International e-Conference on New Horizons And Multidisciplinary Applications In Science And Technology
In Association with International Journal of Scientific Research in Science and Technology
Volume 9 | Issue 6 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)

A Deep Learning Approach for Human Activity Recognition Using Smart Phone

Prof. P. H. Pawar[1], Mr. Jitendra S. Rokade[2]

[1]Assistant Professor, Department of Computer Science & Engineering, Babasaheb Naik College of Engineering, Pusad, Maharashtra, India

[2]Student, Department of Computer Science & Engineering, Babasaheb Naik College of Engineering, Pusad, Maharashtra, India

ABSTRACT

Human Activity Recognition is an emerging field of study with a lot of innovations and applications. With digitalization, mobile development and advancement in technology taking over mankind, Smartphones have become an integral part of our life. We've been so dependent on Science and its innovations, that living without mobile phones is nearly impossible. With advancement in technology, comes the responsibility of providing mankind with efficient, conventional and sustainable resources. Our project aims to implement the idea of "Technology at your fingertips".

The number of elderly people is predicted to elevate over the years, "aging in place" (living at home regardless of age factors and other aspects) is becoming an important topic inthe area of ambient assisted living (AAL). Therefore, we have proposed a human activity recognition system based on data collected from Smartphone motion sensors for daily physical activity monitoring. The proposed approach implies developing a prediction model using mainly two sensors available on a smartphone: accelerometer and gyroscope.

We have chosen to implement our solution on mobile phones because they are ubiquitous and do not require the subjects to carry additional wearable or mountable devices or sensors that might impede their activities. For our proposal, we target six basic human daily activities walking, jogging, sitting, standing, ascending, and descending stairs. We evaluate the solution against two datasets (one using only accelerometer data and the other using only gyroscope data) with great effect. We've also implemented predictive models using Deep Learning approaches LSTM (Long Short- Term Memory) and CNN (Convolution Neural Networks). The predicted results show decent especially good and accurate results obtained for walking, running, sitting, and standing. The proposed system is fully implemented on a smartphone device as an Android application, which proves to be efficient, sustainable, and accurate.

This study revolves around Deep Learning methodologies and techniques, precisely an approach based on semantic analysis mimicking the human ability to perform various activities.

Keywords : Human Activity Recognition, Real Time Activity Recognition, Long Short-Term Memory (LSTM), Deep Learning, Convolutional Neural Network (CNN), Gyroscope, Accelerometer, Smartphone Sensors.

I. INTRODUCTION

Smartphones have become an integral part of human life. With its advancing features and specifications, the demand for owning a smartphone has rapidly grown over the years. This piece of technology has really proved to be a boon to mankind, assisting man throughout his daily activities. According to the survey carried out by Statista, the number of smartphone users worldwide surpasses three billion. More than half of the world's populations are smartphone users and it is forecasted to further grow by several hundred million in the next few years.

Technology definitely has its own pros and cons but taking the utmost advantage of it, can be beneficial to mankind. One such field where smartphones play an important role and is a budding area of research and development, is the Human Activity Recognition (HAR).

AR (Activity Recognition) is one of the most significant innovations in the sector of Health monitoring, mobile health applications and user's activity tracking. Utilizing this important feature of smartphones and their sensors, the study of HAR is developing and evolving over time.

Activity Recognition (AR), identifies the daily activities that is performed by a user, has been intriguing and demanding within ubiquitous and mobile computing. The growth and evolution of smart mobile devices and their astounding features with sensing, processing, and network capacity have opened up a huge range of possibilities for activity recognition. Studies for HAR are extensively using Smartphones due to their unassertive, high end hardware and features, none or low costs of installation and ease of use, making it sustainable and feasible.

The HAR system is executed by taking input from smartphones, whereby exploiting data recovered from inertial sensors and observing the human movement using various approaches. Though smartphones today are loaded with a whole lot of sensors like light sensors, motion sensors, compass sensors, this study specifically aims at using two motion sensors available in every smartphone. The reason being its availability and feasibility, making it available to all and cost effective so that its usability increases.

