Solent University

SCHOOL OF Business, Law and Digital Technologies

Applied Artificial Intelligence and Data Science

2021/2022

Souvenir Oyawale- Q15726185

"Predicting Human Activity using Motion Tracking with Machine Learning"

Supervisor : Dr Shadi Eltanani

Date of submission : September 2022

Acknowledgements

This Master Thesis is in memory of my Father, Chief Engineer Enyinda Nathaniel Okey, you have missed this one but we will raise a glass for you. I appreciate my sons- Jamien, Nathan and Arric for their undying love and support through this journey. I want to specially thank my mother- you are a rock, I love you. My siblings and the BrightonSovy Group, you guys are the best.

I also appreciate Dr Eltanani for spurring me on, providing support and counsel through the course of the research. Overall, I give all glory to God Almighty, to Him all praise is due.

Abstract

The recognition of physical activities performed by humans is first identified as a classification problem that when solved provides an automated approach to recognising what is being done at any given time in respect to motion. The primary aim is to recognise activities using wearable sensors. The ability to recognise human physical activity is important to many sectors such as health care, security, rehabilitation and general well-being. Information extracted from the activities performed enable the systems learn and perform better. In general, this concept is important especially in healthcare as it could be used as an indicator of health status and lifestyle choices therefore the ability to correctly predict activities is very important.

The design of this study is based on the use of three traditional machine learning algorithms in comparison with a deep learning algorithm to determine which algorithm achieves higher prediction accuracy using each model's base parameters as the benchmark for evaluation. Random Forest, Support Vector Machines and K Nearest Neighbor were the traditional algorithms used while Convolutional Neural Network was the deep learning algorithm. All models were cross-validated and evaluated with classification evaluation metrics.

CNN achieved the best classification accuracy of 99.9%. Significant results were also obtained from RF and SVM. All models hyperparameters were tuned with the best parameters to achieve these results. KNN however, performed the least when compared to the other models.

Finally, the thesis is concluded with discussions on the entire research and identifying limitations, ideas for possible research areas to be investigated for future work.

Keywords: Wearable sensors, Human Activity Recognition, Machine Learning

Contents

List of Tables

List of Figures

Abbreviations

AI	Artificial Intelligence
ACC	Accelerometer
AUC	Area Under the ROC Curve
CNN	Convolutional Neural Network
FN	False Negative
FP	False Positive
HAR	Human Activity Recognition
KDD	Knowledge Discovery in Databases
KNN	K Nearest Neighbor
ROC	Receiver Operating Characteristic graph
SVM	Support Vector Machines
TN	True Negative
TP	True Positive

1. Introduction

The term movement as has come to be known is usually used when reference is made to some form of fluid change in position which is in most instances due to a form of activity. Being able to identify the activity being performed also referred to as Activity Recognition (AR) or Human Activity (HA) through the course of this paper could be inferred to mean the ability to identify a form of movement being carried out by an observed participant or more in a defined context.

Questions could be raised on why recognising HA has become quite important in the day-to-day activity of humans in particular; it has been observed that the need for physical activity recognition has risen with the development of wearable technologies driven by sensors (Ghorpade, et al., 2020). However, this does not fully address the questions on the need for recognizing these activities in the first place.

To address this is to identify where these technologies have the most impact. Fitness, senior care and living, mental health and other applications including surveillance easily come to mind. Nevertheless, the importance of human activity recognition cuts across different sectors and with research increasing so has the need for precision in predicting the activities being carried out.

With the advancement and notable success in the use of neural networks and deep learning models, Convolutional Neural Network (CNN) in particular has demonstrated significant prowess in being able to extract the right features from provided data to improve performance (Rawat and Wang 2017). The use of CNN architecture has been seen various classification problems such as activity recognition, object detection, speech recognition, and picture classification with astonishing results. The model's design allows it extract and process data features through several layers (Khan, Afzal, and Lee, 2022).

Sensor provided activity information through gyroscopes, magnetometers and accelerometers are presented as time-series data; the use of machine learning models with modified hyperparameters have been able to make substantial progress in the prediction and recognition of human activities (Bonato 2003).

Newer research however has been focused on the use of deep learning models has been useful in eliminating noise in the signals, reliance on feature engineering and modified feature extraction.

With the increase in the deployment of Artificial intelligence powered applications and devices, understanding the complexity and diversity in human activities is imperative (Lara and Labrador 2013).

Research Problem

As straightforward as motion and activity recognition seems, it is not without challenges, some of which have already been identified. The issue of storage and processing power is not commonly addressed but in processing and modelling might lead to a trade off between efficiency and accuracy.

A process flow of the recognition of human activity through sensors will paint a better picture of the problem this research is trying to address. Data is collected from sensors then processed and transformed for exploration. The next stage is dependent on the sector the data collected belongs to but in totality, it undergoes engineering where the most relevant features are extracted then a model applied to make predictions and then evaluated.

The problem encountered by current methods in activity detection is the ability to identify new activities from previously identified activities. This is due to reliance on user's input so the algorithms would be able to identify and label these activities, and also improve in performance (Cheng et al. 2017).

The research question to be addressed in this study can then be stated as-

How efficient are supervised machine learning algorithms in predicting and recognising human activity when compared to neural network, CNN in particular?

*Algorithms: Support Vector Machine, K-Nearest Neighbors, Random Forest and Convolutional Neural Network.

1.2 Research Objectives

To approach this by picking a particular algorithm to perform the task of prediction and recognition is tricky and complicated. Selecting a couple of models with different algorithms is an approach that serves as a good approach as through the evaluation of the performance of each of these models it becomes evident which model serves as a better option which can be further tuned for optimal performance.

The primary aim of this research is to be able to ascertain which machine learning algorithm performs with the highest accuracy in classifying physical activities and compare this with a neural network, CNN in particular. To achieve this aim, the following objectives would be addressed-

I. Evaluate the accuracy of a machine-learning (ML) algorithm in predicting human activity

II. Test the validity of model outcome using independent sample

III. Identify and interpret features most useful to model predictions

IV. Discover insights that could be used to improve model performance

The models utilized in this research are K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Random Forest (RF) and Convolutional Neural Network (CNN). The performance of each of these models will be evaluated using common classification metrics such as the precision, F1 Score, Accuracy and Recall.

Another metric that would be used to evaluate these models are the Receiver Operating Characteristic Curve (ROC) and the Area Under the Curve (AUC).

1.3 Research Methodology

Data collected for this research is done using secondary research approach and is based on existing data from an earlier experiment performed.

Using the knowledge discovery approach, the experiment will be carried out on the existing dataset using the previously identified models to be able to identify the correct human physical activity being carried out. The first process of this approach is collecting and cleaning the data collected, the next stage is data processing, tuning and evaluating the model comes next and is concluded with checking the predictions, evaluating the models and noting findings.

1.4 Scope and Limitations

This study utilizes a single dataset from a previous study carried out and hereon referred to as the MHEALTH (Mobile HEALTH) dataset. The collected data is based on data extracted from the sensors which were placed on each participant's chest, right wrist and left ankle. Further extrapolation and engineering would have been possible if the raw sensor signals were available.

The experiment performed in the pilot for this study is also used in conjunction with that performed in this study as a benchmark. Due to computing constraints, data resampling was performed. Likewise, a limited number of algorithms were used and manipulated to perform the modelling.

1.5 Project

The project would be carried out using the Python programming language. Other resources used is in the model deployment. The project was carried out over a period of three months in different phases as shown in the Gantt Chart in Figure 1 below. The use of the chart was important to guide the project implementation and ensure timeliness.

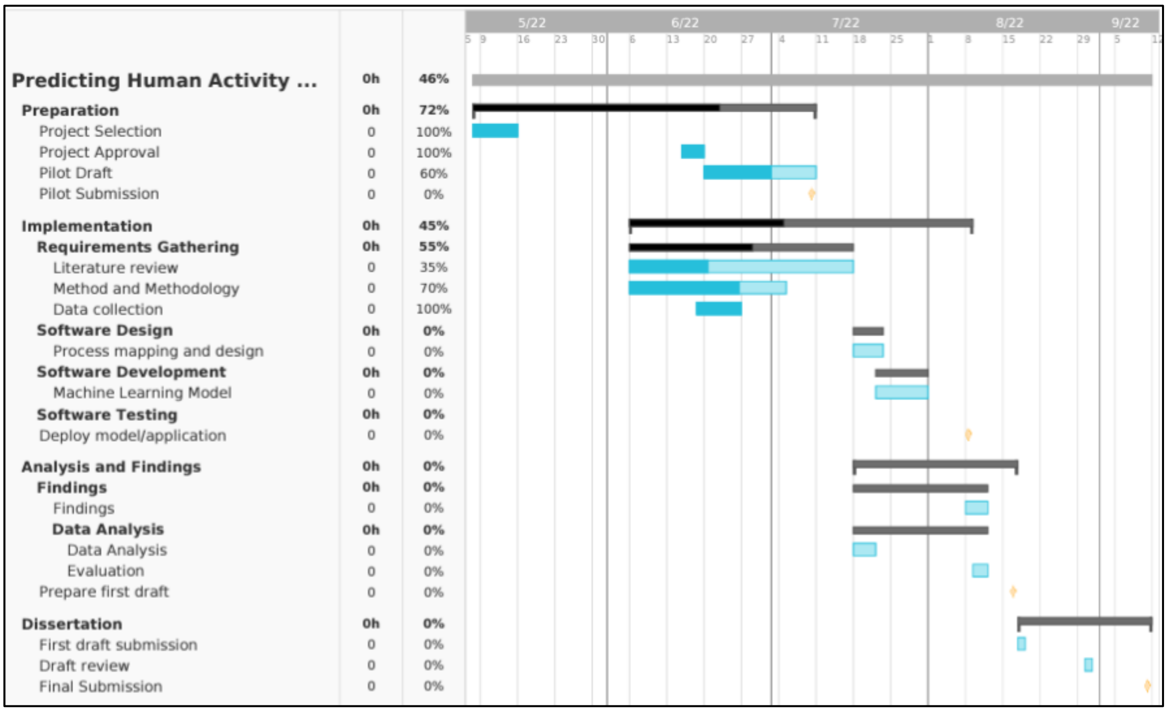


Figure 1 Gantt Chart for project

1.6 Thesis Organisation

This thesis is organized as follows:

- Chapter 2 presents an overview of a literature review on the current related work in activity recognition in humans. The purpose of this chapter is to also review different algorithms and techniques that have been utilised in previous studies and research. Also, to be reviewed are the different metrics that previous researchers have used to evaluate the suitable models adopted in their studies.

- Chapter 3 describes the design used for this study which is modelled on the Knowledge Discovery Process Methodology.
- Chapter 4 presents the implementation of the models and their robustness for activity recognition using the Knowledge Discovery Process Methodology and the results obtained from the experiment evaluated by comparing each algorithm against set evaluation metrics.
- Chapter 5 discusses the results and its importance in the context of the study.
- Chapter 6 concludes the thesis and suggests area for future work.

2. Literature Review

This chapter presents a comprehensive review of relevant literature of approaches and techniques used in the recognition of activities in humans. It discusses the technologies and algorithms employed in the studies as well as where these technologies have been used. It first reviews the concept of human activity recognition then focuses on wearable sensors. It then addresses the use of different machine learning algorithms that can and have been used. Further review is made on prevalent machine learning models in similar research. It will also highlight approaches utilised to enable optimal detection performance in the models. In conclusion, this chapter will highlight and justify the techniques that have in previous studies been identified to provide better detection. An illustration is made below presenting the design of this chapter. This section in summary aims to critically analyse existing studies and approaches and in doing so provide insight on how best to approach current work and optimize existing algorithms in predicting human activity recognition using machine learning models.

