A Machine Learning based Activity Recognition for Ambient Assisted Living

Chandini U
M.Tech: dept. Computer Science and Engineering
Thejus Engineering College
chandiniunnikrishnan@gmail.com

Abstract— Ambient assisted living (AAL) technology is of considerable interest in supporting the independence and quality of life of older adults. As such, it is a core focus of the emerging field of gerontechnology, which considers how technological innovation can aid health and well-being in older age. Human activity recognition plays a vital role in AAL. Successful identification of human activity is crucial for any assistive care services for elderly people living alone in a home. In this paper, a method for activity recognition is proposed which recognizes or classifies activities based on sensor data. The method uses most trending algorithm in deep learning domain, i.e. LSTM to build the model .The proposed method is evaluated using a well known activity sensor dataset.

Keywords- *AAL, Deep learning networks, LSTM, Activity recognition*

*****

I. INTRODUCTION

Ageing is a natural process, which presents a unique challenge for all sections of the society. Population ageing is a global issue, which has been recognized to have implications on the health care and social welfare systems. It has been estimated that the number of people aged over 60 and over will increase to 1.2 billion in 2015 and subsequently to two billion in 2050.[2] Further, by the year 2025, almost 75% of this elderly population will be living in developing nations, which already have an overburdened health-care delivery system.[2]

An ageing population tends to have a higher prevalence of chronic diseases, physical disabilities, mental illnesses and other co-morbidities. As a result, it is necessary to find solutions to improve the living condition and develop more robust, usable, safe but low cost health- care systems to reduce the burden to society. Assisted living is a great intermediate step for seniors who need more help than the family can provide at home, but who don't need the round-the-clock medical care of a nursing facility. Ambient Assisted Living (AAL) can be defined as concepts, products and services which combine new technologies and the social environment in order to improve quality of life.

Human activity recognition is an important aspect of ambient assisted living. Traditional methods used for activity recognition was heavily dependent on the feature extraction and signal processing. Activity recognition has been addressed using methods such as decision trees, Bayesian methods (Naıve Bayes and Bayesian Networks), k-Nearest Neighbors, SVMs, Fuzzy logic, Regression models, Hidden Markov Models and classifier ensembles. The sensor data generally serve as input parameters to activity recognition algorithms. However, these data are time series so classifying it requires a sample window. To achieve this goal, classical approaches have to first generate time and frequency domain statistics for each training window. So, major task lies in data preprocessing and feature extraction. Moreover, the accuracy of common algorithms tends to drop when activities are performed by people not included in the training phase. Recently, there has been growing interest in deep learning. By, taking advantage of existing improvements in artificial intelligence most of the above mentioned problems can be solved. The proposed method uses LSTM RNN for human activity recognition.

This paper is organized as follows: Section II describes the concept of LSTM and architecture. Section II describes the experimental set up which includes the dataset description, model building. Section III contains the result of the experiment with evaluation metrics and at the end in Section IV the conclusions of this research is presented

II. LSTM

An LSTM, or Long Short-Term Memory Network, is a kind of recurrent neural network. It is actually an extremely powerful algorithm that can classify, cluster and make predictions about data, particularly time series and text. Recurrent nets are a type of artificial neural network designed to recognize patterns in sequences of data, such as text, genomes etc. Recurrent networks possess a certain type of memory, and memory is also part of the human condition. Like most neural networks, recurrent networks are old. The vanishing gradient problem emerged as a major obstacle to recurrent net performance. LSTM are designed to alleviate the long-term dependency problem. LSTM remembers information for long period of time. LSTM have a chain like

ISSN: 2454-4248
323 – 326

structure with four layers of neural networks interacting in a special way.

The key to LSTM is the cell state. The LSTM have the ability to add or remove information to the cell state, carefully regulated by structures called gates [12]. Gates are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a point wise multiplication operation. A cell is composed of four main elements, an input gate, a neuron with a self recurrent connection, a forget gate and an output gate. The input gate layer controls the extent to which a new value flows into the cell, the forget gate controls the extent to which a value remains in the cell and the output gate controls the extent to which the value in the cell is used to compute the output activation of the LSTM unit[11]. The connections associated with these gates are recurrent. The weight of these connections is learned during training.