The work described in this paper relies on Android smartphones (used to extract data in preparing the dataset), but the tri-axial accelerometer and gyroscope are basic motion sensors present in all other smartphones and mobile devices.

The aspiration of this project is to use this proposed system to track the behavior of older adults who are in need of constant monitoring, ensuring that they behave in normal parameters. Adding more features and parameters to the application can further detect more activities and give recommendations during inactivity. Hence, we have limited the the exploitation of sensor data from smartphones only, as we believed that human carried sensors and devices would discourage older adults from participating or using the application.

In this paper, we have implemented a system based on Human Activity Recognition using Deep Learning approach and deployed an Activity tracking application, with the hope that it will encourage the development and influence the direction of activity recognition research and bring into existence a feasible and sustainable application.

II. LITERATURE REVIEW

The study of Human Activity Recognition (HAR) began with implementations using the Machine Learning Approach. Over the years, there were many algorithms and methods used for solving real world problems. There is a lot of research work and executions completed using Machine Learning Algorithm. This section gives a brief overview of the transition of HAR from ML to DL, advancing over the years.

In the research conducted in [12], Machine Learning approach was used for the implementation of HAR. Algorithms like Decision Trees, Support Vector Machines, K- nearest neighbors (KNN) and Ensemble classification methods were listed in the paper that were executed using multimode motion sensos. SVM was the most precise approach tested with the highest success rate of 99.4% while the other methods also created effective models.

According to the published work in [13], SVMs have again proved to be the most accurately predicting models. With the introduction of Deep Learning, these ML approaches have lost their limelight.

In recent years, several researchers proposed DL-based solutions for the HAR problem. Though the accuracy rates were fluctuating in both ML and DL approaches, Deep Learning outperforms Machine Learning methods with its intuitive feature selection process. Various studies have also proved that Deep Learning techniques' self-learning capabilities lead to higher accuracy of results and faster processing. The first applications of Deep Learning methods have been implemented in computer vision and natural language processing [14]. The exploration in the field of DL provides scope for significant contribution to HAR and its applications.

Among the DL models, CNNs attracted the attention of several HAR researchers. [15] Use a 1-Dimensional CNN to classify activity data recorded by smartphone sensors. They compare the performance of their proposed model with some shallow ML models e.g., SVMs and DTs. The results indicate that the CNN model is more accurate. [16] use a 2- Dimensional CNN to classify six daily activities recorded from 12 volunteers. They compare their method with traditional ML methods with respect to both accuracy and computational overhead. The results indicate improvement with respect to both measures.

In [17], the authors apply different variants of RNNs (e.g., GRUs and LSTMs) to recognize daily activities and detect abnormal behavior of the elderly people suffering fromdementia. They compare the performance of these models with shallow ML models. The comparison results indicate that RNNs outperform other ML models with respect to most of the evaluated measures (e.g., accuracy, precision and recall), and among the investigated RNN models, LSTMs performed slightly better. [18] use LSTMs to classify human activity data collected by smart-home sensors. They also compare LSTMs with CNNs and traditional ML moldels in [19]. Their evaluations indicate that LSTMs and CNNs outperform other ML models, and CNNs are much faster than LSTMs in training but less accurate. As future work, they propose to combine CNNs and LSTMs to take benefit of both.[4] A study conducted by [4] published a comparative analysis of hybrid deep learning models for HAR, that has given outstanding results.

This is the new advancement to the evolution of Deep Learning.

III. METHODOLOGY

Human activity recognition plays a big role in human-to- human interaction and interpersonal relations. Because it provides information about the identity of an individual, their personality, and mental state, it's difficult to extract. The human ability to acknowledge another person's activities is one among the most subjects of study of the scientific areas of computer vision and machine learning. As a result of this research, many applications, including video surveillance systems, human- computer interaction, and robotics for human behaviour characterization, require a multiple activity recognition system.

