Comparative Analysis of Machine Learning Algorithms on Activity Recognition from Wearable Sensors' MHEALTH dataset Supported with a Comprehensive Process and Development of an Analysis Tool

COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS ON ACTIVITY RECOGNITION FROM WEARABLE SENSORS' MHEALTH DATASET SUPPORTED WITH A COMPREHENSIVE PROCESS AND DEVELOPMENT OF AN ANALYSIS TOOL

N. SHERAZ

MPhil

UNIVERSITY OF BRADFORD

2019

Comparative Analysis of Machine Learning Algorithms on Activity Recognition from Wearable Sensors' MHEALTH dataset Supported with a Comprehensive Process and Development of an Analysis Tool

Nasir SHERAZ

Submitted for the Degree of

Master of Philosophy

Faculty of Engineering and Informatics

University of Bradford

2019

Abstract

Nasir Sheraz

Comparative Analysis of Machine Learning Algorithms on Activity Recognition from Wearable Sensors' MHEALTH dataset Supported with a Comprehensive Process and Development of an Analysis Tool

Keywords: Machine Learning, R-Based, MHEALTH Dataset, Data Analysis, Activity Recognition, Algorithms, Wearable Sensors, Wearable Devices

Human activity recognition based on wearable sensors' data is quite an attractive subject due to its wide application in the fields of healthcare, wellbeing and smart environments. This research is also focussed on predictive performance comparison of machine learning algorithms for activity recognition from wearable sensors' (MHEALTH) data while employing a comprehensive process. The framework is adapted from well-laid data science practices which addressed the data analyses requirements quite successfully. Moreover, an Analysis Tool is also developed to support this work and to make it repeatable for further work.

A detailed comparative analysis is presented for five multi-class classifier algorithms on MHEALTH dataset namely, Linear Discriminant Analysis (LDA), Classification and Regression Trees (CART), Support Vector Machines (SVM), K-Nearest Neighbours (KNN) and Random Forests (RF). Beside using original MHEALTH data as input, reduced dimensionality subsets and reduced features subsets were also analysed. The comparison is made on overall accuracies, class-wise sensitivity and specificity of each algorithm, class-wise detection rate and detection prevalence in comparison to prevalence of each class, positive and negative predictive values etc. The resultant statistics have also been compared through visualizations for ease of understanding and inference.

All five ML algorithms were applied for classification using the three sets of input data. Out of all five, three performed exceptionally well (SVM, KNN, RF) where RF was best with an overall accuracy of 99.9%. Although CART did not perform

well as a classification algorithm, however, using it for ranking inputs was a better way of feature selection. The significant sensors using CART ranking were found to be accelerometers and gyroscopes; also confirmed through application of predictive ML algorithms. In dimensionality reduction, the subset data based on CART-selected features yielded better classification than the subset obtained from PCA technique.

Acknowledgements

I would like to thank my Principal Project Supervisor Professor Rami Qahwaji for his continuous guidance and encouragement throughout the project. Throughout the research phase he had been very kind, patient, accommodative and very helpful. His valuable feedback on my work kept the project going in right direction and at a right pace. Besides, I am also grateful to my associate supervisor Dr Mumtaz Kamala for his active input during the process.

In addition, I would like to thank my wife and whole family for their constant support and patience throughout the phases of this research.

Table of Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
AR	Activity Recognition
CART	Classification and Regression Trees
ECG	Electrocardiogram
GPS	Global Positioning System
GUI	Graphical User Interface
HAR	Human Activity Recognition
HMM	Hidden Markov Model
IMU	Inertial Measuring Unit
KNN	K Nearest Neighbour
LDA	Linear Discriminator Analysis
ML	Machine Learning
NPV	Negative Predicted Value
PC	Principal Component
PCA	Principal Component Analysis
PPV	Positive Predicted Value
RAD	Rapid Application Development
RF	Random Forest
SVM	Support Vector Machine
WBAN	Wireless Body Area Network

Chapter 1
Introduction

1.1 Wearable Sensors Technology

The advancements in computing technologies combined with reduced form-factor defined by the Moore's law have revolutionized portable, handheld and wearable devices. As the wearable gadgets are becoming increasingly common, cheaper and commercially accessible; such technologies are being widely adopted in daily usage. What used to be purpose-built sensors equipment with bulky, wired paraphernalia has now turned into lightweight, small footprint, form-factor and form fitting gadgets. Today we have wearable sensors that exist in the form of accessories such as a smart watch on a player's wrist, a head-mounted display worn by an immersive gamer, a cyclist's helmet having a tiny sensor on it, or a smart garment a runner uses for step counting, speed, calories consumed, altitude and distance tracking [1]. Contrary to these, the conventional in-lab assessments performed at specific time and place cannot capture the entire physiological gamut that is defined by the temporal and environmental conditions, the subject lives in.

In recent years especially, the digital world has witnessed an explosive growth in wearable technology. The concept of a wearable device has become very popular in many applications such as medical, entertainment, security, and commercial fields. They can be extremely useful in providing accurate and reliable information on people's activities and behaviours, thereby ensuring a safe and sound living environment. Developmental Cognitive Neuroscience describes that mobile and wearable technologies enable to capture an individual's data based on physiological, behavioural, psychological and environmental aspects that affect human health. It also helps to deeply understand interactions between humans and its effect on mutual well-being [2]. It would not be wrong to say that smart wearable sensors technology will revolutionize our life, health, social interaction and activities very much in the same way that personal computers have done a few decades back.

Advances in sensors allow deeper measuring capability. It helps the users learn more about themselves, thus changes to their lifestyle can be made under their control. Technology today has empowered patient, consumers to have real-time access to personal information which enables them to make better decisions about their own health and to generate valuable data for broader public health interventions [8]. Data generated by these wearable sensors is considered key element in all the health and well-being domains depending upon its usage and analysis. One of the main usage of this sensors' generated data is "Human Activity Recognition".

1.2 Human Activity Recognition Through Wearable Sensors

Wearable sensor devices have many usages in healthcare and fitness domain, one of these is "Human Activity Recognition". Human activity recognition based on wearable sensor data is considered an important research area due to its application in healthcare and fitness. It is normally defined as a problem of predicting the physical movement of a person by using different sensors. These movements consist of many activities such as standing, sitting, jumping, lying, cycling, etc [75]. Wearable sensors are often located on different locations of a person's body to record the movement data. This sensor data is subsequently analysed using different classification approaches to recognise the different activities.

1.3 Research Problem & Motivation

Activity recognition (AR) systems are typically built to recognize a pre-defined set of common activities. Use of machine learning methods for the purpose of AR using wearable sensors' data has been an attractive area of research. A lot of work has already been done in this domain [84][85][86][87][88]. Very few works done were aimed at giving a comprehensive framework for data analysis and also to develop a custom machine learning based analysis tool for automating the analysis process. The comparative analysis of machine learning algorithms on wearable sensors' data also lacks thorough statistical performance analysis. Another less explored area is the assessment of relative significance of sensors in the Activity Recognition process [91][92].

1.4 Research Scope

1.4.1 Aim

The aim is to conduct a systematic research on wearable sensors data and its analysis as part of activity recognition process. A comparative performance analysis of the selected machine learning algorithms would be carried out on MHEALTH sensors data for the purpose of activities recognition with the help of own-developed interactive data analysis tool.

1.4.2 Objectives of Research

For the achievement of the above-mentioned aim, the following research objectives were identified:

1. To conduct literature survey on state-of-the-art wearable sensors technology, its wide applications in different domains with a focus on HAR and associated data processing implications using classification approaches for the purpose of HAR.
2. Comparative analysis of different Machine Learning approaches used to process wearable sensors data (MHEALTH dataset) for the purpose of activity recognition / classification process.
3. Design and develop an interactive data analysis tool to systemize/ automate the analysis process of wearable sensors' data using different ML algorithms.
4. To compare the performance of selective ML algorithms on classification process in terms of accuracy of prediction and after applying PCA techniques to reduce curse of dimensionality.
5. Assessment on relative significance of different sensors / features viz-a-viz activity recognition process by using ML algorithms.

1.5 Thesis Structure

The remaining part of the thesis is structured as follows: -

Chapter 2 describes the knowledge gained during literature review performed on the subject and dilates upon wearable sensors' introduction, types and usage. It will further discuss the importance of data collected from these sensors and issues related to its processing and analysis. Use of different classification approaches to process the sensors' data will also be discussed along with their limitations and challenges in the sensors data analytics domain. It also highlights knowledge gaps in this field with an attempt to overcome these with own adopted approach.

Chapter 3 narrates own research framework being employed for processing of wearable sensors data. In the light of it, high level architecture, design considerations, development methodology, technologies used, and machine learning algorithms being implemented during design and development will also be discussed in detail. The chapter will also highlight the performance metrics used for evaluation.

Chapter 4 describes the process of experimentation and evaluation performed during this thesis. It includes evaluation of the MHEALTH dataset characteristics such as its source, data summary and usage in other applications. An in-depth analysis of the MHEALTH dataset using three different approaches is carried out after ML algorithm application through the developed tool is carried out. The results of comparative predictive analysis of different ML algorithms achieved thereof are then analysed and presented in great detail using different statistical metrics.

Chapter 5 concludes the thesis by summarising the research work done, and the lessons learnt during the research. It will also highlight the contributions/ limitations it has made to the body of knowledge on the subject and future research directions that can be taken up for advancing this research.

Chapter 2
Literature Survey

The chapter discusses salient concepts explored during the course of literature review. The research topic required to grasp requisite knowledge for thorough understanding ; and in its light, design and develop data science based solution for processing of wearable sensor data. The chapter starts with an overview of wearable sensors technology with emphasis on its scope in various relevant domains (Section 2.1). It is followed by in depth discussion on wearable sensors technology by highlighting its definition, role, attributes, state-of-the-art wearable sensors and devices (Section 2.2). Wearable applications specially potential physical activities that can be monitored using them are discussed in Section 2.3. The process of HAR is elaborated in Section 2.4 with description of own approach. Section 2.5 describes the state-of-the-art approaches adopted in classification process. The chapter concludes by highlighting relevant academic work regarding underlying algorithms being used during activity recognition process followed by its research gaps (Section 2.5 & 2.6).

2.1 Overview of Wearable Technology

The digital world had witnessed an explosive growth in wearable technology in recent years. The concept of wearable devices has become very popular in many applications such as medical, entertainment, security, and commercial fields. These devices are quite useful in providing accurate and reliable information on human activities and behaviours. Today we have wearable sensors that exist in the form of accessories such as a smart watch on a player's wrist, a head-mounted display worn by an immersive gamer, a cyclist's helmet with a tiny sensor , or a smart garment used by a runner for step counting, speed, calories consumed, altitude and distance tracking [1]. The smart wearable sensors technology is going to revolutionize human life, social interaction and activities very much in the same way as personal computers have done a few decades back. Advances in sensors allow deeper measuring capability. Users learn more about themselves, thus changes to their lifestyle can be made under their control [40]. **Table 2.1** illustrates length and breadth of wearable sensors technology in terms of their applications and functions in various commercial wearable products [47].

Table: 2.1 Scope of Wearable sensors technology in relevant domains [47]

Domains	Applications & Functions	Products
Medical	• Vital signs monitoring • Chronic disease management like Diabetes care • Brain / eye movement • Remote EEG, ECG, EMG	• Biofeedback patch • Insulin pump patch • Hearing aid • Wireless ECG Headset
Wellness	• Physiological monitoring like Sleep / Emotion / Stress Tracking • Weight / Energy monitoring for Obesity control • Gait / Posture correction	• Sensor Fitness band • Connected bracket • Interactive belt • Fitness Trackers
Sports	• Sport performance • Fitness monitoring • Virtual coaching • Outdoor navigation / Tracking • Body cooling / Heating	• Bio-harness • Activity tracker • Smart training shoes • GPS ski mask • Heated Jacket

2.2 Wearable Device / Wearable Sensors

2.2.1 Definition

A wearable device is essentially a tiny computer with sensing, processing, storage and communications capabilities. Many wearable devices also include interfaces and actuation capabilities that provide feedback to the user [1]. A sensor is defined as "a device used to detect, locate, or quantify energy or matter, giving a signal for the detection of a physical or chemical property to which the device responds" [9]. Not all sensors are necessarily wearable, but all wearables, must have sensing capabilities.

2.2.2 The Role of Wearables

A wearable device is fundamentally, required to perform the following basic functions or unit operations [1]

i. Sense
ii. Process (Analyse)
iii. Store
iv. Transmit
v. Apply (Utilise)

The specifics of each function will depend on the application domain and the wearer, where all the processing may occur on the individual or at a remote location. A wearable device can detect or sense the required signal and process the information in the form of data which can be stored in the device and can also be transmitted to a local or remote location or application for testing and results (utilisation) [1].

2.2.3 Attributes of Wearable Sensors

The key attributes [1] required of an ideal wearable are shown in Figure 2.1.

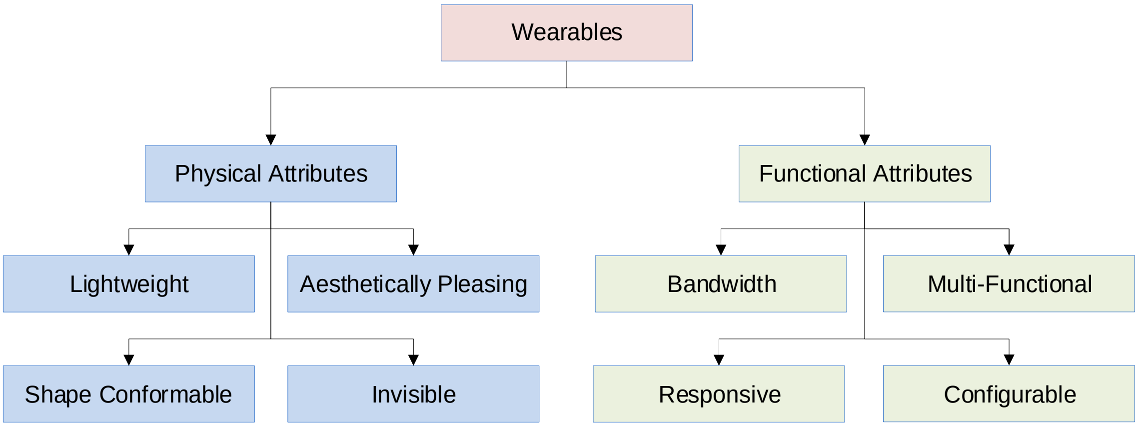


Figure 2.1 Attributes of ideal wearable sensor [1]

A wearable sensor must be lightweight from a physical standpoint with a variable form factor to suit the wearer. In case a user unable to use it "naturally", he is less likely going to adopts or uses the technology. Another key factor in the acceptance and use of any device or technology is how aesthetically it is designed. This is especially important when the device is also seen by others. Therefore, if the device worn is likely to be visible to others, it should have aesthetics in it. With wearable devices increasingly becoming an integral part of everyday lives, it should become a part to look like a “natural” extension of the individual. It must be flexible, shape-conformable and even can behave like the human skin [1].

The wearable device should be multi-functional and easily configurable for desired user application. For real-time data acquisition and control, responsiveness of the devices is quite critical which requires that it must be “always on”. Lastly sufficient data bandwidth is required to enable the degree of interactivity which is a key element in its successful use. These attributes play an important role in overall design of wearable devices [1].

2.2.4 State-of-the-Art Wearable Devices

Over the past few years, there has been a significant growth in demand of the wearable devices. According to a report published by the Transparency Market Research, the wearable devices market in USA was valued at approximately at US$ 2.6 Bn in 2017 and is projected to expand by 17% from 2018-2026 [38]. The report mentions that in Asia-Pacific market, the wearable devices compound annual growth rate in 2018-2026 is expected to be at 18.8%. A wide variety of these wearables have appeared in the market offering different functionalities and wearing options. These modern wearables are sufficiently equipped with advance sensors to measure variety of human attributes as well as the surrounding environment. State-of-the-art wearable devices based upon different Surveys in terms of their applications and functions are categorized and depicted in **Table 2.2** [5].

Table: 2.2 State-of-the-Art of Wearable Devices [5]

Wearable Device	Description	Product
Smart Watches	Wrist-worn devices with a touchscreen displays for fitness tracking.	Apple iWatch Samsung Gear S2 Moto 360 Pebble Time
Wrist Bands	Wrist-worn devices with fitness tracking capabilities or other functionalities, generally without a touch-screen display.	UP by Jawbone Fitbit Flex MOOV NOW Nymi Band
Smart Eyewear	Spectacles or contact lenses with sensing, wireless communication, or other capabilities.	Microsoft HoloLens FUNIKI Ambient Glasses Google Glass
Smart Jewellery	Jewellery designed with features such as health-monitoring and handless-control.	Smarty Ring Kerv Bellabeat Leaf
Straps	Chest straps, belts, arm bands, knee straps equipped with sensors for health tracking or other functionalities.	Shimmer3 IMU MYO Armband Zephyr Bio-harness
Smart Garments	Main clothing items that also serve as wearables such as shirts, pants, and undergarments.	Athos Hug Shirt Solar Shirt Spinovo
Foot / Hand-worn	Shoes, socks, insoles, or gloves embedded with sensors.	Lechal Sensoria Fujitsu Gesture-control Gloves

2.2.5 State-of-the-Art Sensors Used in Wearable Devices

Wearable sensors can be attached to different locations on human body such as waist, chest, wrists and legs, depending on the required relevant attributes [10]. Such sensors are also fitted to clothes or embedded in different accessories [12]. These wearable systems may require the development of a particular design and location as per application domain, i.e. wrist bracelets [13]. Some of the commonly used sensors in Wearable devices are discussed as follows:

Accelerometer

Accelerometers are probably the most common sensor found in wearables. These have sensing capabilities which range from different types of accelerations i.e. Linear and gravity [14]. These versatile sensors can measure the desired inputs and allow monitored data to be programmed for different uses. When user runs, it not only output top speed but its acceleration as well. Accelerometers are also used to monitor sleep patterns, which can be linked to seizures [15]. It is quite obvious from these examples that sport, and medical industries have a great potential for an accelerometer-based wearable because of its diverse range of meaningful data production. Due to flexibility in its positioning, the accelerometer has become quite a multi-functional sensing device [16].

Gyroscope

Gyroscope is another common sensor found in wearables. As compared to accelerometer, it measures angular accelerations exclusively. Both the sensors can be used for rotational accelerations. Sometimes accelerometer is preferred to determine rotational acceleration, whereas sometime the combination of accelerometer and gyro would be used to filter errors. This can increase the accuracy of the monitored data. Gyroscope essentially detects angular velocity on its disk [17].

Magnetometers

Magnetometer, accelerometer and gyroscope are generally combined to form the inertial measuring unit (IMU). These sensors can have three axes each depending on their type used. Its working is quite similar to that of a compass and it helps with coordination. It is normally used with other two sensors and complements them by filtering the orientation of movements [18]. It measures magnetic forces in relation to Earth's magnetic field [19].

Global Positioning System (GPS)

GPS is another commonly found sensor in multiple appliances (smart phones). It is used for navigation and informs user about his location. It synchronizes by sending data to a satellite for measuring precise location and time. It works on the principle of transmitter and receiver where information is fed back into the sensor to update the location [20]. Wearable devices use GPS to measure key data, such as distance, which can be viewed in different ways for different applications [21].

Heart Rate Sensor

Many sensors and techniques are used to measure heart rate. Capacitive sensing is used in one form where the electrode (sensor) and the human skin can be idealized as two components that make a traditional capacitor [21]. The phenomenon of photo plethysmography uses light to measure blood flow and links it to heart beats [22]. Fitbit, a fitness tracker uses this method via a photodiode. There is a constant green light emitting onto the skin of the user, where the photodiode can measure the light absorption. This data is converted so a pulse measurement can be processed [23].

Pedometer

Pedometer are used to count user's steps. It is commonly found in lifestyle-based fitness wearable devices. [24]. There are two versions of pedometers - mechanical and electrical. Movements are recorded and displayed in most pedometers as steps taken (a simple, raw or pure measure of ambulatory activity). Some also have features to estimate energy expended (kcals) and/or distance travelled (miles or kilometres). The simple pedometer can be used equally well by both researchers and practitioners and therefore offers a simple opportunity to bridge the gap between research and practice [96].

