Human Activity Recognition with Smartphones

Diana Darmanin, Christopher Galea, Celine Suban
University of Malta

Abstract
This project aims to expand the work in Human Activity Recognition (HAR) using a smartphone's accelerometer and gyroscope data in order to be able to classify physical activities like walking and jogging. SVM, KNN and LSTM algorithms were implemented on a UCI HAR dataset and a dataset built by the researchers. All 3 algorithms performed with over 90% accuracy on the UCI HAR dataset. On the researchers' dataset LSTM performed with 95% accuracy, KNN with 89% accuracy and SVM with 73% accuracy.

1. Introduction

Activity Recognition (AR) is a general term describing the act of interpreting recorded sensor signals to deduce the activity that caused the signals. AR is essential in the development of many products in healthcare and safety, especially for those individuals requiring assisted living or those having motion-related disorders. Human Activity Recognition (HAR) is an ever-growing research field in which various researchers are coming up with methods to understand and categorise activities based on an individual's motion and physiological signals. Some approaches have adapted various types of sensors, such as a strap worn around the chest [1]. These approaches are often impractical for long periods of time due to the fact that they are uncomfortable for the wearer, as well as create complications when changing in and out of clothes. Smartphones are becoming increasingly ubiquitous in today's society and they come with various built-in sensors such as GPS, proximity sensors and acceleration sensors, thus providing an affordable and comfortable solution to the HAR problem. Figure 1 depicts the advantage of having real-time information transmitted to medical professionals. This is useful for individuals with mobility issues and the elderly, especially in the case of them having a stroke, heart attack, or other such serious issue requiring immediate medical attention. If the individual had a HAR system, such as the one described in this paper, it could sense that the individual had fallen and not gotten up, and hence, alert emergency services in order to get the individual the medical attention they require. Due to a rapidly ageing population in many countries around the world, such systems are sorely needed to help provide care and monitoring of elderly individuals without any privacy concerns, being too expensive, or being invasive or uncomfortable to wear. Our proposed system takes all these issues into consideration by using the individual's smartphone's accelerometer and gyroscope sensors for activity recognition.

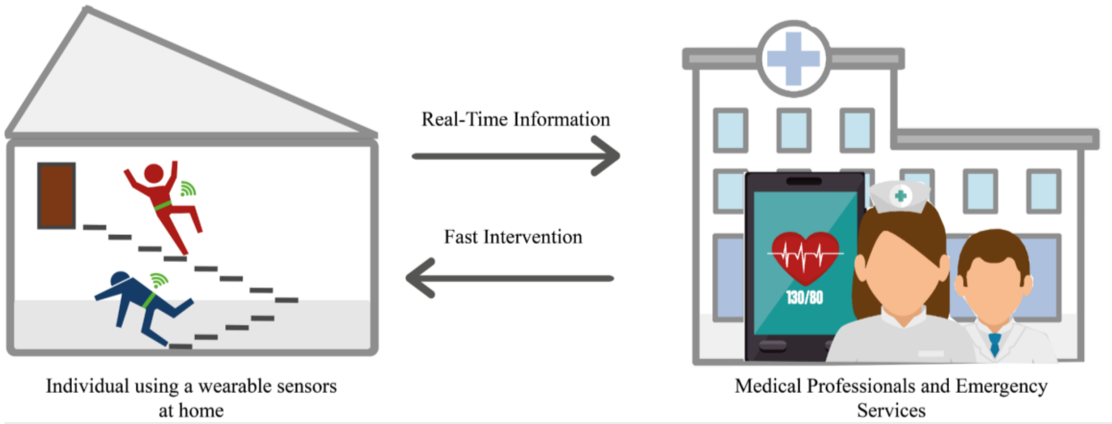

Figure 1: Advantage of having HAR system

This project was based on the work in [1], [2] and [3]. It aims to expand the work in Human Activity Recognition (HAR) using a smartphone's accelerometer and gyroscope data in order to be able to classify physical activities like walking and jogging, among others. In the first part of this project Support Vector Machine (SVM), k-Neighbours Classifier (KNN) and Long Short-Term Memory (LSTM) algorithms were implemented on the UCI HAR dataset [4]. These Machine Learning (ML) models were trained to categorise the actions of the wearer of the system based on the smartphone data.

In the second part of this project, the researchers built a dataset using data collected from smartphones while the subjects performed various physical activities. Due to the limitations posed by the on-going COVID-19 pandemic, data was collected from 3 individuals, the researchers themselves, who performed a series of physical activities in different environments, which simulated different conditions. This data was then pre-processed in the same way as in [1] and was fed into the algorithms that were implemented in the first part of the project. The results produced on the new dataset were then compared to those obtained on the UCI HAR dataset [4].

This project also had several objectives including, implementing the algorithms, building a dataset from smartphone data, pre-processing that data, and comparing the results from the newly built dataset and the UCI HAR dataset. These objectives were achieved and are discussed in more detail in following sections.

2. Literature review

2.1. Sensor Modality

Many approaches have been proposed in literature for HAR, where an important distinction made in the literature is between a video or sensor-based system. A video-based system would perform activity recognition using computer vision techniques applied on video data obtained from cameras placed in the environment. This poses privacy issues for the users of the system, as discussed in [5], [6] and [7]. It is for this reason that recent research has focused more heavily on investigating sensor-based systems, which is the basis of this project.