For many years human action/activity recognition has been studied well. Most of the action recognition methods require to manually annotate the relevant portion of the action of interest within the video or the kind of input data. In recent years it's been studied that the relevant portion of action of interest are often acknowledged automatically and recognize the action. We will review the action recognition methods within the following sections of this paper.

For many years human action recognition has been studied well. Most of the action recognition methods require to manually annotate the relevant portion of the action of interest in the video. In recent years it has been studied that the relevant portion of action of interest can be found out automatically and recognizes the action. We can review the action recognition methods

HAR datasets

We have implemented the current project with two pre available datasets.

1. UCI Machine Learning Repository University of California Irvine [5]
2. WISDM WIreless Sensor Data Mining lab [10]

The main objective of choosing two datasets and implementing Deep Learning approaches separately on each one of them, was to study how the size of the dataset affects the overall performance of the Model.

It was quite evident that we got a better accuracy rate with the larger dataset as input. The results are further mentioned in the paper.

According to the research carried out in [1], their study states that the combination of accelerometer and gyroscope signals, also called multimodal recognition, increases the accuracy in HAR with respect to the use of each signal alone.

The paper consisted of the results of an analysis that was performed in order to compare the effectiveness of machine learning techniques when used separately or jointly on accelerometer and gyroscope signals.

It also states that the results show that the use of deep learning techniques in multimodal mode (i.e., using accelerometer and gyroscope signals jointly) outperforms other strategies of at least 10%. [1]

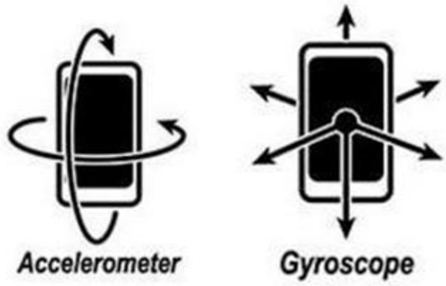
However, we have also previously researched and published a Human Activity Recognition case study with Machine Learning Models which clearly demonstrates the difference between the Approaches used. [2] Sensors based: Accelerometer

HAR research and implementation initially began with the use of accelerometers. Over the years now, the accelerometer has been an integral part of the contribution to the research of Activity Recognition. The accelerometers were used mainly because they were low- cost, low- power and compact sensors. These sensors are practically found in every smartphone today which makes it sustainable and feasible and provide motion-related information. The use of these sensors makes it effortless to recognize activities reliably and robustly.

An accelerometer is an electro mechanical device that will measure acceleration forces. These forces may be static, like the constant force of gravity pulling at your feet, or they could be dynamic which is caused by moving or vibrating the accelerometer. [3]

Sensors based: Gyroscope

Similar to the Accelerometer motion sensor, the gyroscope sensor has been widely used for Activity Recognition. A gyroscope is a device that uses Earth's gravity to help determine orientation. Gyro sensors are devices that sense angular velocity which is the change in rotational angle per unit of time. [3]



Recording Parameters	UCI Dataset	WISDM Dataset
Smartphone mounting	Waist - mounted	Pants' front pocket
Sensors	Accelerometer and Gyroscope	Accelerometer only
Number of subjects	30	36
Subject's age bracket	19 - 48	Not specified
Sampling Rate	50 Hz	20 Hz
Sample Window	128 readings with 50% overlap	200 readings with no-overlap
Different activity	Laying	Jogging
Number of instances	10299	1,098,207

Table -1: Accuracy Table

UCI Dataset

The UCI HAR dataset is an open source 6- activity dataset that contains 3D (x, y, z) raw signals extracted from the accelerometer and gyroscope motion sensors of a smartphone strapped to the waist of a subject [5]. The experiments were carried out with a group of 30 subjects within an age bracket of 19-48 years. Each person performed six basic activities of Walking, Walking Upstairs, Walking Downstairs, Sitting, Standing, and Laying. The experiment captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The sensor signals were preprocessed by applying noise filters and sampled in fixed- width sliding windows of 2.56 sec with a 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. [11]