Pressure Sensor

Pressure sensor works from strain gauge method. A resistance change in the circuit is occurred when forces are applied on the sensor. Mechanical quantities such as force are experienced in multiple ways for sport. These are further converted into an electronic measurement dependent on resistance. Barometric pressure sensor is another widely used element in smart watches and wearables. It measures atmospheric pressure relative to the environment, to determine altitude. It is quite useful when it comes to monitoring elevations that a user goes through during activities involving climb / descend [25].

2.3 Application of Wearable Devices in AR

Wearable sensors are widely used today in many sectors like medical, fitness, communication and sports etc. The advent of these sensors has really revolutionized the quality of human life. Earlier what it took hours to study or monitor an event can now be addressed in minutes or seconds with the help of sensing systems. Wearables are also being used to provide a range of value-added services such as indoor localization and navigation, financial payments, gas sensing, environmental monitoring, monitoring constituents in food products like meat, beverages etc. to name a few [5].

The exponential growth of wearables depicts clearly the utilization of these sensors across many domains but monitoring of health conditions and physiological parameters is one of the most important applications of sensors having a very wide and significant use in medical and fitness domains [93]. Wearable sensors have revolutionized the way the activities of a person are being monitored [10]. They provide the information accurately and efficiently regarding the behaviour and actions of a person. The wearable technologies have overcome limitations of health investigations that are conducted at assigned locations and bounded by time [13]. The work of this research is also focused on use of sensors for human activity recognition by analysing the data generated and recorded by these sensors.

2.3.1 Human Activity Monitoring & Recognition

The use of wearable sensors has recently gained a popular trend for continuous monitoring of different physiological parameters related to an individual. The sensors connected externally to different parts of the body as well as to the garments can specifically detect the parameter it is used for. Some of the sensors are designed in a way to the diagnosis of more than one physiological parameter [93]. Modern wearables are sufficiently equipped with advance sensors to measure variety of human attributes as well as the surrounding environment. Each attribute can be studied individually to understand the anomalies faced by a patient and can be counteracted on [13].

Table: 2.3 Physical Activities Categorisation [94]

Physical Activities Categorisation				
Category	**Simple Physical Activities**			
Sub- Category	**Aerobic Exercises**	**Transportation**	**Sedentary Postures**	**Transitional Activities**
Type	Walking Jogging Running Swimming Climbing Descending	Driving Cycling Taking a bus	Sitting Lying Standing Tilting	Walk-to-run Run-to-walk Sit-to-stand Stand-to-walk
Category	**Complex Physical Activities**			
Sub- Category	**Activities of Daily Life**		**Ball Sports**	
Type	Eating Cooking Cleaning Dressing Brushing teeth Having a party		Playing tennis Playing football	

Physical activities play a fundamental role in human well-being; however, although people are now fully aware of their importance, they still need regular motivational feedback to maintain an active lifestyle. Multiple type of activities can be monitored with the help of wearable sensors ranging from physical to non-physical. Activity recognition based on new wearable technologies (wearable sensors and accessories, smartphones, etc.) is one of the most important challenges. Recognizing and monitoring human activities are fundamental functions to provide healthcare and assistance services to elderly people living alone, physically or mentally disabled people, and children. These populations need continuous monitoring of their activities to detect abnormal situations or prevent unpredictable events such as falls. The new technologies of health and activity monitoring devices range from on-body wearable sensors to in vivo sensors. Human physical activities are categorised in the form of simple and complex activities. A detailed categorisation chart is depicted in **Table 2.3** [94].

2.3.2 Sensors Placement on Human Body for AR

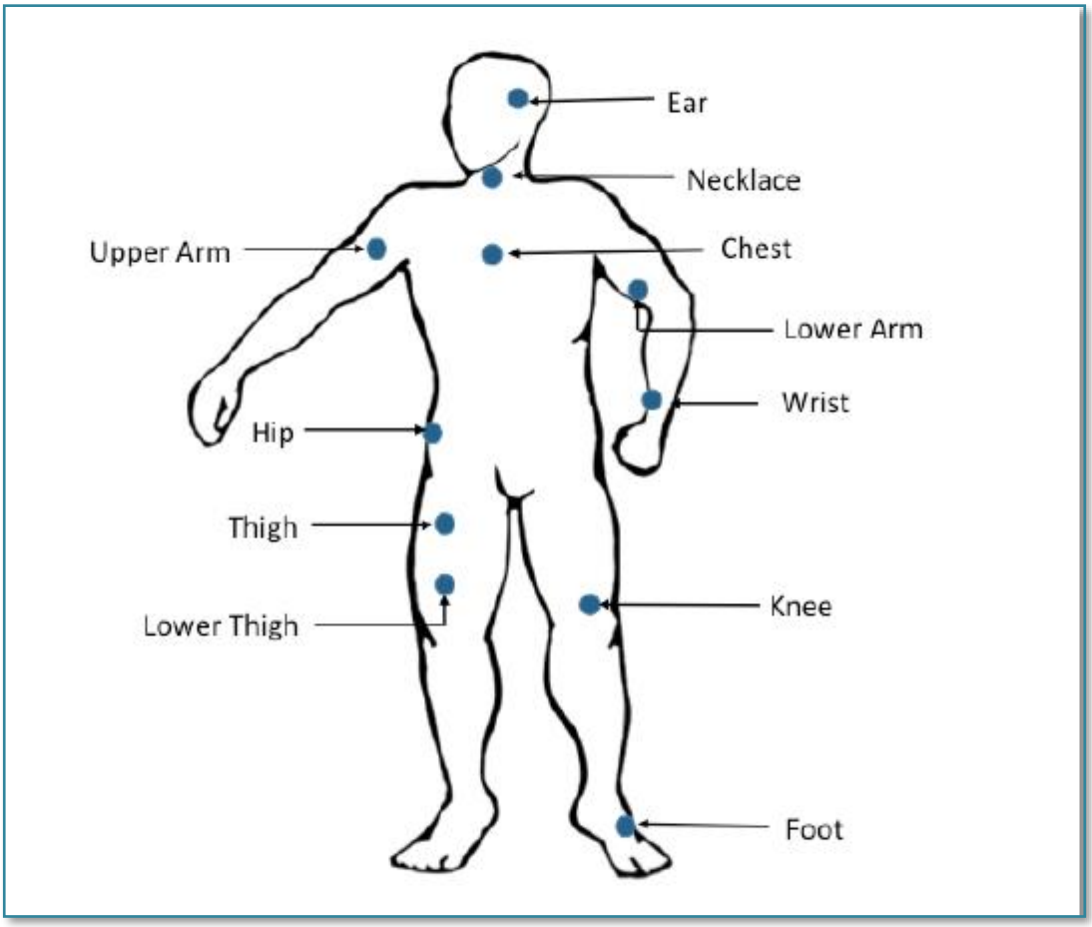


Figure 2.2 Placement of Sensor on Human Body [12]

The placement of wearable sensors is related to the locations where the sensors are placed and how they are attached to those locations. Indeed, wearable sensors placement has a direct effect on the measurement of bodily motions [11], but the ideal sensor location for particular applications is still quite a subject of discussion [12]. **Figure 2.2** depicts how wearable sensors can be placed on different parts of the human body. These sensors are usually placed on the sternum, lower back and waist. Waist-placement of the wearable sensors is considered better to represent most human motions since they are then close to the centre of mass of the human body. Multiple accelerometers, combinations of accelerometer plus gyros and a combination of accelerometer, gyros and magnetometer are used to attach at different locations of the body for the purpose of AR as depicted in **Table 2.4** [95].

Table: 2.4 Sensors Placement on human body for AR [95]

Sensor	Placement / Location	Usage/ Application
Accelerometer	Waist	Recognizing physical activity and posture transition especially for the patients with Parkinson
	Chest Thigh Foot Hip Wrist Lower back	Placement of accelerometers at different locations for detecting daily activities. Recognizing certain physical activities and their intensities
Accelerometer + Gyroscope	Wrist	Recognizing upper limb movements
	Belt/ Upper end of the pelvis	Fall and daily activity detection
Accelerometer + Gyroscope + Magnetometer	Chest Ankle right Thigh right	Identifying daily activity and postures with varied sensor locations

2.3.3 Wearable Data Collection for AR

Wearable Sensors' data is the first material for activity recognition after determining sensor types and sensor deployment. Data collection is considered to be a tedious and cumbersome work. There exist some benchmark datasets in HAR but the specific task of HAR require the ground truth from the different target population over the different time periods. A comprehensive data collection should involve maximum possible target population with diverse age, gender, weight, height and health conditions. The protocol of data collection directly affects the recognition performance. It depends upon the factors like number of activities, number of subjects, activities being performed in a natural way or a constrained way, a controlled environment or a real-home setting. Data collected is often based on the pre-defined activities under controlled environment. The volunteers are normally asked to perform the activities in approximate frequency and intensity or repeat one single activity in one minute or longer time, thereby achieving high accuracy due to the high intra-class similarity [95].

2.4 Human Activity Recognition Process

According to Stephen Bosch et al. [28], an activity recognition process spanning over five major steps (preprocessing, segmentation, feature extraction, dimensionality reduction, classification) is critical to achieve the correct and consistent results in this regard (Figure: 2.3).

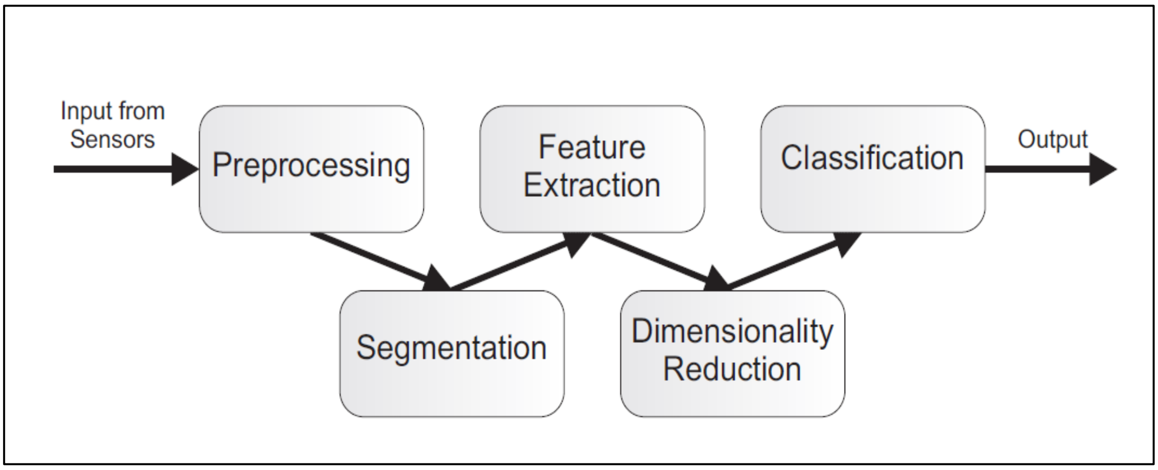


Figure 2.3 Human Activity Recognition Process [28]

2.4.1 Sensors' Data Analysis Process

The generic data analysis process based upon the Stephen Bosch process is used in this research. It flows from data collection to statistical results and visualizations. The data for this study (MHEALTH dataset) is selected from open source repository [3]. The generic post-collection process for data analysis (also applicable to wearable sensors data) is followed and described in ensuing paragraphs and is shown in Figure 2.4.

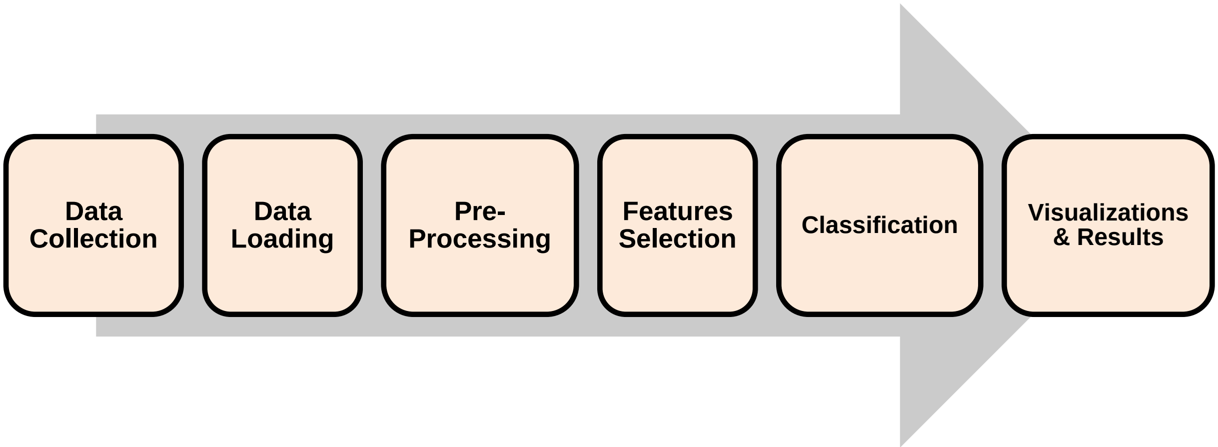


Figure 2.4 Generic Data Analysis Process for Classification

Data Loading

First step after sourcing the relevant data is the ability to load it for further processing. Data may be available in different formats; however, rectangular data formats are generally suitable for analyses. The data files are loaded into computer's memory for further processing.

Data Pre-Processing

This step takes bulk of time and is meant to clean and prepare the data for subsequent analyses. It may involve filtering the dataset for missing values and un-desired values.

Feature Selection

This step is performed to select feature set that includes relevant and valuable information extracted from the original non-reduced dataset. The output of this step serves as an input for the classification algorithms implemented in the next step.

Classification

This is the main step where classification algorithms are applied on the original or subset dataset.

Statistics, Visualizations & Results

This is the final step to compile results and statistics; visualize them as graph and deduce inferences.

2.5 State-of-the-Art Approaches on Activity Recognition

Over the past two decades, a number of studies have been conducted on the activity recognition. The advent of contemporary artificial intelligence techniques has played a significant role in the analysis of health sensors data in these studies. Some reviews on activity recognition [54] [55] described the process adopted in various research studies categorized as data preparation, windowing, features generation, and classification (Figure 2.5). The first two steps i.e. data preparation and windowing can be termed as data pre-processing which is required for subsequent steps. It starts with the sensors data preparation that constitutes cleaning sensor noise by filtering and replacing missing / erroneous values [56]. The next step is windowing, which is undertaken to divide sensors data into smaller time segments [57]. Generally, three types of windowing techniques are used to segment data; sliding windows, event-defined windows and activity-defined windows [55]. It is followed by features generation, output of which is used as an input for classification algorithms.

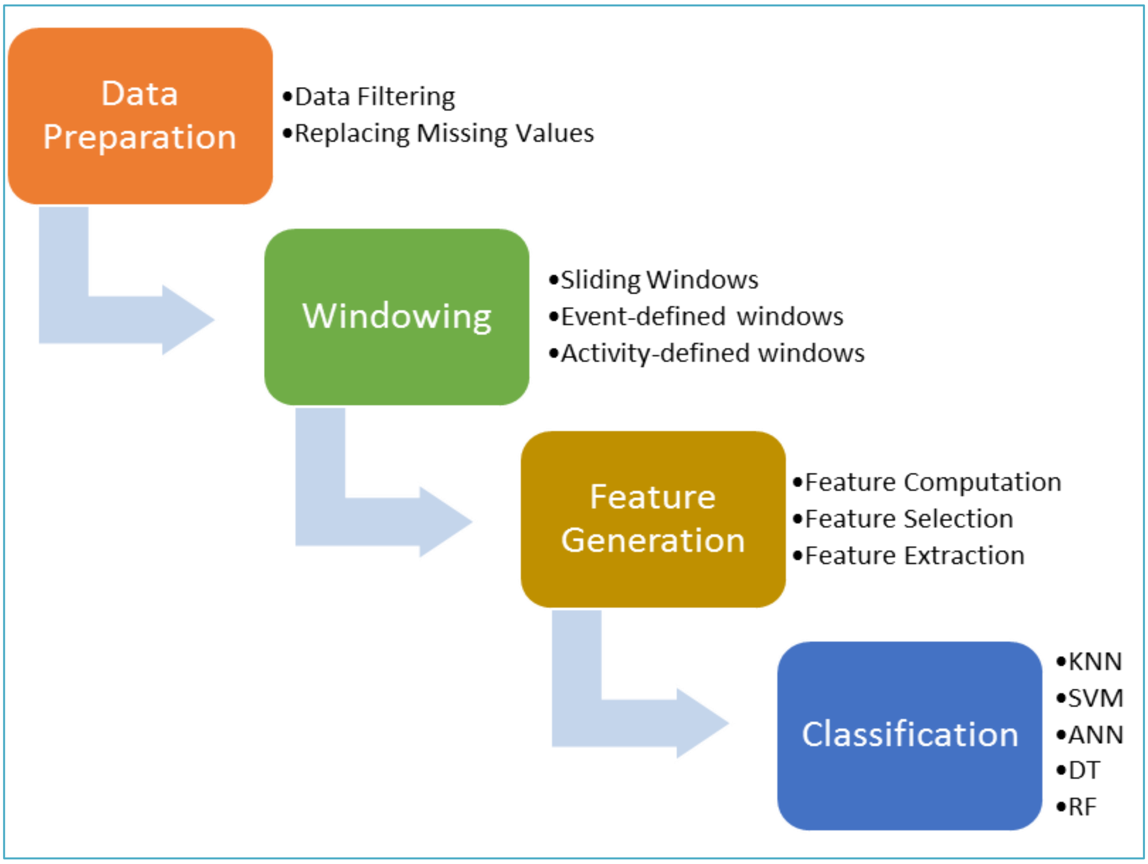


Figure 2.5 Activity recognition process [54] [55]

During the last step of activity recognition process, classification algorithms are employed to measure and characterize the activities. These algorithms fall under the umbrella of Machine Learning, Fuzzy logic and Threshold based rules (Figure 2.6).

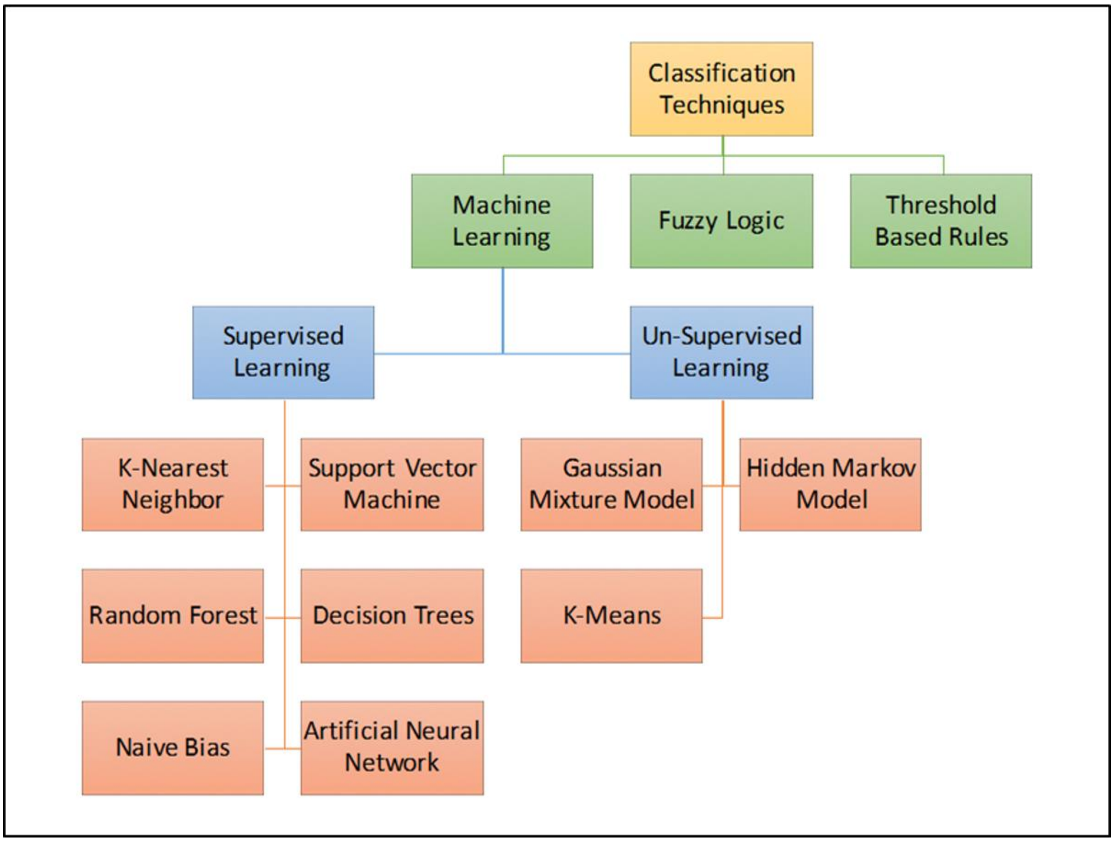


Figure 2.6 Classification techniques for activity recognition [54] [55].

