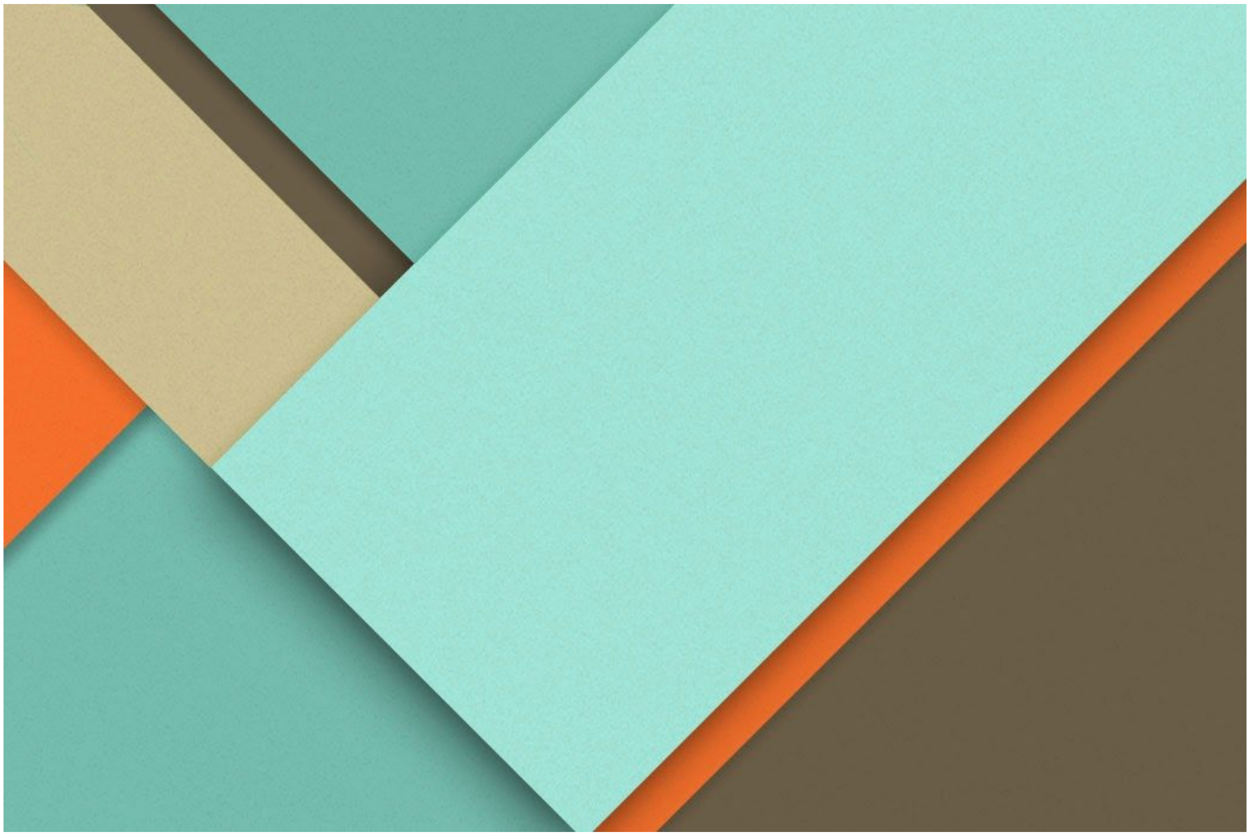


Human Activity Recognition

Udacity Machine Learning Engineer Nanodegree: Capstone Project Proposal

December 30, 2019

Md. Rayed Bin Wahed
Therap (BD) Ltd.
Banani, Dhaka 1213
Bangladesh

Domain Background

The problem of automatic identification of physical activities performed by human subjects is referred to as Human Activity Recognition (HAR) [1]. HAR is an attractive area of research due to its application in areas such as smart environments and healthcare [2], especially in the Intellectual and Developmental Disability (I/DD) space where I operate. If embedded inertial sensors such as accelerometers and gyroscopes found in smartphones and smartwatches can be used to predict activities of daily living (ADL), then patients with I/DD can be accurately monitored and a much better service provided. No such system exists for the industry yet. It is my goal to introduce such a mechanism.

Problem Statement

The goal is to use the processed/raw accelerometer and gyroscope sensor data to classify activities into one of the six basic ADL: three static postures (**standing, sitting, lying**) and three dynamic activities (**walking, walking downstairs and walking upstairs**). The data also includes postural transitions that occurs between the static postures. These are: **stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand**. Hence, given a stream of such data, the goal is to accurately classify the activities with a high degree of accuracy for real-life deployment.

