classification methods such as Boosting, Bagging and Stacking. For classification, binary decision tree is used in which 53.1% of accuracy. When the branching limit is increased to 100, the accuracy is increased to 94.4%. Support Vector machines provides an accuracy of 99.4% that uses hyper dimensional planes. K-NN provides an increased efficiency of 97.1% with k value is set to 3. In Ensemble classifiers there are different accession in which boosting technique has different forms. AdaBoost uses miscalculated probability and classifies with an accuracy of 97.4%. The other way is bagging which is used to get intent results from sensitive learning algorithms with efficiency of 98.1%. Third way is stacking in which 30 stack classifiers are applied and predicted an accuracy of 98.6%.

Akram et al [1] analyses different activities of a person, using which a classification model is built based on the feature selection. In Weka toolkit, Multilayer Perceptron, Random forest, LMT, SVM, Simple Logistic and LogitBoost are compared as an individual and combined classifiers then it was validated using K-fold cross validation. The recognition is analysed as mobile in hand and pocket position. The efficiency is obtained as SVM provides better precision in hand and Random forest dictates highest accuracy. Mobile phone in-hand position, the fusion of Multilayer Perceptron, Logit Boost and SVM classifier yields an accuracy of 91.15% but meanwhile in-pocket position, Multilayer Perceptron, Random Forest and Simple Logistic with 90.34% accuracy. A single triaxial accelerometer was used to obtain accurate recognition of 91.15% on daily activities.

### III. SENSORS IN SMARTPHONE

Smart gadgets have worked with sensors that gauge movement (accelerometers, gravity sensors, and spinner), orientation (magnetometers) and distinctive ecological conditions (markers, photometers, and thermometers). Gear based sensors are physical parts consolidated with a PDA and they gather their unrefined data by explicitly evaluating specific common parameters. Programming-based sensors are not physical devices, in spite of that the way they impersonate equipment based sensors. These are fit for outfitting information with high accuracy and significant in case it is expected to screen three dimensional contraption developments.

The Android sensor framework empowers to get various sorts of sensors and sensor structure uses a standard 3-center point arrange the system to express data. When a contraption is held in its default introduction, the X-hub is level to the ground shows the right, the Y center is vertical and centers up, and the Z turn demonstrates the outside of the screen face. In this structure, orchestrate the screen have negative Z esteem. The figure.1 depicts the sensors available in Smartphone for Human Activity Recognition [HAR].

#### A. Accelerometer

An accelerometer is one of the imperative gear in the smart phones and keen wardrobe. The essential limit is to identify the alterations in the orientation of smart gadget with respect to datum and adjust the introduction to suits the study edge of the client. For example, when scanning for site page with extended width, can get this scene to see from changing the orientation of phone to level. Unique way camera mode moreover changes the scene to representation or picture to scene when there is a change in the orientation of gadget/camera. In the long run, this sensor detects the alteration by 3D (X, Y and Z rotate) [21] estimation of the increasing speed of the device concerning free fall.

#### B. Gyroscope

The utilization of Gyroscope is to keep up and control the position, level or orientation subject to the standard of precise energy. At whatever point 'Gyros' used nearby accelerometer identifies development from 6-Axis for example right, left, up, down, forward and in invert. It distinguishes the move, pitch, and yaw developments. Yaw, Roll, and Pitch are the precise minutes seen from 3-Axis. Using MEMS (Micro Electrical and Mechanical System) innovation, gyroscopic sensors helps in route reason and perceiving the motion structures used in keen mobiles and tablets. Figure 1 depicts the Sensors in Smartphone.
C. Compass Sensor

The Smart Compass is a standard gadget to recognize the heading with respect to the north-south shaft of the attractive field. Compass helpfulness in mobile phones is increasingly unpredictable sensor called a magnetometer. It is used to evaluate the quality and course of alluring fields. By breaking down Earth's magnetic field, the sensor empowers the gadget to choose its orientation with high accuracy [8]. A compass sensor is the buoy number some place in the scope of 0° and 360°. It begins from 0° as outright the north demonstrates the point between present cell phone course and total north in clockwise. The Speedometer and the choice to send GPS masterminds by methods for SMS or email. The information stream came back from compass sensors is a lot of drifting numbers demonstrating the holy messenger, goofy (i.e= D 1, 2, 3,... 0 degree <= campy<= 360 degrees). Compass perusing can be utilized to distinguish the bearing change in the client's movement.

IV. METHODOLOGIES

The following are the machine learning and deep learning methods employed in this paper.