Processing of data from wearable sensors (such as accelerometers, gyros, ECG etc.) using above mentioned techniques have been performed for recognizing different body postures and daily activities including health traumas like elderly falls. During classification, cross-validation is performed to evaluate the system using two methods: between-subjects and within-subject [54]. In between-subjects approach, collective data of all subject is bifurcated into training and test data; to be used for classification algorithm training and evaluation respectfully. While in within-subjects approach, the training and test data sets are made for every subject for algorithm training and testing cycle on every subject followed by averaging to determine overall accuracy [54][76].

2.5.1 Threshold Based Classification

The approach involves selected features to be compared with predetermined threshold values, in order to identify the occurrence of a particular activity. Although simple in design and implementation with low computational resources required; its adaptation to new environments / datasets and ascertaining appropriate threshold values for a particular activity is a tricky process. However, the issue can be resolved to a certain extent by employing threshold values from multiple features simultaneously or by employing dynamic threshold values on the basis of performance of previous threshold-based classification techniques. The approach had proven to be particularly useful to differentiate between static postures, postural transitions and in detection of falls [58].

Major academic research works in this domain include but are not limited to: an accelerometer based mobility monitoring system used to identify durations of sitting, standing, lying and moving positions by using midpoint tolerance values and best estimate tolerance values as an input to threshold based classification techniques with an accuracy of 75% and 93% respectively [59]; a quantified functional mobility progress (lying vs sitting, standing vs walking, activity movements etc.) mechanism for chronic disease management by using absolute torso angles and kinetic energy values to discriminate physical activities [60]; a stair climbing detection approach based on peak angular velocity of shank sourced from miniature gyroscope and was used to differentiate toe-off, heel-strike and foot flat positions through sequential threshold based rules [61] and gait pattern determination (descending stairs, ascending stairs and level walking) through garment based accelerometers; while using vertical acceleration and antero-posterior acceleration signals as an input feature for threshold based rules. The two-step approach was able to achieve sensitivity and specificity of 98.79% and 99.52% respectively for ascending stairs, while that of descending stairs is claimed to be 97.35% and 99.62% respectively [62].

2.5.2 Fuzzy Logic Based Classification

Another paradigm for human activity recognition is fuzzy logic methods. Fuzzy logic takes its origin from fuzzy sets theory. It shows a great potential for activity classification problems. However, fuzzy logic needs to employ methods for constructing proper membership functions as well as the combination and the interpretation of fuzzy rules. It allows mapping from a set of inputs to one or more outputs via a set of if–then statements called rules [55]. For an activity classification problem, features derived from body-worn sensor signals constitute the inputs, with the outputs being fuzzy truths corresponding to each class of activity. Information flows through a fuzzy system via several steps. Firstly, the inputs (or features) are assigned membership to fuzzy sets via appropriate membership functions. Once each input has been assigned membership of a fuzzy class, the rules can be applied to produce a corresponding output. For an activity classification problem, this output is a membership value, or fuzzy truth, ranging from 0 to 1 for each class of activity. The classification result is then normally taken to be the activity with the maximum fuzzy truth [54].

2.5.3 Machine Learning Based Classification

The machine learning approach involves multiple supervised as well as unsupervised learning techniques. Activity recognition process through these techniques is complex to implement and computationally expensive in general. However, the techniques are best suited for unknown and unpredictable sensors data over multiple iterations because of their ability to learn from their environment. Better decisions are made based on rewards collected from previous decisions. Most important of them are Random Forests (RF), Support Vector Machines (SVM), K-Nearest Neighbour (k-NN), K-Means, Artificial Neural Networks, Gaussian Mixture Model, Hidden Markov Model etc [54][55].

Significant academic work related to activity recognition process implementing machine learning based classification techniques include: a tri-axial accelerometer-based approach to differentiate between falls and everyday activities by implementing k-NN with an accuracy of >95% [68]; physical activity detection based on acceleration data using five small biaxial accelerometers on 20 subjects applied K-NN algorithm that showed strong relative performance and was second accurate after decision trees [69]; detection of human motion such as walking, driving and being in a train that used smartphone sensors data and implemented RF that gave highest accuracy (97.71%) compared to other ML techniques such as SVM and Naïve Bayes [70]; activity recognition for discrimination of ambulatory falls from other physical activities using SVM [68] [71]; use of Hidden Markov Models for automatic classification of human activity [72]; daily activities and fall detection using Naïve Bayes that have shown to be relatively less accurate [69] [71].

Use of classification algorithms for online or offline recognition is also narrated in a number of studies [55][75][76][77] for the purpose of Human Activity Recognition from Sensors' data. These studies have mainly used supervised algorithms; except one study [12] which also demonstrated use of unsupervised algorithm for activity recognition.

2.6 Gaps in Existing Literature

Being commonly researched topic in academia as well as industry; there is no dearth of relevant data available on the subject. However, on the basis of literature review in Chapter 2; there are still some gaps that can be highlighted with reference to the topic under research.

In order to make their work more inclusive, most of the researchers have tried to encompass all stages (sensor placement, data acquisition, data cleansing, windowing, feature selection, feature extraction, feature classification) of the activity recognition process within their research [76]. While discussing and implementing 'a bit of everything' may have resulted in an all-encompassing solutions; the approach has devoid the relevant literature from in-depth analysis on feature classification part of the problem. Although the underlying algorithms or set of algorithms being used during feature classification stage have been mentioned along with its results; but very few have discussed statistical comparative analysis of these approaches and that too in cursory manner. Being the most critical part of activity recognition process; its underlying classification algorithms need to be investigated in detail and critically analysed for their effective implementation.

Moreover, comparing different classification algorithms based on existing literature is a difficult task that may lead to inconsistent conclusions. Existing researches have been plagued with wide-scale variability regarding number and type of sensors, experimental objectives, experimental protocols, evaluation criterion, validation procedure etc. Additionally, lack of due implementation details renders these researches ineffective in context of repeatability / replication on same data for further insights or on other data sets [75][76]. Furthermore, any research could not be find that devised any mechanism to automate the process of statistically comparing the classification algorithms.

Other activity recognition related issues like dimensionality reduction or significance factor of specific feature viz-a-viz particular physical activity have been discussed in relevant research works; albeit with sketchy implementation details and no automation mechanism to handle the issue efficiently [75][77]. In the absence of such ready-to-go, comprehensive solutions or tools, preferring one algorithm over the other becomes a time-taking task.

Foregone in view, it would be interesting to concentrate one's research efforts on underlying classification algorithms being used in feature classification stage and come out with an innovative tool that can systemize as well as speed-up the process of comparing relevant algorithms in statistical perspective; thus giving new insights to datasets being processed through them.

2.7 Own Research Framework

The Knowledge gaps found in the existing literature motivated to design a research framework to conduct data analysis of Wearable Sensors Data (WSD) for the purpose of Human Activity Recognition (HAR) using machine learning approaches. It also laid foundations for the development of a handy Data Analysis Tool to automate this whole process which would greatly benefit the research study and can further be utilized for other such studies. The framework is designed based upon the generic data analysis process. It is novel in its concept since it uses a unique idea of choosing all the three techniques of data processing i.e.

using feature ranking, feature sub-setting and finally using PCA techniques for feature selection. Another novel approach used in this study is the detailed results analysis of ML algorithms application on data carried out with the depiction of different graphs of performance statistical metrics.

The conceptualised research framework is discussed in detail in **Chapter-3** whereas a comprehensive comparative analysis of different machine learning approaches on MHEALTH data is presented in **Chapter-4**.

2.8 Summary

The chapter explains background research conducted in the field of wearables sensors technology and includes discussion on; understanding of its meaning and scope in different industry domains, attributes and types of wearable sensors, types of physical activities monitored through these sensors, existing research work in relevant domain and research gaps thereof. Later part of this chapter highlights the classification approaches adopted for human activity recognition. Having covered the literature survey in Chapter 2; next chapter would discuss the research framework designed to conduct data analysis of Wearable Sensors Data (WSD) for the purpose of Human Activity Recognition (HAR) using machine learning approaches.

Chapter 3
Research Methodology & Framework

3.1 Data Analytics

Data analytics is science of exploring raw data and drawing out the useful information and hidden patterns in data. With analytical techniques and available computational power, data scientists and others can analyse huge volumes of data that conventional analytics and business intelligence solutions can't do. Data analysis is being used in different domains like science, business, and social sciences. With ever increasing data and need for its analysis, tools that help in analysing the data and derive conclusions with ease are in great demand [4][5].

As mentioned in Chapter 2, several research studies conducted over the past two decades leveraged wearable sensors data. Many of these studies focused on activity recognition by analysing data collected during experiments [84][85][86][87][88]. However, there remains a gap of following a wholesome process and availability of a tailor-made data analysis tool for activity recognition from wearable sensors data that may aid in analysing the output from different classifier algorithms; comparing them with the help of visual analytics. While one of the main goals of this research is performance comparison of different classifier algorithms on wearable sensors data; the methodology followed is to achieve this by designing an easy to use and reusable data analysis tool for standardized wearable sensors data.

3.2 Research Framework

Based upon the data analysis process as discussed in **Chapter 2 Section 2.4**, a research framework was designed for processing of Wearable Sensors Data (WSD) for the purpose of Human Activity Recognition (HAR). It laid foundations for development of a handy Data Analysis Tool to carry out this research. The conceptualised research framework is depicted in Figure 3.1.

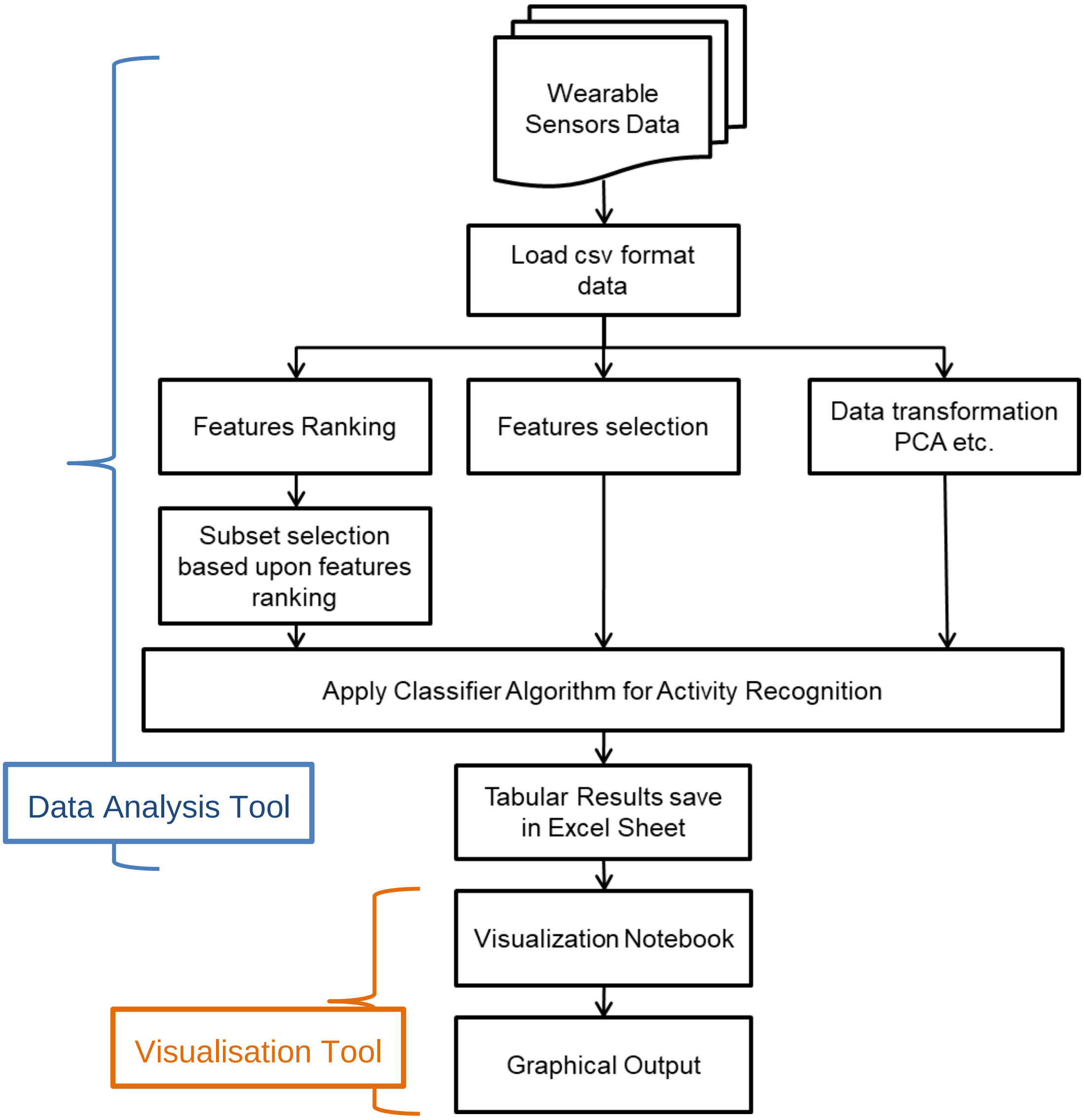


Figure 3.1 Research framework for Activity Recognition through ML Algorithms

3.2.1 Data Loading

The framework begins with sourcing the suitably labelled Wearable Sensors Data in rectangular format and loading it into processing environment.

3.2.2 Data Processing

After Loading the data there are three options to process the data: -

(i) Feature Ranking

Features ranking algorithm is applied on the data to find relative importance of input features and then to select a subset of those features. This technique is quite useful when there are multiple features in a dataset.

(ii) Sub-setting the data

Second option is either to select the whole dataset as input or to select a subset of features based upon domain knowledge on relative importance of features.

(iii) Data Transformation using PCA techniques

Third option is to transform the data to reduce its dimensionality by using techniques like Principal Component Analysis (PCA) and then selecting its subset columns.

3.2.3 Applying Classification ML Algorithm on Data

After sub-setting the data, classifier ML algorithms are applied to get the results. The chosen algorithm with different experimental setting can be used again and again. Final results are saved to the disk for subsequent analyses.

3.2.4 Result Analysis through Novel Visualisations

A visualization script is used to visualize different statistics as graphs while using different performance metrics and to compare the performance of different algorithms using these metrics.

3.3 Novel Research Framework Concept

The framework is designed based upon the generic data analysis process. It is novel in its concept since it uses a unique idea of choosing all the three techniques of data processing i.e. using feature ranking, feature sub-setting and finally using PCA techniques for feature selection. Another novel approach used in this study is the detailed results analysis of ML algorithms application on sensors' data carried out with depiction of different graphs of performance statistical metrics. It was felt to develop a re-usable data science-based tool to automate this whole process which would greatly benefit the research study and can further be utilized for other such studies.

3.4 Design of Data Analysis Tool

As Frederick Brooks put it, “The hardest single part of building a software system is deciding what to build. No other part of the work so cripples the resulting system if done wrong. No other part is more difficult to rectify later" [48]. The desired properties while designing the Data Analysis Tool for Wearable Sensors Data were defined as follows: -

- Ability to Input Rectangular Data for Analysis
- Ability to Choose Factors for Model
- Ability to rank data features for importance in modelling
- Ability to transform data using PCA
- Ability to run ML classification Algorithms
- Repeatability of analyses
- Ease of use
- Ease of extension
- Graphs, Summary Statistics, Confusion Matrix etc.

3.4.1 Data Analysis Tool Implementation - Technology Considerations

R-Language and Python are two most popular free and open-source programming languages used by data analysts. R mainly focuses on statistical analysis and Python is a general-purpose programming language. Tasks involving machine learning, working with large datasets, or creating complex data visualizations can be performed by them with ease. These languages have many built-in libraries which implement a wide variety of statistical and graphical techniques, including linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, and others [53]. The whole process requires a deep understanding and learning of these languages, machine learning concepts /algorithm and a good knowledge of different statistical functions to get the analyses done. Since R is mainly focused on statistical analyses, it was selected for the purpose of this study. Its three main components used in development of tool for this study are as follows: -

- R- Engine - R Engine is the backbone of the application. All the data scrutiny, processing and output will be dealt by R-engine running in the background.
- R- Shiny - Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R.

- R- Studio - R-studio provides the development environment and all the necessary packages for the development of back end and front end of the desired tool.

3.4.2 High Level Architecture

The tool is built on R-Engine, for exploiting R-processing of datasets using its powerful libraries with built-in machine learning functions. For interactive and user-friendly interface, Web-based architecture of R-Shiny is used. The higher-level architecture is illustrated in Figure 3.2.

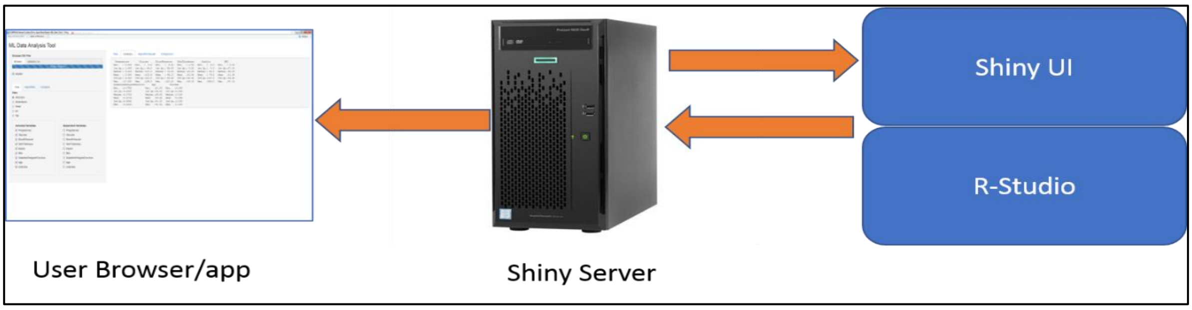


Figure 3.2: Software Architecture of Data Analysis Tool

3.4.3 Development Methodology

RAD or Rapid Application Development process was adopted for the development of data analytics tool in a short span of time [79]. As depicted in Figure 3.3, RAD follows the iterative software development life cycle. Keeping in view the short span, already known requirements which can be improved through the feedback phase, the RAD model was found quite appropriate for the development of the tool.

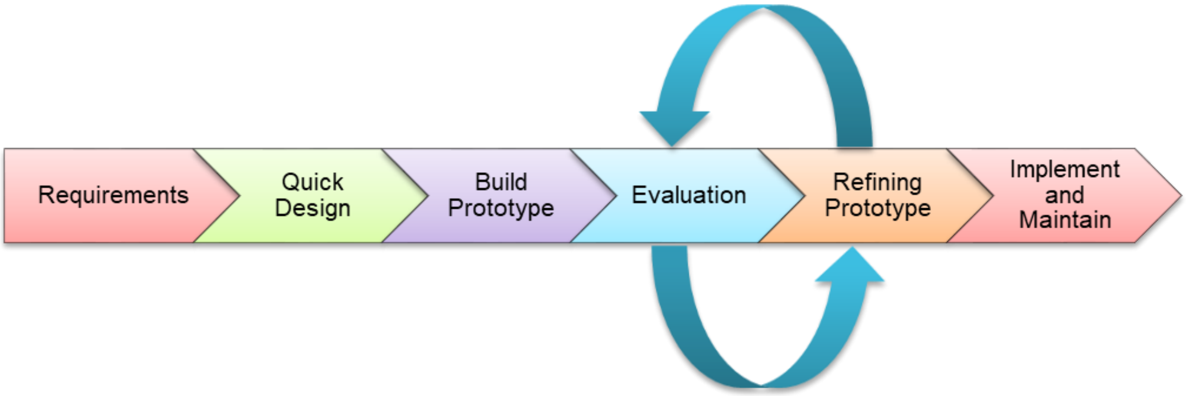


Figure 3.3: RAD Software Methodology [79]

Detailed user requirements for the development of data analysis tool is attached as **Appendix 'A'**, whereas GUI design and its functionality description are attached as **Appendix 'B'**.