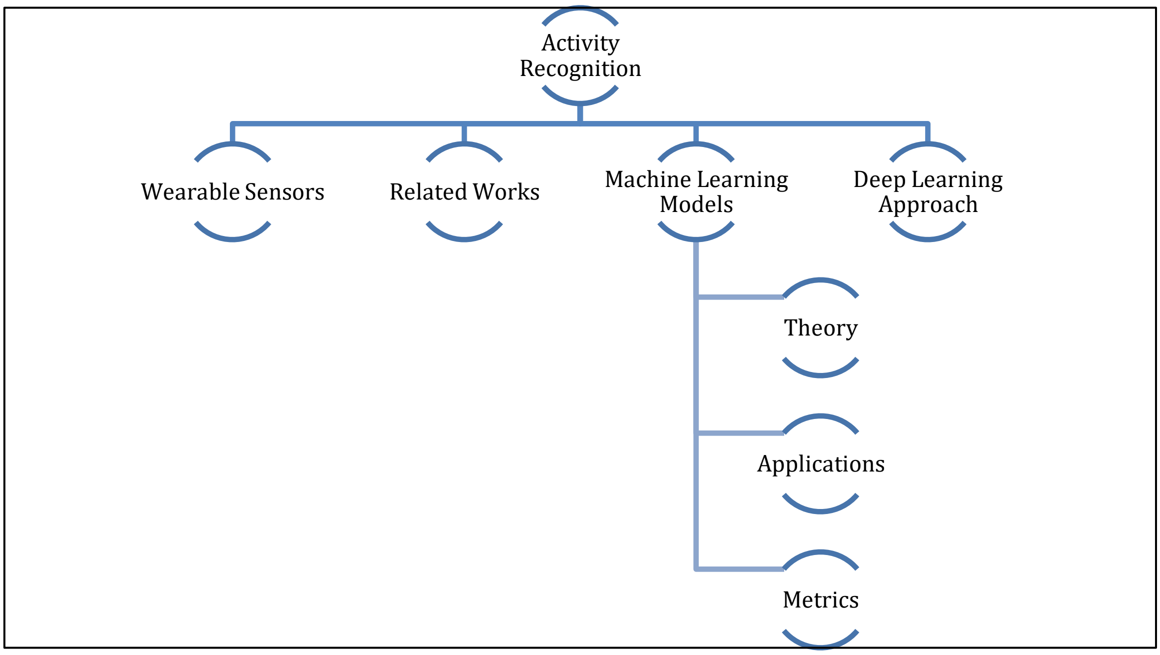


Figure 2 Layout of Chapter 2

Activity Recognition (AR)

Human activity recognition using sensors can be broadly categorised into two main segments. First, stationary sensors which are environment based. Such sensors are security and or surveillance application based and use computer vision technologies. These stationary sensors even with the advancement in technology are sometimes seen as intrusive especially in the health sector. The second category of sensors are the mobile and, in most instances, placed on the body where they initially used accelerometers to detect activities. These sensors are quite robust and small with a long battery life (Chowdhury, et al., 2018)

In recent times, there has been an increase in research and importance in the recognition of activities in humans, activities in this context being physical although the scope of activity recognition covers this and is not limited to mental activities (Moustafa, Luz and Longo 2017) . This increase has been largely due to its application in real world scenarios such as medical and security (Yurur, and Liu 2014) another application in particular is that of wearable sensors whose use has also been on the increase with the advancement in technology (Lara and Labrador 2013). These wearable devices make use of sensors which are non-visual and can be found in smartwatches, fitness bands and smartphones which require constant and regular use. Despite the size of these devices, they are powerful, able to communicate and quite affordable. Their use encompasses monitoring daily health stats, rehabilitation and general well-being (Xu *et al.* 2016).

2.2 Wearable Sensors

Data gathering for activity recognition using sensors is extracted per individual user at a time even where multiple users are involved. Due to the portability and compactness of these sensors, it is easier for them to be worn and for a longer length of time which translates to longer periods of data generation (Bulling, Blanke and Schiele 2014). These sensors are integrated with accelerometers and gyroscopes and have been used in many applications especially in the health sector to aid diagnosis, monitoring and treatment (Powell, Hanson and Lach 2009).

With the advancement in technology, smartphones have provided novel opportunities in the daily research of human activity recognition as its inbuilt sensors which consists of accelerometers, magnetometers and gyroscopes among others make it an easy sensing instrument that is unobtrusive, multiuse and flexible. Wearable sensor devices such as the Mi Band and Fitbit in combination with other health tracking apps that can be downloaded in smartphones make data accessible and this data can be used to the recognition of activities (Xu *et al*. 2016).

However, several studies have indicated that the location and number of the sensors on the body have provided varying results using accuracy as a metric. Using a single sensor placed around the waist to detect similar activities resulted in an accuracy of 93% and 98% respectively (Bonomi *et al*. 2009) (Gupta and Dallas 2014). Nevertheless, an increased accuracy in another study was observed using three sensors with one of the sensors also placed on the waist. Accuracy for this experiment was 100% (Wee-Soon *et al*. 2008).

2.3 Machine Learning

2.3.1 Overview

Advancement in technology has seen an increase in the dependence in machine learning especially with the ease of availability of data. As data is being generated from multiple sources there is need for quicker and more efficient analysis, hence the need for integrating machine learning.

Machine learning also referred to as automated learning can simply be said to be programming computers using algorithms to learn from data fed to them and having the information transformed to knowledge (Shalev-Shwartz and Ben-David 2014). Machine learning could also be defined as using machines to replicate the way human behave particularly the way humans intelligently solve problems or perform intricate activities (Brown 2021).

The essence of this learning which is a subset of artificial intelligence is having the computers with the help of algorithms use data to describe what has happened, what will happen and suggest what should be done without the intervention of a human (Malone, D. 1993) (Malone, T., W., Rus and Laubacher 2020) (Malone, Rus and Laubacher 2020). The choice of algorithms which is a set of mathematical computations needs to be efficient in both time and space, flexible and adaptable to achieve minimal error in its predictive accuracy (Vellampalli 2017).

Several algorithms exist and are in use however, the regular ones are:

Logistic Regression- this model is able to address complex patterns but is easily skewed by outliers in the dataset

Support Vector Machines- this algorithm is known for its excellence in prediction especially for classification problems

Random Forest- the design of this algorithm is similar to that of the Decision Tree as it makes use of various trees and is adept at handling missing information in the dataset

K-Nearest Neighbour- a quite robust algorithm that is suitable to noisy data and large datasets.

2.3.2 Applications

The application of machine learning on the other hand keeps growing with trends in technology advancement and usages. Some of these applications of which predicting human activity using sensors is included are (Shinde and Shah Aug 2018)-

- Computer vision- used in object detection, identification and recognition for security, motion detection and in newer applications logistics
- Information and Semantic analysis- text translation using natural language processing to predict users’ activities
- Prediction- through classification and or regression methods to predict prices, health conditions and actions

Other possible applications are further illustrated in Figure 3.

Figure 3 Applications of Machine Learning

2.3.3 Metrics

Understanding how well a machine learning model has performed is a crucial part in the model learning process as it evaluates the accuracy in prediction, model reliability and efficiency. There are several ways a model can be evaluated. Selecting the right metric is dependent on what the model is trying to achieve. This study is a classification problem and would focus on common classification metrics.

2.3.3.1 Confusion Matrix

This matrix is quite useful in showing how a model has performed. It compares the number of predictions against the number of incorrect predictions. It is better illustrated using a table.

		Predicted	
		Positive	Negative
Actual	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Table 1 Confusion Matrix

There are 4 values reported.

True Positives- this value is the number of positives correctly identified by the model as positives

False Positives- This is the number of observations incorrectly identified as positives but are indeed negative

True Negatives- The number of negatives which are actually negative

False Negatives- The number of incorrectly identified observations which are indeed positive but reported as negative.

2.3.3.2 Accuracy

The accuracy of a model could simply be said to be the division of the number of correct predictions against the total number of predictions made. The result is a ratio between 0 and 1 where 1 would show an accurate prediction.

Although this presents a straightforward metric, it is advised that it is used in conjunction with other metric measurements so as not to result in a misleading metric interpretation.

Accuracy = Correct Predictions/Total Predictions

It can also be written as below for a clearer understanding using the context of the confusion matrix.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

2.3.3.3 Precision

This metric measures the integrity of the model in identifying positive prediction class. The usefulness of this metric is visible in detecting fraud where it would be important in determining that all identified positives are indeed positive although this would defeat the purpose of investigating instances where the model missed to predict all positive classes.

Precision= True Positives/True Positives + False Positives

2.3.3.4 Recall

This metric is usually used together with precision to present a better picture of the model's performance. It illustrates how well the model predicts all positive instances in the data presented to it. A key difference between Recall and Precision is that it is not calculated with the false positives. In other words, it calculates the true positive against the false negatives and true positives.

Recall = True Positives/True Positives + False Negatives

2.3.3.5 F1 Score

This score is the harmonic mean of the both the precision and recall values. Where the score is 1, it would indicate a perfect precision and recall and where 0, it indicates that either of these two metrics are 0.

F1 = 2 x precision x recall/precision + recall

2.3.3.6 AUC Value/ROC Curve

The Receiver Operating Characteristic (ROC) curve and the Area under the Receiver Operating Characteristic Value (AUC) are two important values used in evaluating a model's performance. The ROC is a plot of the true positive against the false positive and is a plot of the model's accuracy and is best used where there is no imbalance in the dataset. A good classifier is highlighted when the curve is closer to the True Positive o the Y-axis.

The AUC on the other hand measures the area under the ROC curve. Where the AUC value is close to 1 it means the classifier is a good fit. The closer the better the classifier. However, this metric should not be used in isolation as it is not able to identify when the model misclassifies data

2.4 Deep Learning Approach

Deep learning is a method of machine learning which uses neural network made up of multiple layers to extract features and transform data. The layers are structured in such a way where those close to the input data learn the basic features of the data and the higher layers use the features already transformed by the lower layers. This shows that deep learning is likely to perform better when used to analyse complex and large datasets.

Just as is the case in traditional machine learning, there are different categories. For the purpose of this paper, 3 major categories would be addressed.

Long Short-Term Memory (LSTM)- this model is able to analyse and exploit patterns in dataset by self-learning. It works using gates and stores information for ease of reference in training.

Recurrent Neural Networks (RNN)- this model though difficult to train is quite good at showing relationships in the neural network between the input and output.

Convolutional Neural Networks (CNN)- excellent in image recognition, classification and detection. Uses layers to extract important features from data to aid classification or prediction.

2.5 Related Studies

A cross-section of relevant work that correspond with the objectives of this study have been reviewed. Using machine learning on data extracted from sensors in a smartphone (Saha *et al*. Jul 2020) were able to achieve an accuracy of 95.99 percent through Logistic Regression Model. In this instance, the sensor being placed close to the waist achieved better precision. Support Vector Machines though modified (Anguita *et al*. 2013) were also used to recommend an efficient approach to human activity recognition using wearable sensors with the focus on smartphones. With an accuracy of 94 percent (Casale, Pujol and Radeva 2011) using Random Forest were able to design a device that was based off 20 features that can identify activities in real-time. An interesting paper is that of (Phan Sep 13, 2014)which was able to mitigate issues with the classifications in recognizing human activity using machine learning models. This was able to be achieved using neural networks and resulted in an improvement of about 10 per cent. Another approach of using an ensemble model with random forest and SVM models resulted in an F1 score of 91% (Chowdhury *et al.* 2018).

In (Arif *et al.* 2014) different types of classifiers have been used to classify six types of physical activities. From their research, it was found that KNN was best suited for their research. They have achieved an average accuracy of about 95%. By comparing five different approaches using machine learning models including random forest, KNN and SVM, activities like sitting, walking, standing and falling were identified with KNN having the best performance (Ghazal *et al.* 2019).

Convolutional Neural Networks (CNN) was used as a model in research to recognise human activity using a dataset similar to the dataset for the experiment to be conducted in this paper (Ignatov 2018). By using this model, the research noted a significant improvement in performance compared with traditional machine learning models. Besides the difference in performance, the research also noted a difference in computational cost, with the CNN model being lower in cost which allows for its use in smartphones and wearable devices. CNNs with one dimension have been demonstrated to be easier to train and have the least computational complexity and still being able to achieve state-of-the-art levels in accuracy (Kiranyaz *et al.* 2021).

There have been numerous studies undertaken to compare algorithms using several techniques in classification problems to recognise activities from data collected from wearable sensors. The studies evaluated and compared four supervised classification algorithms - k-Nearest Neighbour (k-NN), Support Vector Machines (SVM), Gaussian Mixture Models (GMM), and Random Forest (RF), and three unsupervised classification algorithms- k-Means, Gaussian mixture models (GMM) and Hidden Markov Model (HMM). The metrics measured were F-measure, recall, accuracy, and specificity (Attal *et al.* 2015). The results highlighted the efficiency of the k-NN method as it presented the best results among all the algorithms used. In (Bayat, Pomplun and Tran 2014), data was extracted for phone being in hand and in pocket. Applying several algorithms for both positions which included both SVM and Random Forest, an accuracy of 91 per cent was achieved. Another similar dataset was used to conduct research where algorithms utilised included, KNN, Random Forest, Decision Trees and SVM among others, the research was able to achieve a high accuracy of 98.96 per cent with SVM although the SVM kernel was tuned with the RBF kernel (Minarno, Kusuma and Wibowo).

2.6 Summary

As can be seen from previous research reviewed, the use case of activity recognition has become a focus of studies and one of the objectives of this study is to be able to determine the suitability of machine learning to achieve high accuracy in recognising human activity. The next section of this paper would then evaluate machine learning algorithms to ascertain which of these algorithms performs with the highest accuracy in classifying physical activities and then compare the algorithm's result with a neural network, CNN in particular.

2.6 Ethics

Ethics approval and clearance was received for this project in accordance with the University's policy on ethics in research and can be seen in Appendix A. Dataset for the research in this study was obtained from the UCI repository (Dua and Graff 2019).