Datasets and Input

The dataset is the Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set [3]. It was built from the recordings of 30 study participants within the age bracket of 19 to 48 years performing ADL while carrying a waist-mounted smartphone (Samsung Galaxy S II) with embedded inertial sensors. The dataset consists of 10299 samples with 561 attributes. The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 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. 70% of the volunteers were selected for generating the training set and 30% for the test set. For each record in the dataset the following is provided:

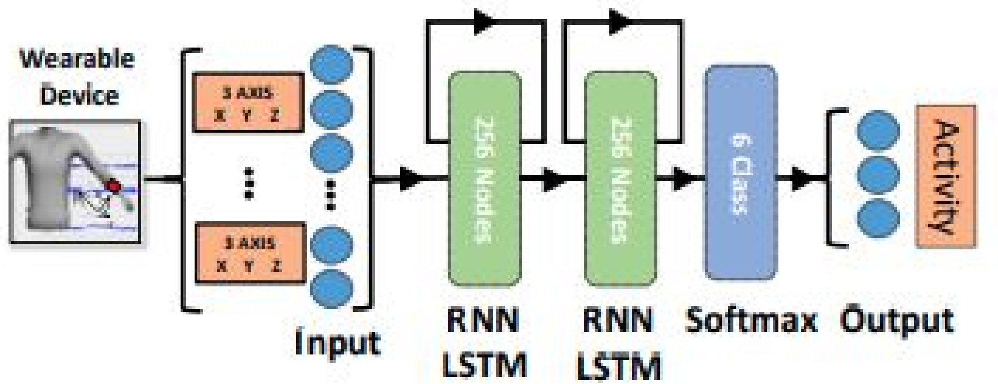
I. Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.
II. Triaxial Angular velocity from the gyroscope.

III. A 561-feature vector with time and frequency domain variables.
IV. Its activity label.
V. An identifier of the subject who carried out the experiment.

Solution Statement

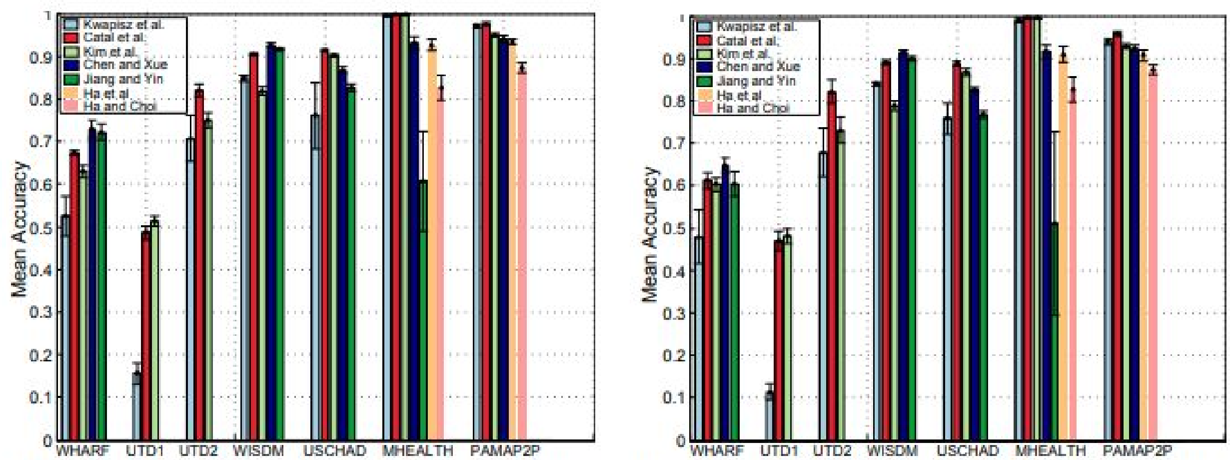
The goal is to develop a state-of-the-art HAR classifier. In the past, CNNs have been used since they have the advantage of being translationally invariant. However, instead of making a classification on a sample of data, I want to learn the underlying trend in these sensor readings. This is because my target audience (I/DD patients) are likely to be very different psychologically and in movement from a fully healthy human being. Hence, I plan on using a RNN since it usually performs better in capturing trends in time-series data.

Fig 1: A possible RNN model [4].



Benchmark Model

Fig 2: Mean Accuracy of different sampling techniques [2].



(a) Results using the semi-non-overlapping-window and cross validation (SNCV) combination.

(b) Results using the full-non-ovelapping-window and cross validation (FNCV) combination.

The diagram shows accuracies obtained on different datasets using different sampling and validation strategies. It is clear that at least 95% accuracy is a realistic target and that is what I am aiming for.

Evaluation Metrics

A simple accuracy measure is sufficient for this task.

Accuracy = (TP + TN) / (TP + FP + FN + TN)

Project Design

I. Load and analyze the training data for feature dependence.

II. Manually extract features or apply PCA. Retain components that preserve 80% of the data variance.

III. Divide the dataset into sequences of length128.

IV. Train a baseline LSTM model.

V. Tune hyperparameters and retrain .

VI. Test accuracy on the test data.

References

1. Antonio Bevilacqua, Kyle MacDonald, Aamina Rangarej, Venessa Widjaya, Brian Caulfield, Tahar Kechadi, “Human Activity Recognition with Convolutional Neural Networks”, arXiv:1906.01935v1 [cs], June 2019.
2. Artur Jordao, Antonio C. Nazare Jr., Jessica Sena, William Robson Schwartz, “Human Activity Recognition Based on Wearable Sensor Data: A Standardization of the State-of-the-Art”, arXiv:1806.05226v3 [cs], February 2019.
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. Patricio Rivera1 , Edwin Valarezo1, 2, Mun-Taek Choi3 , and Tae-Seong Kim1 1 Dept. of Biomedical Engineering, Kyung Hee University, Republic of Korea 2 Escuela Superior Politécnica del Litoral, ESPOL, Guayaquil, Ecuador 3 School of Mechanical Engineering, Sungkyunkwan University, Republic of Korea, “Recognition of Human Hand Activities Based on a Single Wrist IMU Using Recurrent Neural Networks”, International Journal of Pharma Medicine and Biological Sciences Vol. 6, No. 4, October 2017.