3.5 Classification ML Algorithms - Implemented in Research

Based upon our research methodology; five ML algorithms are implemented in this study. These five machine learning models were trained and the compiled compared results will be discussed in the next **Chapter 4**. Repeated cross-validation with 10 folds and 3 repeats; training and test ratio of 70:30, a common standard configuration for comparing models is used. The evaluation metric is accuracy and kappa because these are easy to interpret. The ML algorithms used in the research were chosen for their diversity of representation and learning style. These are the most common ML approaches which are used for classification process in several studies for the purpose of activity recognition from wearable sensors data [84][85][87]. These ML algorithms are briefly described in the following paragraphs.

Classification and Regression Trees

Classification and Regression Trees or CART for short is a term introduced by Leo Breiman [49] to refer to Decision Tree algorithms that can be used for classification or regression predictive modelling problems. It has the ability to predict target by applying simple decision boundaries. Classically, this algorithm is referred to as decision trees, but on some platforms like R they are referred to by the more modern term CART. CART can be used efficiently to assess massive datasets and can provide quick solutions [73].

Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a linear method for classification of predictive modelling problems. Logistic regression is a classification algorithm traditionally limited to only two-class classification problems whereas for more than two classes LDA is the preferred linear classification technique [73].

Support Vector Machine

Support Vector Machines are one of the efficient classification machine learning algorithms. They were extremely popular around the time they were developed in the 1990s and continue to be the go-to method for a high-performing algorithm with little tuning. It can solve linear and non-linear problems and works well for many practical problems. The algorithm creates a line or a hyper plane (decision surface) which separates the data into classes [73].

K-Nearest Neighbour

It is a non-parametric algorithm used for regression and classification. KNN makes predictions using the training dataset directly. Predictions are made for a new data point by searching through the entire training set for the K most similar instances (the neighbours) and summarizing the output variable for those K instances. For regression this might be the mean output variable, and in classification this might be the mode (or most common) class value. When KNN is used for classification, the output can be calculated as the class with the highest frequency from the K-most similar instances. Each instance votes for their class and the class with the most votes is taken as the prediction. Class probabilities can be calculated as the normalized frequency of samples that belong to each class in the set of K most similar instances for a new data instance [73].

Random Forests

Random Forest is one of popular and powerful machine learning algorithms. It is a classification algorithm consisting of many decision trees. It improves the classification performance of a single-tree classifier by combining the bootstrap aggregating (bagging) method and randomization in the selection of partitioning data nodes in the construction of decision tree. The assignment of a new observation vector to a class is based on a majority vote of the different decisions provided by each tree constituting the forest. It needs huge amount of labelled data to achieve good performances [12].

3.6 Statistical Visual Analysis of Algorithms

The culmination of our research is to analyse performance of each algorithm using statistics and to have comparative analysis of these algorithms. These statistics are presented with respect to overall accuracies of algorithms and class-wise statistics like sensitivity, specificity, positive and negative prediction values, detection rates and detection prevalence etc. As it is said that a “picture is worth thousand words”, same would be even more true for hundreds of output numbers in matrices or statistical results. Thus, novel visualization techniques have been used to present performance and statistical comparisons in powerful and intuitive way. The output of this research in the form of graphical comparisons would be an important step towards making meaningful inferences.

3.7 Performance Statistics used for Analysis

A number of statistics can be used for analyses of classifier algorithms for two class or multi-class problems. In fact, multi-class case is generalization of two class problem. The statistics used in this study [74] are defined as follows: -

Accuracy

Accuracy or overall accuracy in terms of a multi-class problem is defined as the ratio of sum of true positive and negatives to the total population count. Mathematically, it is given as,

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Positives + Negatives}$$

Accuracy gives a single metric to judge the performance of classifier algorithm.

Sensitivity, Specificity and Balanced Accuracy

Sensitivity, Detection Rate or Recall are defined as the ratio of True Positives to sum of True Positives and False Negatives i.e. positives in the population.

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Specificity is the ratio of True Negatives to sum of True Negatives and False Positives i.e. actual negatives in the population.

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$

Balanced accuracy is defined as the average of Sensitivity and Specificity.

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2}$$

For a good classifier, both Sensitivity and Specificity should be close to 1.0. However, in case of multi-class classification with large number of classes having similar prevalence, the Specificity may always be closer to 1.0.

Predicted Values

Positive Predicted Value (PPV) is the ratio of True Positives to the sum of True Positives and False Positives i.e. the positively predicted items by the classifier.

$$PPV = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

In contrast to the PPV, Negative Predicted Value (NPV) is defined as the ratio of negatives that have been correctly identified as negatives to the sum of both truly and falsely identified negative values. It is given mathematically as,

$$NPV = \frac{True\ Negatives}{True\ Negatives + False\ Negatives}$$

For a good classifier, both PPV and NPV should be close to 1.0. However, in case of multi-class classification with large number of classes having similar prevalence, the NPV may always be closer to 1.0.

Prevalence, Detection Rate and Detection Prevalence

Prevalence is defined as the ratio of positive class to total population in the dataset. In multi-class cases, the positives would refer to individual class instances and negatives would be all other classes.

$$Prevalence = \frac{Positives}{Positives + Negatives}$$

The Detection Rate is defined same as Sensitivity i.e. ratio of True Positives to the count of all Positives in the population either predicted as positives or negatives. Detection Rate may be seen in the context of Prevalence, because for each class these statistics helps in gauging the rate of detection of an algorithm to actual prevalence of positive class.

$$Detection\ Rate\ = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

In contrast to Prevalence, Detection Prevalence is the ratio of predicted Positive class to the total population.

$$Detection\ Prevalence\ = \frac{True\ Positives + False\ Positives}{Positives + Negatives}$$

When analysed together, these three statistics gives performance estimate of the classifier algorithm. The closer these three values are for each class, the better the algorithm.

Precision and Recall

Precision and Recall are two important statistics to measure performance of classifier algorithms. Mathematically, precision is defined as follows:

$$Precision\ = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Precision tells what fractions of positively identified cases actually belong to the positive class. Recall on the other hand is defined as the ratio of True Positives to actual number of Positives in the data. Mathematically, it is given as,

$$Recall\ = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

In a classification task, a precision score of 1.0 for a class X means that every item labelled as belonging to class X does indeed belong to class X (but says nothing about the number of items from class X that were not labelled correctly) whereas a recall of 1.0 means that every item from class X was labelled as belonging to class X (but says nothing about how many items from other classes were incorrectly also labelled as belonging to class X).

3.8 Selection of MHEALTH Dataset for Research

3.8.1 Data Source and its Choice for Research

The MHEALTH (Mobile HEALTH) data is an open source dataset used during this research downloaded from UCI repository of machine learning databases [3]. Although the available data does not include sensors noise or other garbage; however, it required filtering for un-desired values. This data was collected through external wearable sensors (Shimmer sensors). Ten volunteers participated in the study and they performed a total of twelve physical activities. These volunteers had diverse profiles. The wearable sensors were placed at three locations on the body i.e. the subjects' chest, the right wrist and left ankle. The sensors recorded body motion and vital signs of the subjects. Shimmer sensors on left ankle and right wrist had accelerometer, gyroscope and magnetometer that are used to measure the motion experienced by the body parts, namely, acceleration, rate of turn and magnetic field orientation. The chest sensor had accelerometer, gyroscope and 2-lead ECG measurements, which can be used for basic heart monitoring or looking at the effects of exercise on the ECG [75].

MHEALTH Sensor dataset is one of the most reliable and balanced sensors' datasets that is easily available through open sources [76]. It has been used in many research studies due to its quality and is considered good due to diversity of the sensors used and their placement while compiling this dataset [75][87][88][89][90]. For the purpose of Human Activity Recognition, the MHEALTH dataset facilitated the exploration greatly.

3.8.2 Dataset Description

In MHEALTH data, there are 10 subjects in the dataset who performed 12 activities and three types of sensors were used to record the data. The data collected for each subject is stored in a different log file: 'mHealth_subject.log'. Each file contains the samples (by rows) recorded for all sensors (by columns). The labels used to identify the activities are similar to the abovementioned (e.g., the label for walking is '4') [3]. Data was amassed with incorporation of a frequency about 50Hz, which is quite enough for encapsulating the values regarding the body movements. The meaning of each column is detailed next in Table 3.1:

Table 3.1 MHEALTH Dataset Columns Description (Units: Acceleration (m/s^2), gyroscope (deg/s), magnetic field (local), ecg (mV)) [3].

Column	Meaning
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 12	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)

The activity set performed is listed in the following Table 3.2:

Table 3.2 Activities in the Column 24 [In brackets are the numbers of repetitions (Nx) or the duration of the exercises (min)] [3].

Label	Physical Activity
L1	Standing still (1 min)
L2	Sitting and relaxing (1 min)
L3	Lying down (1 min)
L4	Walking (1 min)
L5	Climbing stairs (1 min)
L6	Waist bends forward (20x)
L7	Frontal elevation of arms (20x)
L8	Knees bending (crouching) (20x)
L9	Cycling (1 min)
L10	Jogging (1 min)
L11	Running (1 min)
L12	Jump front & back (20x)

3.9 Data Cleaning and Preparation

Data cleaning and preparation, also known as pre-processing, is the first and one of the most important steps in data science process and usually takes bulk of time. It may comprise format conversions of raw data, converting it into nice rectangular format, concatenating subjects' data after adding identification fields, filtering the required fields, replacing or removing missing and outlier values etc. The process followed in this research to prepare the MHEALTH data is broadly shown in Figure 3.4.

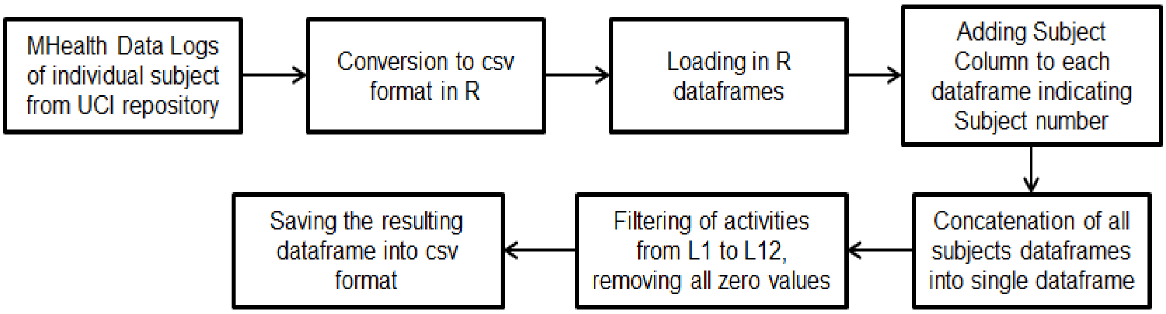


Figure 3.4 Data Cleaning and Preparation Process

The individual log files from UCI repository were converted to the csv format and concatenated together to make a single large file. The data for activities classification was in 24th column and the activities were marked from 1 to 12. This column also contained the value '0' which indicated data outside the experimental setting. This data was filtered and final csv file prepared was used for the purpose of applying ML algorithms.

3.10 Implementation of Perceived Framework

The discussed research framework was effectively utilised in the implementation and experimentation phases. MHEALTH data was first filtered and then loaded in the Data Analysis Tool developed for this study. The MHEALTH dataset was processed using all the three processing techniques i.e. using feature ranking, feature sub-setting and finally using PCA techniques for feature selection. Subsequent to data processing, the pre-defined machine learning algorithms in the developed Data Analysis Tool were run with different experiment settings on the processed data and the results were stored in the Excel sheets for further analysis.

The outputs were subsequently fed into a visualising script developed in Python for depicting the results in the form of visualisation graphs based upon the statistical data analysis. The framework perceived for the research addressed the requirement of MHEALTH data analyses quite successfully. A detailed comparative performance analysis of different ML algorithms used to recognise human activity from MHEALTH data with the supporting custom-made data analysis and visualisation tool is presented in **Chapter 4**.

3.11 Summary

The chapter describes research framework based upon the research methodology adopted in the research from data collection to data analyses. A brief description of supporting tailor-made data analysis tool for wearable sensors data is given along with the classification algorithms used. The next chapter would cover experimentation and evaluation part of the project leveraging the developed tool.

Chapter 4
Analysis of Results

4.1 Results Overview

Experimentation and evaluation are an important phase of any technology solution to examine its suitability for the purpose. It helps in verification and validation of the results. The Data Analysis Tool developed during the research was used to get the results of this experimentation and evaluation activity. The dataset chosen for this study is called ‘**MHEALTH’** data derived from wearable sensors. It is publicly available at UCI’s repository of machine learning databases [3]. The analysis was carried out between the subjects.

4.2 Data Analysis on MHEALTH Data

Data analysis of MHEALTH sensors’ data was done for activity recognition process through application of ML algorithms while using following three approaches for experimentation: -

i. Activity Recognition through ML Algorithms on Complete MHEALTH Dataset using ‘Hold out’ method.
ii. Activity Recognition through ML Algorithms on Reduced Dimensionality (PCA) MHEALTH Dataset
iii. Activity Recognition through ML Algorithms on Selective Sensors from MHealth Dataset.

After processing the data, the machine learning algorithms were applied through data analysis tool and the results are depicted in the forms of different statistical graphs described in the following sections.

4.3 Activity Recognition through ML Algorithms on Complete MHEALTH Dataset

In the first approach the aggregated MHEALTH data from all subjects is filtered from Null activity value and is partitioned into train and test subsets using ‘Hold out’ method. Seventy percent (70%) data i.e. 7 subjects out of total 10 are used as training set while the data for rest 3 subjects was 'Held' as test data for validation. The cleaned and prepared data was loaded into the Data Analysis Tool. The ML algorithms were applied on this data and the results obtained in the form of statistics and confusion matrix are further described for each algorithm in the subsequent pages. The results in the form of graphs depicts all the activities from L1 to L12 on x-axis while y-axis indicates the proportion that can vary between 0 and 1, with 1 being the perfect score.

4.3.1 Classification and Regression Trees (CART)

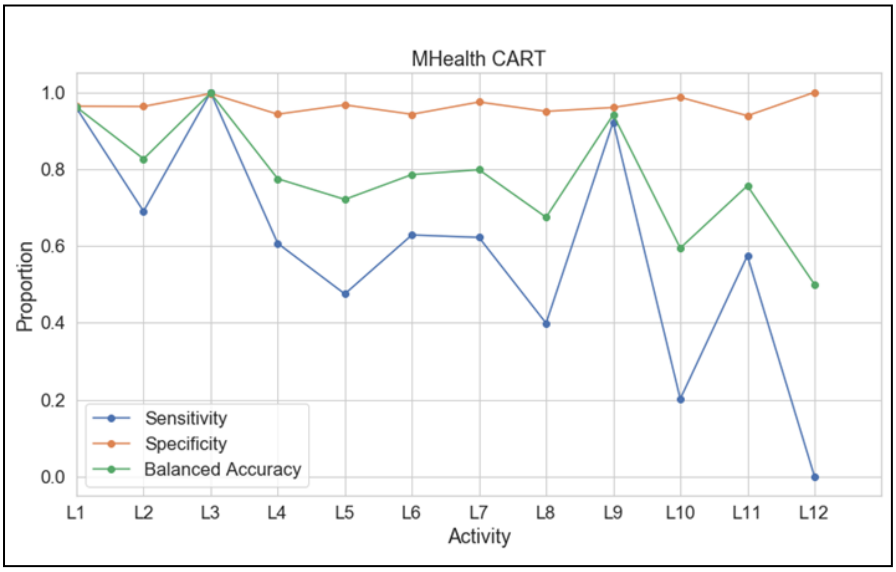
Overall Statistics

Table 4.1 CART Overall Statistics

Accuracy	0.6254
95% CI	(0.6224, 0.6283)
No Information Rate	0.0895
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.5892

Statistics by Class

Sensitivity, Specificity and Balanced Accuracy



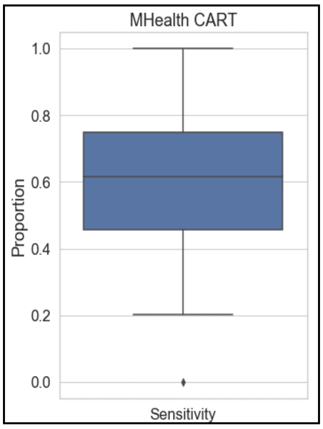
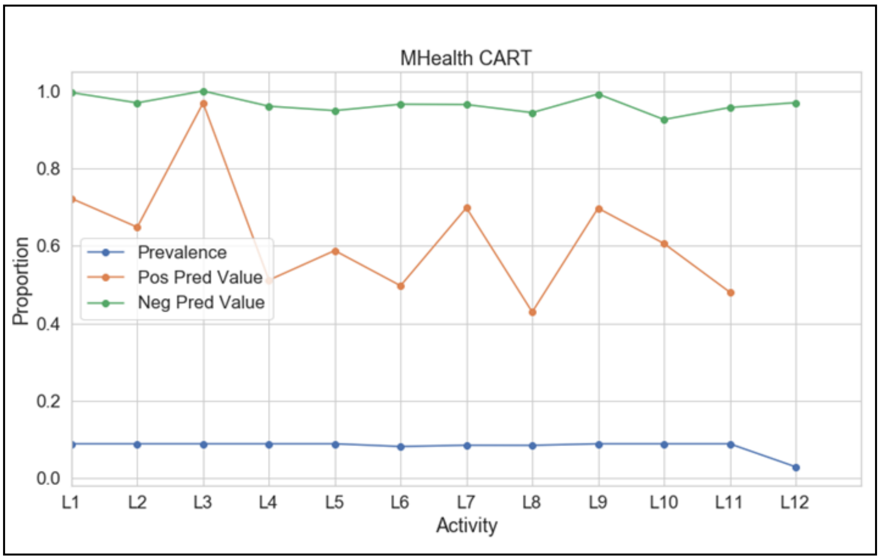


Figure 4.1 Sensitivity, Specificity and Balance Accuracy on CART Algorithm

The sensitivity by class depicted by blue line on test MHEALTH data on CART algorithm varies greatly for the range of activities. It is high for L1 (Standing Still), L3 (Lying Down) and L9 (Cycling), however, it is below 50% for L12 (Jump front and back), L10 (Jogging), L8 (Knees bending), L5 (Climbing stairs) and L11 (Running). The specificity of CART algorithm is quite high for all the classes because of the large number of classes and a lot of negative values for each class. Specificity in this problem does not have much to indicate about the model.

Predicted Values



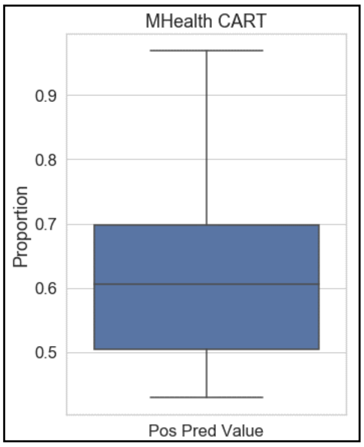


Figure 4.2 Positive and Negative Predicted Values on CART Algorithm

In case of Positive and Negative Predicted Values also, the Negative Predicted Value statistic does not have much meaning due to large data and large number of classes. The prevalence of each class is around 8% for each class except L12 (Jump front & back) which is around 3%. Under such prevalence, the positive predicted value shows 55% - 70% results for all classes except L3 (lying down) which is quite high in CART algorithm. For L12 (Jump front & back) there is not a single positively predicted value by the algorithm.