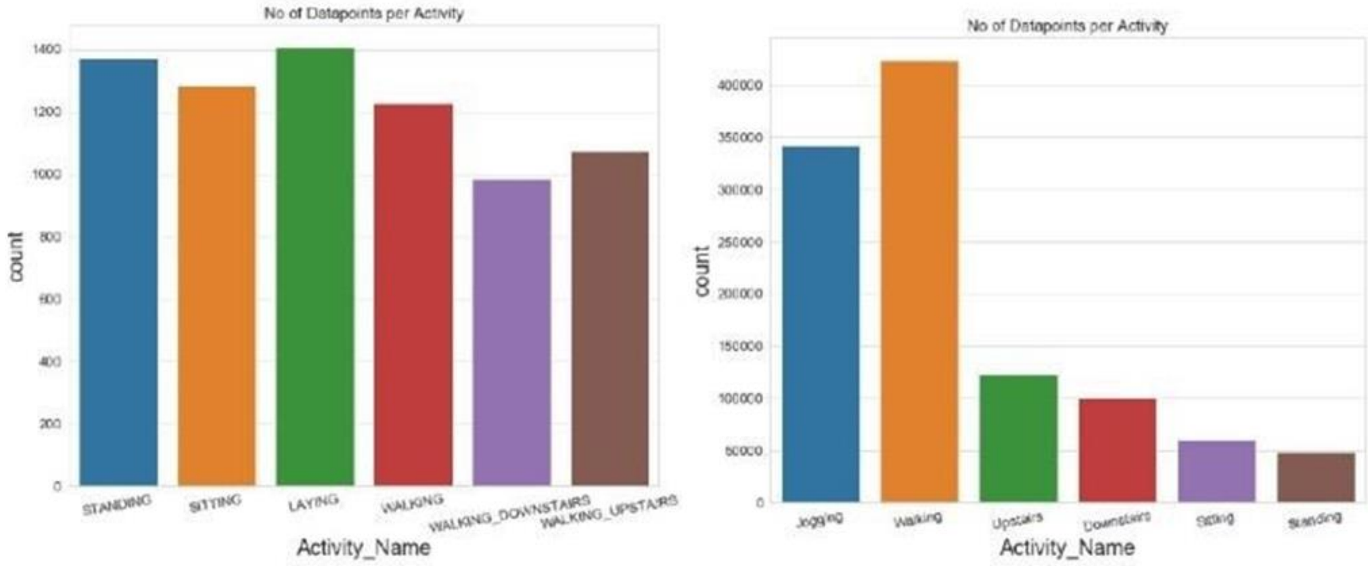
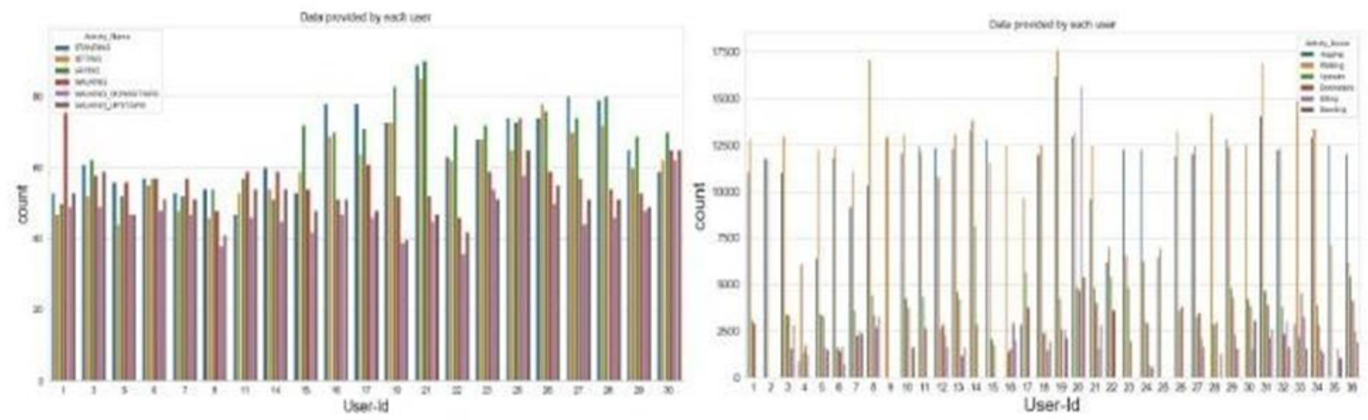