A. Data Collection

The dataset is prepared by monitoring the activities of 30 volunteers within an age of 19-48 years. Each person performing six different kind of activities are listed below:

- WALKING
- WALKING_UPSTAIRS
- WALKING_DOWNSTAIRS
- SITTING
- STANDING
- LAYING

These data are gathered from UCI Machine learning store of Human Activity Recognition [HAR]. Utilizing its embedded accelerometer and gyroscope, captured 3-axial linear acceleration and 3-hub precise speed at a steady rate of 50Hz. The acquired dataset has been arbitrarily partitioned into two sets, where 70% of the volunteers were chosen as preparing information and 30% test information. The noise channels are connected to sensor signals which were pre-handled and after that inspected in fixed-width sliding windows of 2.56 sec and half cover (128 readings/window). The sensor quickening signal, which has gravitational and body movement segments, was isolated utilizing a Butterworth low-pass channel into body speeding up and gravity. The gravitational power is expected to have low recurrence segments and a channel with 0.3 Hz cut off recurrence was utilized. From every window, a vector of highlights was acquired by ascertaining factors from the time and recurrence area. At last dataset comprises of 2947 records with 561 highlights. For each record it is provided as:

- The estimated body and triaxial acceleration from the accelerometer
- Triaxial Angular speed from the gyroscope.
- Its activity label.
- An identifier of the subject who carried out the analysis.

B. K-Nearest Neighbor

K-Nearest Neighbor is a supervised learning algorithm which is standout amongst the most mainstream algorithm for pattern recognition [20]. K-Nearest Neighbor algorithm utilizes neighborhood characterization as expected prediction. The conventional K-NN content classification algorithm has three confinements: (a) computation unpredictability because of the use of all the preparation tests for order, (b) the execution is exclusively subject to the preparation set, and (c) there is no weight distinction between tests. The exactness of K-NN algorithm can be debased in the presence of noise or inappropriate features. In pattern recognition, K-NN is a technique for arranging objects based on nearest training features in the element space.
C. Classification and Regression Tree

Classification and Regression Tree algorithm characterizes a sample as indicated by gatherings of different examples with comparative properties. During preparing, the preparation information is constantly isolated into smaller subsets. At the point when the divisions are done, the examples are grouped together as per their properties [15]. Testing tests are then assessed against specific conditions in every hub and engendered all through the tree. At the point when the example achieves a leaf hub, it is then doted out the class to which the examples in that hub have a place. In this paper, a parallel tree with legitimate conditions was utilized. Trucks are still under bread research and can be utilized as an independent classifier or as a piece of bigger algorithmic structures.

D. Support Vector Machine

In machine learning, Support vector machines are models that are related with learning algorithms for investigating and grouping information. A SVM algorithm makes a model that assigns guides to one or other divisions of class [20]. This model maps the precedents which isolate into the different classes. SVM’s can likewise play out a non-direct order. The activity of SVM depends on the hyper plane. The task of SVM depends on the hyper plane that gives biggest least separation to the preparation models. The attributes are taken after implementing dimensionality reduction technique.

E. Random Forest

Random Forest is an outfit of unpruned demand or descends like bootstrapping algorithm with various decision trees. Each tree depends upon the estimations of the vector picked unpredictably and independently. Random Forest reliably gives an immense improvement than the single tree classifier [21]. Each tree is fabricated using the algorithm. Let the amount of planning cases be N, and the amount of components in the classifier is M.

- The number m of input components to be used to choose the choice at a hub of the tree; m should be significantly not exactly M.
- Choose a preparation set for the tree by picking n times with substitution from all N available getting ready cases. Use whatever is left of the cases to assess the error by envisioning their classes.
- For each hub of the tree, heedlessly pick m factors on which to base the decision at that hub.
- Calculate the best split dependent to these m factors in the preparation set.
- Each tree is totally created and not pruned (as may be done in structure a common tree classifier).

F. Extra Trees

The Extra-Trees algorithm constructs a ensemble of unpruned choice or relapse trees as indicated by the established top-down methodology. Its two primary contrasts with other tree based group techniques are that it parts hubs by picking cut-focuses completely aimlessly and that it utilizes the entire learning test (as opposed to a bootstrap reproduction) to develop the trees. The Extra-Trees part methodology for numerical traits has two parameters: K, the quantity of characteristics arbitrarily chose at every hub and nmin. It is utilized a few times with the unique learning test to create a group model. The parameters K, nmin and M have various impacts: K decides the quality of the trait determination process, nmin the quality of averaging yield noise, and M the quality of the fluctuation reduces the total ensemble model. These parameters could be adjusted to the issue points of interest in a manual or a programmed way.

G. Gradient Boosting

Gradient boosting is an AI method for relapse and order issues, which creates an expectation model as a group of powerless forecast models, normally choice trees. The goal of any directed learning algorithm is to characterize a misfortune work and limit it [20]. Gradient boosting machines are in light of a ensemble of choice trees where numerous weak learner trees are utilized in mix as a group to give preferred forecasts over singular trees. Boost has unrivalled regularization and better treatment of missing qualities and also much improved proficiency.

