A Deep Learning Multi-Class Time Series Classification Approach for Activity Recognition from a Single Chest-Mounted Sensor

Syed Abdul Hadi
George Mason University
Fairfax, Virginia
shadi4@gmu.edu

Vaishnavi Putcha
George Mason University
Fairfax, Virginia
vputcha@gmu.edu

Abstract:
Human Activity Recognition is defined as the problem of identifying the activity being performed by a human being through sensor data that has been acquired by sensors mounted onto a subject or data acquired through visual recording of a subject while performing a task. It is an active area of research as many different techniques are being tested that detect activity. With the recent surge of wearables, the need for a model that can accurately and efficiently recognize activity has grown. Previously work has been done on point classification from sensor data that has been collected at the frequency of 50Hz, and time series classification has been performed over data that has been collected at a frequency of 20Hz. State of the art point classification algorithms on 50Hz data give an accuracy of about 75%. While state of the art models built on 20Hz data give a higher accuracy of close to 90%. Moreover, data collected at 20Hz presents a 150% latency as models requires at least 4 seconds of data to accurately predict, while this problem can be solved using 50Hz data, currently there is no model that surpasses an accuracy of 75%. In this paper we describe the feasibility of a model that employs multi-class time series classification to predict activity from acceleration data collected from 3 chest mounted sensors at the frequency of 50Hz. Moreover, we will also be drawing a comparison with the broader feasibility of point classification algorithms for the purpose of activity recognition, and why these models do not scale well for practical applications.

Keywords --- time series, classification, CNN, LSTM, Hybrid, HAR, activity recognition.

I. Introduction

Activity recognition is the problem of predicting movement and detecting activity of a person from sensor data. These sensors are often accelerometers and gyroscopes, and they capture activity data when mounted as smartphones or wearables (smart watches). It has been an active area of research recently since wearables (particularly for fitness tracking) are gaining popularity. Activity recognition data is collected over time as a subject performs an activity and the sensor tracks movements over fixed intervals of time. The body-mounted sensors allow for ongoing monitoring of a variety of physiological data. This has intriguing implications in healthcare, such as the use of Ambient Intelligence (AmI) to track old people's everyday activity. We are showcasing a solution for mobile inertial sensors that recognizes human physical activity (AR). However, the problem of activity recognition is approached either as single point classification or as time series classification. The advantage in using time series classification is that it allows to model upon the temporal coherence information that is made available as an activity is recorded over a time interval rather than as a single point observation. Current state of the art time series classification algorithms perform well on data collected at 20Hz frequencies (0.05 seconds), however these models fail to perform well on sensor data collected at 50Hz frequencies (0.02 seconds) using a single chest mounted sensor. This presents latency challenges and also raises the question of why this approach cannot be scaled on this particular dataset. In this paper, we will be exploring a time series classification algorithm that beats the current best performing 50Hz AR model having an accuracy of 75%.

II. Motivation

For this project, we will be using a dataset that contains 50Hz accelerometer data from 15 subjects who perform a variety of tasks. Data is collected as acceleration in 3 dimensions (x, y and z axes). We will be expanding on the existing research on this dataset by applying deep learning multi class time series classification algorithms to identify the activity. Previously, the dataset has been used with single point ensemble classification algorithms. We will be exploring the advantage of taking into account the temporal information of this dataset. Existing single point classification algorithms will also be implemented to benchmark the results.

Previously the approach of time series classification has been successful on a more complex dataset (UCI HAR dataset) with accelerometer and gyroscope sensors, the dataset is more complex in terms of the sensor information (having 560 features on localization information of subjects). The goal of this project is to understand whether the added complexity of a multi-layer deep learning time series classification algorithm offers significant benefits in terms of the efficiency and the predictability when only accelerometer data is available and at the frequency of 50Hz.

III. LITERATURE SEARCH

A single point ensemble classification algorithm is applied to predict the activity being performed with accuracy of up to 94%. A second dataset is gathered using an Android-based commercial smartphone in a "wild" setting with obstacles, unfriendly surroundings, and challenging terrain. Results are extremely positive for both systems and datasets, with orders of magnitude lower verification error rates than cutting-edge technologies. Results for both systems and datasets are highly encouraging, with an order of magnitude less verification error rates compared to state-of-the-art technology. Two distinct scenarios using two various wearable devices are suggested as validation. In a free setting, a unique high-performance wearable system is created and used. Ten users' data are collected over the course of several days using this device, and five typical behaviors are recorded: walking, ascending stairs, standing, engaging with the surroundings, and working sitting. A second dataset is obtained from a commercial Android-based smartphone. In this last experiment, twenty participants freely complete seven runs in a "wild" environment that includes challenging terrain, hostile circumstances, congested areas, and impediments. Few drawbacks in this research paper are since there is more feature space that is regarded to be approved, a broad verification system that just considers the patterns of allowed users is not practical. The creation of user profiles that are automatically chosen is a practical solution to this issue. [1]

A single-point classification algorithm is applied to a more complex dataset (UCI HAR) collected using smartphone sensors. Features are, however, truncated for the model and this is shown to have a regularization effect as well, giving a recall of 89%. Activity-Based Computing aims to capture the state of the user and its environment by exploiting heterogeneous sensors. This has appealing use in healthcare applications, example- the exploitation of Ambient Intelligence in daily activity monitoring for elderly people. The author of this research describes a method for recognizing human physical activity (AR) utilizing inertial sensors from smartphones. We provide a cutting-edge, hardware-friendly solution to multiclass classification because these mobile phones have limited energy and processing capability. This approach uses fixed-point arithmetic to reduce processing costs while adapting the traditional Support Vector Machine (SVM). A comparison with the conventional SVM reveals a considerable reduction in processing costs while keeping comparable accuracy, which can help create more long-lasting AmI systems. In the study, authors presented a fresh approach to creating multiclass SVMs using integer parameters. Fixed-point computations are utilized in this option, which may be employed for AR since it uses less memory, CPU time, and power. Additionally, it offers accuracy levels on par with conventional methods like the MC-SVM, which use floating-point arithmetic. The experimental findings show that it is possible to replace the standard MC-SVM even with a decrease of bits equal to 6 for encoding the learnt MC-HF-SVM model parameter. This result has advantages for cellphones since it may help to free up system resources and cut down on energy usage. A publicly accessible AR dataset will be presented in further study to enable other researchers to test and contrast various learning techniques. [2]