3. Methodology and Design

In this chapter focus is on the design and methodology for this study. Approach was to review and study several methodologies as study required data mining. The Knowledge Discovery in Databases (KDD) approach was chosen as the method for the study. This approach is quite interactive and cuts across various disciplines (Fayyad and Uthurusamy 1996). This methodology uses data mining processes to extract useful information and this chapter will go through each of these processes separately for better understanding. The processes are iterative and could be seen to reoccur depending on the result to be achieved. Nevertheless, it is worth to note that this methodology requires an understanding of the use-case, applicable domain and overall knowledge by the user (Azevedo, Ana Isabel Rojão Lourenço and Santos 2008). This is sometimes confused with the Cross Industry Standard Process for Data Mining (CRISP-DM) which is also a data mining methodology that is quite similar to KDD and according to (Piatetsky 2014) quite preferred but not widely used.

First process in this methodology is the Application Domain understanding phase; this stage requires prior knowledge relevant to the business use case and identifying what the customer would like to achieve. The next process is the creation or selection of a dataset for the project. This could also be sampling of an existing dataset for further analysis or processing. The dataset is explored element by element for quality and any initial statistical findings. The next stage which is the third is the data cleaning stage. This involves removing 'noise' from the data if applicable, duplicates, handling missing data and further cleaning is applied. This is done to create a stable dataset that would be used in the next stages. The fourth stage is the extraction of useful features in the dataset dependent on what is to be achieved in the project. Transforming or reducing the data dimensions are applied as applicable dependent on the number of variables presented in the dataset. The next stage is selecting the right data mining method that aligns with the goal of the project. This could either regression, classification or clustering but not limited to these three. The sixth process is the model selection phase. This includes selecting the appropriate model and algorithms, searching for patterns in the data, experimenting with the settings and matching the method selected with the goal of the project. The seventh is the actual mining of data; looking for interesting patterns. The correctness of the previous steps determines how well this step performs. The eight stage is the interpretation of the mined information. This is done through evaluating metrics, visualising the patterns extracted. The ninth and final process is using the information or knowledge extracted either by incorporating it into another process, document and report findings or resolving discrepancies from initial knowledge extracted (Fayyad and Uthurusamy 1996).

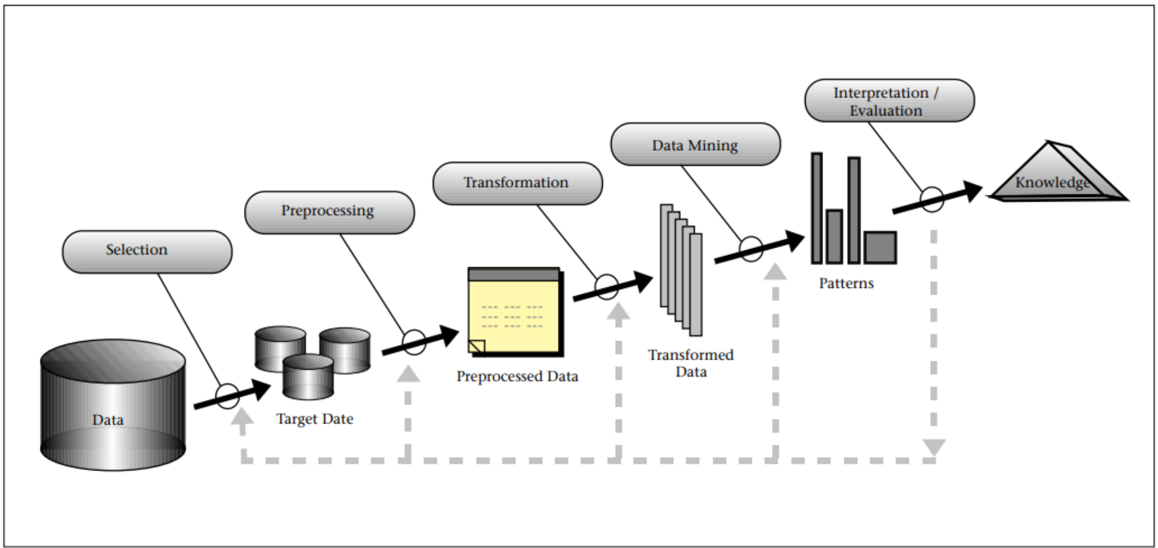


Figure 4 An Overview of the KDD Process

3.1 Domain Understanding

Recognising human activity has been identified to have significant impact across various sectors as identified in previous chapters. One of the motives of this study is to aid decision making using historical information extracted from monitoring physical activities identified by the sensors in the devices worn by patients and their general well-being. Providing the correct information on the activity or activities carried out is critical as it could determine the type of care that needs to be provided should an emergency arise. To achieve this, there is need for high accuracy in the machine learning model employed. Multiple models from those reviewed in Chapter 2 will be utilised and analysed for the best fit.

It is assumed in the process of this study that being able to provide a system with a high prediction accuracy would result in better health decision making.

Several constraints arise and a significant one to note is the computing capabilities. The specifications of the machine used for this project is a 64- bit Intel i7 processor with 16GB of RAM. The coding will be done using Python programming language in Jupyter notebook in PyCharm version 3.8. Python is a robust and versatile programming language that is compatible with different systems. As an academic project with limited computing power, the experiment is limited in its ability to use certain algorithmic configuration which require more power.

3.1.1 Objective

This study centers on the ability of the sensors to provide correct and timely information that would aid decision making by healh professionals and also to improve general well-being in daily living. This requires not just high accuracy but also precision as a wrong information could lead to wrong decision making that may negatively impact lives where the information is especially used by health practitioners. Bearing these in mind, the objective of the experiment in this study is to ascertain the best fit machine learning model that would provide high prediction accuracy in classifying the activities carried out.

In view of this objective, the identified success criteria is in the replicating ability of the experiment.

3.2 Dataset Selection and Understanding

The dataset used in this study also referred to as the MHEALTH (Mobile HEALTH) dataset was retrieved from a previous experiment conducted by (Banos, et al., 2014) recording twelve activities carried out by ten volunteers. Sensors were fitted on the chest, right wrist, and left ankle to record acceleration, rate of turn, and magnetic field orientation. The sensors captured the participants carrying out normal everyday activities without constraints and recorded using a video camera (Dua and Graff 2019). The list of activities is in the Figure below.

1: Standing still (1 min)
2: Sitting and relaxing (1 min)
3: Lying down (1 min)
4: Walking (1 min)
5: Climbing stairs (1 min)
6: Waist bends forward (20 repetitions)
7: Frontal elevation of arms (20 repetitions)
8: Knees bending (crouching) (20 repetitions)
9: Cycling (1 min)
10: Jogging (1 min)
11: Running (1 min)
12: Jump front & back (20 repetitions)

Figure 5 List of activities performed

As seen in the figure below, the signals extracted which formed the features of the dataset are labelled for each signal on the rate of turn for each sensor-accelerometer, magnetometer and gyroscope.

Column 1: acceleration from the chest sensor (X axis)
Column 2: acceleration from the chest sensor (Y axis)
Column 3: acceleration from the chest sensor (Z axis)
Column 4: electrocardiogram signal (lead 1)
Column 5: electrocardiogram signal (lead 2)
Column 6: acceleration from the left-ankle sensor (X axis)
Column 7: acceleration from the left-ankle sensor (Y axis)
Column 8: acceleration from the left-ankle sensor (Z axis)
Column 9: gyro from the left-ankle sensor (X axis)
Column 10: gyro from the left-ankle sensor (Y axis)
Column 11: gyro from the left-ankle sensor (Z axis)
Column 13: magnetometer from the left-ankle sensor (X axis)
Column 13: magnetometer from the left-ankle sensor (Y axis)
Column 14: magnetometer from the left-ankle sensor (Z axis)
Column 15: acceleration from the right-lower-arm sensor (X axis)
Column 16: acceleration from the right-lower-arm sensor (Y axis)
Column 17: acceleration from the right-lower-arm sensor (Z axis)
Column 18: gyro from the right-lower-arm sensor (X axis)
Column 19: gyro from the right-lower-arm sensor (Y axis)
Column 20: gyro from the right-lower-arm sensor (Z axis)
Column 21: magnetometer from the right-lower-arm sensor (X axis)
Column 22: magnetometer from the right-lower-arm sensor (Y axis)
Column 23: magnetometer from the right-lower-arm sensor (Z axis)
Column 24: Label (0 for the null class)

Figure 6 Variables description

The term ‘Subject’ was used to describe the participant carrying out the stated activities and labelled 1 to 10 for each participant.

After collecting data, the next step is to understand it better by exploring it to see its dimensions and confirm its structure is intact and the features of the dataset do not contain any discrepancies and have been extracted in the right format. The easiest way is to review the statistics of the dataset as it summarises the overview of each feature. This provides an easy way to quickly identify any disparity between the feature variables and the target feature.

The next step would be visualising the data as this helps understand its degree of skewness, outliers and relationships. The current research problem has been identified as a Classification problem. The dataset has been identified to have a total of 22 features made up of extracted information from the different signals in a tri-axial form.

3.3 Data Preprocessing and Transformation

This is a crucial stage in data mining. Data integrity is determined at this phase as the insights extracted here could mar data interpretation. This process involves manipulating and preparing the data for subsequent stages in the KDD process. It is expedient that resultant data at this stage is clean; clean here meaning data is void of missing information or noisy. To achieve this data undergoes the tasks in the following subsections.

3.3.1 Cleaning

Data cleaning is primarily carried out on the entire dataset to assess missing information. Missing data could be represented as no entry in the dataset or null or N/A. The occurrence of this in a dataset could cause the wrong insights to be extracted. They are usually investigated to understand the reasons why they occur. The researcher determines at the start of the experiment how missing data would be addressed.

The process also involves checking for outliers and or noise. Outliers are best visualised and details noted. They may be genuine dependent on the nature of data collected or caused by errors in the collection process. Noise in data on the other hand could be caused by wrong labelling or by the features. This noise is filtered out or by choosing a type of model that are not affected by noise. Another action that may be carried out is to check for correlation and in doing so remove redundant data especially if this is a large amount of data. Using correlation coefficient, it is clear to see what features have little or no effect on the target variable.

3.3.2 Transformation

Transformation as the name implies is changing the data form using computations so it can be fit for application. Several ways exist to apply transformation and the suitable method is dependent on the data type and the type of experiment.

A form of transformation applied is to have the data normalised to make it fit into a value range using a scale of the values present in the features to be normalised. The essence of doing this is to ensure similarity in the type of values presented to the algorithm for easy modelling. Another form of transformation is on the features of the dataset. Feature engineering creates new features which are added to the dataset to provide an elaborate understanding and possibly increase model accuracy. This is not without its drawbacks. Encoding the data labels is another form of transforming data. This is common in classification problems and make it easy for the model to read and interpret data input.

3.3.3 Reduction

Reducing data is applicable when computing capability is limited. An important aspect of the technique is to ensure no loss of significant information in the data. The plus side to this is the increase in analytical speed. Techniques applied are dimensionality reduction, principal component analysis, feature selection and sampling.

In sampling, a section of the data is selected for analysis, this can be systemic or random. The other techniques include looking for the best set of features or dimensions to represent the entire dataset without losing critical information.

3.3.4 Splitting

This is the final process in this stage as it prepares the data for the model by creating two sets of the dataset. It separates the data into the training set and testing set. The training set would be fed to the models or algorithms and then evaluated with the test set. A common split is either the 70/30 split ratio or the 80/20 split for the train and test set respectively. It is important that these two sets are kept apart so the model is not exposed to them and result in wrong metrics.

After these stages are completed, the data is now ready to be modelled by the algorithm chosen.

3.4 Data Mining Modeling

Modelling the data requires the use of algorithms to extract gainful insights from the processed and transformed data. The models to be used will be tuned and might have its parameters reviewed for optimal performance. Several models and algorithms have been identified in the literature review that have been used in the recognition of human activities and from that review, the following algorithms will be used for this experiment- K-nearest neighbour, Support Vector Machines, Random Forest and Convolutional Neural Networks.

3.5 Interpretation and Evaluation

The confusion matrix is a common evaluation tool used in classification problems as it presents several metrics in a tabular form and is easy to interpret and read. The model is evaluated for precision, accuracy and recall against the different classes of the target variable.

3.6 Software Development

The software development life cycle consists of five main steps which encompass the process in the development of a software. Each of these steps or phase require a set of actions for implementation. See Figure 7 below.