Sensor based systems can be classified according to the sensor modality used into four main areas, as shown in [8]:

> ***Wearable Sensor***. A wearable sensor should be small, non-intrusive and low-powered in order to be used for long periods of time if it is intended for Assisted Living. Many wearable sensors measure the physiological parameters of interest in an accurate and reliable manner. These parameters usually include body temperature and heart rate. Wearable sensors can be mounted on various body parts, such as, the waist, arm, ankle, and wrist in order to recognise human activities.
> *Accelerometers* can be used to measure the acceleration along a particular axis, thus allowing the researcher to draw conclusions on the individual's posture or in order to detect the difference between moving activities, such as running and walking. *Gyroscopes* can measure orientation and angular velocity. This device is usually integrated with an accelerometer, mounted on the same body part. Other wearable sensors include *magnetometers, electromyography sensors,* and *electrocardiography sensors.*
>
> ***Ambient Sensor***. These types of sensors are embedded in the environment in order to capture the interactions between the subject and the environment. These types of sensors include *Wi-Fi, Radio-Frequency Identification (RFID)* and *Radar*.
> ***Object Sensor.*** These types of sensors are useful for detecting composite activities such as eating, cooking, cleaning that an individual performs by interacting with objects. RFID can be used in this scenario by attaching RFID tags to the objects and an RFID Reader attached to the individual performing the tasks.
> ***Other Modalities*** include audio sensors and pressure sensors.

Wearable sensors can be placed on various parts of the body. These can be uncomfortable and intrusive, or small and unobtrusive, as is the case in this project and others using smartphones, such as, [1], [2], [3] and [9]. As discussed in [8], these sensors provide accurate information while often being cheaper than other sensor modalities, as only a few sensors would be needed on the subject in question. They are also more robust as they are unaffected by environment changes caused by the subject, allowing for continuous uninterrupted monitoring.
Ambient and Object sensors can sometimes be more accurate than wearable sensors at classifying certain types of activities. However, they can also be more expensive as they would need to be present throughout the environment the subject is in to be effective, and if the subject leaves this environment the sensors are unable to do their job.
Different sensor modalities could also be used in combination for a more robust HAR system, as was done in [10].

Given the above reasons, our project focused on a sensor-based system having wearable sensors, specifically a user's smartphone. Following from previous research in [8], [11] and [12], the sensors used in this project are an accelerometer and a gyroscope, both of which are found in today's smartphones.

2.2. Types of Human Activity

The term ***activity*** is a very broad description which can be used to describe various types of actions including *physical activities,* such as, walking, running, climbing stairs, standing, sitting and laying. These types of activities are mainly focused on the posture of an individual, as well as their movement. Another category of activities is *complex activities,* such as, Activities of Daily Living (ADL). ADL refers to an individual's daily self-care activities and can be further split into Basic ADLs (BADL), such as, feeding, dressing, and showering, and Instrumental ADLs (IADL), such as, transportation, shopping, house cleaning, communication, and meal preparation.

There is no universal definition regarding which set of activities HAR aims on detecting, but a majority of research in this area tends to focus on physical activities and BADLs, due to their relative simplicity. As surmised in [8] body worn sensors, specifically accelerometers and gyroscopes are commonly utilized when classifying physical activities and ADLs. Hence, another reason why these sensors, and this sensor modality was used in this project, which aims to classify physical activities.

2.3. Approaches used by Different Works

Various ML methods have been employed in past works for HAR including Gaussian Naive Bayes [13], Markov Chains [14], Decision Trees [15], KNNs [16], [17], SVMs [18], [19], CNNs [20] and RNNs [21]. The research done in [1], [2] and [3], which are the basis of the UCI HAR dataset used in this project, focused primarily on SVMs to perform HAR using smartphone's accelerometers, due to their success as a classical ML algorithm in HAR. However, more recently Deep Learning (DL) methods have been increasingly investigated for HAR as discussed in [8], [9], and [22].

[8] and [22] state DL methods make it possible to perform high level feature extraction automatically, while also using the original signal as input to the models, often resulting in improved performance. Hence, no pre-processing or hand-crafted features are required, saving considerable time. [8] and [22] also show that RNN, LSTM, and CNNs are well suited for HAR, where wearable sensors are commonly used with such DL methods to perform HAR, as shown in [23], [24], [25], [26] and [27].

[16] used a smartphone's accelerometer and concluded that comparable results on Random Forest, Decision Trees and KNN classifiers could be obtained using only raw accelerometer data, where KNN performed the best of all the models. The implications of this are that hand-crafted features, which are very time consuming to create, are not needed for these ML classifiers. [16] also found that RNN and CNN had a higher accuracy on a balanced dataset and when data was collected from multiple sensors located on different parts of the body. These DL methods were also found to perform better than traditional ML classifiers that did not use hand-crafted features.