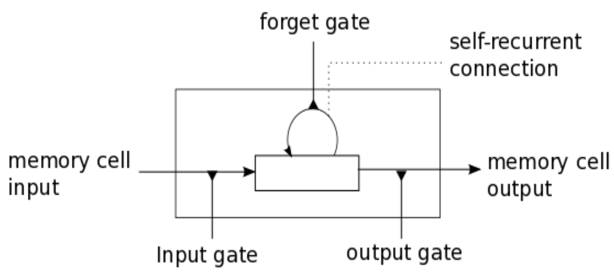


Fig.1.LSTM memory cell illustration

The following equations depict updates to a memory cell at every time step t.

$W_i, W_j, W_c, W_e, U_i, U_f, U_o, U_c, V_o$ are weight matrices

$$i^{(t)} = \sigma(W^{(i)}x^{(t)} + U^{(i)}h^{(t-1)}) \quad (1)$$

$$f^{(t)} = \sigma(W^{(f)}x^{(t)} + U^{(f)}h^{(t-1)}) \quad (2)$$

$$o^{(t)} = \sigma(W^{(o)}x^{(t)} + U^{(o)}h^{(t-1)}) \quad (3)$$

$$\tilde{c}^{(t)} = tanh(W^{(c)}x^{(t)} + U^{(c)}h^{(t-1)}) \quad (4)$$

$$c^{(t)} = f^{(t)} \circ \tilde{c}^{(t-1)} + i^{(t)} \circ \tilde{c}^{t} \quad (5)$$

$$h^{(t)} = o^{(t)} \circ \tanh(\mathrm{c}^{(\mathrm{t})}) \quad (6)$$

Equation 1 depicts a sigmoid layer (input gate layer). The input word $x^{(t)}$ and past hidden state $h^{(t-1)}$ is used to determine whether or not the input is worth preserving and thus is used to gate new memory function. Equation 2 depicts the forget gate layer to determine whether or not the past memory cell is useful for the computation of the current memory cell. Equation 3 depicts the output gate layer which makes assessment regarding what part of final memory need to be exposed or hidden. Equation 4 and 5 represents new memory and final memory cell respectively. Equation 6 represents the present state information.

III. Experimental Set Up

A. *Dataset Description*

The data for the research is collected from accelerometer sensors. Every modern Smartphone has a tri-axial accelerometer that measures acceleration in all three spatial dimensions. An LSTM Neural Network is trained for activity recognition from accelerometer data. Here, two datasets are used for research. Wireless Sensor Data Mining (WISDM) dataset and UCI HAR (Human Activity Recognition) dataset is used for the research work. The WISDM dataset contains 1,098,207 rows and 6 columns. There are no missing values. There are 6 activities in the dataset: Walking, Jogging, Upstairs, Downstairs, Sitting, Standing. The HAR dataset also includes 6 activities such as walking, walking_upstairs, walking_downstairs, sitting , standing and laying. The model is trained and tested using both the datasets.

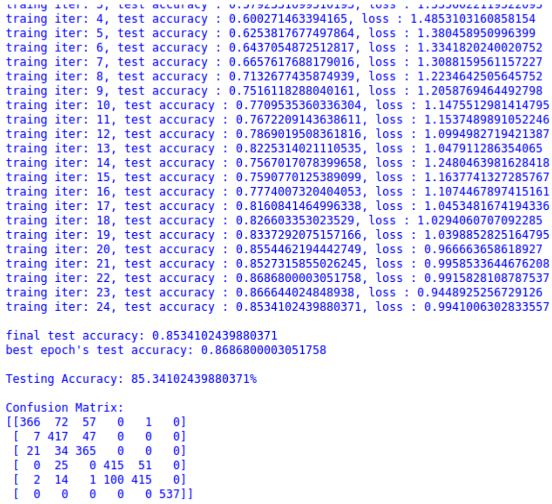
B. Building the model

The major aim of the project is to build an activity recognition model from the smart-phone's sensor data. For this purpose, a deep network is built by using two fully connected RNN with 2 LSTM layers stacked on top of each other. Stacked multiple layers of LSTM are used so that various complex patterns could be recognized at each layer and this will solve the vanishing gradient problem. A single layer RNN cannot capture all the structure of sequences, hence 2 RNN are used. The whole model is built using the tensor flow framework. A loss function is also calculated to provide a stable solution.L2-norm loss function which is also known as Least Squared Error is used to minimize the difference between the target value and the estimated value by taking the sum of absolute differences between them.