Localization data (in terms of x, y and z coordinates) is used to detect the state of an individual while also making use of the temporal information from sensors. In order to maintain the freedom of elderly persons living alone at home, a multi-agent system is presented in this study. The system is made up of seven groups of agents that work together to provide a dependable, strong, and flexible monitoring by detecting the user in the environment, reconstructing the user's position and posture to create physical awareness of the user in the environment, responding to critical situations, requesting assistance in an emergency, and issuing warnings if unusual behavior is noticed. Several online demos have served as testing grounds for the system. The multi-agent system we've described in this study is made up of seven groups of intelligent agents: the sensor, refinement, reconstruction, interpretation, preventive, and cognitive groups. The agents inside groups are placed vertically, offering even more abstract situational awareness, and horizontally, contributing multiple interpretations of the circumstance. Each agent in a group has strengths and shortcomings of their own, but sophisticated combination and integration overcomes each deficiency and brings together many elements into a trustworthy interpretation. The outcomes of the fall detection experiment demonstrate that context-dependent reasoning may comprehend complicated events that acceleration-based systems would mistake. Additionally, the early findings on handicap detection are positive, suggesting a possibility for the early identification of a health issue that might result in dangerous situations. A presentation to EU reviewers was made online after hundreds of hours of testing, and it operated flawlessly in both the predefined scenarios and the ones that were created on the spot. However, the project's final year, when the system is implemented and tested in senior citizens' homes, the author shows how really applicable the system is. [3]

The A deep learning CNN-LSTM approach is used for activity recognition with the UCI HAR dataset, giving a 92% accuracy. Although somewhat comparable, this is an improvement from the conventional single-point classification algorithms. In order to increase the precision of human activity recognition, usually presented a CNN-LSTM method of activity recognition by utilizing the feature robustness the extraction of a CNN network while utilizing the work an LSTM model does for forecasting time series and classification. This CNN-LSTM model is geographically and temporally robust. when it was deeper in time and performed better contrasted to other deep learning methods that rely on raw data a signal's input data. We assessed this model in terms of predictability and Soft max loss on both internal and external (UCI HAR) dataset and an openly accessible (ISPL) dataset. in each. In several instances, especially on the ISPL, it outperformed the other models. Dataset with better precision (over 1%) than its closest competitor and less than 2% fewer Soft max losses. The time it took to run the various models and produce predictions was another parameter that was not assessed in this paper but was shown in the experiments. Compared to the other models, it took a lot longer using our suggested method. By testing it on different datasets, we intend to apply this model to more intricate tasks to address additional deep learning and HAR difficulties. We'll compare our strategy to the most recent findings for the UCI dataset and other publicly accessible datasets as well. [4]

IV. Proposed Approach

For the purpose of activity recognition using the accelerometer data of subjects, we will primarily be testing multi-class time series classification algorithms. We will be testing out convolutional neural networks (specifically 1D CNNs) and recurrent neural networks (specifically LSTMs). Additionally, a bi-directional LSTM will also be tested. The results of these models will be compared to single point classification approaches (KNN and ensemble random forests). For the purpose of this experiment, these fixed-point classifiers will also be created to validate on the results on the same data that is prepared for the time series classification models.

V. Preliminary Results

A. *Data and Pre-processing*

The data is collected from an accelerometer mounted on chest of a person. The uncalibrated accelerometer data are collected from 15 participants performing 7 different activities (Working at computer, standing up, walking, and going up\downstairs; standing, walking, going up\downstairs, walking, and talking with someone, talking while standing) at a sampling frequency of 52 Hz. The dataset is designed for use in studies on activity recognition. Null values are found within the target variable which are removed. Outliers are not handled because doing so may result in gaps within the sensor data which is essential and to be modeled upon.

The data can be used as is for single point classification techniques, however, it needs to be transformed in a 3-dimensional form with each observation being a time series array of x, y and z axis sensors data. The window of this series is to be determined as a trainable parameter. Additionally, these windows may also overlap, the length of this overlap is also a parameter which will be tested over various models.

B. *Exploratory Analysis*

Activities within the dataset appear to be imbalanced with more observations for four of the activities and the remaining three being underrepresented.

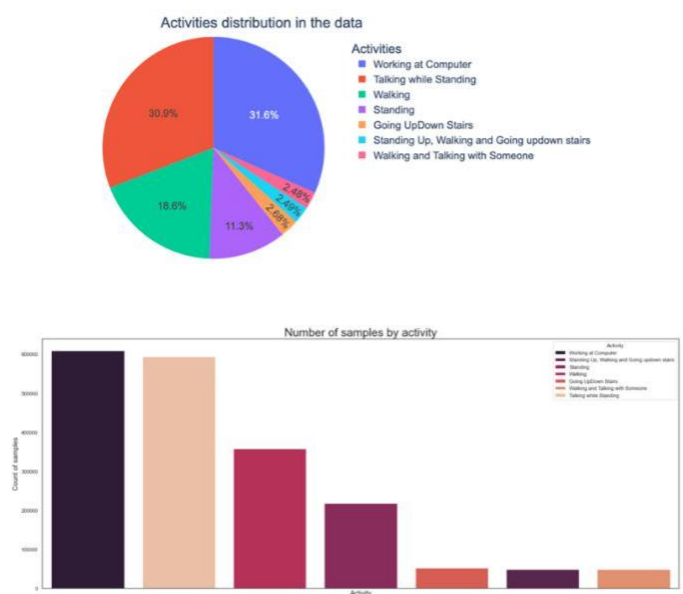


Fig 4. Samples count of each activity

From Fig4 we can see a bar graph with a significant class imbalance indicating the majority of the samples having class-label ‘working at computer’ and ‘talking while standing’. ‘walking and talking with someone' and ‘Standing up, walking and going up/down stairs’ and 'going up/down stairs' activities have the most minor representation in the dataset. From the above figure we can conclude saying that “walking at computer” has the highest recorded activity when compared to others.