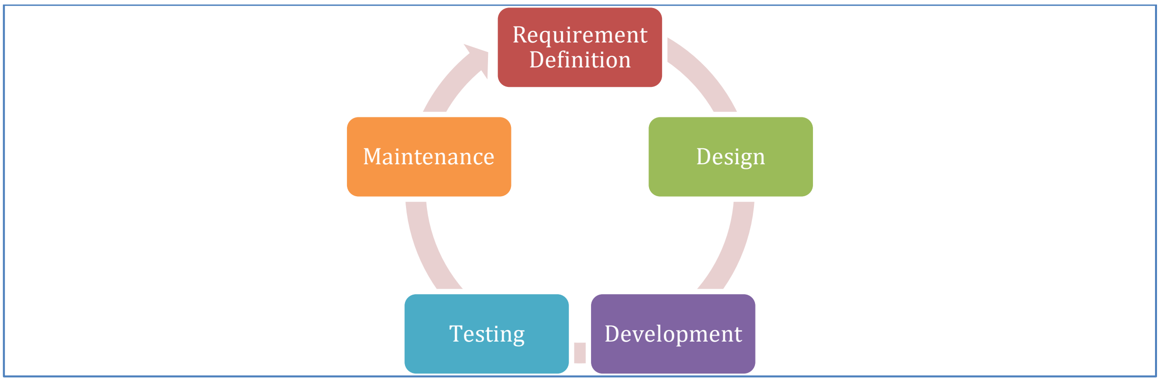


Figure 7 Software Development Life Cycle

The software for this project would follow the outlined steps per the figure. However, there is no set requirement given for a software for this project. Notwithstanding, the software is required to have a display interface that would illustrate the model(s) capability per the data input from the sensors. The design, development and testing phases of the life cycle has been implemented using

Streamlit framework. See Appendix B for screen shots of the codes for the software's visualisation.

However, considering this is not a full deployment in a live situation, the software is a prototype and would be modified as the use case is determined by the end user and based off the results of a User Acceptance Test (UAT). The final phase in the SDLC would be after modifications and live deployment.

The waterfall model approach was used in the development of this software as input to the model was totally reliant on the completion of the selected model for prediction (Mani, et al., 2018).

The ideal suggestion is an executable software for this type of dataset and machine learning problem is the creation of an android app considering that most smartphones incorporate the three sensors from which data was extracted- accelerometer, magnetometer and a gyroscope. However, considering time constraints and computing limitations, an android app could not be created.

3.7Summary

Models were chosen from different classes of machine learning to show robustness and versatility of the models. A detailed overview of the experiment to be carried out has been produced in this chapter. An understanding of the scope and use case of this experiment discussed. This is followed on by the exploration, cleaning and transformation of the dataset in preparation for the modelling. Evaluation metrics then employed to determine which model is the best fit for this classification problem. A major constraint was identified being the computational capability due to the magnitude of the data with computational time for modelling being a significant constraint.

In the next chapter, the implementation of this proposed design and its results would be detailed.

4. Implementation and Results

This chapter aims to answer the research questions by describing how the experiment was executed and implemented. It would do this by working through the sequence of the already described methodology in the previous chapter.

4.1 Data Collection and Understanding

Data was collected into separate files for the individual participants for all sensors; this data was then collated into a single file and saved in a format for ease of analysis and processing. The initial process was to understand how the data was framed and its dimension. Data contained a total of 1,215,745 rows and fit into 23 columns which are labelled with the feature names in Figure 8 below.

```
0:  'acc_chx',
1:  'acc_chy',
2:  'acc_chz',
5:  'acc_lax',
6:  'acc_lay',
7:  'acc_laz',
8:  'gyr_lax',
9:  'gyr_lay',
10:  'gyr_laz',
11:  'mag_lax',
12:  'mag_lay',
13:  'mag_laz',
14:  'acc_rax',
15:  'acc_ray',
16:  'acc_raz',
17:  'gyr_rax',
18:  'gyr_ray',
19:  'gyr_raz',
20:  'mag_rax',
21:  'mag_ray',
22:  'mag_raz',
23:  'activity'
```

Figure 8 Feature names

An understanding of the feature names is explained in the table below, with 'x', 'y' and 'z' being the triaxial readings for each sensor. There is only an accelerometer reading for the chest.

Sensor		Location	
acc	accelerometer	ch	chest
gyr	Gyroscope	la	Left ankle
mag	magnetometer	ra	Right lower arm

Table 2 Explanation of feature names

The statistical description of the dataset was reviewed to gain further insight but due to the size of the dataset not much insight was derived from this. An overview of the dataset was then displayed to give a quick understanding of how the dataset appears in tabular form. The distribution of the data and activity performed across participants was then visualised.

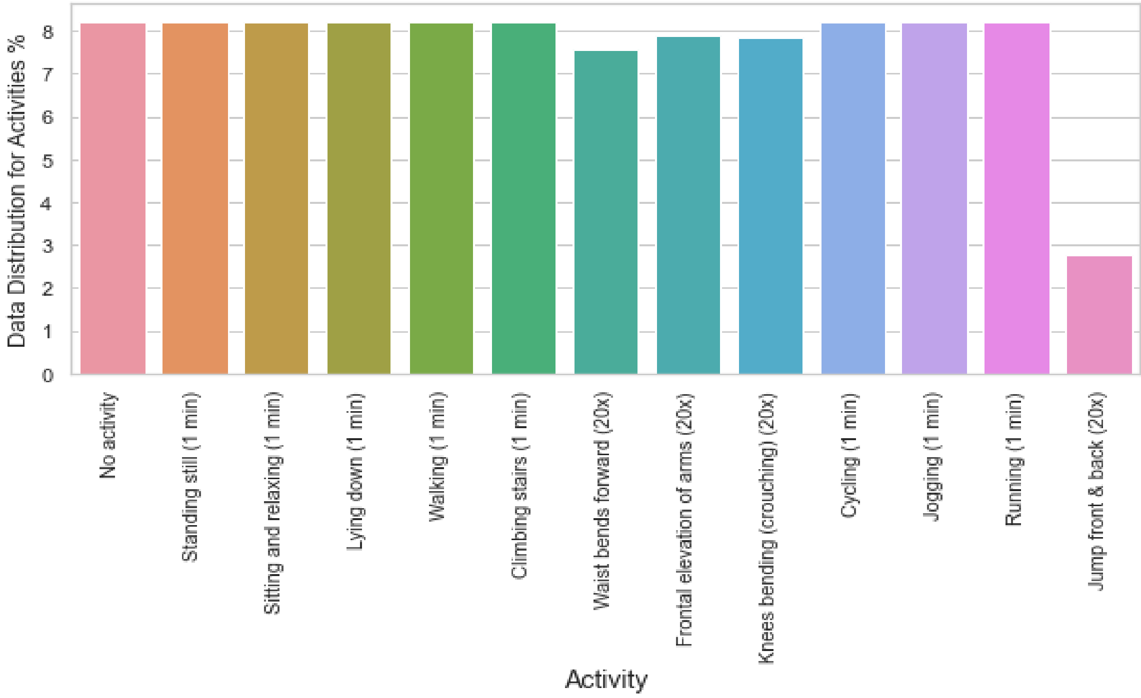


Figure 9 Histogram of activities performed in the experiment

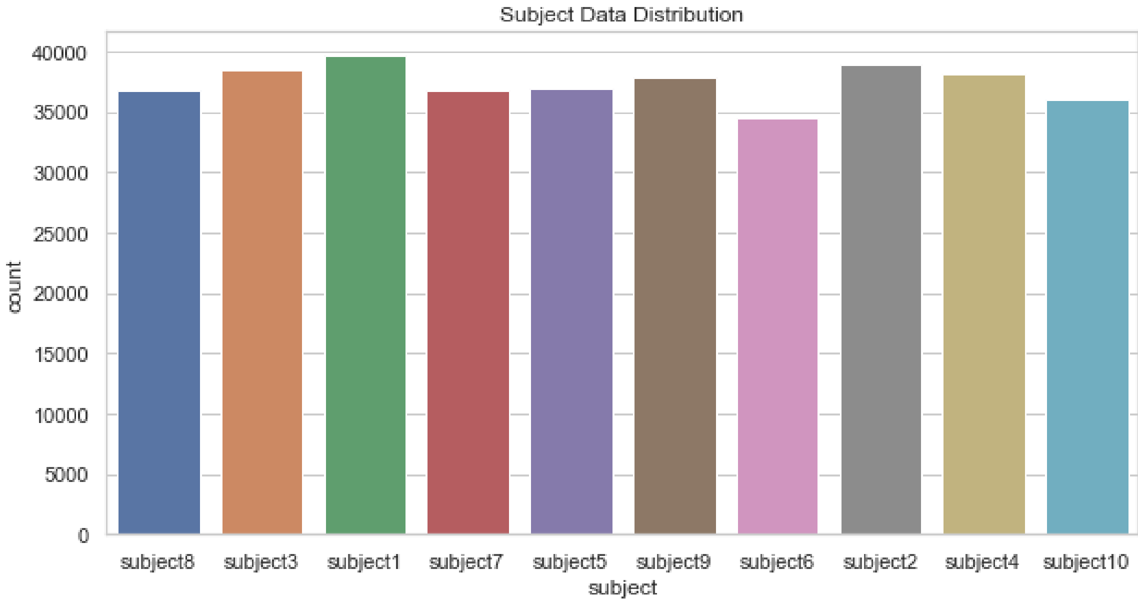


Figure 10 Distribution of activities per participant

Dataset was also checked and there was no missing data. It was also checked for unique value counts for uniformity in analysis. A unique value count of 30,720 was observed per activity with exception of ‘null activity’. Data was resampled for activity ‘0’ to allow for uniformity across all activities.

4.2 Data Processing

There is a total of 12 activities carried out by 10 participants which was recorded separately and compiled into a zip file. The recording of these activities was then integrated into a single csv file as each file contained the same number of dependent and independent variables for the same number of activities. On review and as evident in Figure 8 above, subject 1 carried out the most count of activities. This fact was exploited for visualisation of the dataset.

Datatype is as float for all data extracted from the participants and would require no encoding. Feature names is composed to reflect position of sensor and type of sensor. Due to computational limitations, the dataset was resampled using a value count of 3,000 which is approximately 10 per cent of the total value count. This was checked for loss or any negative impact on the computational analysis of the dataset by viewing the distribution of the entire dataset against the sampled dataset.

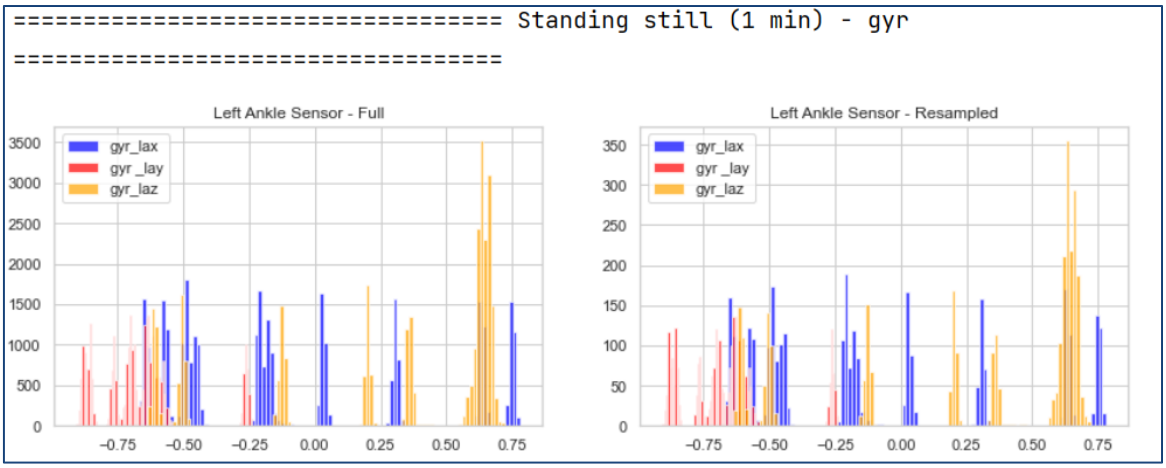


Figure 11 Comparison of distribution of full against sampled dataset

Outliers were observed in the dataset especially in the reading from the magnetometers. To address this, features outside a 98 percent confidence interval level were dropped for all readings for ease of analysis.

The resampling of data was used to make it easier to compute data for visualisation. Correlation was checked across features to check for relationship between features and visualised in both tabular and matrix forms.