Prevalence, Detection Rate and Detection Prevalence

CART model has higher detection prevalence than detection rate which means false positives impose a problem for any prediction. For classes L3 (lying down) and L10 (Jogging) the difference is less. For L10 (jogging) and L12 (jump front & back) the detection rates are much lower than the prevalence. In the nutshell, the detection rate for CART on MHealth data is quite lower than prevalence for all classes except L1 (standing still), L3 (lying down) and L9 (cycling). The box plot indicates the variability among class on Detection Rate and Detection Prevalence.

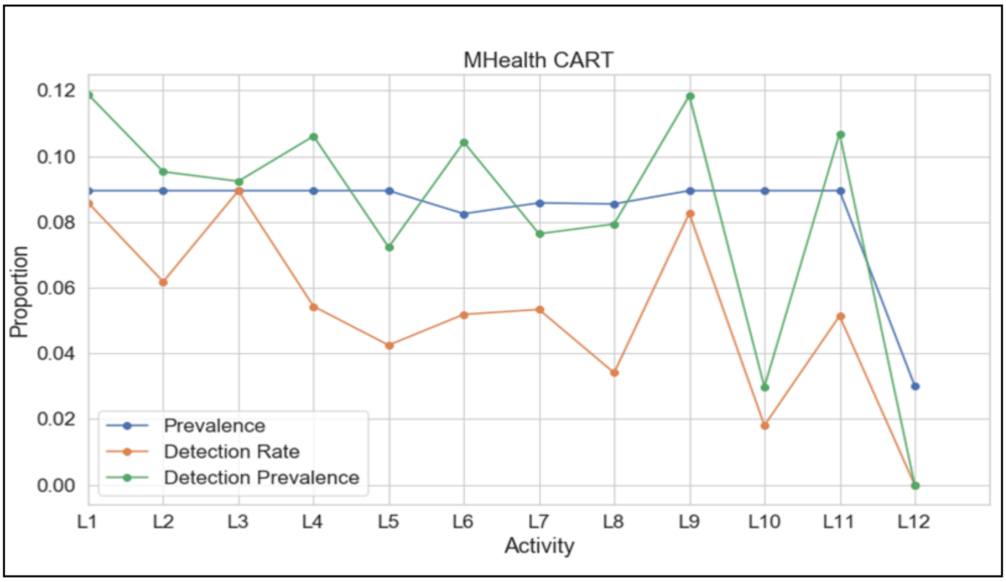


Figure 4.3 Prevalence, Detection Rate and Detection Prevalence on CART Algorithm

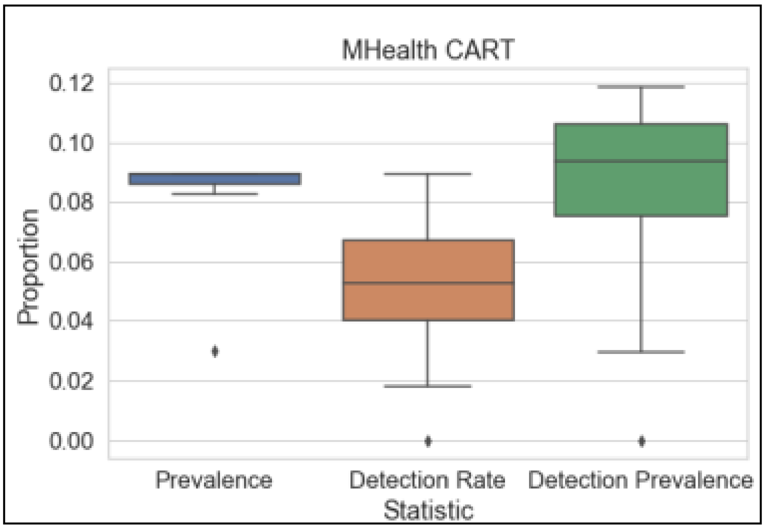


Figure 4.4 Prevalence, Detection Rate and Detection Prevalence on CART Algorithm

Confusion Matrix

Figure 4.5 Confusion Matrix of CART Algorithm

The confusion matrix gives even more detail about the correct and false predictions by CART algorithm. As evident L12 (Jump front & Back) has not been predicted by CART at all while Sensitivity and PPV for L3 is highest among all. The overall accuracy for CART is 62.5%.

4.3.2 Linear Discriminant Analysis (LDA)

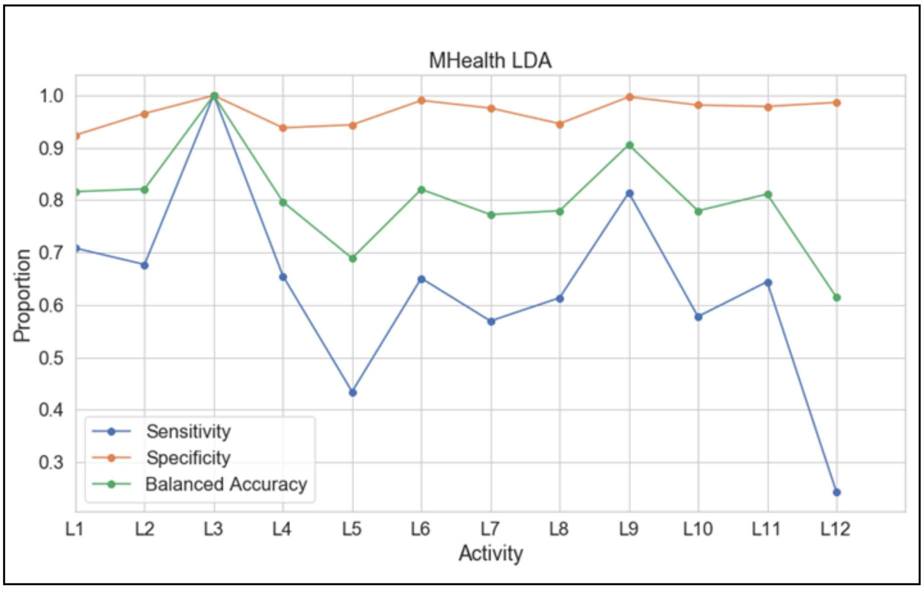
Overall Statistics

Table 4.2 LDA Overall Statistics

Accuracy	0.6576
95% CI	(0.6557, 0.6595)
No Information Rate	0.0895
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.6249

Statistics by Class

Sensitivity, Specificity and Balanced Accuracy



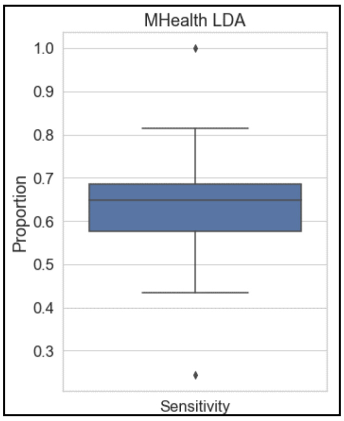
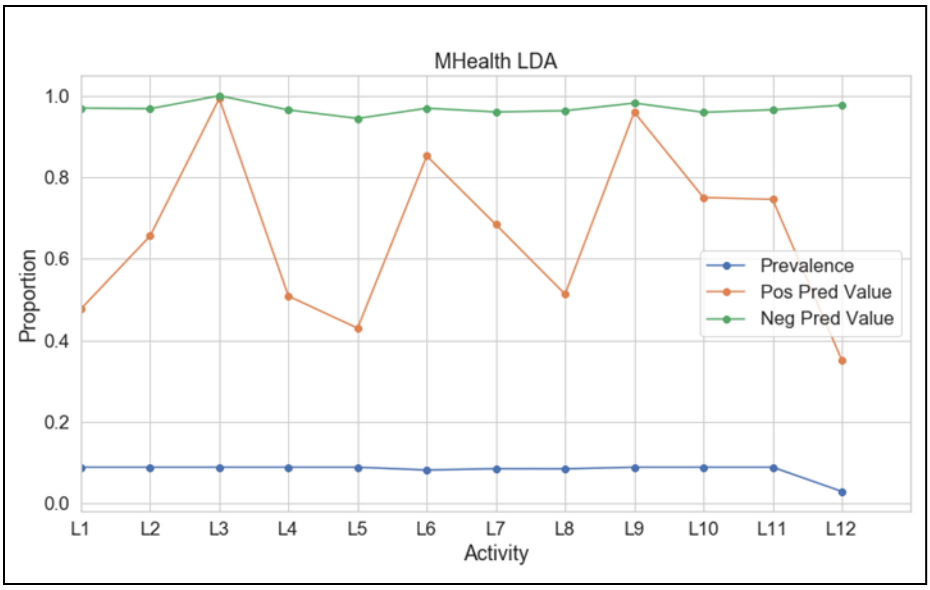


Figure 4.6 Sensitivity, Specificity and Balance Accuracy on LDA Algorithm

Like CART, the sensitivity by class on LDA algorithm also varies for the range of activities. L3 and L12 are outliers in sensitivity. Similar to CART, LDA has high sensitivity for L3 (Lying Down) and better for L9 (Cycling), however, the median is at around 65%. The specificity is not much applicable as stated above. Balanced

Accuracy is only average of Sensitivity and Specificity. The worst identified class is L12 (jump front & back) as was the case in CART, however, it may be attributed to less prevalence as compared to all other classes.

Predicted Values



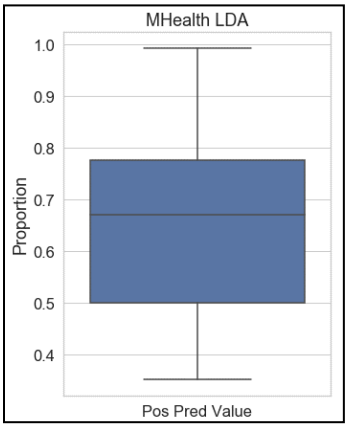


Figure 4.7 Positive and Negative Predicted Values on LDA Algorithm

There is a great variation among classes in Positive Predicted Values as shown in the Box plot. The Line Plot also shows that for L3 (lying down) and L9 (cycling) the positive prediction is high, but the median is around 0.67 with IQR close to 0.3.

Prevalence, Detection Rate and Detection Prevalence

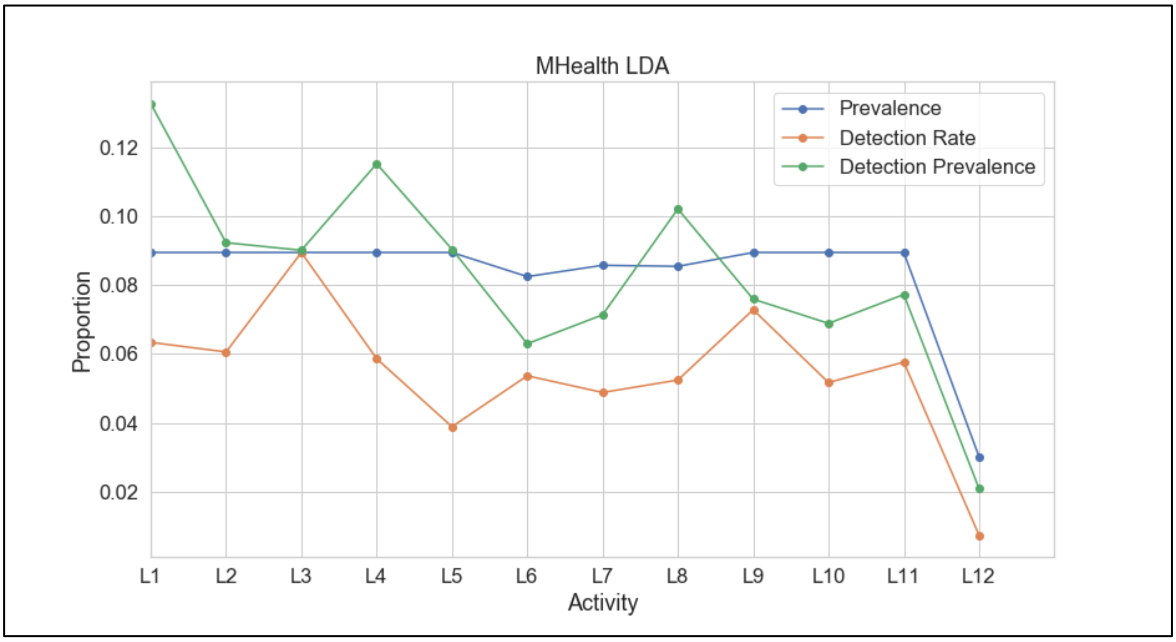


Figure 4.8 (a) Prevalence, Detection Rate and Detection Prevalence on LDA Algorithm

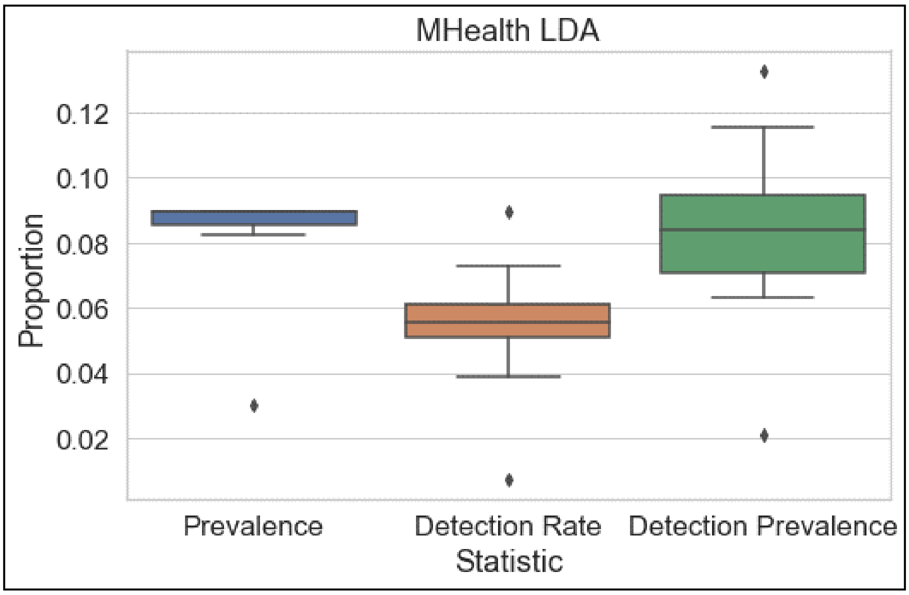


Figure 4.8 (b) Prevalence, Detection Rate and Detection Prevalence on LDA

The Detection Prevalence is higher (median ~ 85%) as compared to Detection rate (median ~ 58%) which shows a lot of false identifications for classes. L3 (lying down) is best identified class with all Prevalence, Detection Rate and Detection Prevalence coinciding.

Confusion Matrix

Figure 4.9 Confusion Matrix of LDA Algorithm

The confusion matrix shows greater details with and overall accuracy of 65%. L12 (jump front & back) is only 24% correctly identified with rest going to almost all other classes (activities).

4.3.3 Support Vector Machine (SVM)

The SVM algorithm for classification in this study uses radial-basis kernel (Gaussian). Since R language implementation ksvm is used to build this model, the hyperparameter kpar is set to "automatic" which uses the heuristics in sigest function to calculate a good sigma value for the Gaussian RBF.

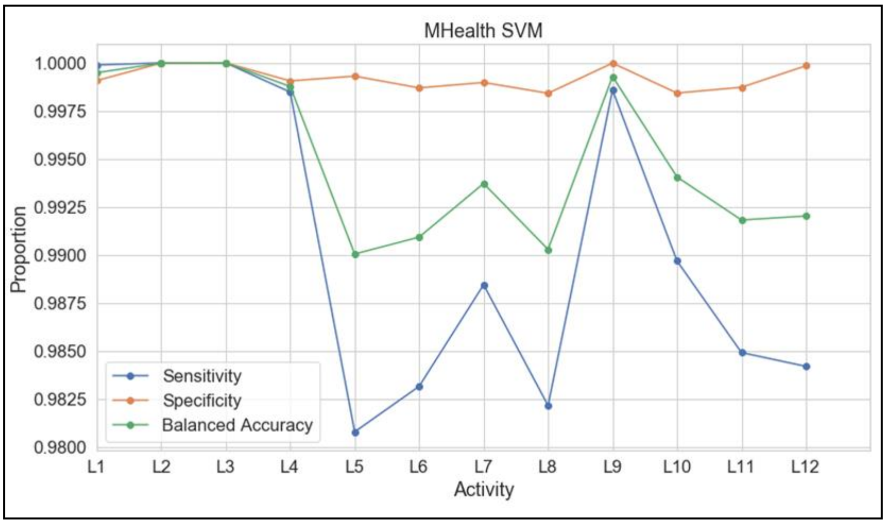
Overall Statistics

Table 4.3 SVM Overall Statistics

Accuracy	0.9914
95% CI	(0.9908, 0.9919)
No Information Rate	0.0895
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.9905

Statistics by Class

Sensitivity, Specificity and Balanced Accuracy



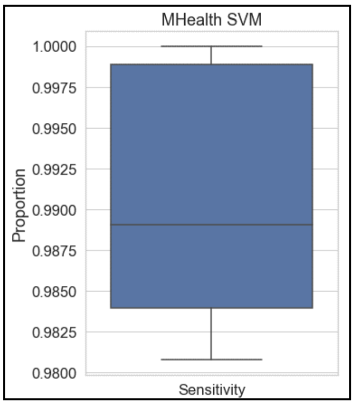
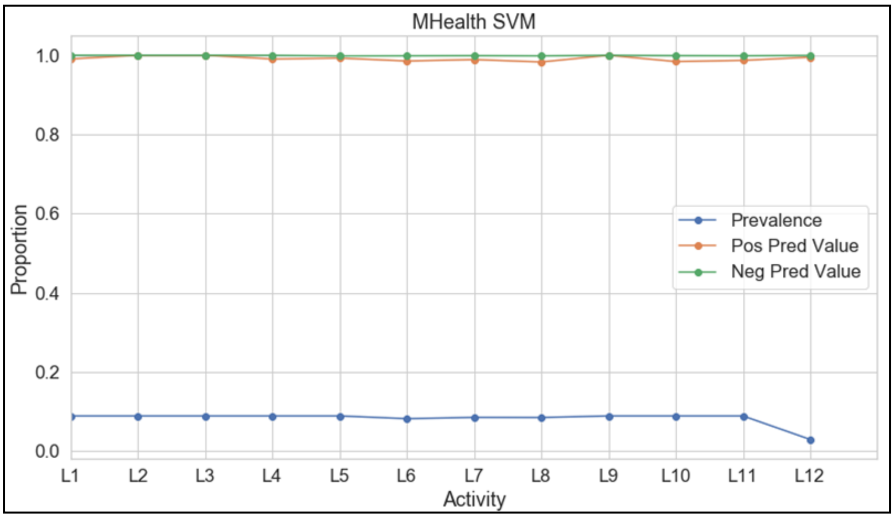


Figure 4.10 Sensitivity, Specificity and Balance Accuracy on SVM Algorithm

SVM is a very high-performance algorithm on MHEALTH test data. It has an accuracy of greater than 99%. The sensitivity is also very high. Even the minimum value is greater than 98% for L5 (Climbing Stairs). L1 (standing still), L2 (sitting and relaxing), L3 (lying down) and L9 (cycling) classes have very high sensitivity i.e. better prediction of the class close to 100%.

Predicted Values



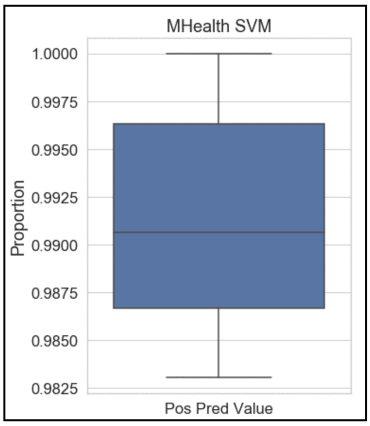


Figure 4.11 Positive and Negative Predicted Values on SVM Algorithm

The Positive Predicted Value Statistic also has 99% median for all classes. The negative predicated values in such multiclass problem with almost identical prevalence of each class would understandably be very high as was the case with specificity. Considering a prevalence of 8.5% or less for all classes, SVM shows very good performance on MHealth data.