3. Methodology

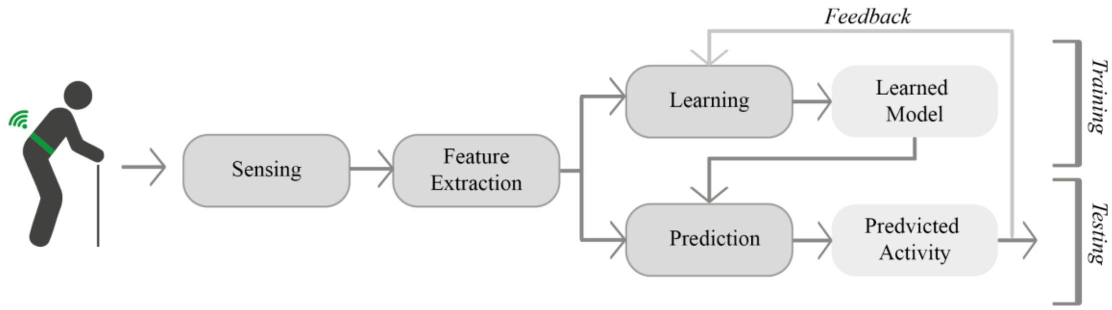

Figure 2: Human Activity Recognition Process Pipeline

3.1. Implementation of ML Models on UCI HAR Dataset

In order to fully understand the concept of our experiment, as well as envision the kind of data that needed to be collected, an ipynb file was created in order to carry out the intended experiment on the UCI HAR Dataset [4]. For the collection of this dataset, a group of 30 participants performed a total of six activities, recorded using the embedded accelerometer and gyroscope of a Samsung Galaxy SII mounted on the participants' waists. In this python file (See Appendix), the pre-processed training dataset and testing dataset are imported. This data is analysed for its characteristics. This was useful in order to gain a better understanding on what our dataset should look like. Then SVM, KNN, and LSTM models are fit to the training data and a prediction is made on the test data. This pipeline is depicted in Figure 2 aside from sensing and feature extraction, which were already carried out by the researchers who collected the data. The accuracy of each model is then evaluated and compared to the other models, in later sections.

3.2. Data Collection

Due to the limitations posed by the on-going COVID-19 pandemic, the researchers were unable to collect data from multiple individuals. This was due to the fact that a 4-person limit was imposed by the government during the period of collection. In order to compensate for this restriction, the data was collected from 3 individuals, the researchers, aged 19-20 years performing a series of physical activities in different environments, which simulated different conditions. Each subject mounted their smartphone on their left arm using a smartphone armband.

Each subject performed a total of two activities as follows:

Subject 1 Jogging and Falling
Subject 2 Walking and Sitting
Subject 3 Laying and Walking Upstairs/ Downstairs

These activities were recorded using the Android application 'Sensor Logger' [30] with the Accelerometer and Gyroscope sensors selected. Each activity in each environment was recorded 5 times for an average time period of *13 seconds*, with the exception of the falling activity which had an average time period of *4 seconds*.

Figure 3 illustrates the various environments taken into account. The labels of each environment are arbitrary as they were used by the researchers to keep track of what data was collected.

Label	Environment	Activity	Additional Notes
0	Ground - rubble	*Walking, Jogging*	
1	Ground - grass/ dirt	*Walking, Jogging*	
2	Ground - tarmac	*Walking, Jogging*	
3	Ground - tiles	*Walking, Jogging, Walking Upstairs/Downstairs*	This is considered as a regular walk up and down stairs
6	Wearing Shoes - Heels/Football Shoes	*Walking, Jogging, Walking Upstairs/Downstairs*	
7	Wearing Shoes - Boots	*Walking, Walking Upstairs/Downstairs*	
9	Wearing Shoes - Flip Flops	*Walking, Jogging, Walking Upstairs/Downstairs*	
10	Wearing Shoes - Barefoot	*Walking, Jogging, Walking Upstairs/Downstairs*	
11	Pace - Fast Pace	*Jogging*	It is assumed that all other Walking and Jogging activities are done at a medium pace
12	Pace - Slow Pace	*Walking*	
4	Added Weight - 2kg	*Walking, Jogging, Falling, Walking Upstairs/Downstairs*	In the case of falling - the weight is added to env. #25 In the case of walking upstairs /downstairs, the weights were added to envs. #3, #6,#7,#9,#10 as well The weights were placed in a backpack which the subject wore while performing the activity.
5	Added Weight - 4kg	*Walking, Jogging, Falling, Walking Upstairs/Downstairs*	
8	Added Weight - 10kg	*Walking, Jogging, Falling, Walking Upstairs/Downstairs*	
13	On the ground	*Laying, Sitting*	
14	On the sofa	*Laying, Sitting*	
15	On the bed	*Laying, Sitting*	
16	On a bench	*Laying, Sitting*	
17	On one’s side	*Laying*	
18	Facing Up	*Laying*	
19	Comfortable	*Laying, Sitting*	
20	Slouched	*Laying, Sitting*	
21	Straight Down	*Falling*	For these environments, the smartphone was dropped cushioned by two pillows, therefore no additional weight situations were taken into consideration.
22	Stumble Down	*Falling*	
23	Straight Down with added force	*Falling*	
24	Slide Down	*Falling*	
25	Free Fall with weight of body	*Falling*	

Figure 3: List of the different activities and environments

3.3. Saving of Raw Data

The root folder (See Appendix) consists of three subfolders: 'Jogging and Falling', 'Laying and Walking up-down Stairs' and 'Sitting and Walking'. These folders then contain a number of subfolders, one for each record log. Lastly, each record log folder (for example '`Jogging_1.3_(1)`') contains the output produced by the Sensor Logger application, i.e. three comma-separated value (CSV) files: '`metadata.csv`', '`accelerometer.csv`' and '`gyroscope.csv`'.

3.4. Pre-processing of Raw Data

The data collected from the Sensor Logger application included `time in UNIX epoch nanoseconds`, `time in seconds` and `x, y` and `z` values for both the accelerometer sensor and the gyroscope sensor. The researchers contacted the developer of the application [30], Kelvin Choi, to find out the sample frequency of the data collected. Choi replied saying that the application is set to log data at the fastest rate possible for the hardware that it is running on. Since the data was collected on three different smartphones, the sampling frequency was calculated separately for each smartphone by finding the average number of samples per second. This produced the following:

Smartphone	***Activities***	***Sampling Frequency Average of Time Domain Signals***
One Plus 6T	Jogging and Falling	200Hz
Samsung Galaxy A8	Walking and Sitting	100Hz
Samsung Galaxy A51	Laying and Walking Up/Downstairs	120Hz

Figure 4: Sampling frequencies for each activity

The data was first passed through a median filter with a kernel size of 3, then passed through a third order low pass Butterworth filter with a cut-off frequency of 20Hz in order to remove noise. This was the same procedure followed in [1] which aimed to do the same task with smartphone data. Next, the acceleration and gyroscope signals were separated into body acceleration and gravity acceleration by applying another low pass Butterworth filter, with a cut-off frequency of 0.3Hz.

