Human Activity Recognition using Long Short-Term Approach

Ankit Kashyap
Department of Computer Science, Bangalore Institute of Technology, Bangalore, India
ankitkashyap367@gmail.com

Animesh Banathia
Department of Computer Science, Bangalore Institute of Technology, Bangalore, India
animeshbanathia11@gmail.com

Ankit Ranjan
Department of Computer Science, Bangalore Institute of Technology, Bangalore, India
ankit.rj876@gmail.com

Rishabh Sipani
Department of computer Science, Bangalore Institute of Technology, Bangalore, India
rsipani9@gmail.com

Dr. Harish Kumar B T
Assistant Professor, Dept. of computer Science, Bangalore Institute of Technology, Bangalore India
harish.bitcse82@gmail.com

ABSTRACT

Human Activity Recognition is one of the most popular area of research for machine learning and deep learning enthusiasts. It is a classification problem which aims to learn predicting the class of activities one is performing. This paper takes Long short-term memory approach for this time series classification problem. We try to classify six activities which are standing, sitting, jogging, walking, walking upstairs and walking downstairs. We build and train the model using Google's TensorFlow library. Our experiment was done with dataset by WISDM (Wireless Sensor data mining) dataset. LSTM approach shows that it performs better than traditional algorithm with 97% accuracy.

Keywords: Machine Learning, Deep learning, HAR, Human Activity recognition, LSTM, RNN, TensorFlow, Time Series Analysis

1. INTRODUCTION

Human Activity Recognition aims to classify and detect different activities being performed like walking, sleeping, lying down, jogging, walking upstairs and downstairs. Researchers use any one of two ways of human activities recognition. First way is to use data collected from different sensors and other one is using computer vision system. The former method is widely popular because of various reasons that include portability, easier mounting at different places and low power consumption. While later method suffers due to blurriness in videos and images.

Human Activity Recognition aims to classify activities based on data received from set of sensors present in smartphones. In the era of digitalization, almost everyone now has access to smartphones. Smartphone have become ubiquitous around the world. These smartphones come with built in sensors that include accelerometer, gyroscope, proximity sensors etc. These sensors response to movement in body carrying smartphones.

Accelerometer is one of the popular sensors that senses the change in position or movement. It responds dynamically to sense movements or vibrations. Accelerometer measures acceleration in m/s^2 along all the three axes x, y and z. Accelerometers can also sense orientation. Another popular sensor used in activity recognition is gyroscope. It basically adds another dimension to data supplied by accelerometer adding angular velocity or rotation axis. Many of the researchers have used data from both sensors or only from either one.

With the availability of many wearable device that may have more other sensors than these two like heart beat monitors, HAR is progressing in applications to make life healthier. Practically thinking it would not be prudent to use external sensors. However, it would be wise to leverage in-built sensors in smartphones which is easier and comfortable to carry. In this paper we have used deep learning techniques called long short-term memory (LSTM) to train a model from publicly available dataset containing data only from accelerometer. The learnt model then will in turn try to predict activities in real time.

LSTM: We worked out this project using special deep learning technique called LSTM. Long short-term memory (LSTM) is a man-made recurrent neural network (RNN) architecture utilized in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can't only process single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks like unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic or IDS's (intrusion detection systems). A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and therefore the three gates regulate the flow of data into and out of the cell.

2. RELATED WORKS

Cheng et al., [1] formulated the problem of activity recognition using wearable sensors. The authors used three machine learning methods namely support vector machine, hidden markov model and artificial neural network. While conducting the experiment they collected their data by putting accelerometer on their body at different places. For each classification they also used two strategies to test the validity: one vs all and one vs own. In one vs own authors have used the data from same subject on which they were trained. The accuracy for these three algorithms were all above 90% but accuracy for one vs all testing method was below satisfactory. Authors argue that many discrepancies may cause this such as gender, age and weight.

Haihua Gong et al., [2] studied the pattern mining based on human activity recognition and location services. Authors have developed android application to collect inertia data. They applied raw sensor data as input to xgboost model to classify 8 different activities. They also applied one vs all strategy to improve the accuracy. The testing accuracy of their model was 92%. They also mined the pattern of their activities such as careers and consumption and amount of exercise level.