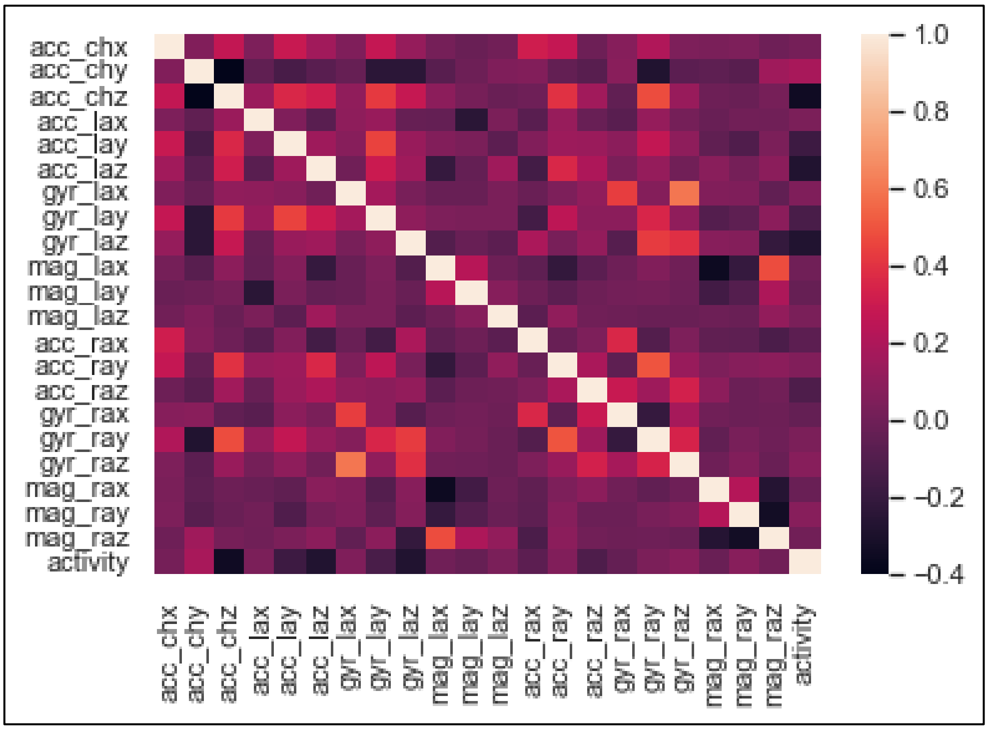


Figure 12 Data Correlation between features

Feature importance was performed to identify which features were more important to the algorithms and models for optimal performance. The data was split into two categories, test and train after it was X variables were scaled using a standard scaler function. The sensor on the chest and right arm appeared to be most relevant to the algorithm.

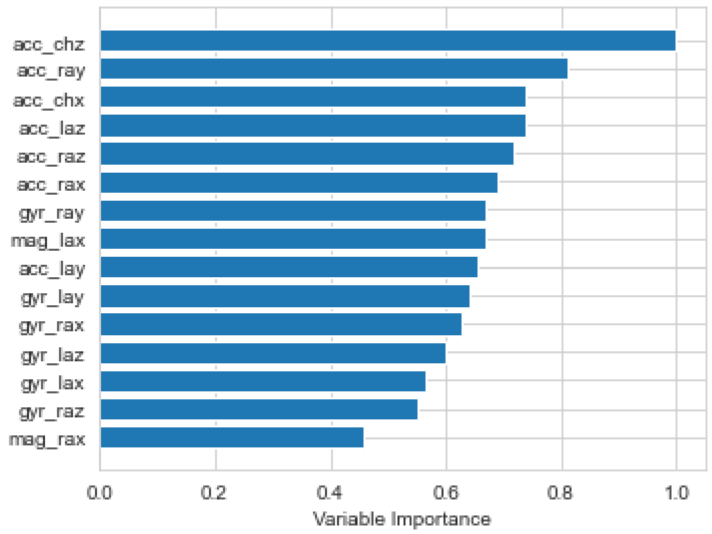


Figure 13 Feature importance

Using a different method to extract important features. A total of 12 features were identified as relevant.

```
features = X.columns[(sel.get_support())]
print(len(features))
features

12

Index(['acc_chx', 'acc_chz', 'acc_lay', 'acc_laz', 'gyr_lax', 'gyr_laz',
       'mag_lax', 'acc_rax', 'acc_ray', 'acc_raz', 'gyr_rax', 'gyr_ray'],
      dtype='object')
```

Figure 14 Identified relevant features

This information was used to create a different dataset to check how it affects model performance.

4.3 Data Modelling

Data was fed into the models after being split into the train and test datasets. The Y variable is fed with the Activity variable and the X variable fed with the rest of the data less the activity and subject variables. The split was done with 75 percent of the dataset to be used to train the model and 25 percent withheld to test the dataset independently. This dataset is the resampled dataset and not the total dataset due to computational limitations earlier discussed.

Each model is fed with this resampled dataset without any modification to the default parameters of the model. This is done to discover how well the models perform as a baseline against modifications to the model's parameters. A cross validation value of 5 was used on the first model in determining the best parameters for the model using the grid search function. Support Vector Machines was the first model to be trained. A lot of time was used in training this model due to cross validation as this requires the model to take 5 different sets of data for training and validation. The best parameters selected for this model is illustrated in Figure 15 below.

```
grid.best_params_

 {'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
```

Figure 15 SVM best parameters

The KNN model was the next to be trained; first with the default settings and then with cross validation to determine the best 'k' value for the model. A 'k' value of 5 was initially used but better result was determined using a value of 3.

```python
# KNN with k=3 (and shuffled data sets)
knn = KNeighborsClassifier(n_neighbors=3)
cv_scores = cross_val_score(knn, X_scaled_shuffled, y_shuffled, cv=5)
# Print cross val scores along with average
print(cv_scores)
print(f'cv_scores mean:{np.mean(cv_scores)}')
```

```
[0.94153846 0.94320513 0.94128205 0.9375641  0.94076923]
cv_scores mean:0.940871794871795
```

Figure 16 KNN values using k-neighbour=3

Random forest classifier is the final traditional machine learning model used in this experiment. A common occurrence across all models is the long training time. Using a random state of 1, default parameters were used to first train this model. With the cross validation being 5, the best parameters for this model were then extracted.

```
{'criterion': 'gini',
 'max_depth': 30,
 'max_features': 'auto',
 'n_estimators': 500}
```

Figure 17 Random Forest best parameters

These models were then stacked together to form an ensemble. The ensemble was used first with base parameters and then with earlier identified best parameters and a feature filtered dataset. The stacked ensemble was further passed through Gradient boost classifier as the final ensemble.

The final model was created to compare with the traditional machine learning models which had been trained. CNN model was then passed through a similar process like the traditional models. Using the same percentage of split for the train and test datasets, a sequence transformation was applied using 25 timesteps. Model was then cross validated to produce better results.

4.4 Model Evaluation

Models are evaluated without the target variable as this is the variable that needs to be predicted by the models. To make this prediction, the predict function is used for each model and the test dataset. For ease of evaluation, a confusion matrix is visualised. This matrix illustrates the actual and predicted values of the dataset that has been fed into the model.

The models are evaluated for accuracy, precision, recall and the f-1 score. Evaluation results are in the tables below for each model configuration. Table 3 shows the only accuracy for each model using base configuration while Table 4 shows the complete evaluation for the hyperparameter tuned models.

Model	Accuracy
SVM	91.95
KNN	91.95
RF	95.84
Stacked	95.79
CNN	91.8

Table 3 Accuracy for models' base configuration

Model	Accuracy	Precision	Recall	F-1 Score
SVM	95.6	95.5	95.73	95.45
KNN	93.76	93.72	93.61	93.39
RF	95.99	96.07	96.12	95.77
Stacked	95.28	95.37	95.35	95.34
CNN	99.93	99.93	99.92	99.92

Table 4 Tuned Models Evaluation

A further evaluation that was performed for the model is the ROC and AUC. This evaluation was performed on the CNN and RF models. A macro average ROC AUC score of 0.99 was achieved for both Random Forest and CNN.

A thorough review of the evaluation of each of these models is discussed in the next chapter as against the research aims and objectives.

5. Discussion

In this section, the experiment is reviewed in depth with the results obtained and illustrated in the previous chapter. Each model, the approach and corresponding result is discussed separately and an overview of the complete design reviewed. The limitations and strengths of this experiment would also be discussed in this chapter.

To give a robust understanding of this chapter, discussion would be structured per the chosen KDD methodology.

5.1 Domain Understanding

To best understand how this experiment should be approached, it is imperative that the researcher understands the impact of this experiment and how it can be applied. This understanding was derived from mainly the literature review and from the possession of sensor-based devices like a smartwatch and a smartphone. An extensive review of previous research showed the extensive use of sensors in both healthcare, security and daily living. This review also highlighted the importance of precision and accuracy in model prediction and helped with the selection of the models for the experiment.

5.2 Dataset Understanding

This process aided with how data should be prepared for processing. It showed the complexity in the structure in terms of features and data size and how these might affect the computational ability of the models being utilised. The data structure showed how data was captured by the sensors and that the target variable was labelled as activity. Also evident was the balance of data across activities; this information is critical as it forms the basis on which samples was taken for training the models. Subject 1 had the highest recording of activities performed across all activities listed and was used to visualise the dataset. However, the sampling of the dataset was based on a resampling method which was set at approximately 10 per cent of activity counts. The resampling was done to ease up processing time and was checked for distribution uniformity per activity and compared with the full dataset for same.

5.3 Data Processing

The variables of the dataset were extracted from the three sensors placed on three different parts of each participant- right arm, chest and left ankle. The features were measured across three different axis- x, y and z for the three sensors- magnetometer, accelerometer and gyroscope.

Feature selection and correlation was an important step in the data processing. The concept is to better understand how the features of the dataset interrelate and may possibly affect modelling. Furthermore, feature selection can also serve as a means of reducing the dimension of the dataset. Correlation analysis showed that slightly over 40 percent of the features had correlation above 50 per cent. The features with low correlation were from the sensors placed on the wrist of the right arm.

Using feature selection to extract a new dataset from the sampled data was performed in the stacked model ensemble. The dataset comprised of 12 features selected by the algorithm as important.

5.4 Modelling

All models used in the experiment were subjected to cross validation. This technique requires each model to trained and tested on different groupings of the dataset, this is performed five times to show results from the model when replicated can be relied on. Random sampling technique was used in creating the test and train datasets.

Default model parameters and settings were used to create the models for the dataset prediction. Support vector machine was the first model used. It was used with its default settings for prediction and then checked for the best parameters which was then evaluated against the test dataset. It can be seen from Figure 18 that the model struggled to differentiate between jogging and running. While SVM is notably a very good algorithm when used with the right parameters, it struggles with a large dataset and where no clear distinction exists in the classes to be predicted.

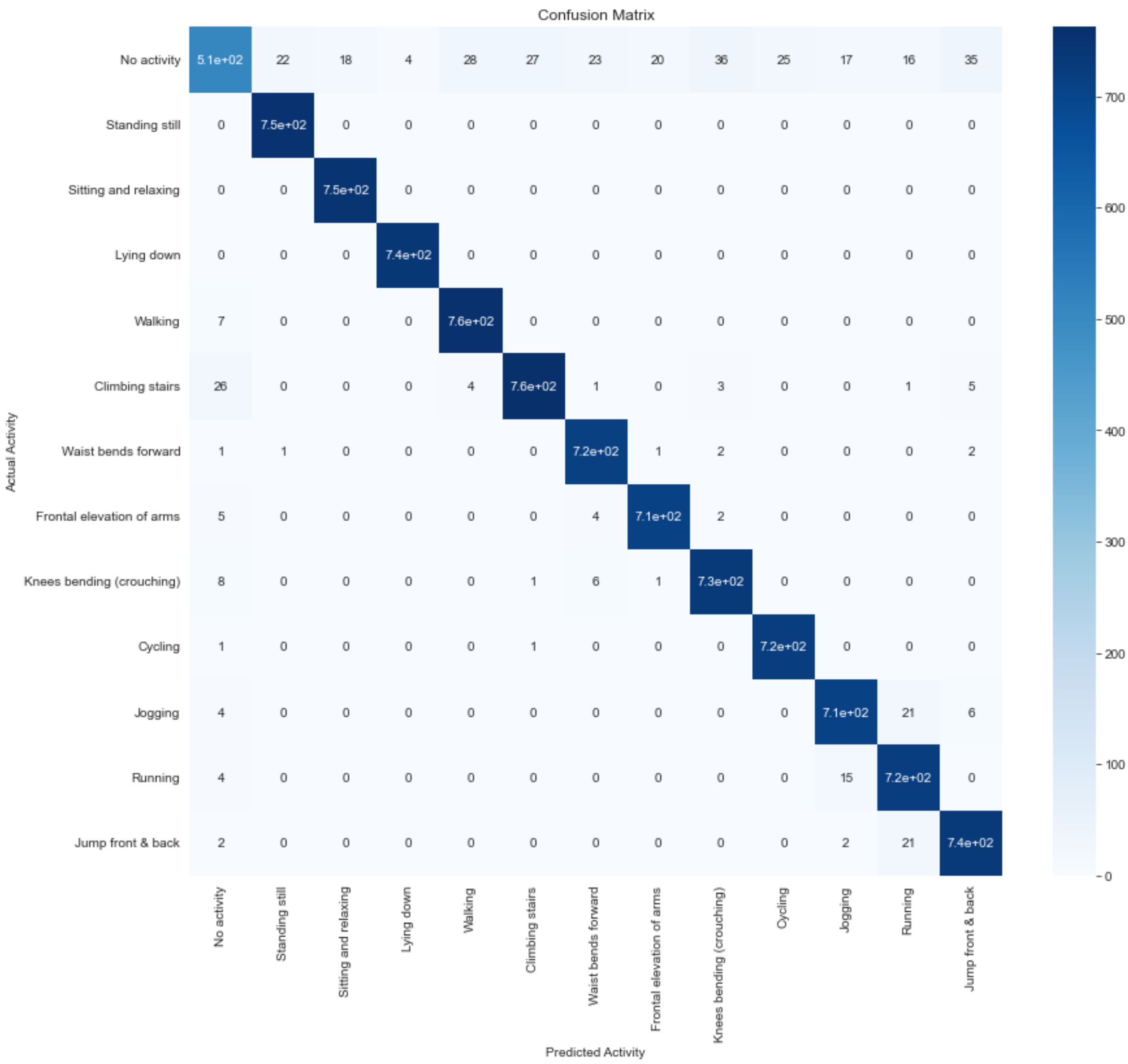


Figure 18 SVM performance- Confusion Matrix

K-nearest neighbors was the next model used in the experiment. From previous studies and research, this model was expected to provide significant results compared with other models. The algorithm was fitted with optimised parameters. Predictions for the different activities in this model showed confusion between climbing the stairs and walking. Figure 19 shows that the model was able to clearly classify stationary activities apart from mobile activities, the number of neighbours used for this algorithm was 3. Similar to the previous model, computational time was long.

