CS522 Assignment 3 Report

Team Members: Hsun Jui Chang, Jaachinma Okafor, Timo Salisch

November 11, 2022

Data Collection Process:

For each class, we used the Accelerometer and Gyroscope of a smartphone (the recorder) to record the motion signals. For android users, the recorder used is an application called AndroSensor. For iPhone users, the recorder used is SensorLogger.

Climbing Down:

An iphone was held in the recorder's hand during the recording of climbing down the stairs. The recorder starts the recording session at the highest step of the stairs at an apartment building and climbs down till reaching the lowest steps. Each session ends as soon as the recorder reaches the lowest step of the stairs.

Walking:

An android smartphone was used to record walking. We made 20 sessions of the recording.. The data was recorded by the same phone and the same person to maintain the consistency.

Standing Up:

An iphone was used to record 50 sessions of standing up movement. For each session, the recorder starts recording at the seating position, and stops the session after she is done with the standing up movement.

Rowing:

An android smartphone was used for a single 22-minute session of the exercise using a rowing machine. The phone was placed in the right pocket of the recorder. The record session starts right after the 1st row was done and stops after 22 minutes of rowing exercise.

Climbing Up:

An android smartphone was used for a single 22-minute session of the exercise on a climbing stair machine in a gym. The phone was placed in the right pocket of the recorder. The record session starts right after the recorder turns on the machine. was done and stops after 22 minutes of the climbing exercise.

NB: The collection of each class was done across multiple days to account for variations in the sampling instances.

Data Processing:

Due to the collection of data from different sources (phones), the data were not stored in the same format. Some data cleansing was required. The data recorded by the phones were stored as csv files in different formats and with different column names. After reading in the data using our pipeline, we had to rename some columns, drop some columns, and recalculate some columns to get all the class data in the same format to be able to feed our classification model.

Also, as part of the pre-processing, we dropped the first-5 seconds and last-5 seconds of instances recorded for more than 5 mins. This could not be done across all classes because some of the classes took shorter than 10 seconds to record. Removing the first and the last 5 seconds will leave little to no amount of data to work with. For instance, in the case of standing up, each recording lasted only 3 seconds because the act of standing up is short. For the case of climbing downstairs, it took around 7 seconds for a single recording, which is also shorter than 10 seconds. We implemented this across the other classes that have longer recording time because the data in these early and late stages tend to include noise from movement of the phones or non-activity.

Rationale for features:

When the different motion activities are recorded using an accelerometer and a gyroscope, the motions are shown to possess different features. It has been shown that a computer model can be trained to distinguish differrent activities using these features.

Due to some of the activities being recorded over a long time, we had to apply windowing. The window size was chosen according to the size of the data collected. For walking, a window size of 2 minutes was chosen. For rowing and climbing up, we used window size of 1 minute because they have less data than walking. For the classes that had less than a minute for each recording. No windowing was applied to them. After testing with different overlap sizes, we chose to have 80% as the overlap size for the best performance. The results from different overlaps will be discussed in the next section.

After that, we buffered the incoming time-domain signal from the sensors and computed the Fast Fourier Transform (FFT) with FFT size = 1024. In applying FFT, we experienced issues occurring due to the different sample rates of the data recorded. Due to the recording of the data being from different devices, they were not recorded at the same frequency, and this presented an issue in applying FFT to the data. Our first approach was to convert all the data to the same frequency, but this caused some of the classes to not have enough data for building a machine model. To solve this problem, for the classes with nperseg greater than the FFT size, we used nperseg = FFT size and for those that had length of the data points less than the FFT, we used nperseg equal to the length of data points. From the FFT, we created bins of 45 sizes each for the time and the frequency bins. This bin size was chosen after trying out a combination of sizes. The result of the different bin sizes is discussed in the next session. Figure 1 shows how these bins vary across the different classes. Binned spectrograms were the only features we applied for all the data points (acceleration x, acceleration y, acceleration z, gyroscope x, gyroscope y and gyroscope z) as it was enough to capture the similarities between each class and the differences between different classes.