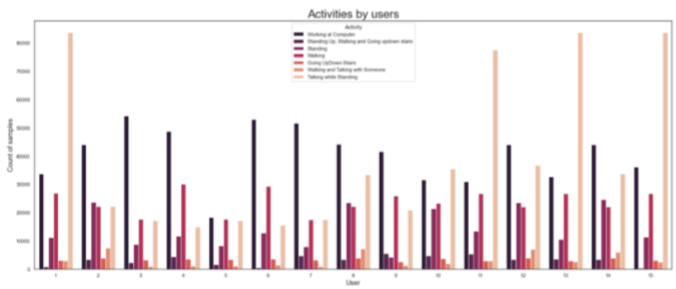
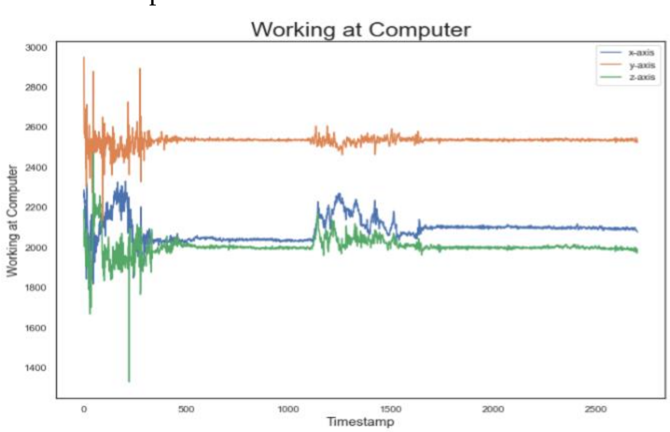
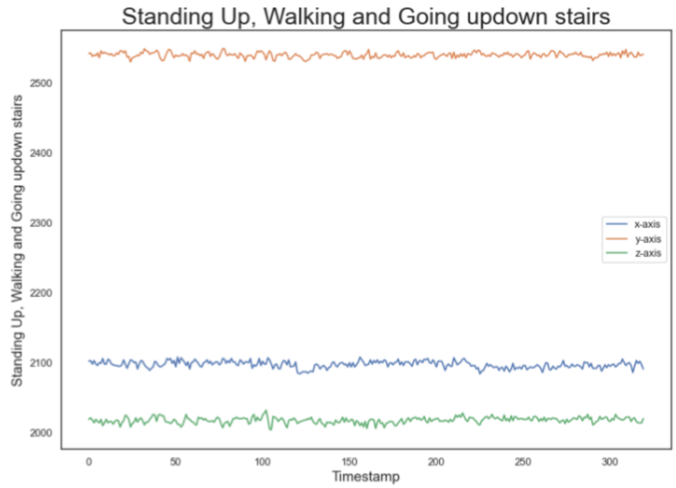


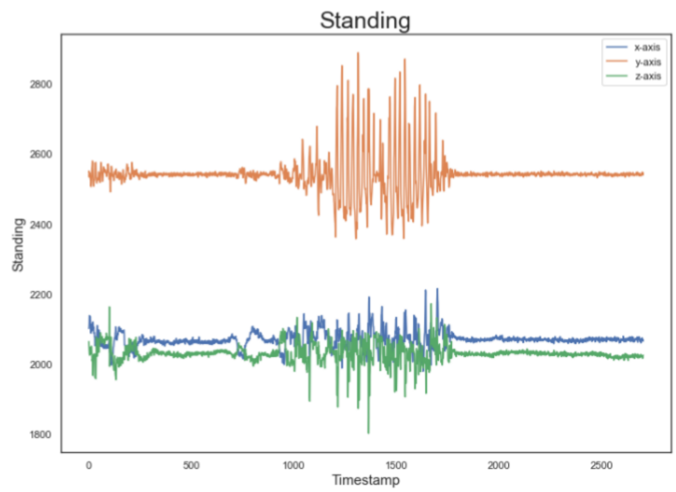
Fig 5. All user’s activity

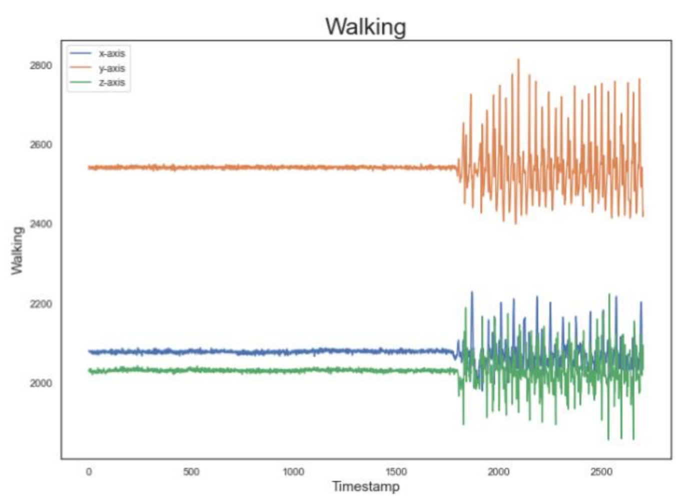
From above fig.5 It can be observed that not all the users are performing all the activities equally. The time for which they complete each activity is also disproportionate. Regardless, this won’t be affecting our analysis as we have a sufficiently large number of data samples and large combinations of data samples. So, we can conclude saying that all the users are alike so it has less effect on our data samples.