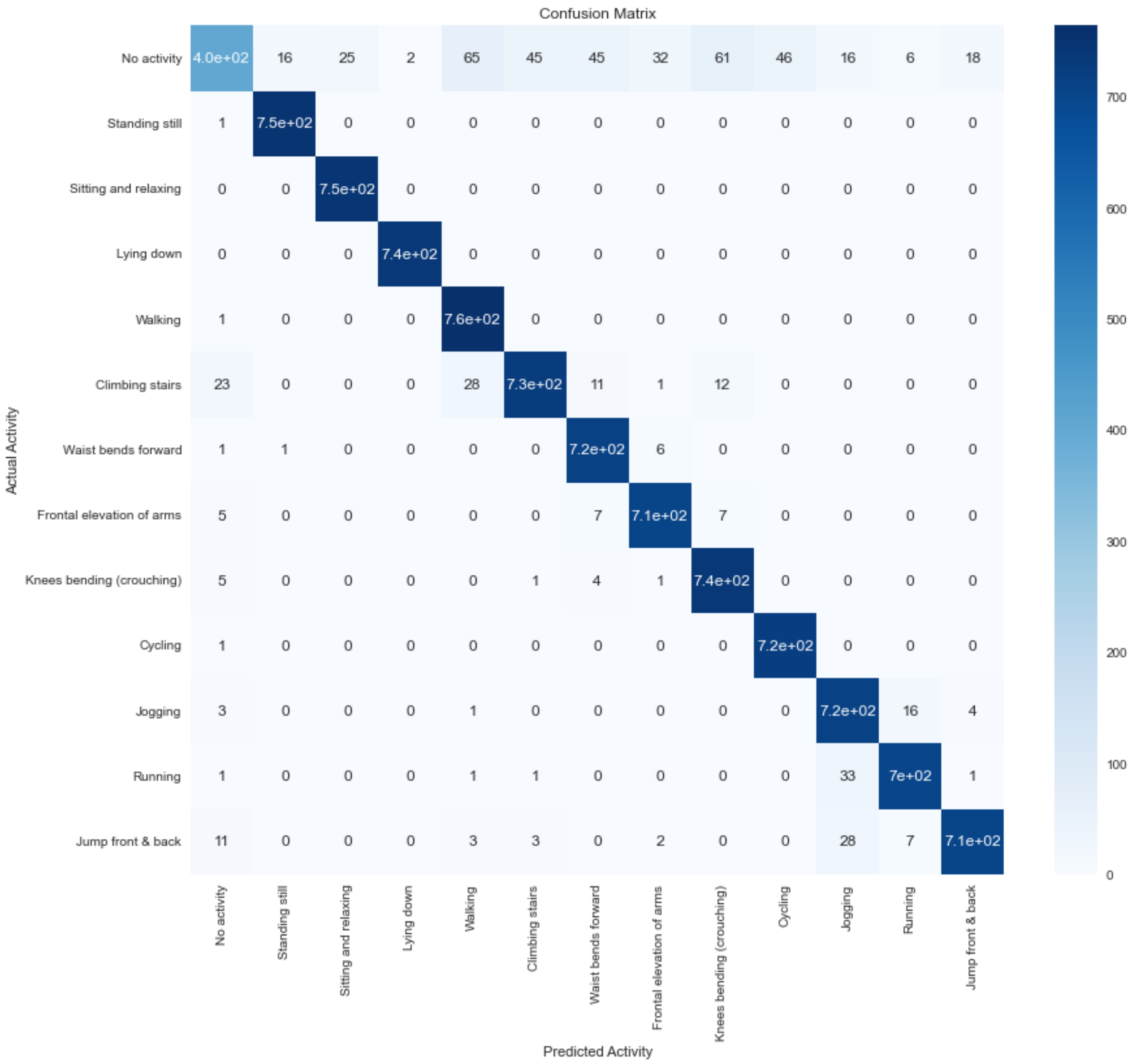


Figure 19 KNN Confusion Matrix

Random forest was the final traditional machine learning model used for the experiment. This model was selected for several reasons, its strength lies in being able to prevent overfitting. Figure 20 shows the results obtained in using this model. the best overall accuracy was observed using this model. this model took the longest to compute but nevertheless, the best parameters were utilised in training and testing this model.

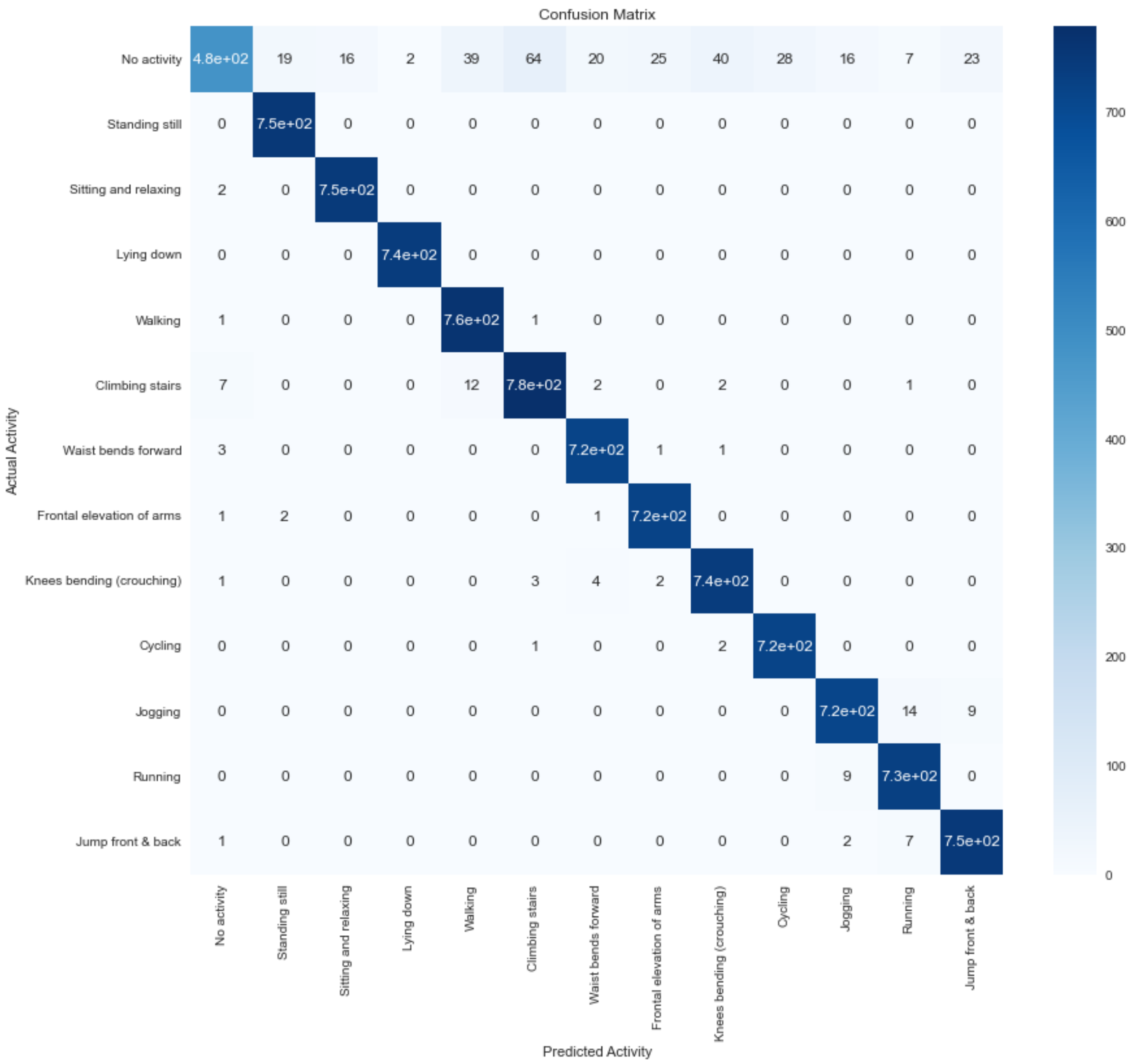


Figure 20 Random Forest Confusion Matrix

The research went further to explore the use of a stacking ensemble using the three models and same optimal parameters. This ensemble's result was similar to that of the random forest model.

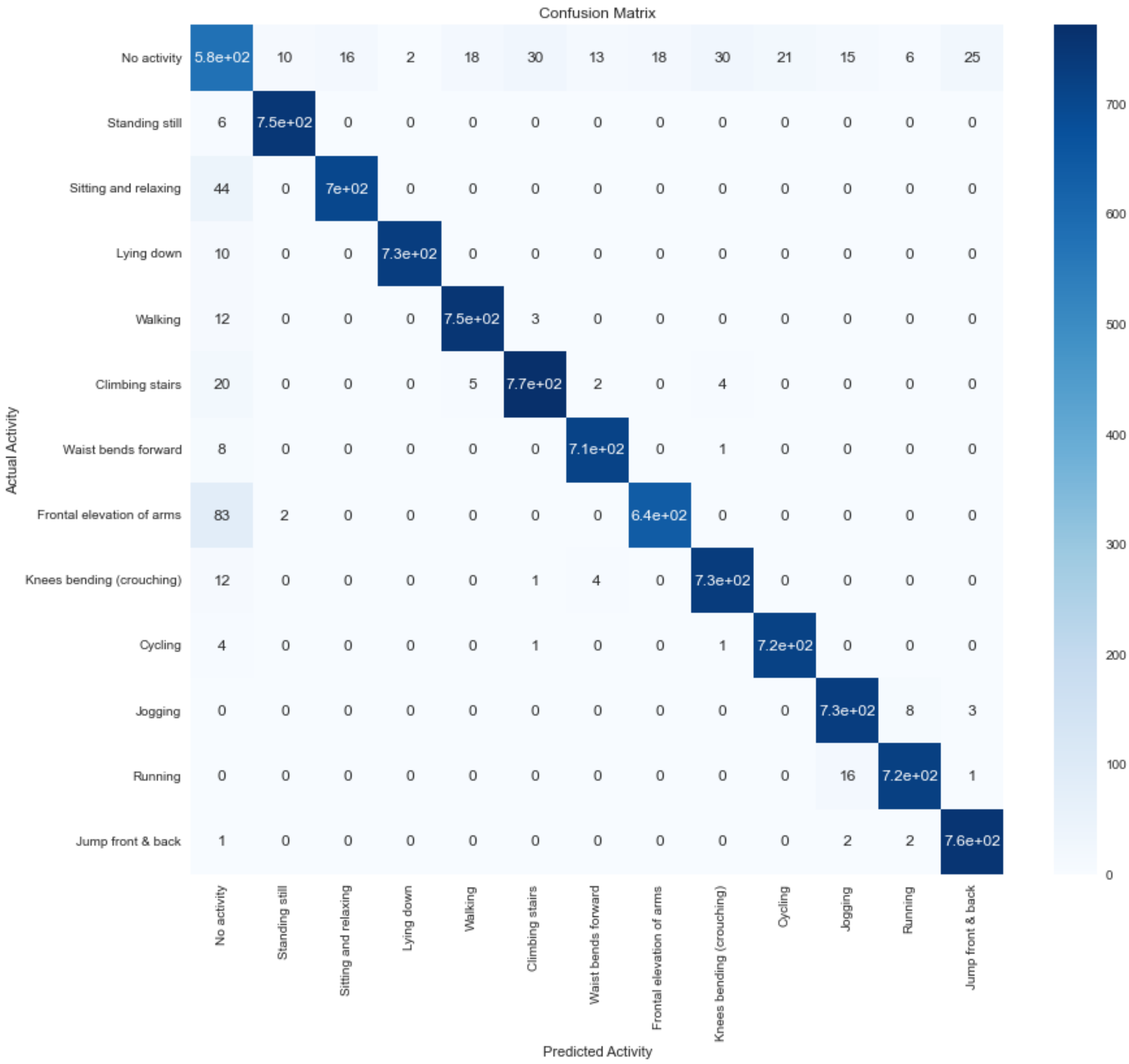


Figure 21 Stacking Ensemble Confusion Matrix

Convolutional neural network was the final algorithm used in this experiment. The one-dimension layer of this algorithm was used. From the literature review in chapter 2, it can be gathered that the best results were obtained using this algorithm in the recognition and prediction of human activities. It was evident that having the dataset shuffled after the creation of samples increased the performance of the model this is likely due to the fact that data shuffling alters the data sequence. There were fewer instances of misclassification using this algorithm.