In order to obtain jerk signals, the body linear acceleration and angular velocity were derived with respect to time. The Euclidean Norm, which is the square root of the sum of the squared components, was computed over time domain signals and jerk signals to obtain the magnitude.

The original `tBodyXYZ` signals and `tBodyJerkXYZ` signals were then passed through a Fast Fourier Transform (FFT) to obtain frequency-based components. The magnitude of these signals was then obtained using the Euclidean Norm. Figure 5 depicts the flow of this process.

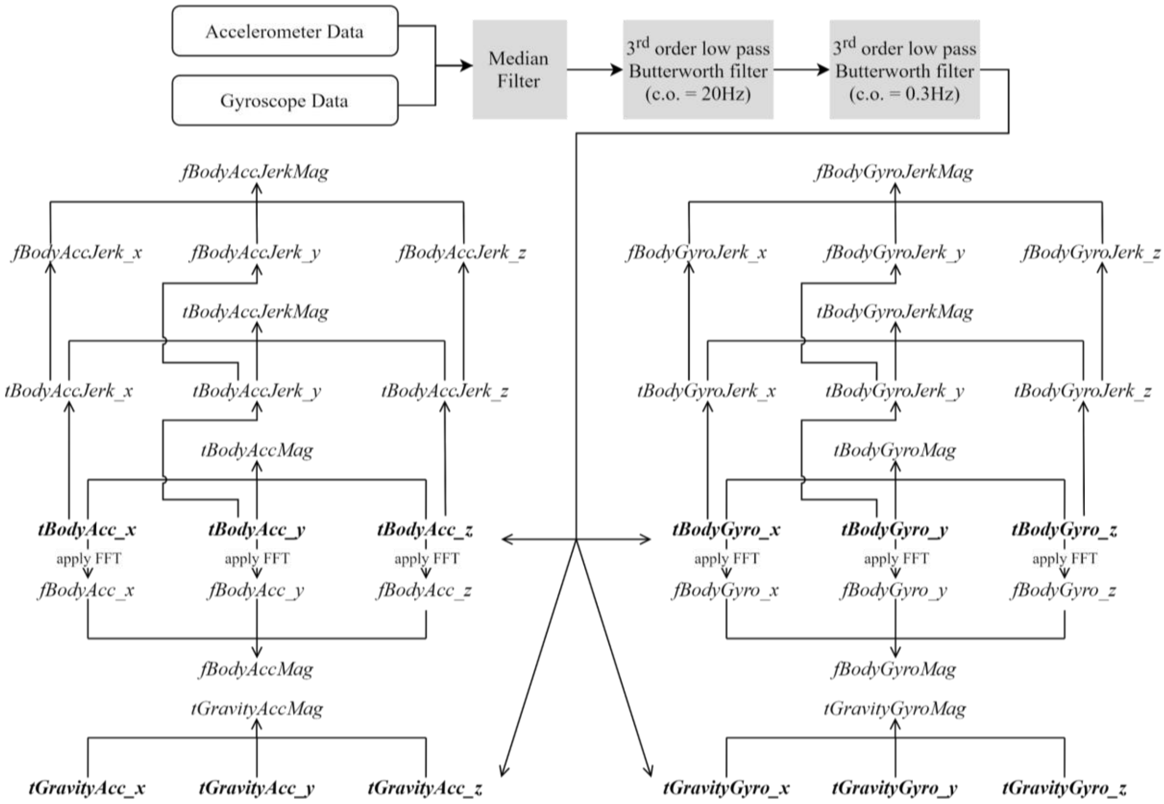

Figure 5: Feature Extraction

After having produced all the above signals, a number of variables were computed to create features. The variables obtained include the mean, standard deviation, min and max value and the interquartile range (IQR). These variables were produced using the Pandas' `describe()` method. Additionally, the median absolute deviation (MAD), skewness, kurtosis and entropy were all calculated in order to increase the variety of features. The `Pandas` library was used to obtain the first three whereas the `Scipy` library was used to obtain entropy. This created a total of 279 features which were calculated for the researchers to be able to better compare the results with those obtained in [1].

With reference to the structure of files mentioned in section 4.3, all the files were loaded using the following method. The paths of all files found were saved to a list (for example: '`..\Data\Jogging _ Falling\Falling_1.4 (4)\Accelerometer.csv`'). The files named '`metadata.csv`' were excluded since the data they contain is not needed. The filename was important to extract the activity and environment. First the filename was tokenized using the '`\`' character. Then the third element of the resulting array was tokenized using the '`_`' character. The first element of the resulting array is then set as the name of activity and the second element is set as the environment.

The data was split into training data and testing data using Pandas' `df.sample()` with a fraction of 0.8 and a random seed value of 250. This resulted in a total of 300 entries for the training data and 75 entries for the testing data. A bar chart depicting the distribution of activities was plotted as seen in Figure 14.