Observing the data over time shows that there is clear differentiation between the activities and from empirically observing these plots, it is evident that the pattern over time should enable us to obtain better results from predictive model.









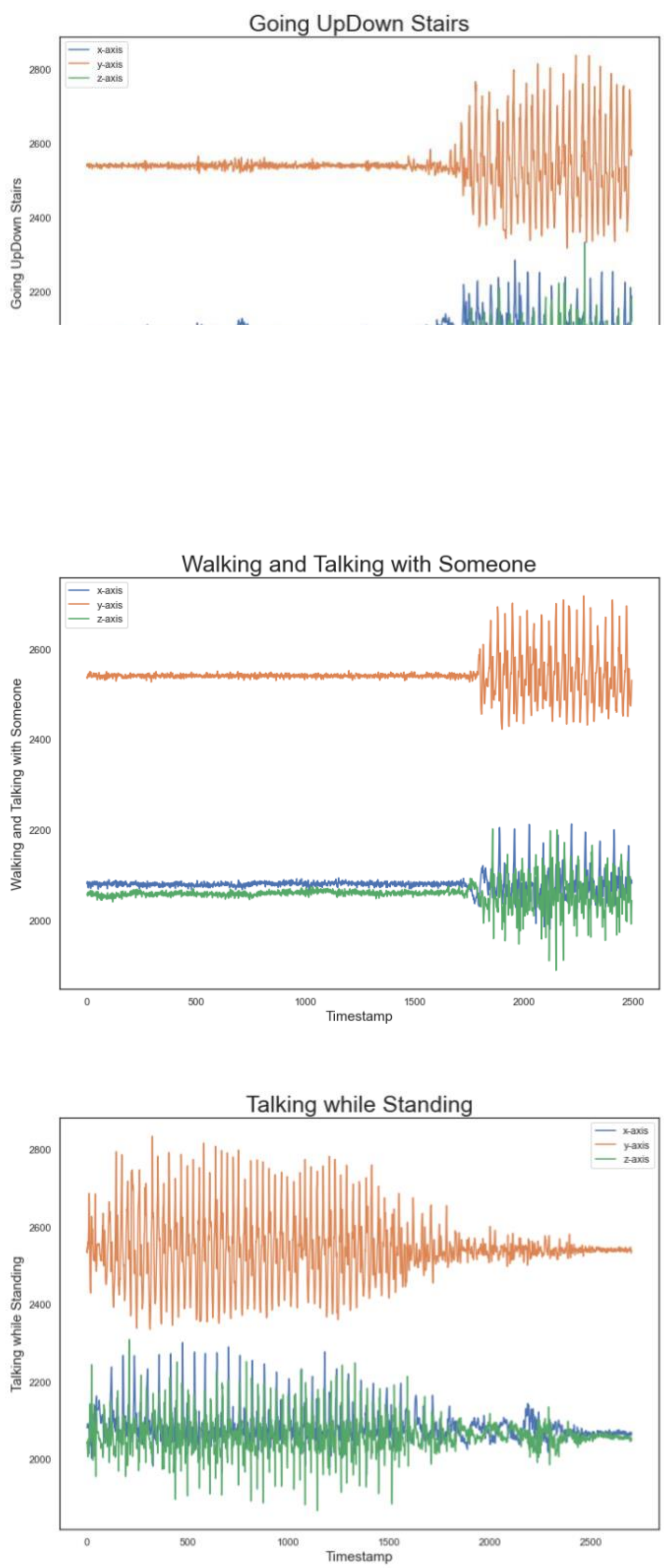


Fig 6. Each activity line plot with timestamp

For all the above visualizations in fig6, we have considered a subset of 2500 samples. This is equivalent to 0.02 secs of activity (as the frequency of data collection was 52Hz). As we can see, the signal moves relatively little during stationary actions like sitting and standing up, walking up and down stairs, and standing while chatting. In contrast, the signal behaves periodically throughout these stationary activities. Time-series data can be used to model these signals.

The accelerations in the x, y and z directions can also be visualized, being significantly different for activities that are moving intensive compared to activities that are not.

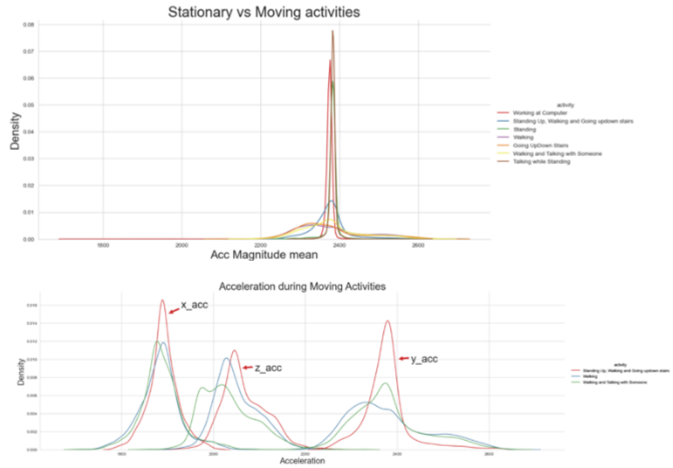
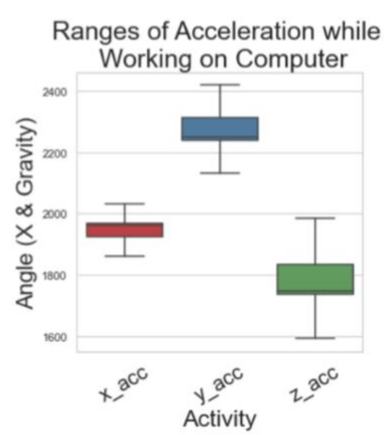