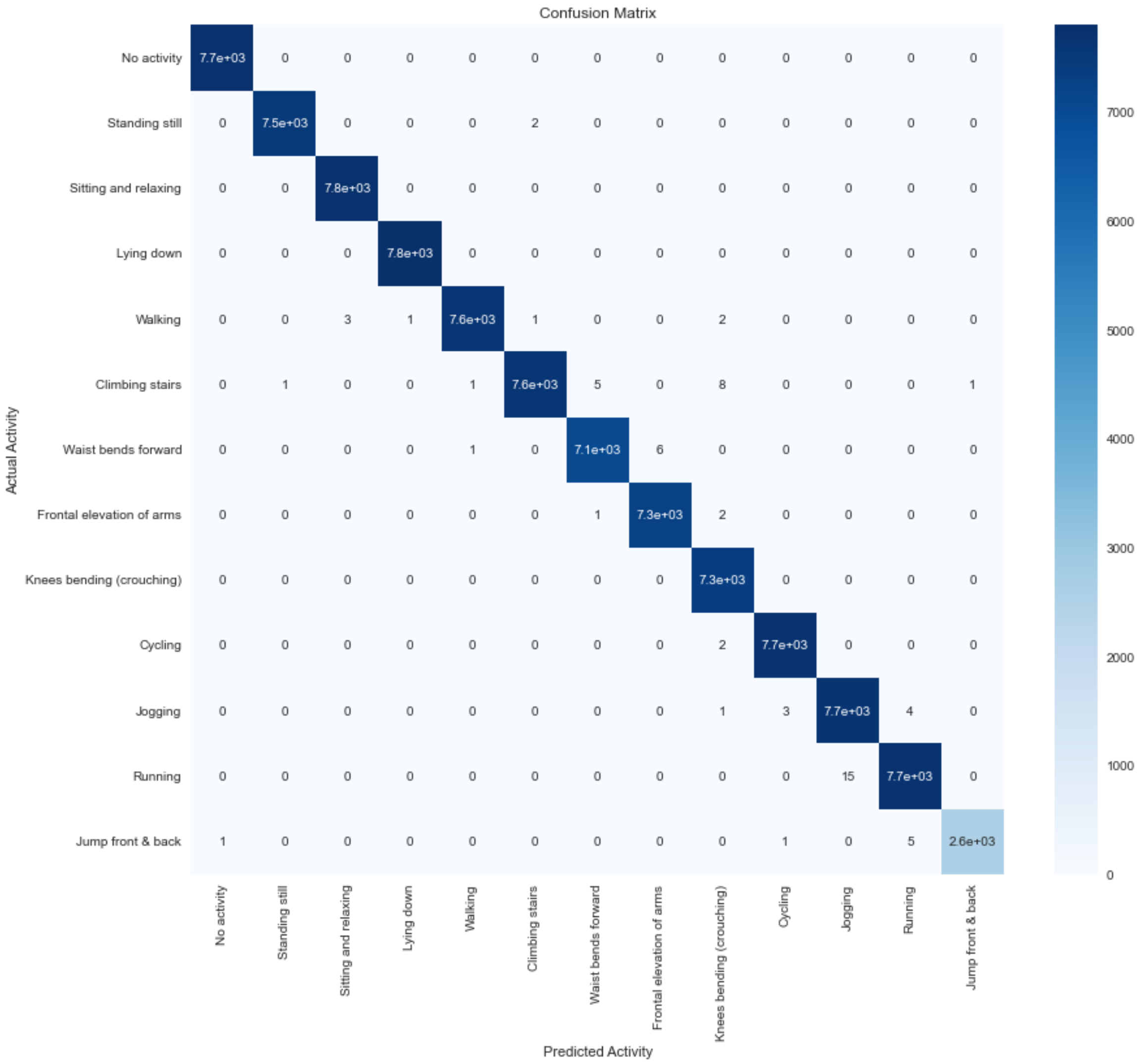


Figure 22 CNN Confusion Matrix

An observation among the traditional machine learning models was the similarity in prediction accuracy. This goes to show how well suited these models are in handling this type of classification problem and with cross validation integrated the results can be replicated.

A significant limitation in this study as had been earlier mentioned is the computational time for the training and testing of each model. another limitation was in the classification of activities. Differentiating between activities such as jogging and running appeared to be a problem for the models.

In summary, this chapter has reviewed the experiments carried out and the corresponding results. Checking that the results were able to answer the research question of how well machine learning models are able to predict human activities in comparison with CNN and also identify its strengths and limitations.

Furthermore, while CNN was able to exceed the traditional machine learning models used- KNN, SVM and random forest, the performance of each model was analysed along with its results to understand how well it performed compared to previous research.

It was observed that there was no significant impact in using feature selection as a method of dimensionality reduction in training the model. it however provided clearer understanding on what features were significant to the model and possible significance of sensor position for optimal performance. The evaluation methods chosen for the models also helped in better interpreting the models results.

In the next chapter, conclusions will be drawn on the findings of this study and suggests area of improvement and further possible research.

6. Conclusions

This chapter concludes the study and restates the research question and how it was addressed by walking through the research objectives and findings. Possible further research and recommendations are made at the end of the chapter.

Primary aim of this research was to be able to recognise human activity with the aid of machine learning algorithms using data extracted from sensors. The concept of this domain and an understanding of the technology and procedures used in recognising these activities was undertaken.

Following the collection of data to be used in this study, the next area of concentration was to be able to process the data to make it suitable for modelling. Several techniques were used with correlation done to understand the relationship between the features and their dependency on one another. Another analysis carried out was the feature selection method. This method identifies the most relevant features that would make an impact on how the selected algorithm would perform. These methods are techniques that could be employed to the study and help reduce the features in the dataset.

The goal of the study was to be able to identify the best model or algorithm to be used in identifying human physical activities and achieve this with high accuracy.

As identified from previous research, several models have been used with different techniques. Some of them worked with the default parameters of the algorithms and some with an ensemble. A few used deep learning techniques and most of them used traditional machine learning models. However, this research's focus was to determine the best model using some of the well-known traditional machine learning models based on previous research and compare this to a deep learning model. The question this research is trying to answer is –

> *How efficient are supervised machine learning algorithms in predicting and recognising human activity when compared to neural network, CNN in particular?*

*Algorithms: Support Vector Machine, K-Nearest Neighbors, Random Forest and Convolutional Neural Network.

6.1 Research Approach

To try answer the research question the following processes/stages were followed-

- A review of current and past research on Activity Recognition through a literature review. Identify commonly used models and their performance
- Design a solution to address the research question. This being find the best model determined by the set evaluation metrics
- Implement the designed solution and tweak as required to obtain optimal performance
- Evaluate models using metrics for classification problems
- Make recommendations and identify areas of future research

Using knowledge obtained through the literature review, the solution designed integrated the use of feature reduction by selecting the key features but noticed no significant improvement was obtained from this using the stacked ensemble made up of three machine learning algorithms – random forest, support vector machine and the K nearest neighbor. The models were primarily evaluated using accuracy in classification as it gives a quick understanding of the model's performance. The CNN model showed the best performance in all measured metrics with 99.928% in accuracy after tuning the parameters of the model. The next algorithm to give a good performance is the Random Forest model with a performance of 95.99% followed by the SVM model with an accuracy of 95.6%. The worst performing model was the KNN with an accuracy of 93.76%.

6.2 Research Contributions and Findings

Several contributions from this research are-

- An understanding of Activity Recognition in Humans through an in-depth literature review
- A statistical and quantitative understanding of the dataset that could serve as a baseline for future studies
- Identified CNN as the best model for use in similar cases of activity recognition
- Demonstrated that traditional machine learning models can be used to in human activity recognition as well as deep learning models and produce significant results in prediction
- Demonstrated that there was no significant improvement in predictive abilities in a stacked ensemble using feature selection as a dimension reduction technique

6.3 Recommendations and Future Research Areas

There are no limits on how the direction of the study could have been approached and data manipulated to obtain different results or insights. The following recommendations are made-

I. GridSearchCV was used in obtaining best parameters for the models with a cross validation technique of 5. However, other methods exist that could have been used to obtain the best parameters and validation process.

II. Other dimension reduction techniques apart from feature selection could have been used such as Principal Component Analysis (PCA) and Exploratory Factor Analysis could have been tested.

III. Another approach to be considered is to run the test by individual subject and not collate results into single file as in practical use case sensor is per user.

Models can be trained and tested for each user in loops to compare how well results are replicated.

7. Reference list / Bibliography

Alessia, S., Antonio, C. & Aldo, Q., 2017. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. *Frontiers in Aging Neuroscience,* Volume 9.

ANGUITA, D., *et al.*, 2013. *Training Computationally Efficient Smartphone-Based Human Activity Recognition Models.* Springer Berlin Heidelberg, pp.426

Ao, Y. et al., 2019. The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. *Journal of Petroleum Science and Engineering,* Volume 174, pp. 776-789.

ARIF, M. *et al.,* 2014. Better Physical Activity Classification using Smartphone Acceleration Sensor. *Journal of medical systems,* 38(9), 95

ATTAL, F. *et al.,* 2015. Physical Human Activity Recognition Using Wearable Sensors. *Sensors,* 15(12), 31314-31338

AZEVEDO, ANA ISABEL ROJÃO LOURENÇO and M.F. SANTOS, 2008. *KDD, SEMMA and CRISP-DM: a parallel overview.* Available from: https://explore.openaire.eu/search/publication?articleId=od______2595::e3a84b91c9fc5298ec939f0d80353d81

Banos, O., Garcia, R. & Saez, A., 2014. *UCI Machine Learning Repository: Center for Machine Learning and Intelligent Systems.* [Online] Available at: https://archive.ics.uci.edu/ml/datasets/MHEALTH+Dataset

BAYAT, A., M. POMPLUN and D.A. TRAN, 2014. A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. *Procedia Computer Science,* 34, 450-457

BONATO, P., 2003. Wearable sensors/systems and their impact on biomedical engineering. *IEEE engineering in medicine and biology magazine,* 22(3), 18-20

BONOMI, A. *et al.,* 2009. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine and science in sports and exercise,* 41(9), 1770-7

BROWN, S., 2021. *Machine learning, explained* [viewed Aug 14, 2022]. Available from: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

BULLING, A., U. BLANKE and B. SCHIELE, 2014. A tutorial on human activity recognition using body-worn inertial sensors. *ACM computing surveys*, 46(3), 1-33

CASALE, P., O. PUJOL and P. RADEVA, 2011. *Human Activity Recognition from Accelerometer Data Using a Wearable Device*. Springer Berlin Heidelberg

Cheng, L., Guan, Y., Zhu, K. & Li, Y., 2017. *Recognition of human activities using machine learning methods with wearable sensors.* s.l., s.n., pp. 1-7.

Chowdhury, A., Tjondronegoro, D., Chandran, V. & Trost, S., 2018. Physical activity recognition using posterior-adapted class-based fusion of multi accelerometers. *IEEE Journal of Biomedical and Health Informatics,* p. 99.

DUA, D. and GRAFF, C., 2019. *MHEALTH Dataset Data Set* [viewed June 18, 2022]. Available from: https://archive.ics.uci.edu/ml/datasets/MHEALTH+Dataset#

EMC Education Services, 2015. *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.* Indianapolis: John Wiley & Sons, Inc.

Fan, J., Han, F. & Liu, H., 2014. Challenges of Big Data analysis. *National Science Review,* 1(2), pp. 293-314.

FAYYAD, U. and R. UTHURUSAMY, 1996. Data mining and knowledge discovery in databases. *Communications of the ACM,* Nov 01, 24-26

Ghazal, S. et al., 2019. Human activity recognition using 2D skeleton data and supervised machine learning. *IET Image Processing,* Volume 13, p. 2572–2578.

Ghorpade, M., Chen, H., Liu, Y. & Jiang, Z., 2020. *SMART: Emerging Activity Recognition with Limited Data for Multi-modal Wearable Sensing.* s.l., s.n., pp. 1316-1321.