3.5. Fitting ML Models to Pre-processed Data

The data was loaded as a `Pandas` dataframe. It was ensured that no null values were present in this data. Several diagrams and graphs were plotted in order to better understand the distribution of the data. The models used were an SVM [28], a KNN [28], and an LSTM model [29]. The data was adjusted so that it matches each models' input type and then was split into *x* and *y*. One occurrence in '*x*' holds the different features of one occurrence of one activity while the corresponding '*y*' holds the activity name. The SVM and KNN models were implemented using the `sklearn` library, while the `Keras` and `sklearn` libraries were used for the implementation of the LSTM. The SVM model used is a 'support vector classifier'. For the KNN model no parameter was given, so the sklearn library takes 5 as the default value for 'k'.

The LSTM model consisted of the following 4 layers.

- LSTM layer with 128 units and a ReLu activation function
- Dropout of 0.2 and Batch Normalisation
- LSTM layer with 64 units and a ReLu activation function
- Dropout of 0.2 and Batch Normalisation
- Dense layer with 32 units and a ReLu activation function
- Dropout of 0.2 and Batch Normalisation
- Dense layer with 6 units and a SoftMax activation function

Finally, the model was compiled using categorical cross entropy and Adam optimiser.

4. Results

On the UCI HAR dataset, the best performing model was that of the LSTM with an accuracy of 96%, whereas the model with the worst accuracy was KNN with an accuracy of 90%. It should be noted that this is still overall a good accuracy result. On the collected data, the models produced a wider range of results. The best performing model was the LSTM model with an accuracy of 94%. The model with the worst accuracy was the SVM with an accuracy of 73%, while the KNN performed with an accuracy of 89%. These values are depicted in Figure 6. The LSTM Model was run 10 times and it produced a variety of accuracy results ranging from 61% - 94%. The average accuracy was that of 76%. A confusion matrix and a Classification Model were produced for each model, as seen in Figures 7 - 12. These results are discussed further in the Evaluation section.

	Results using UCI HAR Dataset	**Results using Data Collected**
SVM	95.05%	73.33%
KNN	90.02%	89.33%
LSTM	93.69%	94.67%

Figure 6: Comparison of the final results

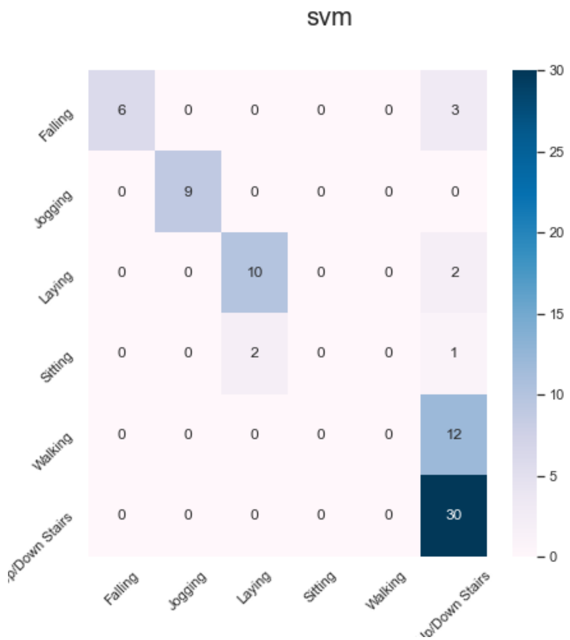

Figure 7: SVM Confusion Matrix

Name: svm, Accuracy: 73.333, Precision: 62.333, Recall: 73.333

	precision	recall	f1-score	support
Falling	1.00	0.67	0.80	9
Jogging	1.00	1.00	1.00	9
Laying	0.83	0.83	0.83	12
Sitting	0.00	0.00	0.00	3
Walking	0.00	0.00	0.00	12
Up/Down Stairs	0.62	1.00	0.77	30
accuracy			0.73	75
macro avg	0.58	0.58	0.57	75
weighted avg	0.62	0.73	0.66	75

Figure 8: SVM Classification Report

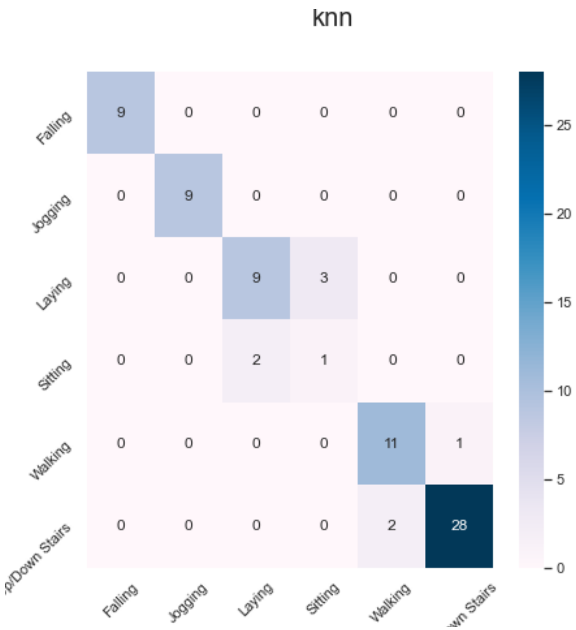

Figure 9: KNN Confusion Matrix

Name: knn, Accuracy: 89.333, Precision: 90.250, Recall: 89.333

	precision	recall	f1-score	support
Falling	1.00	1.00	1.00	9
Jogging	1.00	1.00	1.00	9
Laying	0.82	0.75	0.78	12
Sitting	0.25	0.33	0.29	3
Walking	0.85	0.92	0.88	12
Up/Down Stairs	0.97	0.93	0.95	30
accuracy			0.89	75
macro avg	0.81	0.82	0.82	75
weighted avg	0.90	0.89	0.90	75