H. Convolutional Neural Network with LSTM

The CNN Long-Short-Term Memory Network design includes utilizing Convolutional Neural Network (CNN) layers for feature extraction on input information joined with LSTMs to help arrangement prediction. CNN LSTMs were created for visual time arrangement forecast issues and the utilization of producing literary depictions from successions of images. It has characterizing two sub-models: the CNN Model for highlight extraction and the LSTM Model for deciphering the highlights crosswise over time steps. A 2D convolutional system as contained Conv2D and Max Pooling 2D layers. Apply the CNN model to each information picture and pass on the yield of each information picture to the LSTM as a solitary time step. Wrapping the whole CNN input model (one layer or more) in a Time Distributed layer. This layer accomplishes the ideal result of applying a similar layer or layers on different occasions. They have demonstrated extremely powerful on recognizing and automatically portraying the content.

I. Recurrent Neural Network with LSTM

Recurrent Neural Networks (RNN) are an amazing and vigorous sort of neural systems and have a place with the most encouraging algorithms since they are the main ones with an inside memory. In a RNN, the data pushes through a circle. When it makes a choice, it considers the present feature and furthermore what it has gained from the previous sources. Recurrent Neural Networks add the quick past to the present. RNN's can delineate to many, numerous to many (interpretation) and numerous to one. There are two noteworthy hindrances RNN's needed to manage

- An inclination is a fractional subordinate concerning its inputs. A slope estimates how much the yield of a capacity change, on the off chance that you change the sources of info a tad.
- The higher the angle, the more extreme the slant and the quicker a model can learn. If the incline is zero, the model stops to learning.
A slope just estimates the adjustment in all loads concerning the adjustment in blunder.

A long Short-Term Memory (LSTM) system fundamentally broadens their memory. The units of a LSTM are utilized as structure units for the layers of a RNN, which is then regularly called a LSTM, enables RNN to have their inputs for a prolonged period. LSTM's empower RNN's to recall their data in a memory.

![Diagram of data collection and storage process](image)

**Fig. 2 Process of collecting and storing data from a smartphone**

V. SYSTEM STRUCTURE

A. Human Activity Recognition [HAR]

Human Activity Recognition is monitoring of activities performed by the people who includes following activities such as: walking, walking_upstairs, walking_downstairs, sitting, standing and laying. The flow works like as shown in figure 2 which depicts the process of collecting and storing data from smartphone.

- The person carries Smartphone with them and gathered data will be stored in their memory of smartphone.
- The gathered informed, passed on to server and saved.
- Every activity has predefined training set will be compared with template by classification process.

The accuracy of the HAR model depends on the classifier deployed.

B. Proposed System

The Proposed HAR is designed as two models. One model used CNN with LSTM and other model is deployed as RNN with LSTM. Initially, the dataset called Human Activity Recognition [HAR] is collected from UCI machine learning Repository. The dataset is preprocessed using noise filters. After Preprocessing, the data is splitted as fixed windows. Feature engineering technique is applied to window data. The window data is splitted as 70% training set and 30% testing set. Feature Engineering has following sections.

Primarily, the dataset is loaded which has three main types of signals such as total acceleration, body acceleration and gyroscope. Then, develop an LSTM network model. LSTM can learn and remember over long sequences of data. After modeling, need to define, fit and evaluate an LSTM. Then, Convolution Neural Network [CNN] and Recurrent Neural Network [RNN] is applied with LSTM model to make predictions. The Proposed System working flow is shown in figure 3.
Smartphone Sensor Based Human Activity Recognition using Deep Learning Models

Fig. 3 Working Flow of Designed Model

Fig. 4 Accelerometer Signals for each activity
VI. EXPERIMENTAL RESULTS

In this experiment, the human activity is recognized based on their movements is initiated with preprocessing. The ConvLSTM and RNNLSTM is used as classifiers. The accelerometer data for all the six activities are displayed in figure 4. The performance of the model is validated based on Accuracy, Mean Absolute Error [MAE], Root Mean Square Error [RMSE] and Mean Absolute Percentage Error [MAPE].