Guo, G. et al., 2003. *KNN Model-Based Approach in Classification.* Berlin, Springer, pp. 986-996.

GUPTA, P. and T. DALLAS, 2014. Feature Selection and Activity Recognition System Using a Single Triaxial Accelerometer. *IEEE transactions on biomedical engineering,* 61(6), 1780-1786

IGNATOV, A., 2018. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied soft computing,* 62, 915-922

Joosen, P. et al., 2019. A smartphone-based solution to monitor daily physical activity in a care home. *Journal of Telemedicine and Telecare,* 25(10), pp. 611-622.

Khan, I. U., Afzal, S. & Lee, J. W., 2022. Human Activity Recognition via Hybrid Deep Learning Based Model. *Sensors (Basel),* p. 323.

KIRANYAZ, S. *et al.,* 2021. 1D convolutional neural networks and applications: A survey. *Mechanical systems and signal processing,* 151, 107398

LARA, O.D. and M.A. LABRADOR, 2013. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications surveys and tutorials,* 15(3), 1192-1209

MALONE, D., 1993. Expert Systems, Artificial Intelligence, and Accounting. *Journal of education for business,* 68(4), 222-226

MALONE, T., W., RUS, D. and LAUBACHER, R., 2020. *Artificial Intelligence and Accounting The Future of Work.* Available from: https://workofthefuture.mit.edu/wp-content/uploads/2020/12/2020-Research-Brief-Malone-Rus-Laubacher2.pdf

Mani, Kumar, Y. & Mohammed, S., 2018. A REVIEW PAPER ON SOFTWARE DEVELOPMENT LIFECYCLE MODELS. *SSRN Electronic Journal,* 5(2), pp. 107-113.

Martinez, J., Black, M. J. & Romero, J., 2017. *On human motion prediction using recurrent neural networks.* s.l., s.n., pp. 4674-4683.

Marx, V., 2013. The big challenges of big data. *Nature,* Volume 498, pp. 255-260.

MINARNO, A.E., W.A. KUSUMA and H. WIBOWO, *Performance_Comparisson_Activity_Recognition_using_Logistic_Regression_and_Support_Vector_Machine.*

MOUSTAFA, K., S. LUZ and L. LONGO, 2017. *Assessment of Mental Workload: A Comparison of Machine Learning Methods and Subjective Assessment Techniques*. Springer International Publishing, pp.30

Nguyen, L. T., Zeng, M., Tague, P. & Zhang, J., 2015. *Recognizing new Activities with Limited Training Data*. Osaka, s.n., pp. 67-74.

PHAN, T., Sep 13, 2014. Improving activity recognition via automatic decision tree pruning. ACM, pp.827-832

PIATETSKY, G., 2014. *CRISP-DM, still the top methodology for analytics, data mining, or data science projects* [viewed Aug 14, 2022]. Available from: https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html

POWELL, H.C., M.A. HANSON and J. LACH, 2009. On-Body Inertial Sensing and Signal Processing for Clinical Assessment of Tremor. *IEEE transactions on biomedical circuits and systems,* 3(2), 108-116

Rad, N. M. & Marchiori, E., 2021. Chapter 9 - Machine learning for healthcare using wearable sensors. *Digital Health: Exploring Use and Integration of Wearables*, pp. 137-149.

RAWAT, W. and Z. WANG, 2017. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural computation,* 29(9), 2352-2449

Saboor, A. et al., 2020. Latest Research Trends in Gait Analysis Using Wearable Sensors and Machine Learning: A Systematic Review. *IEEE Access*.

SAHA, A. *et al.,* Jul 2020. Human Action Recognition Using Smartphone Sensors. IEEE, pp.238-243

San , P. P. et al., 2017. Deep learning for human activity recognition. In: C. Hsu, C. Chang & H. Hsu, eds. *Big Data Analytics for Sensor-Network Collected Intelligence*. s.l.:Elsevier Science, pp. 186-204.

SHALEV-SHWARTZ, S. and S. BEN-DAVID, 2014. *Understanding Machine Learning*. Cambridge University Press

SHINDE, P.P. and S. SHAH, Aug 2018. A Review of Machine Learning and Deep Learning Applications. IEEE, pp.1-6

Subrahmanya, S. V. G. et al., 2021. The role of data science in healthcare advancements: applications, benefits, and future prospects. *Irish Journal of Medical Science*.

VELLAMPALLI, H., 2017. *Physical Human Activity Recognition Using Machine Learning Algorithms*. Dublin Institute of Technology Available from: https://explore.openaire.eu/search/publication?articleId=od______1248::ebfcef387bd0e17a5d195a9608b17b69

WEE-SOON *et al.*, 2008. Ambulatory_monitoring_of_human_posture_and_walking_speed_using_wearable_accelerometer_sensors. *30th Annual International IEEE EMBS Conference * Vancouver, British Columbia, Canada:

XU, H., *et al.*, 2016. *Wearable Sensor-Based Human Activity Recognition Method with Multi-Features Extracted from Hilbert-Huang Transform*. MDPI AG

YURUR, O., W. and C.H. LIU, 2014. *A_survey_of_context-aware_middleware_designs_for_human_activity_recognition*.

8. Appendices

8.1 Appendix A: Ethics Approval

Ethical clearance for research and innovation projects

Get Help

Project status

Status

Approved

Actions

Date	Who	Action	Comments
18:04:00 05 July 2022	Femi Isiaq	Supervisor approved	
17:56:00 05 July 2022	Souvenir Oyawale	Principal investigator submitted	

8.2 Appendix B: Software

```
import ...

#Load model
#model= load_model('model.h5')

#configuration of the page
st.title("HAR Web App")
st.sidebar.title("Classification Web App")
st.markdown("Can ML predict activity?")
st.sidebar.markdown("This app allows you to test different machine learning algorithms to classify activity")

st.set_option('deprecation.showPyplotGlobalUse', False)

#Loading the data


data = pd.read_csv('mhealth_complete.csv')

data_activity_0 = data[data['activity'] == 0]
data_activity_else = data[data['activity'] != 0]
data_activity_0 = data_activity_0.sample(n=30720, random_state=1)
data = pd.concat([data_activity_0, data_activity_else])

if st.sidebar.checkbox("Show raw data", False):
    st.subheader("HAR Data Set (Classification)")
```

soft

8.3 Appendix C: Data Collection

```python
import pandas as pd
```

```python
# data sourced from UCI Machine Learning Repository:  http://archive.ics.uci
 .edu/ml/machine-learning-databases/00319/MHEALTHDATASET.zip

combined_df = pd.DataFrame()
```

```python
# Create loop to combine observation of all subjects and create new headers

for i in range(1, 11):
    df = pd.read_csv(f'Dataset/mHealth_subject{i}.log', header=None, sep='\t')
    df = df.loc[:, [0, 1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,
     23]].rename(columns= {
        0: 'acc_chx',
        1: 'acc_chy',
        2: 'acc_chz',
        5: 'acc_lax',
        6: 'acc_lay',
        7: 'acc_laz',
        8: 'gyr_lax',
        9: 'gyr_lay',
```

8.4 Appendix D: Data Exploration and Analysis

```python
# Import relevant libraries
import ...
```

```python
# read csv for analysis and visualisation
data = pd.read_csv('mhealth_complete.csv')
```

```python
data.shape
```

```
(1215745, 23)
```

```python
data.head()
```

	gyr_rax	gyr_ray	gyr_raz	mag_rax	mag_ray	mag_raz	activity	subject
›0	-0.44902	-1.0103	0.034483	-2.35000	-1.610200	-0.030899	0	subject1
›5	-0.44902	-1.0103	0.034483	-2.16320	-0.882540	0.326570	0	subject1
›0	-0.44902	-1.0103	0.034483	-1.61750	-0.165620	-0.030693	0	subject1
›0	-0.45686	-1.0082	0.025862	-1.07710	0.006945	-0.382620	0	subject1
›0	-0.45686	-1.0082	0.025862	-0.53684	0.175900	-1.095500	0	subject1

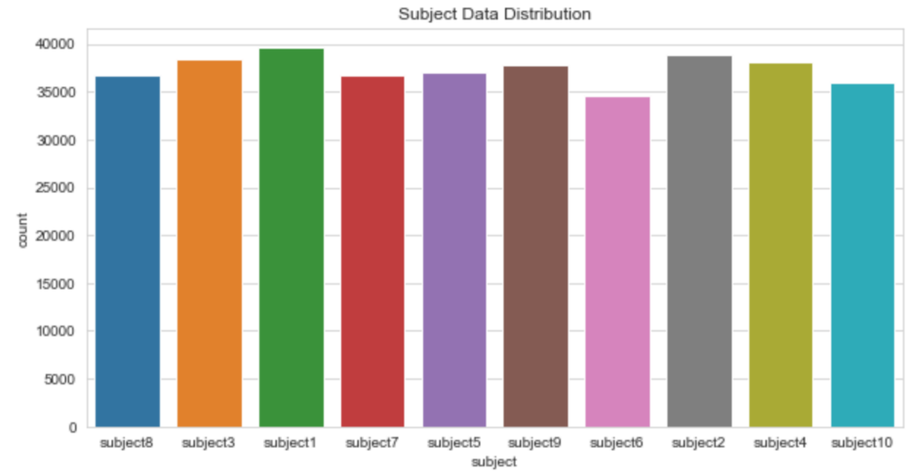
5 rows × 23 columns Open in new tab

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1215745 entries, 0 to 1215744
Data columns (total 23 columns):
 #   Column    Non-Null Count    Dtype
---  ------    --------------    -----
 0   acc_chx   1215745 non-null  float64
 1   acc_chy   1215745 non-null  float64
 2   acc_chz   1215745 non-null  float64
 3   acc_lax   1215745 non-null  float64
```

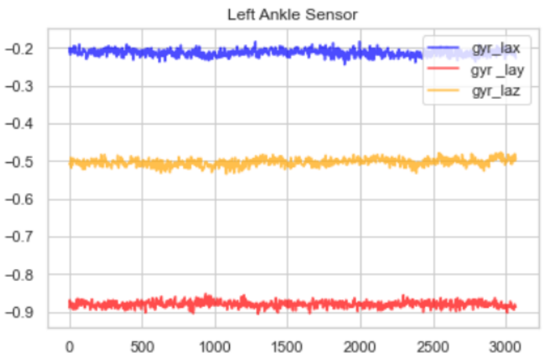
```python
# check for missing and duplicate values

print("The number of missing values in Data:" , data.isnull().values.sum())
print("The number of duplicate values in Data:" , data.duplicated().sum())

```

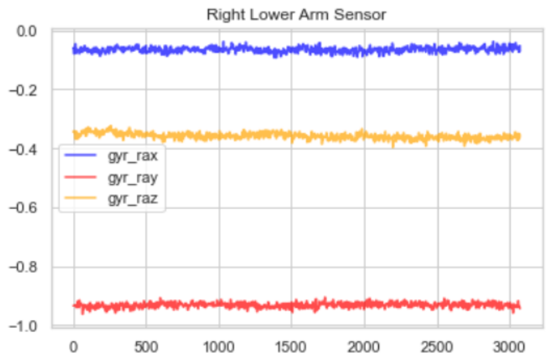
```
The number of missing values in Data: 0
The number of duplicate values in Data: 0
```

```
plt.title('Subject Data Distribution')
sns.countplot(x='subject' , data=data )
plt.show()
```

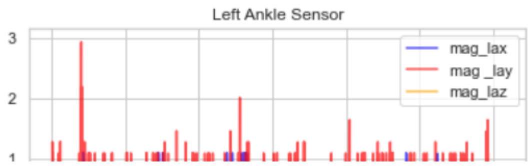


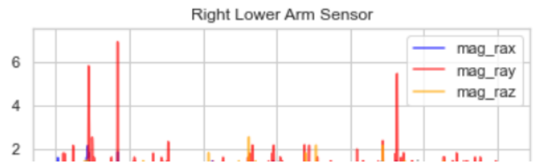

```
=================================== Standing still (1 min) - gyr
===================================
```






```
=================================== Standing still (1 min) - mag
===================================
```





```
data.describe().T
```

	count	mean	std	min	25%	50%	75%	ma:
acc_chx	373915.0	-7.601898	5.550768	-22.4380	-9.717500	-8.89570	-5.636150	
acc_chy	373915.0	-0.150189	2.732111	-20.1880	-1.292300	-0.30029	0.946115	
acc_chz	373915.0	-0.954781	4.506699	-18.4010	-3.583100	-0.88401	1.047700	
acc_lax	373915.0	1.769769	4.172731	-22.1460	0.158155	1.36360	2.882850	
acc_lay	373915.0	-9.131362	5.097113	-19.6190	-10.080000	-9.60720	-7.739550	
acc_laz	373915.0	-0.730620	6.348790	-19.3730	-3.370000	0.29697	1.756600