Chart -1: Datapoints plotted from the two datasets

1.3.3. WISDM Dataset

The WISDM dataset is composed of raw accelerometer data of human activities collected from 36 users. The data was recorded using a smartphone while performing six daily human activities: Walking Downstairs, Jogging, Sitting, Standing, Walking Upstairs, and Walking. The dataset has a total of 1,098,207 instances. Data transformation was made with a sampling rate of 20 Hz, which whereby 10 s worth of accelerometer samples (200 lines of the raw data) were taken to transform them into a single tuple of proposed composite features. The measured sensor data is a 3-axisaccelerometer data for each activity along with the timestamp and the user-id [10]. Before calculating any feature, the raw accelerometer data was also preprocessed to reduce noise using median filter of order 5 in each dimension separately. The total acceleration signal in time domain, ‘A’, captured by accelerometer is known to be the sum of gravity (‘g’) and body (‘B’) accelerations. A 3rd order Butterworth low pass filter is used with a cutoff frequency of 0.3 Hz to separate the acceleration signal into gravity and body acceleration signals. [8][9]

The published work in [8] shows the study based on the datasets used for Human Activity Recognition.



(a) UCI HAR Data Set (b) WISDM Data Set

Chart -2: Data points by each user plotted from the two datasets

1. HAR using Deep Learning

Human Activity Recognition (HAR) is an emerging field of research and study, which has implementations of both Machine Learning and Deep Learning. Due to advancement in the areas of Science and Technology, the introduction of many new possibilities and studies have sprouted. Though Machine Learning and Deep Learning are both emerging fields in Data Science, Deep Learning proves to be more advanced and is still advancing over time. Machine learning (ML) has been at the core of research into human activity recognition for a very long time [5]. However, since AlexNet [6] won the ImageNet competition in 2012, deep learning has seen successful applications in a multitude of domains.

Computer vision, natural language processing, speech recognition, etc. This success led several researchers to use various deep learning approaches in solving the HAR problem [7].

The biggest advantages of Deep Learning models over ML models are their capability of learning complex features from raw data and the automatic feature extraction, improving the performance along with the accuracy rate.

Hence, we have opted to use the Deep Learning approach to eliminate the need of pre-knowledge and handcrafted feature extraction.

In this section, we have given a brief explanation of the DL models with respect to HAR, used in our experimental study.

CNN

CNNs are deep learning models widely used in computer vision. The architecture of a CNN is very similar to that of a visual cortex in the human brain. Through some filters, CNNs are able to extract features (i.e., spatial and temporal dependencies) and distinguish the objects within the input image. The filters constitute the convolution layers, which are usually followed by some fully connected layers responsible for the classification task. Other than being good at learning features, through some pooling layers, CNNs can scale to massive datasets. In fact, the purpose of pooling layers is reducing the dimensionality of input data and also extracting dominant features, which are invariant with respect to rotation and position.