Ahmad Jalal et al., [3] used concept of hierarchical feature extraction for human movements from data of accelerometer. Hierarchical features are employed which include signal magnitude, min/max components, mean and standard deviation. Authors leveraged the fact that SVM offers multiclass classification while being binary classifier. Authors trained and tested for activities in indoor and outdoor environment with recognition accuracy hierarchical feature extraction being around 82%.

Tahmina Zebin et al., [4] in their research used a deep learning method called Convolution neural Network (CNN). They also studied the effect of different hyperparameters such as number of convolution layer and kernel size on performance of CNN. The result showed that CNN provided significant speed up in computation and marginal improvement of accuracy over classification techniques such as SVM.

Qingzhong Liu et al., [5] used sensor-based data for classification of activities using deep learning and machine learning technique. There were two categories of data collected motion based like walking and phone movement like clockwise or anti clockwise. Authors also performed feature extraction on raw sensor data. Based on these features extracted, they were analysed using Machine Learning technique such as SVM classifier with linear kernel and fisher linear discriminant. Second method used is CNN to analyse raw sensor data from accelerometer and gyroscope after normalization. Their study showed that when both the sensors data i.e. accelerometer and gyroscope are used together performs well compared to individual sensors data.

3. DATASETS

Datasets that we used for our experiment was sourced from Wireless Sensor Data Monitoring (WISDM) which collected these data in controlled, laboratory conditions [6]. Datasets comprises of six attributes with no missing values. The six class of activities includes Walking, Jogging, Sitting, Standing, upstairs and downstairs. Since LSTM model requires fixed length input, we transformed the datasets into batch of 200 data so that each generated sequence contains 200 set of data from accelerometer. One hot encoding was also applied during pre-processing for label of activities. One hot encoding was done using sklearn library. Fig 1 depicts the distribution of training examples by activities.

Figure 1. Distribution of training examples by activities

4. PROPOSED METHODOLOGY

The model has been built using LSTM architecture and is trained on only accelerometer data to classify 6 activities. Our model contains 2 fully-connected and 2 LSTM layers with 64 units each. The hidden layers use ReLU as activation function. In the context of artificial neural networks, the rectifier linear Unit is an activation function defined as the positive part of its argument, mathematically as depicted in equation (1).

$$F(x) = \max(0, x) \quad \text{----------------} \quad (1)$$

ReLU is a very popular activation function because of its property that it does not activate all the neurons at same time. We use SoftMax as activation function at the output layer since SoftMax layer provides the desired probability distribution for each classification.

Google's TensorFlow library helped us to implement these in easy way using *BasicLSTMCell* method. We trained the model with batch size of 1024 and for epoch of 25. As we are well aware that every machine learning algorithm needs to be optimized in order for it to learn all the optimal parameters such as weight and learning rate. ADAM optimizer is one such algorithm which can be used to update neural networks weights based on training data. We optimized our model while training the model using ADAM optimizer. Architecture of the proposed system is shown in Fig 2.

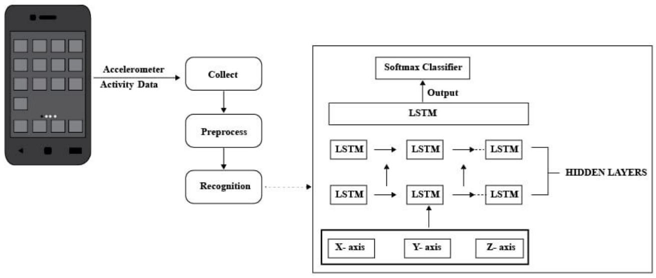

Figure 2. Architecture of the Proposed System

5. RESULTS

After training the model on training set, we test the model on test set which it hasn't seen while learning. We evaluated the model over training accuracy, training loss, test accuracy and test lost over numerous iterations. Fig 3 shows the values of these accuracies and loss over 25 epochs on which the model was trained.

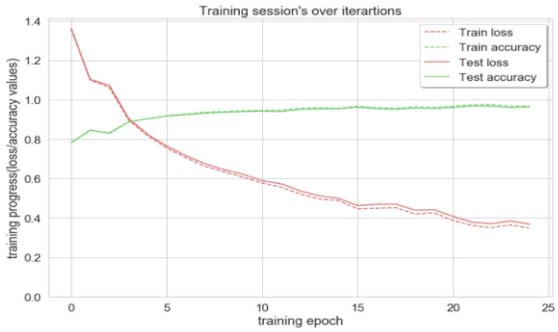

Figure 3. Progress over training session

Fig 4 shows the confusion matrix for classification. Confusion matrix is the best way to study and analyze the classification results. It shows what proportions of class being considered is well classified or misclassified. In our confusion matrix downstairs activity is most misclassified. Our proposed model performs comparatively better for other activities.