Fig 7. Histogram of Acceleration

VI. MODELING AND STRATEGY ITERATION

To set a baseline model, we implemented 2 different point classification models and one time series classification model that is known to be successful on 20Hz frequency data. A decision tree model and a KNN model were implemented as baseline for point classification. There is some evidence within the data that points towards the success of baseline models and why they do not entirely fail. There are significant empirical differences between observations during various (stationary and moving activities), and these differences are apparent in the ranges of values for the acceleration along each axis, as seen in fig8. [11]



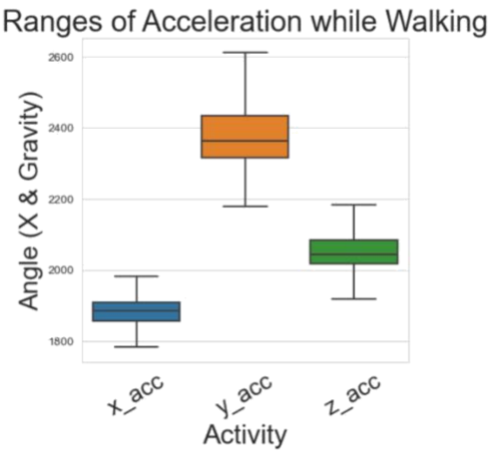


Fig 8. Acceleration of each axis while different activities

Decision Tree Model (Point Classification):

Splits were tested initially for variable tree depth and based on two criterion, Gini impurity and Entropy. The highest validation accuracy of 74.03% was obtained with Entropy based splitting with a max_depth of 15.

```
Accuracy with Splitting based on Gini Impurity:

GINI Impurity based Splitting with Max_depth = 2  is  43.92287773375347 %
GINI Impurity based Splitting with Max_depth = 3  is  48.81997524932664 %
GINI Impurity based Splitting with Max_depth = 4  is  52.29775684023337 %
GINI Impurity based Splitting with Max_depth = 5  is  59.20485861957799 %
GINI Impurity based Splitting with Max_depth = 6  is  62.076768685198466 %
GINI Impurity based Splitting with Max_depth = 7  is  64.75587308520264 %
GINI Impurity based Splitting with Max_depth = 8  is  66.75131812934826 %
GINI Impurity based Splitting with Max_depth = 9  is  68.94872034858932 %
GINI Impurity based Splitting with Max_depth = 10  is  70.81937208165643 %
GINI Impurity based Splitting with Max_depth = 11  is  71.98639752909244 %
GINI Impurity based Splitting with Max_depth = 12  is  72.66423319710064 %
GINI Impurity based Splitting with Max_depth = 13  is  73.4889089944779 %
GINI Impurity based Splitting with Max_depth = 14  is  73.91341424099667 %
GINI Impurity based Splitting with Max_depth = 15  is  74.04153537370397 %
GINI Impurity based Splitting with Max_depth = 16  is  73.97373100801798 %
GINI Impurity based Splitting with Max_depth = 17  is  73.69689784627545 %
GINI Impurity based Splitting with Max_depth = 18  is  73.32355785729885 %
GINI Impurity based Splitting with Max_depth = 19  is  72.73182957393483 %
```

```
Accuracy with Splitting based on Gini Impurity:

ENTROPY based splitting with Max_depth = 2  is  44.01751266132135 %
ENTROPY based splitting with Max_depth = 3  is  48.816439438846075 %
ENTROPY based splitting with Max_depth = 4  is  52.28361359831113 %
ENTROPY based splitting with Max_depth = 5  is  59.09628843893967 %
ENTROPY based splitting with Max_depth = 6  is  61.734211046287925 %
ENTROPY based splitting with Max_depth = 7  is  65.2415270541499 %
ENTROPY based splitting with Max_depth = 8  is  66.756725839495 %
ENTROPY based splitting with Max_depth = 9  is  68.98823823043085 %
ENTROPY based splitting with Max_depth = 10  is  70.20122921411412 %
ENTROPY based splitting with Max_depth = 11  is  72.01135619130815 %
ENTROPY based splitting with Max_depth = 12  is  72.90362836551961 %
ENTROPY based splitting with Max_depth = 13  is  73.6998096902006 %
ENTROPY based splitting with Max_depth = 14  is  73.97477095227696 %
ENTROPY based splitting with Max_depth = 15  is  74.0053453134912 %
ENTROPY based splitting with Max_depth = 16  is  73.84269803138552 %
ENTROPY based splitting with Max_depth = 17  is  73.54257011824166 %
ENTROPY based splitting with Max_depth = 18  is  72.99951122619828 %
ENTROPY based splitting with Max_depth = 19  is  72.28444555371833 %
ENTROPY based splitting with Max_depth = 20  is  71.49616780540563 %
```

```
Criterion selected as ENTROPY and max depth 15 will give us an accuracy score of  74.0053453134912
```