LSTM

LSTMs are an extension of RNNs, which perform much better than standard RNNs when it comes to remembering dependencies for a long time. This capability is due to the structure of the repeating module in these networks. In LSTMs, the repeating module comprises four interacting layers. These layers include a layer called the cell state, together with three other layers called gates. Cell state acts as the RNN memory. Gates are ANN layers responsible for controlling the information added to/removed from the cell state. In other words, these gates allow more relevant information to flow to the cell state and prevent the flow of less relevant information. [4]

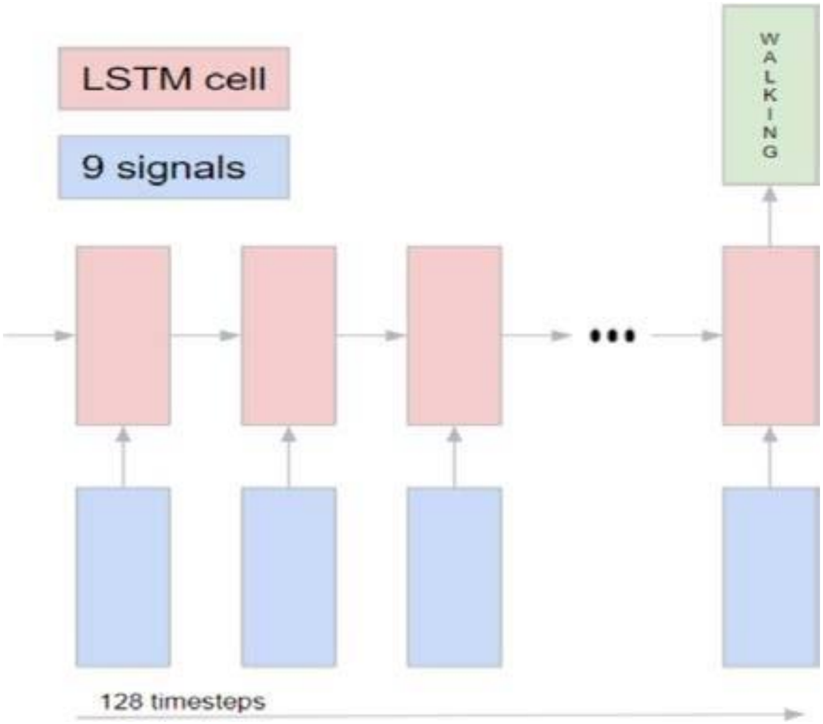


Fig -1: LSTM Model

CNN-LSTM

The figure below depicts the CNN-LSTM model, combination of the LSTM and CNN, Deep Learning models combined to make the Hybrid model.

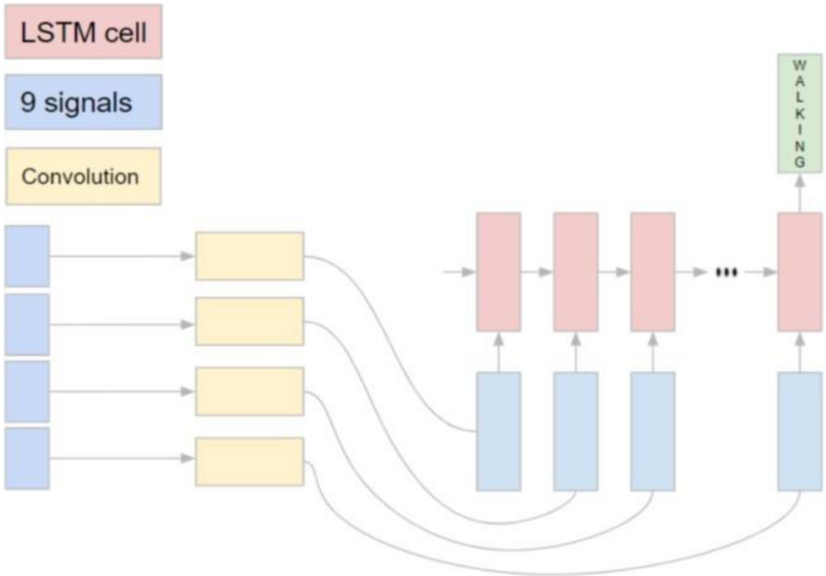


Fig -2: CNN-LSTM Model

Implementing CNN LSTM in Keras

A CNN LSTM model is expected to be trained jointly in Keras. A CNN LSTM can be defined by adding CNN layers on the front end followed by LSTM layers with a Dense layer on the output.