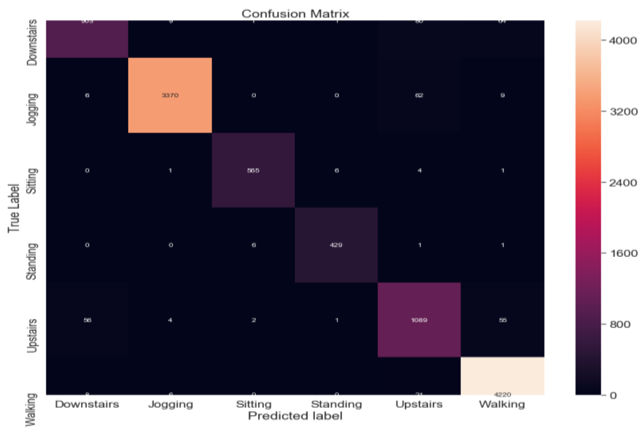

Figure 4. Confusion Matrix for the classification

6. CONCLUSION AND FUTURE WORK

Human Activity recognition has become important part of research among machine learning enthusiasts. Also, now people are moving from traditional machine learning algorithm to deep learning and neural networks which shows promises of computing and learning complex tasks well. We have built an LSTM model that can predict human activity from 200 time-step sequence with over 97% accuracy on the test set. The model was exported and used in an Android app. Our model performs only reasonably well when provided real-time data from smartphones sensors as compared to its performance on datasets while training and testing.

Although consequences of our experiments with model are satisfactory, there is still place for improvements. More parameter and hyper parameters tuning are required. For instance, due to unavailability of powerful system we restricted ourselves to only for 25 epochs. Increasing the epoch will make the system perform better than this. Experiment could be performed on other datasets which may build better models. We have restricted ourselves to usage of single sensor of smartphones i.e. accelerometer. It can be experimented to build the model with datasets containing data from multiple sensors like accelerometer and gyroscope together.

REFERENCES

1. L. Cheng, Y. Guan, Kecheng Zhu and Yiyang Li, "Recognition of human activities using machine learning methods with wearable sensors," *In IEEE Seventh Annual Computing and Communication Workshop and Conference (CCWC),* Las Vegas, NV, 2017, pp. 1-7, doi: 10.1109/CCWC.2017.7868369.
2. H. Gong, K. Xing and W. Du, "A user activity pattern mining system based on human activity recognition and location service," *In IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),* Honolulu, HI, 2018, pp. 1-2, doi: 10.1109/INFCOMW.2018.8406918
3. A. Jalal, M. A. K. Quaid and A. S. Hasan, "Wearable Sensor-Based Human Behavior Understanding and Recognition in Daily Life for Smart Environments," *In International Conference on Frontiers of Information Technology (FIT),* Islamabad, Pakistan, 2018, pp. 105-110, doi: 10.1109/FIT.2018.00026.K.
4. T. Zebin, P. J. Scully and K. B. Ozanyan, "Human activity recognition with inertial sensors using a deep learning approach," *In IEEE SENSORS, Orlando, FL,* 2016, pp. 1-3, doi: 10.1109/ICSENS.2016.7808590Qing Wang, Wenfang Shi, Kinetics of thermal decomposition of epoxy resin containing flame retardant component, *Polymer Degradation and Stability,* 91(8), 1747-1754 (2006).
5. Qingzhong Liu, Zhaoxian Zhou, Sarbagya Ratna Shakya, Prathyusha Uduthalapally, Mengyu Qiao, and Andrew H. Sung, "Smartphone Sensor-Based Activity Recognition by Using Machine Learning and Deep Learning Algorithms," *In International Journal of Machine Learning and Computing* vol. 8, no. 2, pp. 121-126, 2018
6. Jennifer R. Kwapisz, Gary M. Weiss and Samuel A. Moore, “Activity Recognition using Cell Phone Accelerometers”, *In Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data (at KDD-10),* Washington DC.