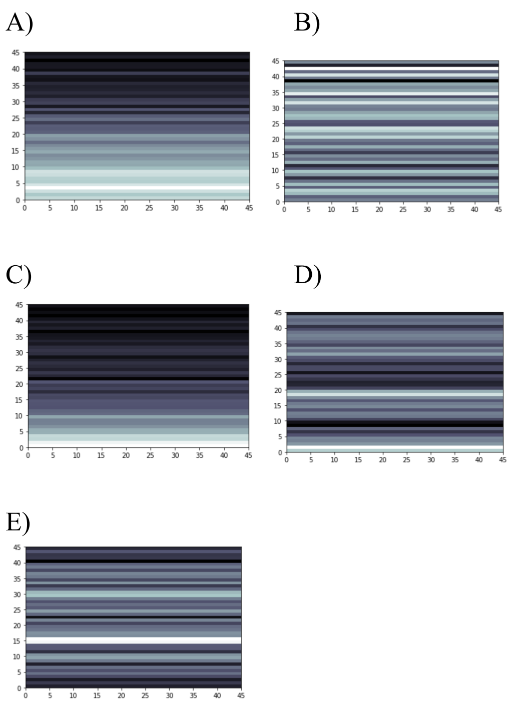


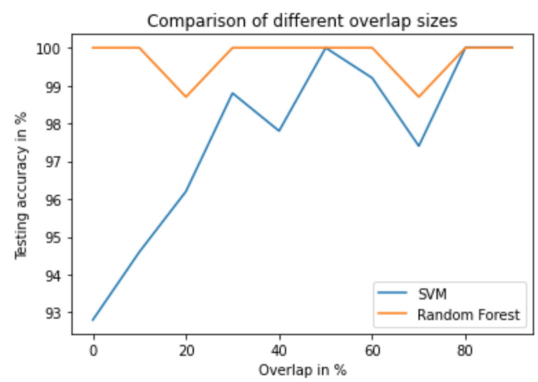
Figure 1: Bins of the acceleration x of different classes A) Climbing Down B) Walking C) Standing Up D) Rowing E) Climbing Up

Results:

We developed two models: an SVM and a Random Forest model. For the SVM model, we had to normalize our features to improve accuracy but that is not required for the random forest. Also, we had to use the linear kernel for the SVM as it has the highest accuracy among all the kernel types, which signifies the features have a linear relationship with the classes.

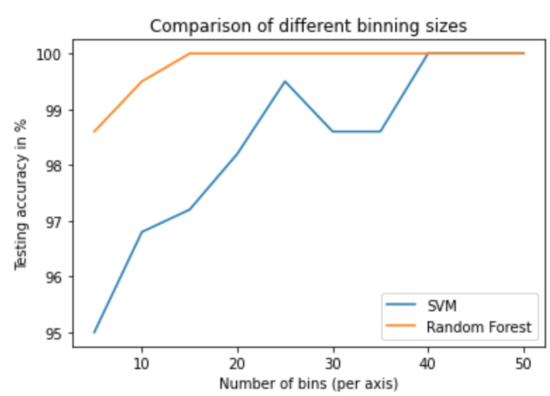
We also compared overlapping sizes for the windows approach and realized they produce different accuracy levels between the SVM and the random forest (Figure 2). The highest accuracy is shown to occur at around 80% of overlapping size for both classification models.

Figure 2: Comparison of Different Window Sizes



In the feature engineering, we tried different sizes of bins. The result of the different sizes of the bins is shown in Figure 3

Figure 3: Comparison of Different Bin Sizes



From Figure3, we can conclude the bin size of 40-50 produces the best accuracy for the SVM and random forest model. Therefore, we chose a bin size of 45 for our model.

We did not do under-sampling or over-sampling for our data even though the classes had varying sizes of data as we have very high accuracy without using them. The confusion metric and testing accuracy scores in Figure 4 shows that there is no bias of classification of the labels due to the different number of samples of the classes. Therefore, we chose to not include under-sampling or over-sampling for the data.

Figure 4: Metric Evaluation

```
Average Cross Validation Score from Training:
0.9980392156862745


Confusion Matrix:
[[135   0   0   0   0]
 [  0  33   0   0   0]
 [  0   0  13   0   0]
 [  0   0   0  24   0]
 [  0   0   0   0  13]]


Test Statistics:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       135
           1       1.00      1.00      1.00        33
           2       1.00      1.00      1.00        13
           3       1.00      1.00      1.00        24
           4       1.00      1.00      1.00        13

    accuracy                           1.00       218
   macro avg       1.00      1.00      1.00       218
weighted avg       1.00      1.00      1.00       218


Testing Accuracy: 1.0
```

In conclusion, we chose the window approach with a binned spectrogram. We developed models with accuracy of 100% with linear SVM and Random Forest Model.