The architecture can be divided as two models:

1. The CNN Model for feature extraction and

2. The LSTM Model for interpreting the features across time steps.

The CNN LSTM model can be defined in Keras by first defining the CNN layer or layers, wrapping them in a TimeDistributed layer and then defining the LSTM and output layers.

There are two ways to define the model that are equivalent.

1. Defining the CNN model first, and then adding it to the LSTM model by wrapping the entire sequence of CNN layers in a TimeDistributed layer.
2. An alternate approach is to wrap each layer in the CNN model in a TimeDistributed layer when adding it to the main model.

7.Results and Accuracy

ACCURACY OF ALL THE MODELS IMPLEMENTED			
DEEP LEARNING MODELS → DATASET ↓	LSTM	CNN	CNN-LSTM
UCI (Dataset One)	88%	93%	88%
WISDM (Dataset Two)	97%	94.70%	94.78%

Table -2: Accuracy Table

The above table (Table 2) enlists the various approaches implemented and their testing accuracy.

8. HAR Application: Real Time Activity Detection

The need of Activity Detection in Real Time comes into picture when an efficient predictive model-based application needs to be deployed. We have successfullyimplemented and deployed an application using the LSTM model that stood out best in our research. The application is Android based and predicts activity in Real timeaccurately and efficiently.

IV. CONCLUSIONS

Taking into consideration our work of research, and the researches and studies reviewed for this paper, we come to a few conclusions:

1. Larger datasets do affect the accuracy of the models irrespective of the approach used.
2. The more the number of sensors, the more signals and increased data, this helps make Activity Recognition more accurate.

As an exception we do notice that our paper gets the highest accuracy from the LSTM model implemented on the WISDM dataset that comprises of data from signals of a single motion sensor.

We intentionally chose to show out this difference, since we felt the need to make technology feasible and available to all with minimum requirements. Since the main objective of this project was to design a prediction model and activity tracker for the older adults, we assumed that a normal smartphone with minimal sensors would serve the purpose. And as shown our model performs the best from the rest. This theory may not be applicable to all sectors, but with improvement and advancement with technology comes the responsibility to cater to human needs. That's what been our focus all throughout the development of this study and deployment of the application. The paper [21] is a review of all the various tools and techniques which can be used in human activity recognition that included machine learning algorithms and neural network techniques. The survey concludes by deducing that there is no single method which is best for activity recognition of any type. There are various factors that are affecting that need to be taken into consideration while choosing the appropriate approach. Although there are numerous methods, some of the challenges still remain open and have to be resolved.

3. The idea of Hybrid Models is an advantage to the implementation of Deep Learning Approach. Though we implemented a CNN- LSTM model and got a accuracy rate less than a simple LSTM, we would eagerly try implementations with more hybrid models.

V. ACKNOWLEDGEMENT

We sincerely thank our mentor and guide Prof. P.H.Pawar for her constant guidance and help in our research. Without his tutelage, we wouldn't be able to complete this papert. He has always been encouraging us to bring about the best and be zealous in everything we do. Her constant enthusiasm, enduring spirit and endless efforts has really motivated us to work on this project under her mentorship.

We are also grateful to our college, Babasaheb Naik College of Engineering,Pusad for giving us this opportunity to work with them and providing us with the necessary resources for the project. Working on this project also helped us to do lots of research and we came to know about many new aspects in the field of Data Science.

We are immensely grateful to all involved in this project, who helped us throughout as without their inspiration and valuable suggestions it would not have been possible to develop this project within the prescribed time.

VI. REFERENCES

[1]. Varma T, John S, Joy S, James J, Menezes A. An Analytic Study on Human Activity Recognition using Smartphones.