<table>
<thead>
<tr>
<th>CLASSIFIER</th>
<th>ACCURACY (%)</th>
<th>Mean Absolute Error [MAE]</th>
<th>Root Mean Square Error [RMSE]</th>
<th>Mean Absolute Percentage Error [MAPE] (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K- Nearest Neighbor [K-NN]</td>
<td>61.893</td>
<td>4.87</td>
<td>5.33</td>
<td>58</td>
</tr>
<tr>
<td>Classification and Regression Trees [CART]</td>
<td>72.14</td>
<td>2.987</td>
<td>3.08</td>
<td>49.4</td>
</tr>
<tr>
<td>Support Vector Machine [SVM]</td>
<td>76.96</td>
<td>2.68</td>
<td>2.96</td>
<td>46.7</td>
</tr>
<tr>
<td>Random Forest</td>
<td>84.66</td>
<td>0.4</td>
<td>1.245</td>
<td>15.8</td>
</tr>
<tr>
<td>Extra Trees</td>
<td>86.90</td>
<td>0.02</td>
<td>0.0554</td>
<td>6.90</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>87.61</td>
<td>0.0128</td>
<td>0.0336</td>
<td>6.78</td>
</tr>
<tr>
<td>Convolutional Neural Network with Long Short Term Memory [ConvLSTM]</td>
<td>92.24</td>
<td>0.0007</td>
<td>0.0044</td>
<td>5.91</td>
</tr>
<tr>
<td>Recurrent Neural Network with Long Short Term Memory [RNNLSTM]</td>
<td>93.89</td>
<td>0.0004</td>
<td>0.0024</td>
<td>4.78</td>
</tr>
</tbody>
</table>

Table 1: Performance Metrics of the Classifier

Fig. 5 Comparison of Accuracy for different models

A. Performance Metrics

MAE is that the normal over amassed dataset tests. It offers total variety between forecast and genuine perception. Root Mean Square Error (RMSE) is a quadratic reviewing rule estimating the normal greatness of the mistake. It's the square base of the normal of squared varieties among forecast and genuine perception. MAE esteem ought to be not exactly or equivalent to RMSE. MAPE is utilized to quantify the exactness of expectation technique. Regularly, the littler these qualities are, the better the expectation. The Table 1 demonstrates the performances of the designed models.

B. Results

The experiments have been done on Python of version 3. The results for proposed model are shown in table 1, figure 5 and 6. The result shows that Recurrent Neural Network with Long Short Term Memory [RNNLSTM] provides better accuracy with lower mean absolute percentage error [MAPE]. RNNLSTM have a better balanced accuracy of 93.89%, mean absolute error [MAE] of 0.0004, Root Mean Square Error [RMSE] of 0.0024 and Mean Absolute Percentage Error [MAPE] of 4.78%. Due to that, RNNLSTM can be used to recognize the human activities in real world applications inorder to reduce loss of lives. The figure 5 shows the comparison of accuracy for different classifiers and figure 6 shows the comparison of MAE, RMSE and MAPE.

VII. CONCLUSION

Smart phones are pervasive and winding up increasingly modern. This has been changing the scene of individuals' day by day life and has opened the entryways fascinating information mining applications. Human action acknowledgment is a center structure hinder behind these applications. It takes the crude sensor's perusing as sources of info and predicts a client's movement action. This paper exhibits an extensive overview of the ongoing advances up to 93.89% on different movement acknowledgment with PDA's sensors. The information of various classifiers was utilized for assessing acknowledgment execution. Consolidating several classifiers, best classifiers utilizing the normal of probabilities strategy ended up being the best classifier for movement acknowledgment, outperforming all other classifiers. Further
demonstrated the acknowledgment strategy can identify exercises autonomous of cell phone’s position. For future work, movement acknowledgment task in a few different ways. To start with plan to perceive extra exercises. Second, want to gather information from more clients of different ages. Third, intend to remove more highlights that could more likely segregate various exercises.

REFERENCES


Fig. 6 Comparison of MAE, RMSE and MAPE for different models

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication
Retrieval Number: A1385058119/19©BEIESP
2747
AUTHORS PROFILE

Ms. S. Roobini was born in Coimbatore, Tamil Nadu, India, in 1994. She received M.E degree from PSG College of Technology in the year 2017, Coimbatore. She has secured TANCET scholarship during her P.G. course. She pursued Best Project Award in P.G Course. She received B.E degree from SNS College of Technology. Currently she is serving as Assistant Professor at Sri Krishna College of Engineering and Technology, Coimbatore. Her research areas are cloud computing, data mining and IoT. She has published research papers in International Journals and more than five papers in National and International Conferences. She is a member in Indian Society of Technical Education [ISTE].

Ms. J. Fenila Naomi was born in Coimbatore, Tamil Nadu, India, in 1993. She received M.Tech degree from PSG College of Technology in the year 2018, Coimbatore. She has secured TANCET scholarship during her P.G. course. She received B.E degree from Hindusthan College of Engineering and Technology. Currently she is working as Assistant Professor at Sri Krishna College of Engineering and Technology, Coimbatore. Her research areas are cloud computing, data mining and IoT. She has published research papers in International and National Journals.