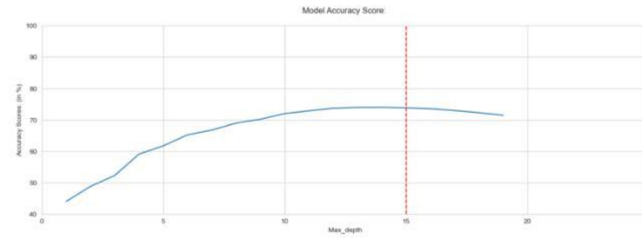


Fig 9. Decision Tree model

This model was then tested on training data, giving an accuracy of 74%. The model results, although promising, indicate the opportunity for improvement.

```
Accuracy Score for Normal Decision Tree Classifier:  74.0390395074824
Classification Report for Normal Decision Tree :
                                       precision    recall  f1-score   support

                  Working at computer      0.372     0.111     0.171     12829
Standing up, Walking and going updown      0.593     0.456     0.516     54282
                             Standing      0.458     0.149     0.225     12026
                              Walking      0.746     0.841     0.791    148534
                  Going upDown stairs      0.636     0.713     0.672     89288
                  Walking and talking      0.495     0.189     0.274     11771
               Talking While Standing      0.861     0.902     0.881    152065

                             accuracy                          0.740    480795
                            macro avg      0.595     0.480     0.504    480795
                         weighted avg      0.721     0.740     0.723    480795

Confusion Matrix:
[[  1421   2435     25   2363   5183    212   1190]
 [   887  24772    214  13349  11584    548   2928]
 [    48    598   1792   2304   2481     27   4776]
 [   215   5750    313 124955  10179   1138   5984]
 [   625   5157    422  12373  63642    265   6804]
 [   477   1311     42   4952   2236   2225    528]
 [   147   1730   1107   7124   4710     78 137169]]
```

Fig 10. Classification report of Decision Tree

KNN (Point Classification):

A KNN model was also tested for point classification. Validation results suggest an optimal k of 25 with an accuracy of 75.13% as shown in fig 12.

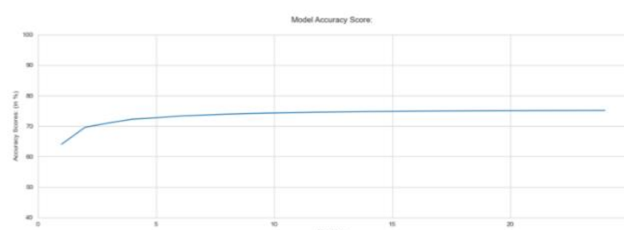


Fig 11. KNN Model accuracy graph

```
Accuracy:  64.01938456098753 % for  2 nearest neighbours
Accuracy:  69.60970891960191 % for  3 nearest neighbours
Accuracy:  71.0327686436007 % for  4 nearest neighbours
Accuracy:  72.25886292494722 % for  5 nearest neighbours
Accuracy:  72.70728688942273 % for  6 nearest neighbours
Accuracy:  73.30421489408168 % for  7 nearest neighbours
Accuracy:  73.61411828326003 % for  8 nearest neighbours
Accuracy:  73.91757401803262 % for  9 nearest neighbours
Accuracy:  74.13388242390208 % for  10 nearest neighbours
Accuracy:  74.30880104826382 % for  11 nearest neighbours
Accuracy:  74.42485882756684 % for  12 nearest neighbours
Accuracy:  74.56608325793738 % for  13 nearest neighbours
Accuracy:  74.66841377302177 % for  14 nearest neighbours
Accuracy:  74.76304870058965 % for  15 nearest neighbours
Accuracy:  74.80506244865275 % for  16 nearest neighbours
Accuracy:  74.89803346540626 % for  17 nearest neighbours
Accuracy:  74.93921525806216 % for  18 nearest neighbours
Accuracy:  74.98247693923605 % for  19 nearest neighbours
Accuracy:  75.02033091026321 % for  20 nearest neighbours
Accuracy:  75.03655404070342 % for  21 nearest neighbours
Accuracy:  75.06546449110327 % for  22 nearest neighbours
Accuracy:  75.08272756580247 % for  23 nearest neighbours
Accuracy:  75.11600578209008 % for  24 nearest neighbours
Accuracy:  75.13701265612163 % for  25 nearest neighbours

 The optimum k for KNN =  25
```

Fig 12. KNN Model

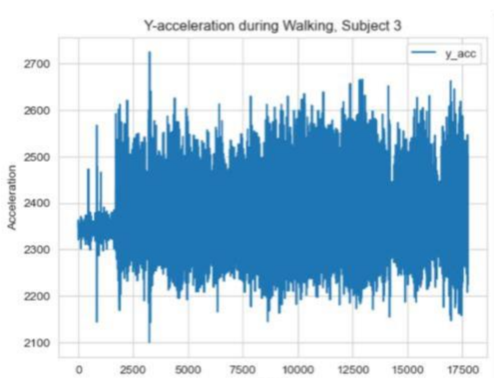
There is a slight improvement from the decision tree algorithm.

```
Accuracy for KNN, k = 25:  0.7513701265612163
Classification Report:
                                               precision    recall  f1-score   support

                          Going UpDown Stairs       0.41      0.13      0.20     12829
                                     Standing       0.61      0.47      0.53     54282
Standing Up, Walking and Going updown stairs       0.57      0.18      0.28     12026
                       Talking while Standing       0.76      0.84      0.80    148534
                                      Walking       0.64      0.74      0.69     89288
             Walking and Talking with Someone       0.53      0.20      0.29     11771
                          Working at Computer       0.87      0.91      0.89    152065

                                     accuracy                           0.75    480795
                                    macro avg       0.63      0.50      0.52    480795
                                 weighted avg       0.74      0.75      0.74    480795

Confusion Matrix:
 [[  1670   2296     13   2309   5370    168   1003]
 [  1052  25519    187  12463  11863    533   2665]
 [    42    603   2209   1934   2585     19   4634]
 [   351   5878    301 125412  10049   1057   5486]
 [   384   4659    232  11745  66046    264   5958]
 [   484   1320     36   4785   2303   2345    498]
 [   126   1583    898   6715   4654     35 138054]]
```

Fig 13. KNN Accuracy

<u>Time Series Classification with LSTM:</u>

A thorough analysis of visual difference suggests a strong correlation between the temporal patterns for the same activities across various subjects as shown in fig 14.