Figure 10: KNN Classification Report

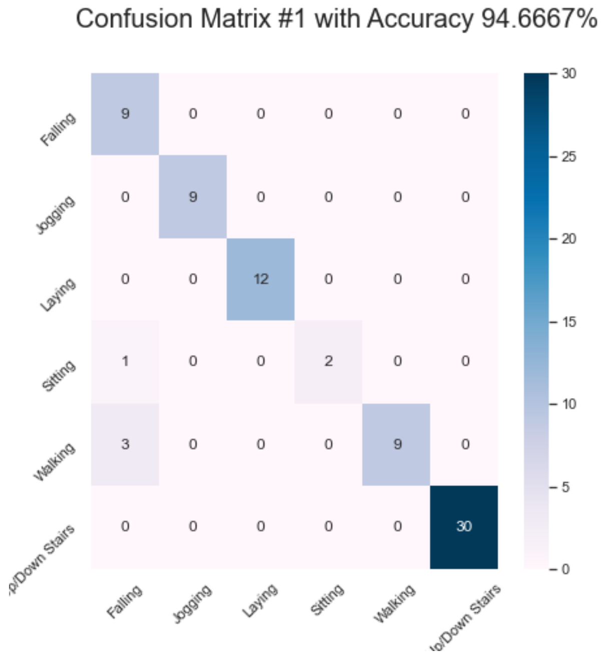

Figure 11: Best LSTM Confusion Matrix

```
LSTM: 94.66666666666667%
                precision    recall  f1-score   support

       Falling       0.69      1.00      0.82         9
       Jogging       1.00      1.00      1.00         9
        Laying       1.00      1.00      1.00        12
       Sitting       1.00      0.67      0.80         3
       Walking       1.00      0.75      0.86        12
Up/Down Stairs       1.00      1.00      1.00        30

      accuracy                           0.95        75
     macro avg       0.95      0.90      0.91        75
  weighted avg       0.96      0.95      0.95        75
```

Figure 12: Best LSTM Classification Report

Figure 13: Activity Distribution for Kaggle Dataset

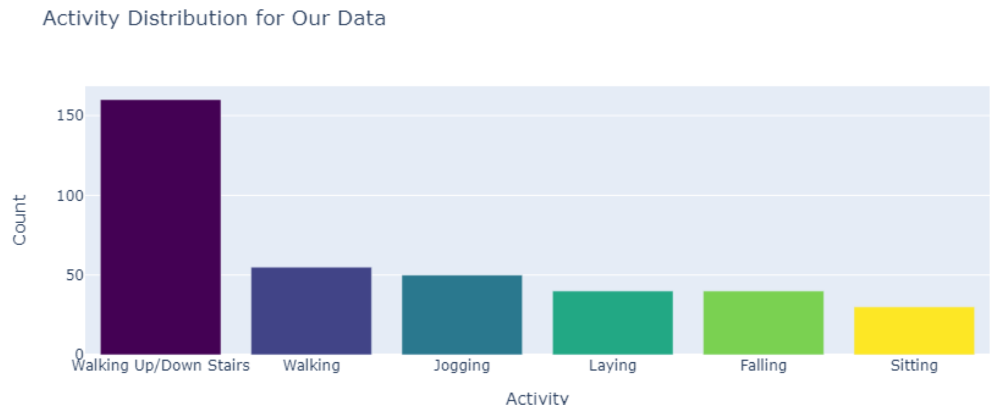

Figure 14: Activity Distribution for Our Dataset

5. Evaluation

The model in [1] was an MC-SVM implemented on the UCI HAR Dataset, which produced an accuracy of 89%. On the same dataset, the LSTM implemented in this project performed the best whereas the KNN performed the worst. It should be noted that they all performed well, with an accuracy of over 90%.

Figure 14 depicts the distribution of the data in terms of activities. The data collected was distributed unevenly with the majority of it being the ‘Walking Up/Down Stairs’ Activity due to the fact that it was grouped as one activity but was recorded as 2 separate activities. As a result, all models predicted this activity very accurately each time. On the other hand, ‘Laying’, ‘Falling’ and ‘Sitting’ were the activities with the least amount of data collected. This was due to the fact that the researchers were unable to add variety to the environments in which they were performed (adding weights and wearing different footwear is not applicable when sitting or laying as these factors will not affect the signals). The lack of data for these activities resulted in them rarely ever being predicted accurately, as can be seen in the confusion matrices of the models. The UCI HAR dataset [4] was more evenly distributed (Figure 13) and this substantially affects the performance of the models. It should therefore be noted that when performing an experiment of this sort, apart from having a large dataset, it is also ideal that this dataset is balanced and that there is not a large gap between the activity which has the most data collected and that which has the least.

The confusion matrices as well as the classification results depict more clearly which activities were misclassified. It should be noted that Jogging is, in all three models, classified correctly with 100% precision and recall. Also, walking up/downstairs is mostly classified with 100% recall, except for the KNN model which predicts this activity with 93% recall. In the case of SVM, the recall was 100%, but the precision was 62%. This means that although walking up/downstairs was classified correctly in all cases, other activities were mis-classified as walking up/downstairs. While the LSTM classified sitting and walking with 100% precision and ~70% recall, the SVM model never classified these activities correctly resulting in 0% precision and 0% recall. The LSTM Model classified 25% of the test cases for walking as falling. This is peculiar because the activities are nowhere near similar to each other. It would make more sense had this activity been mistaken with another moving activity, such as, jogging or walking up/downstairs.