The model is trained using tensor flow API with appropriate epochs and batch sizes to avoid over fitting. The data is split in such a way that 80% of data is used for training and 20% is used for testing.

C. Results

The model has been trained for 25 epochs in case of HAR dataset and achieved an overall accuracy of 85% with loss of over 0.9%.It is likely to achieve more than 95% accuracy in training if the number of epochs is increased by 50 epochs. The WISDM dataset was trained for 50 epochs and achieved 97% accuracy overall. The figure below shows the training session's progress of both the datasets.



```
traing iter: 4, test accuracy : 0.600271463394165, loss : 1.4853103160858154
traing iter: 5, test accuracy : 0.6253817677497864, loss : 1.380458950996399
traing iter: 6, test accuracy : 0.6437054872512817, loss : 1.3341820240020752
traing iter: 7, test accuracy : 0.6657617688179016, loss : 1.3088159561157227
traing iter: 8, test accuracy : 0.7132677435874939, loss : 1.2234642505645752
traing iter: 9, test accuracy : 0.7516118288040161, loss : 1.2058769464492798
traing iter: 10, test accuracy : 0.7709535360336304, loss : 1.1475512981414795
traing iter: 11, test accuracy : 0.767220914363861l, loss : 1.1537489891052246
traing iter: 12, test accuracy : 0.7869019508361816, loss : 1.0994982719421387
traing iter: 13, test accuracy : 0.822531402111O535, loss : 1.047911286354065
traing iter: 14, test accuracy : 0.7567017078399658, loss : 1.2480463981628418
traing iter: 15, test accuracy : 0.7590770125389099, loss : 1.1637741327285767
traing iter: 16, test accuracy : 0.7774007320404053, loss : 1.1074467897415161
traing iter: 17, test accuracy : 0.8160841464996338, loss : 1.0453481674194336
traing iter: 18, test accuracy : 0.826603353023529, loss : 1.0294060707092285
traing iter: 19, test accuracy : 0.8337292075157166, loss : 1.0398852825164795
traing iter: 20, test accuracy : 0.8554462194442749, loss : 0.966663658618927
traing iter: 21, test accuracy : 0.8527315855026245, loss : 0.9958533644676208
traing iter: 22, test accuracy : 0.8686800003051758, loss : 0.9915828108787537
traing iter: 23, test accuracy : 0.866644024848938, loss : 0.9448925256729126
traing iter: 24, test accuracy : 0.8534102439880371, loss : 0.9941006302833557

final test accuracy: 0.8534102439880371
best epoch's test accuracy: 0.8686800003051758

Testing Accuracy: 85.34102439880371%

Confusion Matrix:
[[366  72  57   0   1   0]
 [  7 417  47   0   0   0]
 [ 21  34 365   0   0   0]
 [  0  25   0 415  51   0]
 [  2  14   1 100 415   0]
 [  0   0   0   0   0 537]]
```

Fig.2. A sample of Training session using HAR dataset

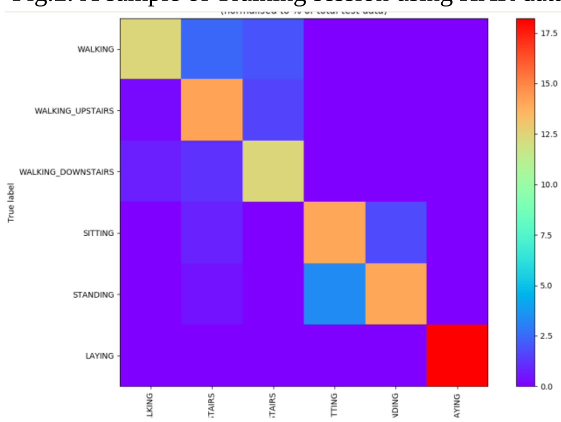


Fig.3. Confusion matrix for HAR dataset (25 epochs)