Prevalence, Detection Rate & Detection Prevalence

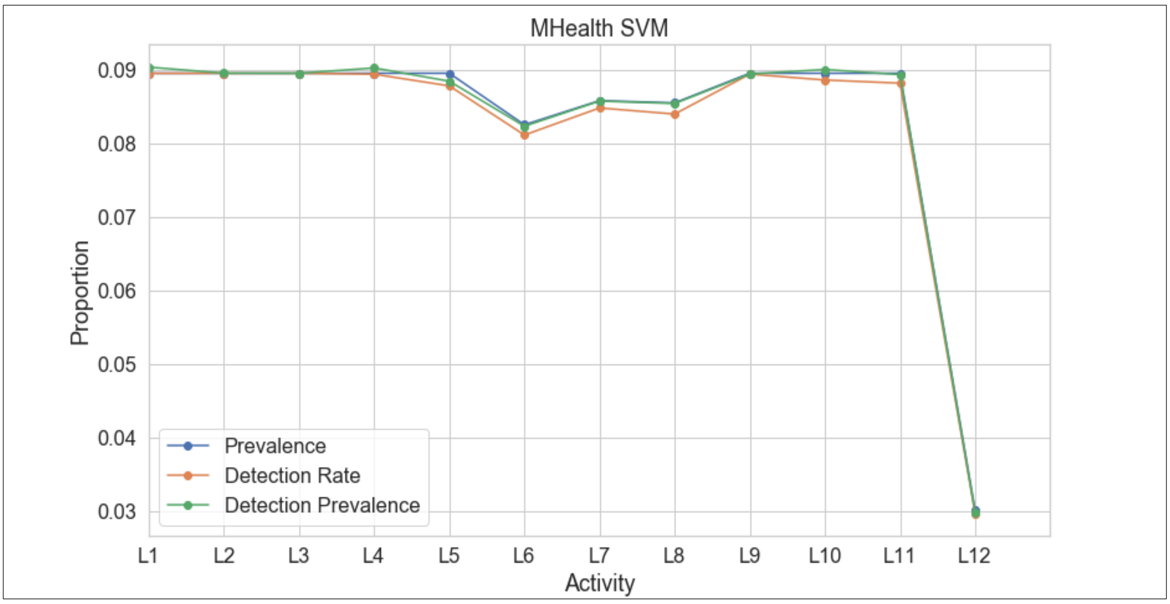


Figure 4.12 (a) Prevalence, Detection Rate & Detection Prevalence on SVM Algorithm

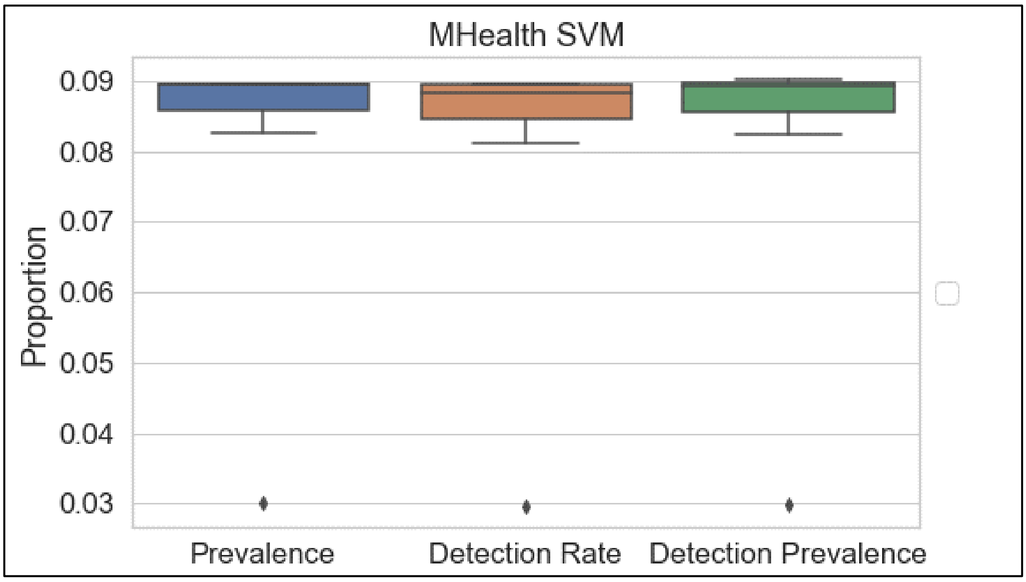


Figure 4.12 (b) Prevalence, Detection Rate & Detection Prevalence on SVM

Prevalence, Detection Rate, Detection Prevalence values are very close to each other which shows the power of model to predict the outcomes. The box plot shows and outlier which is L12 having prevalence close to 3%.

Confusion Matrix

Figure 4.13 Confusion Matrix of SVM Algorithm

SVM shows nearly diagonal matrix with some sparse values otherwise. The confusion matrix shows the details of correct prediction model by SVM with an accuracy of 99.14% on test data.

4.3.4 K-Nearest Neighbour (KNN)

In the implementation of k-Nearest Neighbours, K=5 neighbours are considered to build the model.

Overall Statistics

Table 4.4 KNN Overall Statistics

Accuracy	0.9914
95% CI	(0.9908, 0.9919)
No Information Rate	0.0895
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.9905

Statistics by Class

Sensitivity, Specificity and Balanced Accuracy

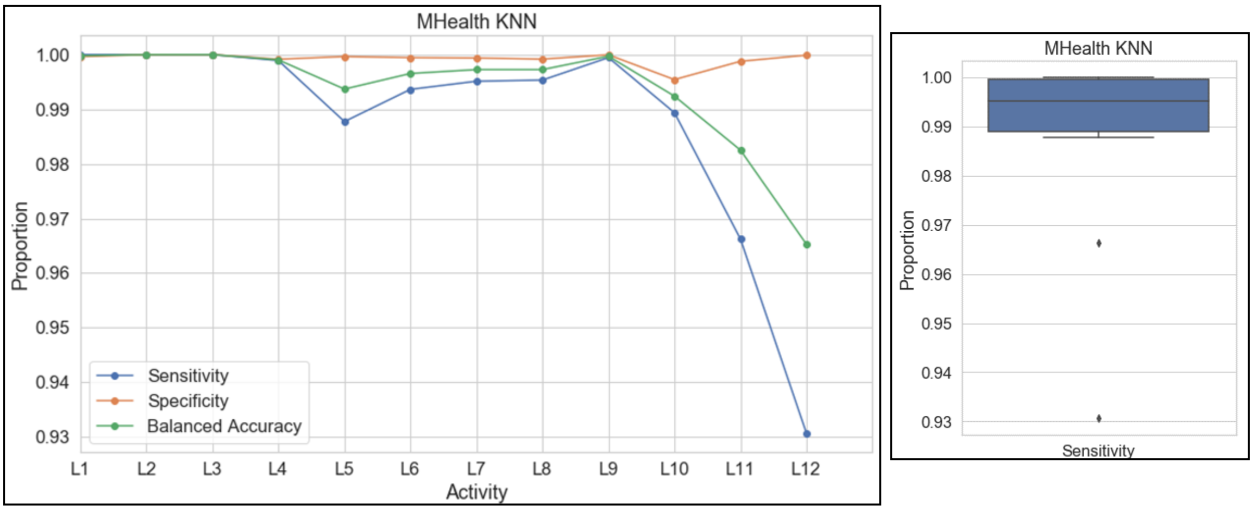


Figure 4.14 Sensitivity, Specificity and Balance Accuracy on KNN Algorithm

KNN model also predicts a very accurate and sensitive model like SVM. The sensitivity is quite high for all the classes. Even the lowest measure of sensitivity is 93% for L12 (jump front and back) and 96.7% for L11 (running).

Predicted Values

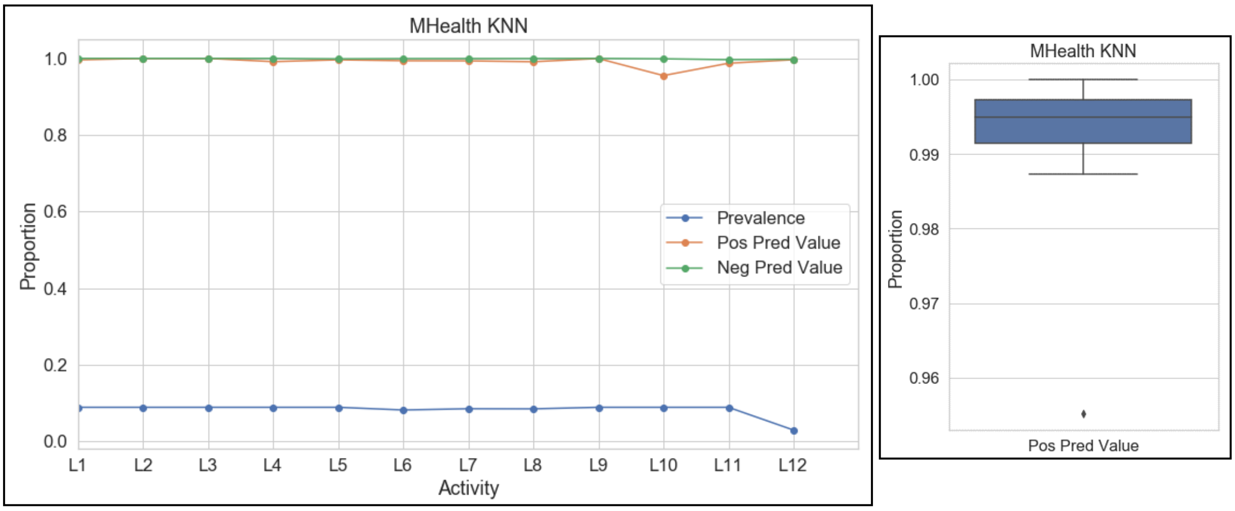


Figure 4.15 Positive and Negative Predicted Values on KNN Algorithm

The positive predicted value for all classes is quite high given small prevalence of classes in overall data. The median is greater than 99% showing very less False Positives.

Prevalence, Detection Rate and Detection Prevalence

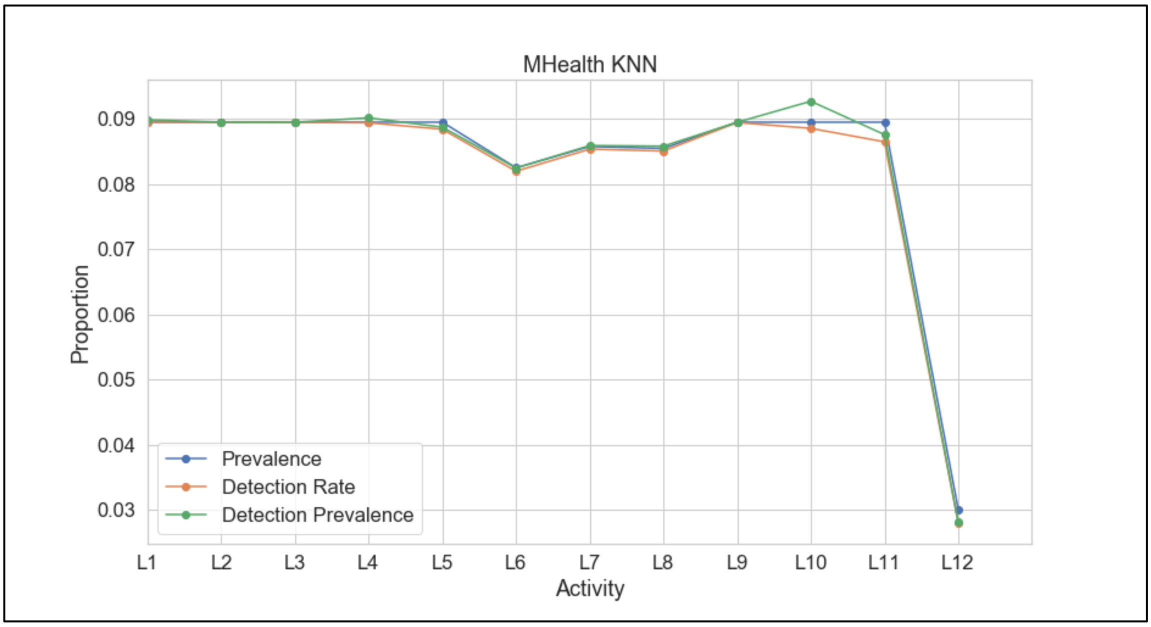


Figure 4.16 (a) Prevalence, Detection Rate and Detection Prevalence on KNN

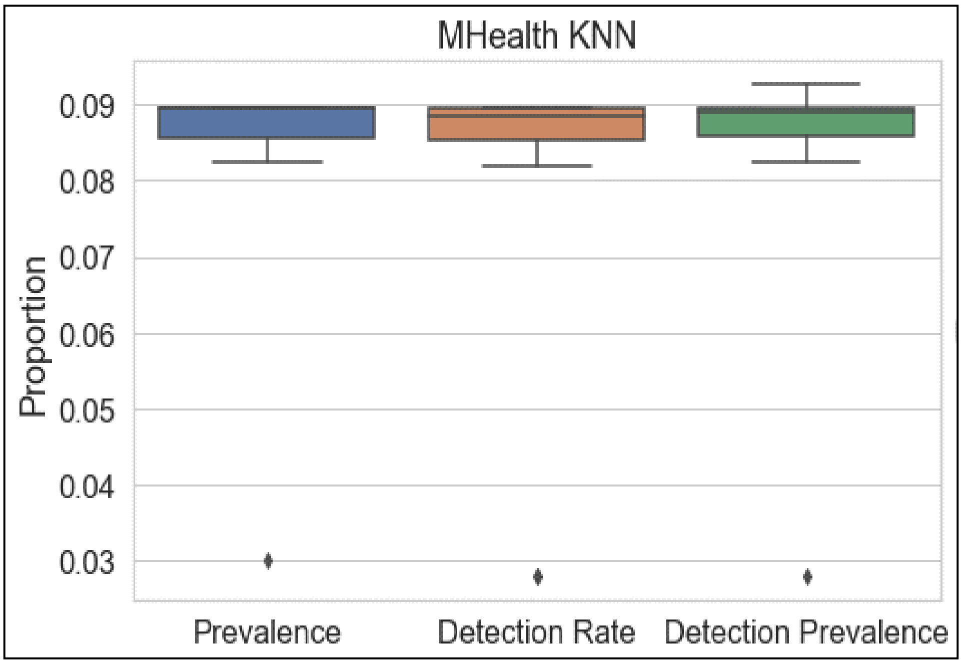


Figure 4.16 (b) Prevalence, Detection Rate and Detection Prevalence on KNN

The Detection Prevalence is slightly higher for L10 (Jogging) and L11 (Running) than the Detection Rate, however, for the rest of the classes these are quite close showing good quality of predictive models.

Confusion Matrix

Figure 4.17 Confusion Matrix of KNN Algorithm

The confusion matrix of KNN confirms the statistics with sparse diagonal matrix shape, resulting in high accuracy, high PPV and high sensitivity.

4.3.5 Random Forest (RF)

In the implementation of RF algorithm, 200 trees are grown in this study to build the model.

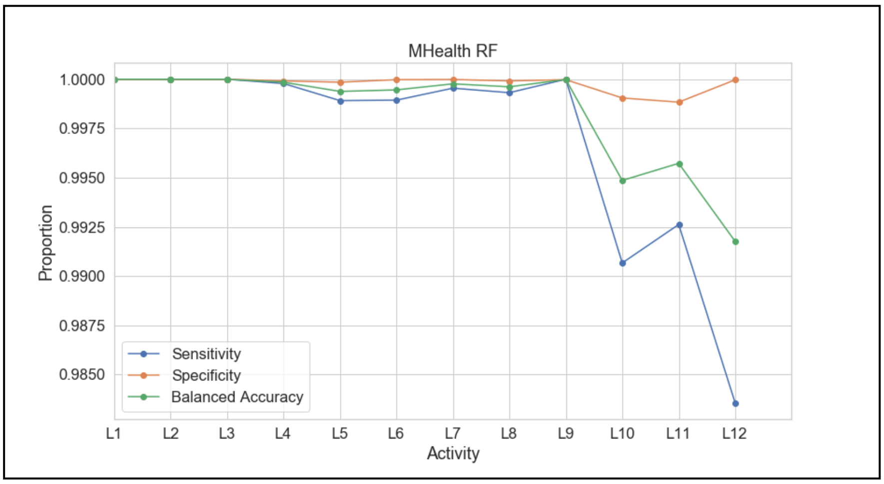
Overall Statistics

Table 4.5 RF Overall Statistics

Accuracy	0.9977
95% CI	(0.9974, 0.9979)
No Information Rate	0.0899
P-Value [Acc > NIR]	< 2.1e-16
Kappa	0.9975

Statistics by Class

Sensitivity, Specificity and Balanced Accuracy



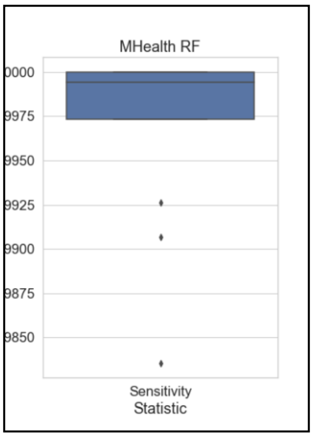
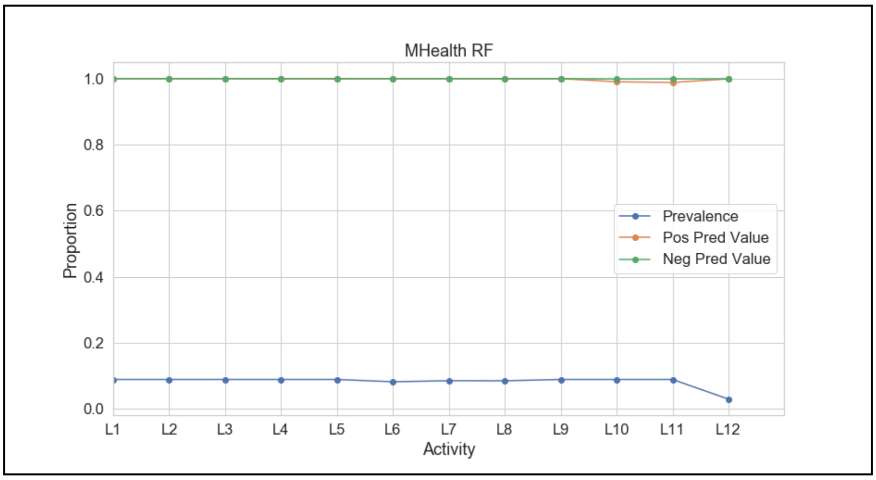


Figure 4.18 Sensitivity, Specificity and Balance Accuracy on RF Algorithm

RF model predicts very accurately and has greater sensitivity comparable or better than SVM or KNN. The sensitivity is quite high for all the classes. Even the lowest measure of sensitivity is 98.5% for L12 (jump front and back) and 99% for L10 (Jogging).

Predicted Values



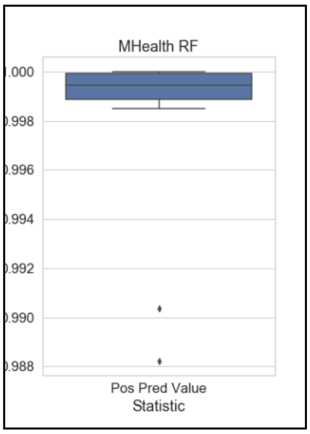


Figure 4.19 Positive and Negative Predicted Values on RF Algorithm

The positive predicted value for all classes is quite high given small prevalence of classes in overall data. The median is greater than 99.8% showing very less False Positives.