Upon comparing these classification results with [9], as well as those of the UCI HAR dataset, it was noted that the misclassification of walking with walking up/downstairs is a common issue. It is worth mentioning that the precision for these activities was still above 92%, for both [9] and the UCI HAR Dataset.

The data was not highly processed owing to the fact that not many hand-crafted features were extracted for the researcher’s dataset, therefore the SVM produced a relatively *low* accuracy of 73%. When the same model was implemented on the UCI HAR dataset an accuracy of 95% was produced. This reduced performance is expected since SVMs classify the data based on the number of complex, hand-crafted features, where the larger the amount of these features, the more accurate the result.

The KNN, much like the SVM, is a very easy model to implement. This model performed quite well when taking into consideration the imperfections of the dataset, following from the conclusions made in [16]. The activities that were misclassified as each other were ‘sitting and laying’ and ‘walking and walking up/down stairs’. This is logical since the movements, or lack of, are very similar.

As noted in the Results, the LSTM produced a different result with each run. This is due to the fact that Keras randomly initializes all the weights in the model, and it is therefore unlikely that the weights will be the same with each run. Despite the model performing well with the best accuracy being of 94%, the average accuracy produced is comparable to that of the SVM. The program was run various times and despite the best accuracy changing each time, the average remained the same throughout.

As previously stated, the COVID situation imposed certain limitations on our data collection process. Possible future improvements include a dataset which is substantially larger in size and contains data from a number of subjects varying in ages. The dataset should also be more balanced, having the same amount of data for each activity being performed. The activity of walking up and down stairs could also be treated as two separate activities. Furthermore, the data should be collected on an application that records data at a constant sampling rate and one device should be used for all activities, for a more standardized approach.

The smartphone’s location on the arm, despite not hindering the individual’s ability to perform tasks, is still slightly uncomfortable and unrealistic in a natural setting. In the future one could try this experiment while having the phone in the individual’s pant pockets, as this is a more common, convenient location for the individual and might also provide more accurate results. This could also be combined with other body worn sensors, like a smartwatch. The system could then be modified to be able to fuse the sensor readings from both the smartphone and the smartwatch, allowing a greater accuracy when it comes to classifying the physical activities, following from [16]. This inclusion would also be ideal as the individual would be able to use the smartphone freely while the smartwatch will always be fixed to their wrist throughout the day regardless of the activity being performed.

With regard to the ML models, future research could include a CNN model and compare it to the models investigated in the above project, since CNN is another DL model that might more accurately classify the activities, following from [8], [16] and [24]. It was also found that the SVM performs better with more complex, hand-crafted features, hence, increasing the number of features extracted would result in improved performance. This could include adding features, such as, Signal Magnitude Area (SMA), energy measure and autoregression coefficients.

6. Conclusion

This project aimed to expand the work in Human Activity Recognition (HAR) using a smartphone's accelerometer and gyroscope data in order to be able to classify physical activities like walking and jogging. SVM, KNN and LSTM algorithms were implemented on a UCI HAR dataset with over 90% accuracy and a dataset built by the researchers with results ranging from 73% to 94% for SVM and LSTM respectively. The researchers can therefore conclude that with a large, balanced dataset, an SVM is ideal, whereas with a smaller scale dataset collected from just three subjects, an LSTM is a more robust choice able to handle the flaws in the dataset and still perform at a high accuracy.

This project provides an affordable, comfortable, accurate and accessible system to perform HAR, without the problems of privacy. The above system obtained a comparable performance to other similar systems, while having a greater number of advantages.

References

[1] D. Anguita, A. Ghio, L. Oneto, X. Parra Perez, and J. L. Reyes Ortiz, "A public domain dataset for human activity recognition using smartphones." 2013, [Online]. Available: https://recercat.cat/handle/2072/220865.

[2] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. Reyes-Ortiz, "Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine," 2013. [Online].

[3] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. Reyes-Ortiz, "Energy Efficient Smartphone-Based Activity Recognition Using Fixed-Point Arithmetic," 2013.. [Online].

[4] "Human Activity Recognition with Smartphones," *kaggle.com*, Dec. 2012. https://www.kaggle.com/uciml/human-activity-recognition-with-smartphones.

[5] S. Azar, M. Ghadimi Atigh, A. Nickabadi, and A. Alahi, "Convolutional Relational Machine for Group Activity Recognition," 2019. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/papers/Azar_Convolutional_Relational_Machine_for_Group_Activity_Recognition_CVPR_2019_paper.pdf.

[6] J. L. Reyes-Ortiz, A. Ghio, D. Anguita, X. Parra, J. Cabestany, and Andreucataì, "Human Activity and Motion Disorder Recognition: Towards Smarter Interactive Cognitive Environments," 2013. [Online]. Available: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-11.pdf.

[7] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, vol. 181, pp. 108–115, Mar. 2016, doi: 10.1016/j.neucom.2015.08.096.

[8] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep Learning for Sensor-Based Activity Recognition: a Survey," *Pattern Recognition Letters*, vol. 119, pp. 3–11, Mar. 2019, doi: 10.1016/j.patrec.2018.02.010.

[9] C. A. Ronao and S.-B. Cho, "Human Activity Recognition with Smartphone Sensors Using Deep Learning Neural Networks," *Expert Systems with Applications*, vol. 59, pp. 235–244, Oct. 2016, doi: 10.1016/j.eswa.2016.04.032.

[10] E. M. Tapia, S. S. Intille, L. Lopez, and K. Larson, “The Design of a Portable Kit of Wireless Sensors for Naturalistic Data Collection,” in *Lecture Notes in Computer Science*, 2006, pp. 117–134.