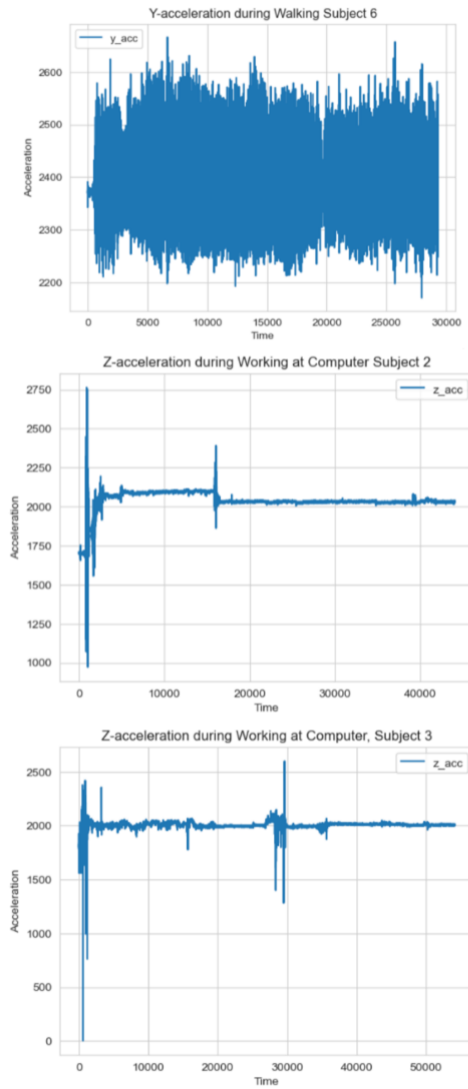


Fig 14. Time-series classification

We initially compared three time series classification techniques on the data after transforming it into a 3-dimensional shape. The length of the window for each observation is initially set to 128 with a step of 64 (giving a 50% overlap between consecutive observations).

Following are the results of three algorithms, 1D-CNN, LSTM and Bi-Directional LSTM.

Model	Window	Step	Performance
1D - CNN	128	64	28%
LSTM	128	64	40%
Bi Directional LSTM	128	64	44%

Fig 15. Performance of different models

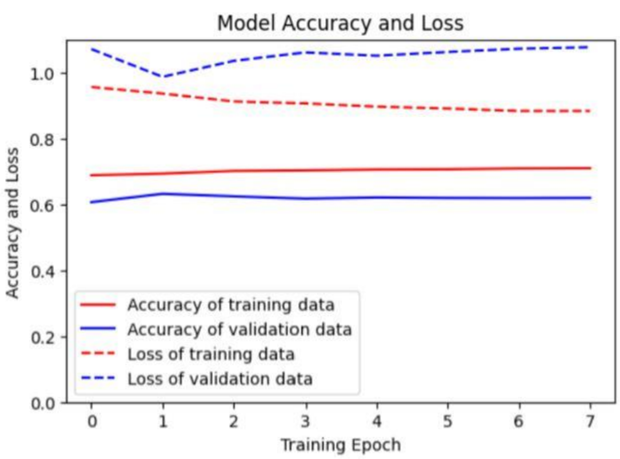
Although the bi-directional LSTM model performs better than the other two, the accuracy does not surpass that of single point classification.

We changed the parameters of the Bi-LSTM model to find the model with the best performance. It appears that the best performance is achieved when window is set to 12 with a step of 5, having an accuracy of 72%.

Model	Window	Step	Performance
Without Scaling	25	25	44%
With Scaling	25	25	60%
With Scaling	3	3	63%
With Scaling	8	5	70%
With Scaling	12	5	72%

Fig 16. Bi-directional LSTM

However, we still haven't been able to exceed the performance of single point classification. The reason for this is the ambiguity within the target class, as can be observed by the classification report for the model.




```
9117/9117 [==============================] - 88s 10ms/step
              precision    recall  f1-score   support

           0       0.69      0.03      0.06      8305
           1       0.56      0.04      0.07     34056
           2       0.67      0.00      0.00      8094
           3       0.60      0.87      0.71     78488
           4       0.73      0.89      0.80     56306
           5       0.00      0.00      0.00      7320
           6       0.84      0.90      0.87     99157

    accuracy                           0.72    291726
   macro avg       0.58      0.39      0.36    291726
weighted avg       0.69      0.72      0.65    291726
```

Fig 17. Bi-LSTM Classification Report

The recall for 3 of the most well differentiated classes comes out reasonably well. These classes being: Talking While Standing, Going Up Downstairs and Walking. It appears that the remaining activities are nested (multiple activities combined into one). The dataset has the following activities which we are modeling upon. It can be observed that some activities are not as well differentiated from other activities which can present a challenge for modeling.

1: 'Working at Computer',

2: 'Standing Up, Walking and Going updown stairs',

3: 'Standing',

4: 'Walking',

5: 'Going UpDown Stairs',

6: 'Walking and Talking with Someone',

7: 'Talking while Standing'

Activity 2 contained segments of activities 3, 4 and 5. Moreover, activity 3 and 7 are not high differentiated. Similarly, 4 and 6 are not differentiated either. These challenges are reflected within the model results but the model recall is a good metric for validating our hypothesis that time series classification enables higher predictability of activity from time series sensor data.

Finally, to validate our hypothesis that the challenge is a result of ambiguous target variable, we use our best model (Bi-LSTM) on raw human activity acceleration data which has been made available by the Wireless Sensor Data Mining Lab. This dataset is collected for 36 individuals during 6 activities at a frequency of 20Hz. The subjects' data was recorded during the following activities:

1. Walking
2. Jogging
3. Going Upstairs
4. Going Downstairs
5. Sitting
6. Standing

The activities are well differentiated and upon closer look, the temporal patterns over each activity are intuitively representative of the activity.