Prevalence, Detection Rate and Detection Prevalence

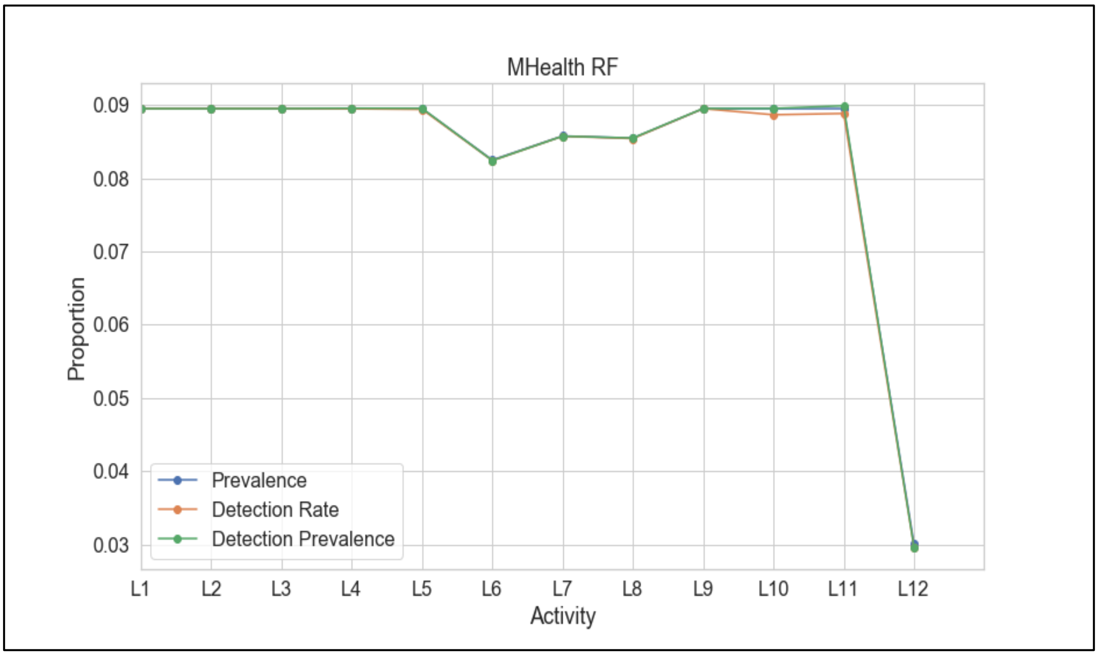


Figure 4.20 (a) Prevalence, Detection Rate and Detection Prevalence on RF

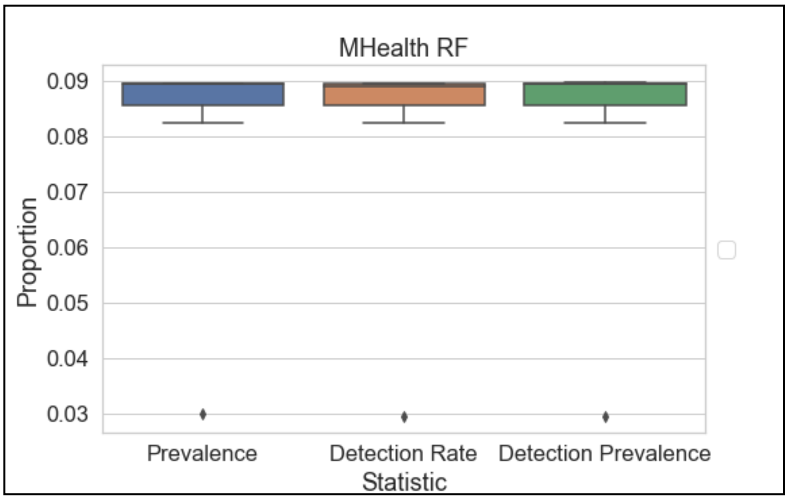


Figure 4.20 (b) Prevalence, Detection Rate and Detection Prevalence on RF Algorithm

The Detection Prevalence is very slightly higher for L10 (Jogging) than the Detection Rate, however, for the rest of the classes these are quite close to perfect quality for predictive models.

Confusion Matrix

Figure 4.21 Confusion Matrix of RF Algorithm

The confusion matrix of RF confirms the statistics with sparse diagonal matrix shape, resulting in high accuracy, high PPV and high sensitivity.

4.4 Activity Recognition through ML Algorithms on Reduced Dimensionality (PCA) MHEALTH Dataset

MHEALTH dataset has 24 dimensions where Column # 24 has activity information in it. The first 23 columns have sensors information in it. The dimensionality of Wearable Sensors Data can be very high owing to three axes information from accelerometers, gyros and magnetometers. For each Shimmer sensor, nine additional variables may be obtained. Due to this fact, the curse of dimensionality may arise which adds both computational time and complexity to classification algorithms. The dataset may be reduced through features selection or features extraction. Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA itself is designed to maximize the variance of the first 'k' components out to total 'p' components and minimize the variance of the last 'p−k' components, compared to all other orthogonal transformations. The purpose of this transformation was to be able to select first 'k' components explaining a threshold variance. The MHEALTH dataset was thus transformed using principal components analysis (PCA) technique resulting in 23 transformed principal components.

4.4.1 Selection of PCA Components for Analysis

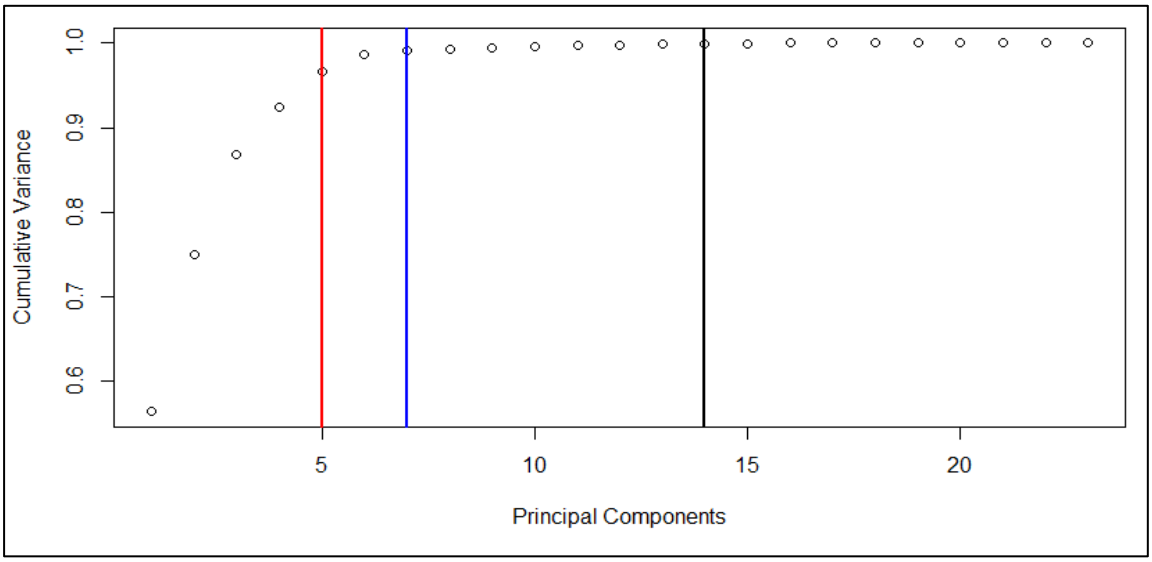


Figure 4.22 Cumulative Variance Explained by Principal Components. **Red** = 95.0%, **Blue** = 99.0%, **Black** = 99.9%.

For the purpose of analysis on reduced dimensionality dataset, PCA components were selected based upon the cumulative variance explained. As shown in the Figure 4.22, first 5 principal components explain 95% variance in the dataset, while 7 explain 99%. For 99.9% cumulative variance, 14 principal components would be required out of 23. These three thresholds were selected to compare the classifier performance of the algorithms. The reduced dimensionality datasets thus obtained are referred as PC5, PC7 and PC 14 respectively. The original dataset without any reduction is also shown for comparison in figures below and is referred as ALL_VARS. The statistics used for this comparison are overall accuracy as well as class-wise sensitivity, precision and recall.

4.4.2 Overall Accuracy

In terms of overall accuracy of all the algorithms, the following figure shows the impact of using three subset datasets passed through five classifier algorithms along with performance of these algorithms on original MHEALTH dataset with all variables included.

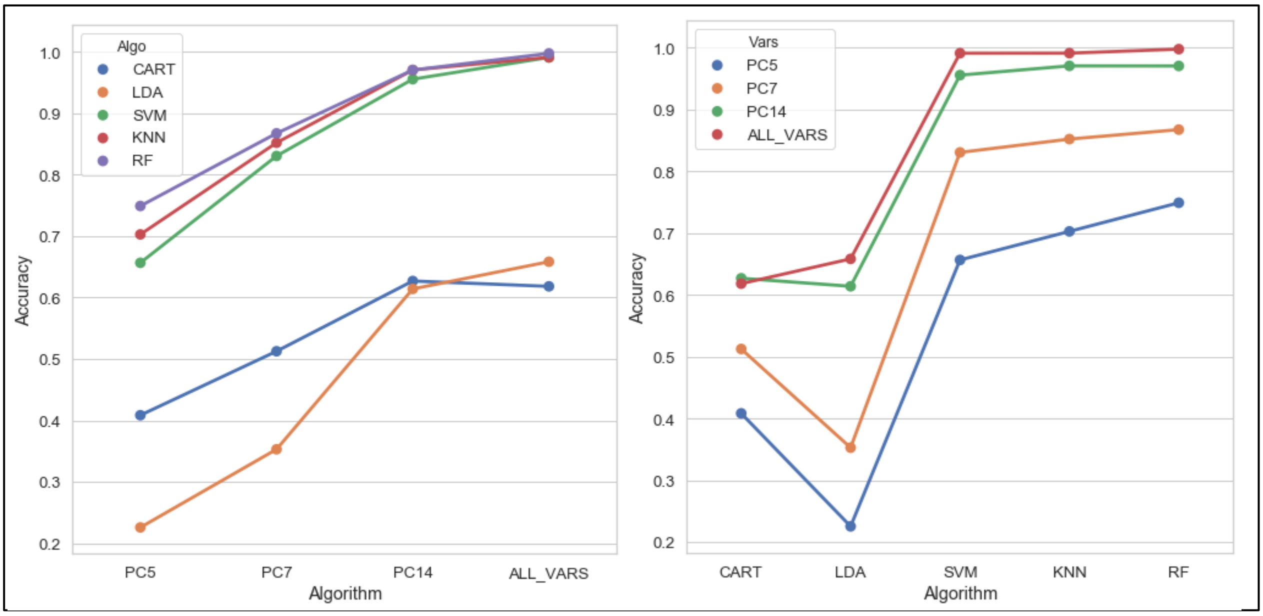


Figure 4.23: (a) Impact on Algorithmic Accuracy on using Principal Components and comparison with original dataset. **Figure 4.23: (b)** Alternate view for better algorithmic comparison on accuracy.

Figure 4.23(a) indicates RF to be best overall classifier with 5 PCs only, however, as the PCs increase the gap closes with KNN and SVM which also show better accuracies. In original dataset, these three high performance algorithms give nearly identical performance. In case of CART, the accuracy on 14 PCs dataset is slightly more than with using all the variables. **Figure 4.23(b)** flips the Variables and Algorithms and more clearly indicates the difference of performance of algorithms on four datasets. Although, using 14 PCs increases the accuracy closer to the original dataset, however, with exception of CART, the performance of all other algorithms are consistently better as more principal components are added for classification.

4.4.3 Class-wise Sensitivity on Different Algorithms

Figure 4.24 and 4.25 show the class-wise sensitivities attained by five classifier algorithms on three Principal Components datasets having 5, 7 and 14 PCs respectively. Both LDA and CART have missed some activities completely in PC5 and PC 7 datasets (CART: Climbing stairs, front elevation of arms, knees bending and jump front / back; LDA: Walking, waist bend forward, front elevation arms, knees bend, jump front / back). In all the cases, Jump front / back has the least activity recognition for all datasets while standing still, lying and cycling have best comparative recognition.

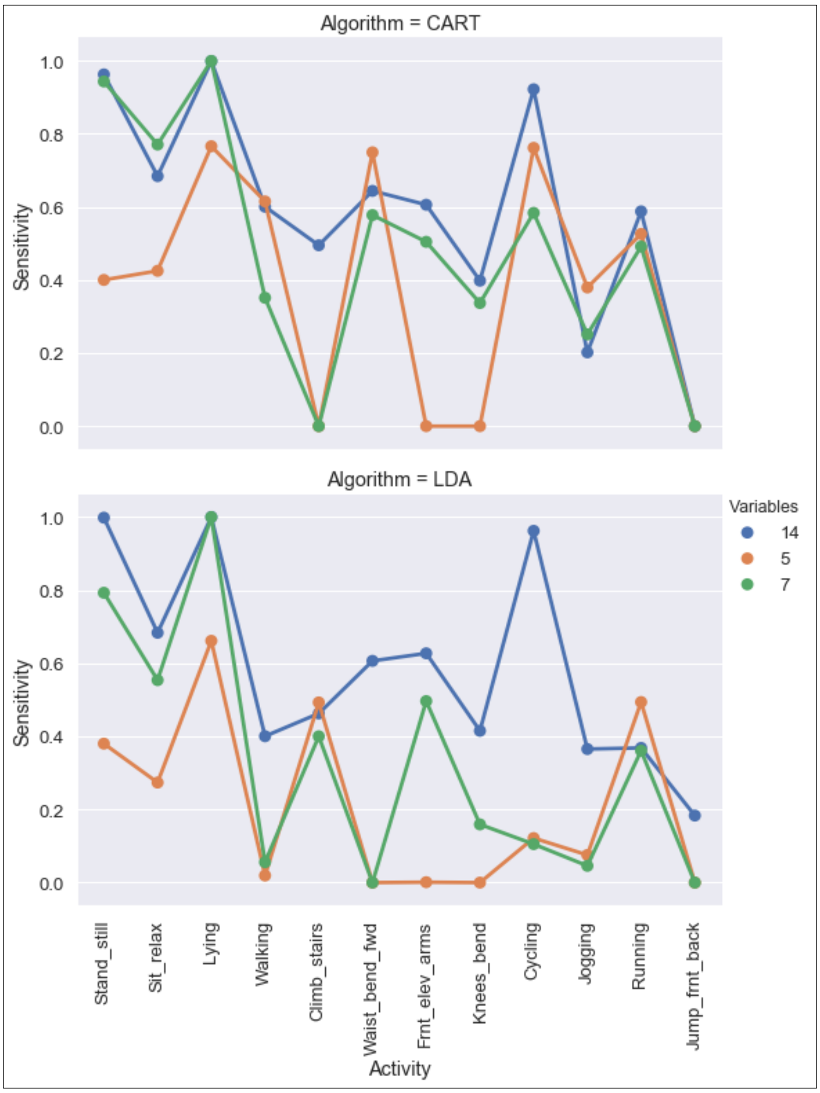


Figure 4.24 Class-wise comparison of sensitivity on CART and LDA Algorithms on PC Datasets

In Figure 4.25, the performance on higher performance algorithms i.e. SVM, KNN and RF are compared for the PCs datasets. PC14 has consistently outperformed PC7 which has outperformed PC5. This indicates that there is addition of quality information for activity recognition with increase in principal components. PC14 gives almost perfect sensitivity for number of activities in all three algorithms; however, KNN and RF have an edge over SVM. The worst-case recognition in these algorithms is also Jump front / back.

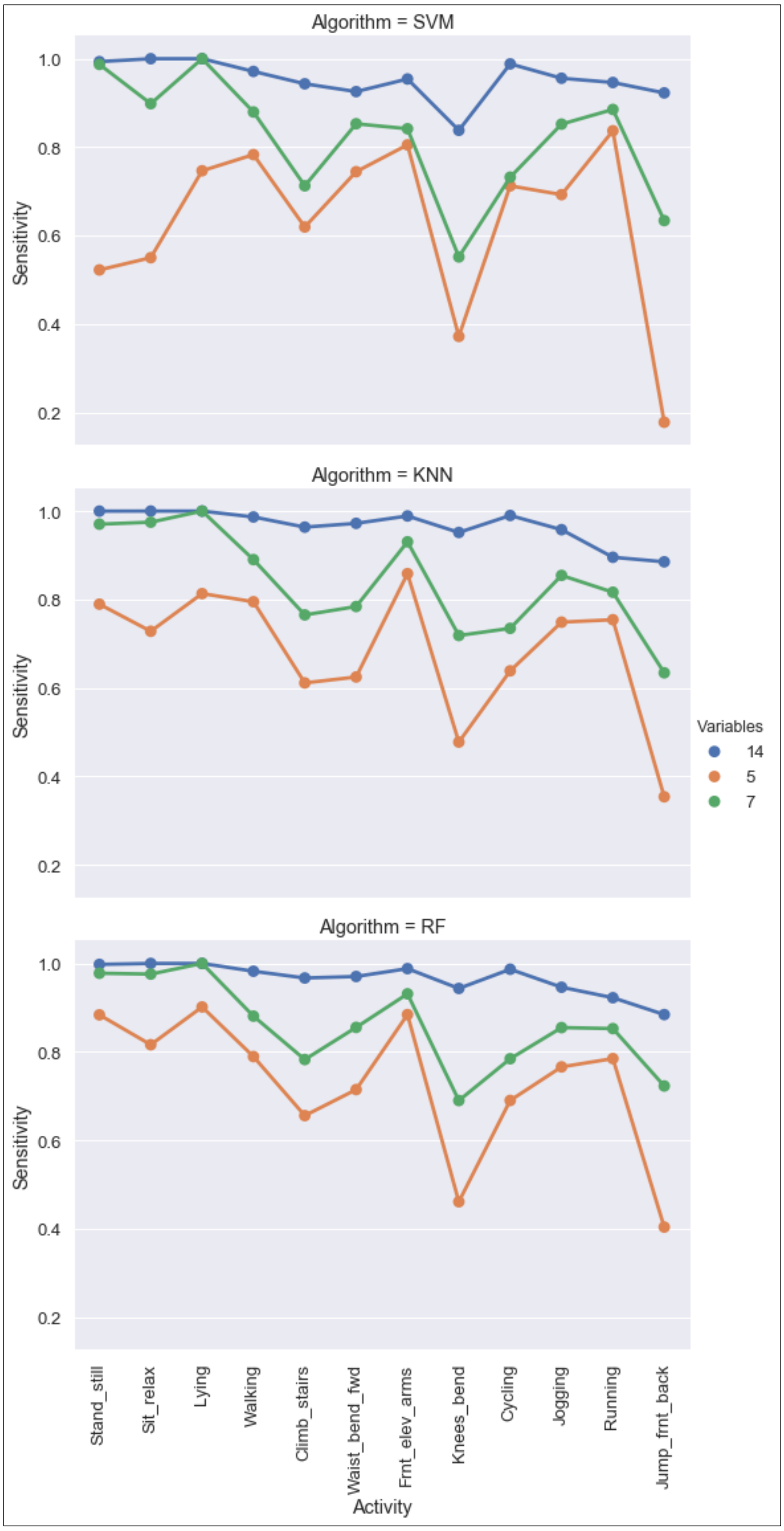


Figure 4.25 Class-wise Sensitivity of SVM, KNN and RF Algorithms on PC Datasets

4.4.4 Class-wise Precision and Recall

Precision and Recall are analysed for three higher performance classifiers i.e. SVM, KNN and RF on principal components datasets in Figure 4.26. Recall is more of a problem than precision in PC 5 dataset and is more pronounced in SVM and RF which means lesser sensitivity as compared to positive predicted value. In cases of PC 7 and PC 14, both precision and recall improve indicated by movement of points towards the upper right corner.