[11] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and G. J. Norman, “Classification Accuracies of Physical Activities Using Smartphone Motion Sensors,” *Journal of Medical Internet Research*, vol. 14, no. 5, p. e130, 2012, doi: 10.2196/jmir.2208.

[12] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, “Fusion of Smartphone Motion Sensors for Physical Activity Recognition,” *Sensors*, vol. 14, no. 6, pp. 10146–10176, Jun. 2014, doi: 10.3390/s140610146.

[13] J. Shen and H. Fang, “Human Activity Recognition Using Gaussian Naïve Bayes Algorithm in Smart Home,” *Journal of Physics: Conference Series*, vol. 1631, p. 012059, Sep. 2020, doi: 10.1088/1742-6596/1631/1/012059.

[14] M. H. Kabir, M. R. Hoque, K. Thapa, and S.-H. Yang, “Two-Layer Hidden Markov Model for Human Activity Recognition in Home Environments,” *International Journal of Distributed Sensor Networks*, vol. 12, no. 1, p. 4560365, Jan. 2016, doi: 10.1155/2016/4560365.

[15] N. R. Nurwulan and G. Selamaj, “Human Daily Activities Recognition Using Decision Tree,” *Journal of Physics: Conference Series*, vol. 1833, no. 1, p. 012039, Mar. 2021, doi: 10.1088/1742-6596/1833/1/012039.

[16] S. R. Shakya, C. Zhang, and Z. Zhou, “Comparative Study of Machine Learning and Deep Learning Architecture for Human Activity Recognition Using Accelerometer Data,” *International Journal of Machine Learning and Computing*, vol. 8, no. 6, Dec. 2018, doi: 10.18178/ijmlc.2018.8.6.748.

[17] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity Recognition from Accelerometer Data,” in *In Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence*, 2005, pp. 1541–1546.

[18] Z. Chen, Q. Zhu, Y. C. Soh, and L. Zhang, “Robust Human Activity Recognition Using Smartphone Sensors via CT-PCA and Online SVM,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3070–3080, Dec. 2017, doi: 10.1109/tii.2017.2712746.

[19] H. Qian, Y. Mao, W. Xiang, and Z. Wang, “Recognition of Human Activities Using SVM Multi-Class Classifier,” *Pattern Recognition Letters*, vol. 31, no. 2, pp. 100–111, Jan. 2010, doi: 10.1016/j.patrec.2009.09.019.

[20] X. Jiang, Y. Lu, Z. Lu, and H. Zhou, “Smartphone-Based Human Activity Recognition Using CNN in Frequency Domain,” *Web and Big Data*, pp. 101–110, 2018, doi: 10.1007/978-3-030-01298-4_10.

[21] D. Singh *et al.*, “Human Activity Recognition Using Recurrent Neural Networks,” *Lecture Notes in Computer Science*, pp. 267–274, 2017, doi: 10.1007/978-3-319-66808-6_18.

[22] F. Ordóñez and D. Roggen, “Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition,” *Sensors*, vol. 16, no. 1, p. 115, Jan. 2016, doi: 10.3390/s16010115.

[23] M. Baccouche, F. Mamalet, C. Wolf, C. Garcia, and A. Baskurt, “Action Classification in Soccer Videos with Long Short-Term Memory Recurrent Neural Networks,” *Artificial Neural Networks – ICANN 2010*, pp. 154–159, 2010, doi: 10.1007/978-3-642-15822-3_20.

[24] C. Chen, R. Jafari, and N. Kehtarnavaz, “UTD-MHAD: A Multimodal Dataset for Human Action Recognition Utilizing a Depth Camera and a Wearable Inertial Sensor,” 2015.

[Online]. Available: http://jafari.tamu.edu/wp-content/uploads/2019/06/ICIP2015-Chen-Final.pdf.
[25] M. Gjoreski, S. Kalabakov, M. Luštrek, M. Gams, and H. Gjoreski, "Cross-dataset Deep Transfer Learning for Activity Recognition," *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, Sep. 2019, doi: 10.1145/3341162.3344865.
[26] C. Ito, X. Cao, M. Shuzo, and E. Maeda, "Application of CNN for Human Activity Recognition with FFT Spectrogram of Acceleration and Gyro Sensors," *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, Oct. 2018, doi: 10.1145/3267305.3267517.
[27] A. Jordao, A. Nazare, J. Sena, and W. Robson Schwartz, "Human Activity Recognition Based on Wearable Sensor Data: A Standardization of the State-of-the-Art," 2018. [Online]. Available: https://arxiv.org/pdf/1806.05226.pdf.
[28] J. Brownlee, "Evaluate Machine Learning Algorithms for Human Activity Recognition," *Machine Learning Mastery*, Sep. 18, 2018. https://machinelearningmastery.com/evaluate-machine-learning-algorithms-for-human-activity-recognition/
[29] khan-farhan, "khan-farhan/HAR-LSTM-keras," *GitHub*, Mar. 2019. https://github.com/khan-farhan/HAR-LSTM-keras/blob/master/src/LSTM%20.ipynb
[30] K. Choi, "Sensor Logger," *Kelvin Choi*, Mar. 2021. https://www.tszheichoi.com/sensorlogger.

Appendix

https://github.com/chrisgalea06/HAR-using-smartphones

Distribution of Work

The work was distributed such that all members contributed equally to the creation of the dataset, as well as the video and the presentation.

Christopher and Diana did the pre-processing of the data, as well as most of the implementations of the ML models. Celine aided in the debugging process.

Celine was responsible for the research and literature review, while Chris and Diana were responsible for the Methodology and Evaluation sections. All members contributed equally to the rest of the paper.