[2]. Abbaspour S, Fotouhi F, Sedaghatbaf A, Fotouhi H, Vahabi M, Linden M. A comparative analysis of hybrid deep learning models for human activity recognition. Sensors. 2020 Jan;20(19):5707.

[3]. D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes- Ortiz, "A Public Domain Dataset for Human Activity Recognition using Smartphones," Proc. of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, 24–26 April, 2013.

[4]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, 25(2), pp. 1097–110, 2012.

[5]. Mutegeki R, Han DS. A CNN-LSTM approach to human activity recognition. In2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) 2020 Feb 19 (pp. 362-366). IEEE.

[6]. Bilal M, Shaikh FK, Arif M, Wyne MF. A revised framework of machine learning application for optimal activity recognition. Cluster Computing. 2019 May;22(3):7257-73.

[7] . Khare S, Sarkar S, Totaro M. Comparison of Sensor- Based Datasets for Human Activity Recognition in Wearable IoT. In2020 IEEE 6th World Forum on Internet of Things (WF- IoT) 2020 Jun 2 (pp. 1-6). IEEE.

[8] . Gary M. Weiss Jennifer R. Kwapisz and Samuel A. Moore. Activity recognition using cell phone accelerometers. Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data (atKDD-10), Washington DC., 2010.

[9] . Lichman M. UCI machine learning repository.

[10] . Bulbul E, Cetin A, Dogru IA. Human activity recognition using smartphones. In2018 2nd international symposium on multidisciplinary studies and innovative technologies (ismsit) 2018 Oct 19 (pp. 1-6). IEEE.

[11] . Subasi A, Fllatah A, Alzobidi K, Brahimi T, Sarirete A. Smartphone-based human activity recognition using bagging and boosting. Procedia Computer Science. 2019 Jan 1; 163:54-61.

[12] . LeCun Y, Bengio Y, Hinton G. Deep learning. nature. 2015 May;521(7553):436-44.

[13] . Ronao CA, Cho SB. Human activity recognition with smartphone sensors using deep learning neural networks. Expert systems with applications. 2016 Oct 15; 59:235-44.

[14] . Zebin T, Scully PJ, Ozanyan KB. Human activity recognition with inertial sensors using a deep learning approach. In2016 IEEE SENSORS 2016 Oct 30 (pp. 1- 3). IEEE. Arifoglu D, Bouchachia A. Activity recognition and abnormal behaviour detection with recurrent neural networks. Procedia Computer Science. 2017 Jan 1; 110:86-93.

[15] . Singh D, Merdivan E, Psychoula I, Kropf J, Hanke S, Geist M, Holzinger A. Human activity recognition using recurrent neural networks. InInternational cross- domain conference for machine learning and knowledge extraction 2017 Aug 29 (pp. 267- 274). Springer, Cham.

[16] . Zebin T, Scully PJ, Ozanyan KB. Human activity recognition with inertial sensors using a deep learning approach. In2016 IEEE SENSORS 2016 Oct 30 (pp. 1-3). IEEE.

[17] . Arifoglu D, Bouchachia A. Activity recognition and abnormal behaviour detection with recurrent neural networks. Procedia Computer Science. 2017 Jan 1; 110:86-93.

[18] . Singh D, Merdivan E, Psychoula I, Kropf J, Hanke S, Geist M, Holzinger A. Human activity recognition using recurrent neural networks. InInternational cross-domain conference for machine learning and knowledge extraction 2017 Aug 29 (pp. 267- 274). Springer, Cham.

[19] . Singh D, Merdivan E, Hanke S, Kropf J, Geist M, Holzinger A. Convolutional and recurrent neural networks for activity recognition in smart environment. InTowards integrative machine learning and knowledge extraction 2017 (pp. 194-205). Springer, Cham.

[20] . Jobanputra C, Bavishi J, Doshi N. Human activity recognition: A survey. Procedia Computer Science. 2019 Jan 1; 155:698-703.