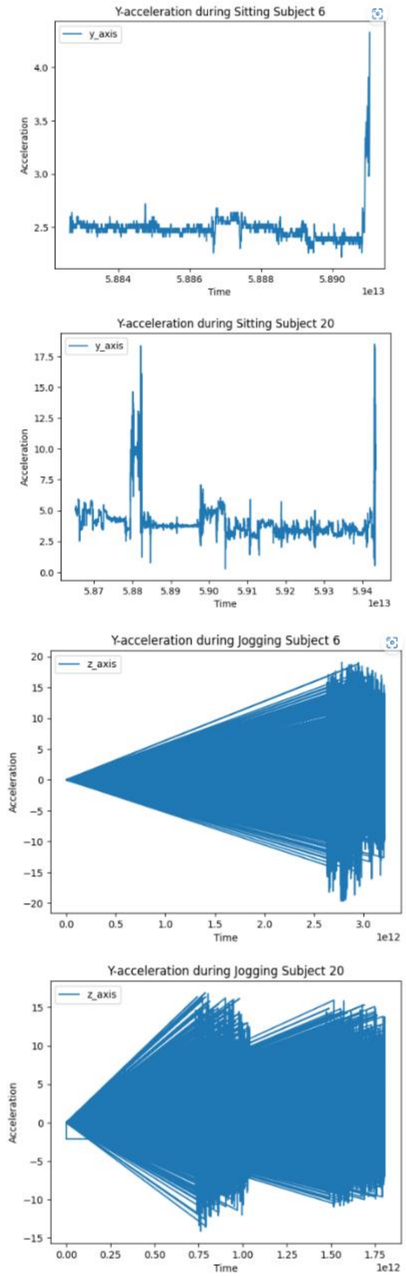
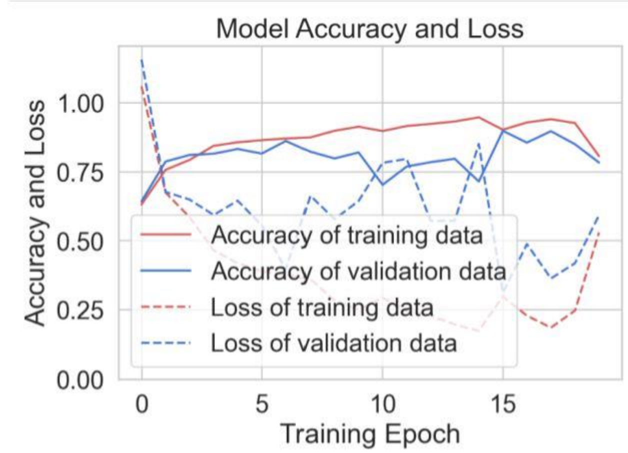


Fig 18. Patterns of different activities

VII. RESULTS

The Bi-Directional LSTM is finally trained on the WISDM data with varying window lengths and steps. The model gives an accuracy of 88% with a recall greater than 90% for 4 of the 6 activities.



	precision	recall	f1-score	support
0	0.65	0.48	0.55	2068
1	0.97	0.98	0.97	6743
2	0.95	0.96	0.96	1143
3	0.94	0.90	0.92	957
4	0.81	0.65	0.72	2566
5	0.86	0.96	0.91	8684
accuracy			0.88	22161
macro avg	0.86	0.82	0.84	22161
weighted avg	0.88	0.88	0.88	22161

Fig 19. Bi-directional LSTM model accuracy

VIII. CONCLUSION

When the results of the time series classification and single point classification are tested alongside each other, there are some empirical findings which are useful results of this projects. A single point classification algorithm like KNN appears to work almost reasonably but most of the relatively 'high' accuracy comes from overrepresented classes within the dataset. Additionally, the data is collected at 50Hz which resolves to a lag of 0.02 seconds between consecutive data points. An individual performing an activity naturally is unlikely to change their acceleration significantly in $2/100^{th}$ of a second which means a KNN algorithm is able to put an out-of-sample data point precisely back into the dataset where it was originally taken from, leading to the hypothesis that these models will not scale well when tested on data that is collected in an uncontrolled environment, and over different individuals that perform an activity uniquely. Time series classification algorithms on the other hand are able to dissect patterns within the activities and hence give the ability to scale these models over various individuals performing an activity in uncontrolled environments.

IX. FUTURE WORK

It would be a useful expansion to validate the empirical results of this paper on a dataset that contains acceleration data of subjects performing a variety of activities in an uncontrolled environment. One of the limitations of this paper is that both the datasets used, although contain raw acceleration data, have been created by recording subject activity in a lab-controlled environment.

X. References

[1] A. K. A. K. T. H. Katsunori Ohnishi, "Recognizing Activities of Daily Living With a Wrist-Mounted Camera," 2016. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/Ohnishi_Recognizing_Activities_of_CVPR_2016_paper.html.

[2] A. K. B. J. N. Lei Gao, "Sensor Positioning for Activity Recognition," 2015. [Online]. Available: https://www.researchgate.net/profile/Alan-Bourke/publication/290575901_Sensor_positioning_for_activity_recognition_using_multiple_accelerometer-based_sensors/links/5ede9266a6fdcc4768909224/Sensor-positioning-for-activity-recognition-using-multiple-acceler.

[3] A. B. J. N. Lei Gao, "A comparison of classifiers for activity recognition using multiple accelerometer-based sensors," 03 April 2014. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6782169.

[4] A. Godfrey, "Activity classification using a single chest mounted tri-axial accelerometer," 2011. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1350453311001111.

[5] Daily Activity Recognition based Principal Component Classification, https://www.researchgate.net/publication/235701636_Daily_Activity_Recognition_based_Principal_Component_Classification.

[6] R. Jothi, "Clustering Time-Series Data Generated by Smart Devices for Human Activity Recognition," 12 April 2019. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-16657-1_66.

[7] R. Sharma, "Human activity recognition in cyber-physical systems using optimized machine learning techniques," 17 08 2022. [Online]. Available: https://link.springer.com/article/10.1007/s10586-022-03662-8.