

Human Activity Recognition Report

With the increase of public interest in fitness during recent years there has been many rapid innovations in the field of fitness technology. A very popular item that is being used is the fitness watch. Allowing users to track their daily activity and how it impacts their health has caused this piece of technology to stand at the forefront of this field. The main feature that these fitness watches provide is the ability to determine what kind of physical activity a user performs and how much is performed. This feature stems from the watch's use of accelerometer and gyroscope sensors. The data we used for our project was captured in an experiment conducted using a group of 30 volunteers between the age group of 19-48 years old who had smartphones strapped to their waists. These volunteers performed six physical activities.

The dataset was converted in a CSV file and imported into a pandas data frame. The only preprocessing required was to combine the subjects and activity columns to their respective windowed features. Upon exploration, it was found that the dataset did not have any missing values. The data set was explored further to understand the various features and their effects on the activities. The training dataset has a total of 7352 observations or windows of data. It has a total of 561 time and frequency features where each observation corresponds to one of the 6 ambulatory class activities. These activities were walking, walking upstairs, walking downstairs, standing, sitting and laying. We wanted to create a deep learning model to accurately label the activity being done using the collected data from the experiment.

The data was collected using the accelerometer and gyroscope sensor embedded within the smartphones. The data collected was divided into measurements of the 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz and used a video recording to manually label the data. The sensor signals 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. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of features was obtained by calculating variables from the time and frequency domain. The obtained data was partitioned at random into two sets where 70% of the data was used for the training set and 30% was used for the testing set.

Recurrent Neural Networks usually have a short memory and are extended by LSTM units to extend the memory of the network. It enables the network to remember input over a longer period of time which makes it an essential unit in the layers of the recurrent neural network. It provides the capabilities to absorb more information from even longer sequences of data. This helps to boost the precision of the prediction by taking into account more data. As mentioned above the data was taken in 128 timestep intervals so we decided to use LSTM as our model choice for this sequence. After declaring our input layer and dropout layer which we set at

0.5 we declared our dense layer. In this dense layer we decided to use the sigmoid activation function. The reason we chose to use sigmoid instead of relu is because sigmoid allows us to predict the probability of our outputs and allows us to predict a single activity. This is beneficial because using the provided data there is a possibility that the model will get confused between multiple activities. For instance, different forms of walking could be determined from a data interval that may not be the correct determination. In our second model we decided to add LSTM layers with true return sequences and adjust the dropout layer in hopes of increasing the accuracy of the prediction. We kept our dense layer as is with the same activation function. Both of our models were compiled using the rmsprop optimizer and categorical cross entropy loss function.

Our model accuracy plots show that the test and train plotted fairly balanced and sat around 90%. Even with the extra LSTM layers and the change in the dropout layer, our accuracy stayed consistent. Next we created a confusion matrix to see quantitatively how accurate our model predicted the activities when compared to the true values. An interesting result we saw was that even though only about 90% of the predictions were accurate, the remaining 10% were falsely labeled for activities with similar acceleration and velocities. For instance false predictions in walking, walking upstairs and walking downstairs stayed within that group. Likewise a similar outcome was seen for laying, sitting and standing. Being able to better differentiate physical activities that have similar vector qualities is one opportunity we saw that we could use to improve our model.

Human activity recognition is used in many applications such as surveillance, anti-terrorists, and anti-crime securities as well as life logging and assistance. Environment-based sensors are used to detect the users' interaction with the environment. Interaction with objects that are also equipped with sensors.

Most of the existing studies in this field failed to efficiently describe human activities in a concise and informative way as they introduce limitations concerning computational issues. The gap of a complete representation of human activities and the corresponding data collection and annotation is still a challenging and unbridged problem. In particular, we may conclude that despite the tremendous increase of human understanding methods, many problems still remain open, including modeling of human poses, handling occlusions, and annotating data.

Using the captured data from an accelerometer and gyroscope we were able to design a deep learning model that is able to predict what sort of physical activity an individual is performing based on their 3-axial linear acceleration and 3-axial angular velocity. Using an LSTM model we were able to predict the action correctly with a 90% accuracy. There is opportunity to expand data collection into more complex movements, transitional movements and even anatomical movements. Measuring basic physical actions is merely the starting point for researching human interactions with the environment.