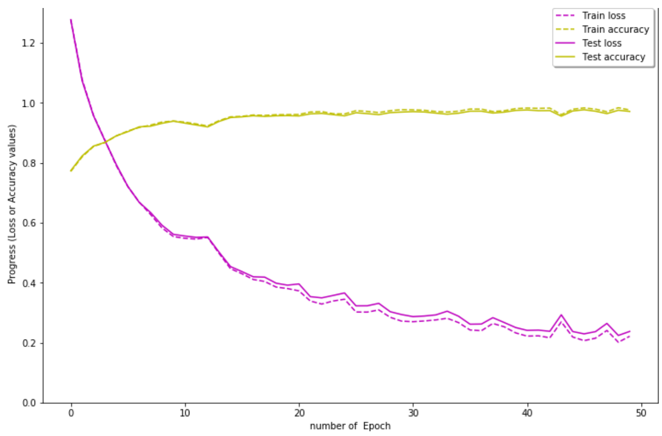


Fig.4. Training progress for WISDM dataset (50 epochs)

IV. CONCLUSION

An LSTM model that can recognize human activity is built using two datasets and it is observed that both of the datasets achieved over 90% accuracy using LSTM model. HAR dataset achieved an accuracy of over 85% for 25 epochs and WISDM dataset achieved an accuracy of 97% for 50 epochs. The dataset's accuracy variations at some moments during training depend upon the neural network's weight initialization and the number of epochs used in training. Compared to classical approaches, using a Recurrent Neural Networks (RNN) with Long Short-Term Memory cells (LSTMs) require no or almost no feature engineering. Data can be fed directly into the neural network which acts like a black box to model the problem correctly. Other research on the activity recognition domain mostly used a big amount of feature engineering, which is rather a signal processing approach combined with classical data science techniques. The approach here is rather very simple in terms of how much did the data was preprocessed. The only one limitation of the approach is the environmental setup. It requires high processing power such as GPU to train the model for more number of epochs in order to achieve a better accuracy. Even though the time required to train the model in CPU is high, it is able to produce an accuracy over 90%.

REFERENCES

[1] A.M. Khan, Y.-K. Lee, S.Y. Lee, and T.-S. Kim, "A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer", IEEE Trans. on Information Technology in Biomedicine, vol. 14, pp. 1166-1172, 2010.

[2] C.A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," Expert Syst.Appl. vol. 59, C, pp. 235-244, October 2016.

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

[4] J. C. T. Hsueh and Y. T. Wang, "Living Arrangement and the Wellbeing of the Elderly in Taiwan," International Journal of Welfare for the Aged, vol. 23, pp. 87–107, 2010.

[5] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers,"ACM SIGKDD Explorations Newsletter, vol. 12, no. 2, pp. 74–82, 2011.

[6] N. Y. Hammerla, S. Halloran, and T. Ploetz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables,"arXiv:1604.08880, 2016.

[7] O. D. Lara and M. A. Labrador, "A mobile platform for real-time human activity recognition," in Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '12), pp. 667–671,IEEE, 2012.

[8] V. R. Jakkula and D. J. Cook, "Detecting Anomalous Sensor Events in Smart Home Data for Enhancing the Living Experience," Proceedings of the 7th AAAI

Conference on Artificial Intelligence and Smarter Living, pp. 33-37, 2011.

[9] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," in Proc. of the 23rd ACM Int'l Conf. on Multimedia (MM '15). ACM, New York, NY, USA, pp. 1307-1310, 2015.

[10] World Health Organization. Towards policy for health and ageing.Availablefrom:http://www.who.int/ageing/publications/alc_fs_ageing_policy.pdf

[11] https://deeplearning4j.org/lstm

[12] http://colah.github.io/posts/