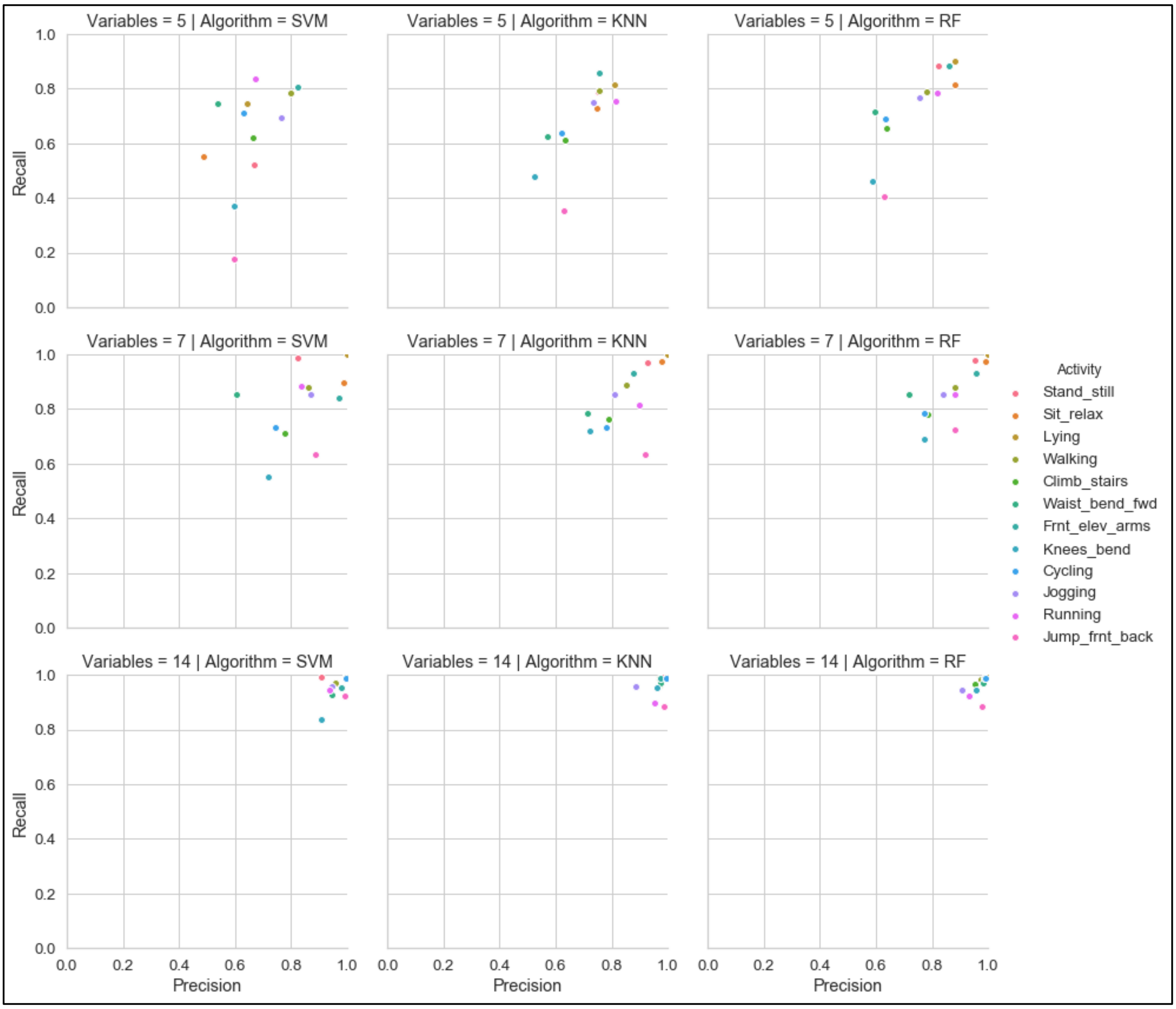


Figure 4.26 Class-wise Precision and Recall Comparison of SVM, KNN and RF Algorithms on Principal Components Dataset

4.5 Activity Recognition through ML Algorithms on Selective Sensors of MHealth Dataset

The other method for feature selection in this study is selecting a subset of sensors data for activity recognition task. Features selection helps to reduce the dimensionality of dataset and the computational time required by classifier algorithms to build models. Features selection may be obtained through filter methods or wrapper methods. Filter methods do not use any classifier in the selection process while wrapper methods, which often yield better results, use a classifier to evaluate the selected subsets based on their predictive accuracies.

4.5.1 Selection of Sensors Using Ranking by CART Algorithm

To calculate a variable importance score, CART looks at the improvement measure attributable to each variable in its role as a either a primary or a surrogate splitter. CART recursively divides the observations space and defines a piecewise constant function on the partition induced, function called predictor or classifier as appropriate. The importance score measures a variable's ability to perform in a specific tree of a specific size either as a primary splitter or as a surrogate splitter. The scores reflect the contribution each variable makes in classifying or predicting the target variable, with the contribution stemming from both the variable's role as a primary splitter and its role as a surrogate to any of the primary splitters.

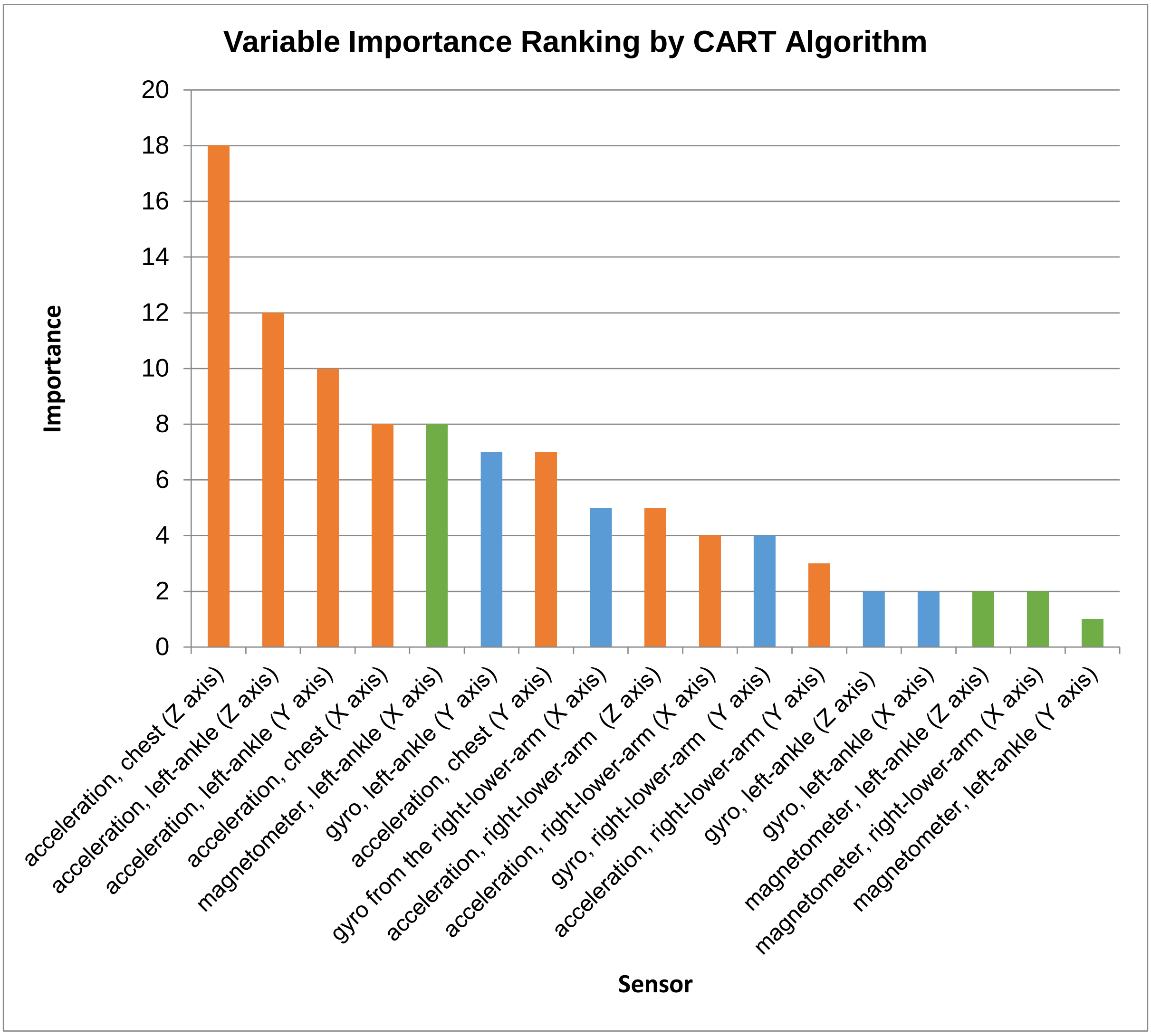


Figure 4.27 Variable Importance by CART Algorithm, Accelerometer Columns = Orange, Gyro Columns = Blue, Magnetometer Columns = Green.

Figure 4.27 shows the Variable importance attained by CART algorithm in our experimental setting. It indicates that accelerometer variables are the most important in activity recognition followed by gyros. Based upon these results two types of subsets were formed:

- Accelerometer data from chest, right wrist and left ankle sensors (total 09 Columns) abbreviated as ACC in figures.
- Accelerometer and Gyros data from right wrist and left ankle sensors (total 15 Columns) abbreviated as ACC_GYR in figures.

The classifier algorithms were applied on these datasets and also compared with full MHEALTH (23 Columns – abbreviated as ALL_VARS in figures) results described in section 4.5.

4.5.2 Overall Accuracy

The overall accuracy of five classifier algorithms on three datasets is shown in Figure 4.28. The accuracy of combine accelerometers and gyros data (15 columns) is quite close to the 23-column full MHealth dataset. This indicates that for the 12 activities recognition (table 4.2) using three Shimmer sensors, accelerometers and gyros provides the most signification information.

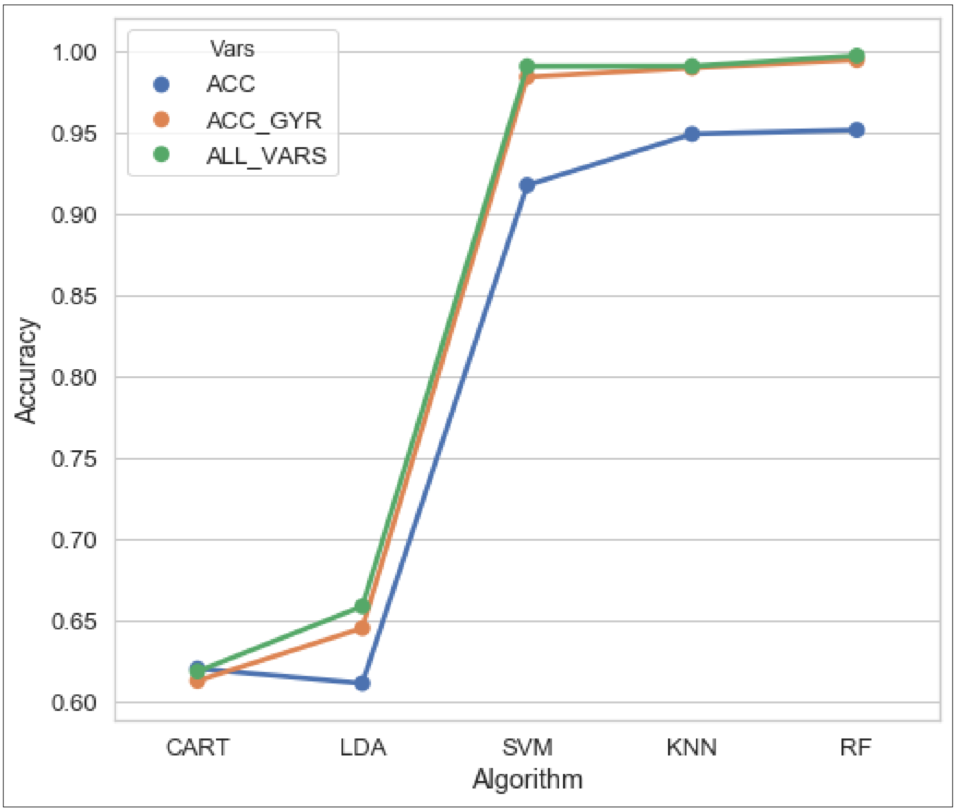


Figure 4.28 Comparison of Algorithmic Accuracy on Reduced Sensors Datasets

4.5.3 Class-wise Sensitivity on Different Algorithms

The comparison of class-wise sensitivity achieved via different algorithm on three datasets is shown in Figures 4.29 and 4.30. CART has similar sensitivity on all classes except in Sitting & Relaxing, Walking, Jogging and Running. In case of

Jump front / back, the sensitivity is 0.0 on CART in all three datasets. In case of LDA, ACC_GYR and ALL_VARS datasets have almost similar sensitivity on all classes.

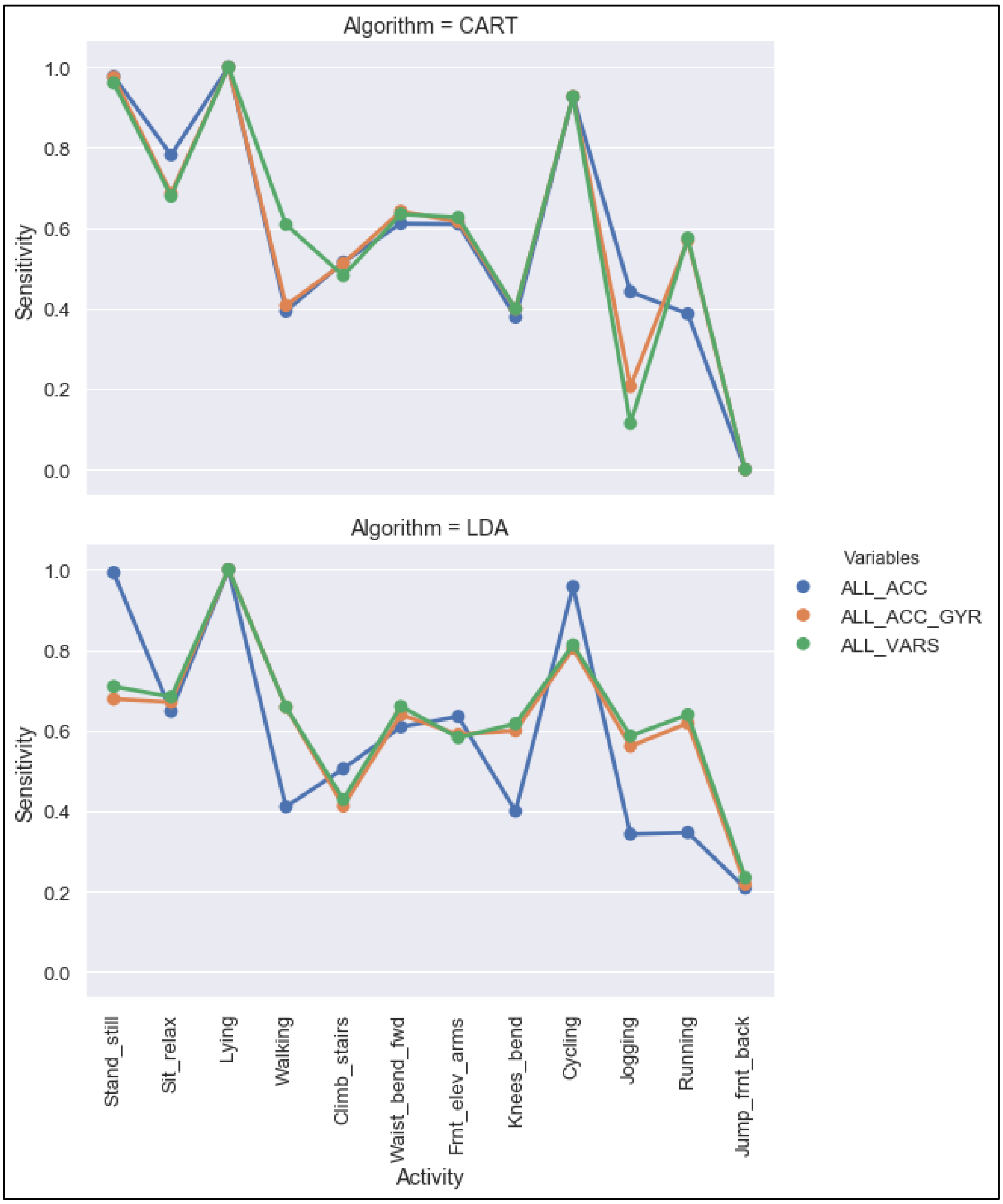


Figure 4.29 Class-wise Comparison of Sensitivity on CART and LDA Algorithms on Reduced Sensors Datasets

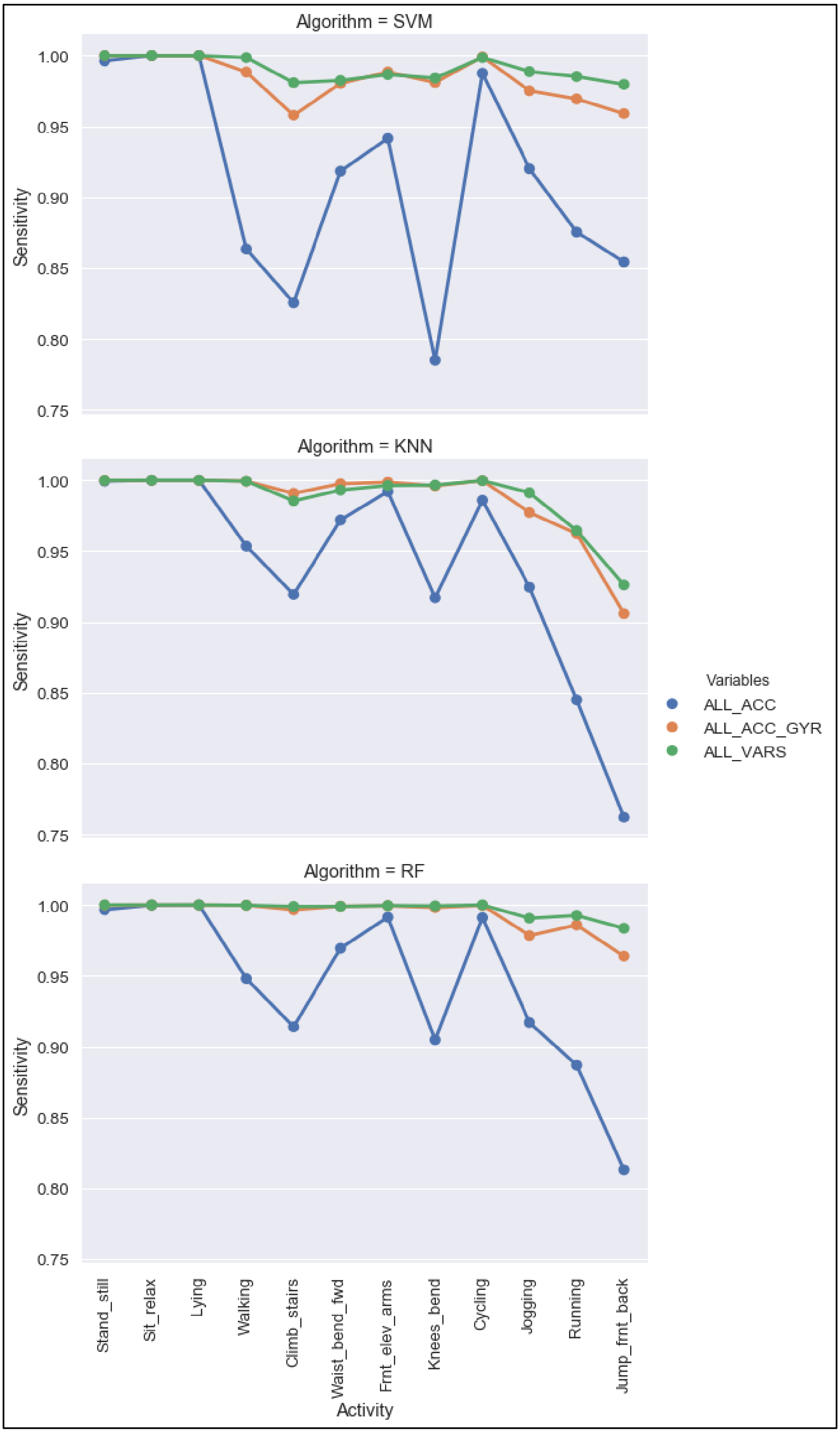


Figure 4.30 Class-wise Comparison of Sensitivity on SVM, KNN and RF Algorithms on Reduced Sensors Datasets

In Figure 4.30, the same comparison is given on higher performance SVM, KNN and RF algorithms. ACC data has obvious lesser sensitivity in many classes as compared to ACC_GYR and ALL_VARS data which have very similar sensitivities in almost all classes.

4.5.4 Class-wise Precision and Recall

Class-wise Precision and Recall are analysed for SVM, KNN and RF on ACC, ACC_GYR and ALL_VARS datasets in Figure 4.31.

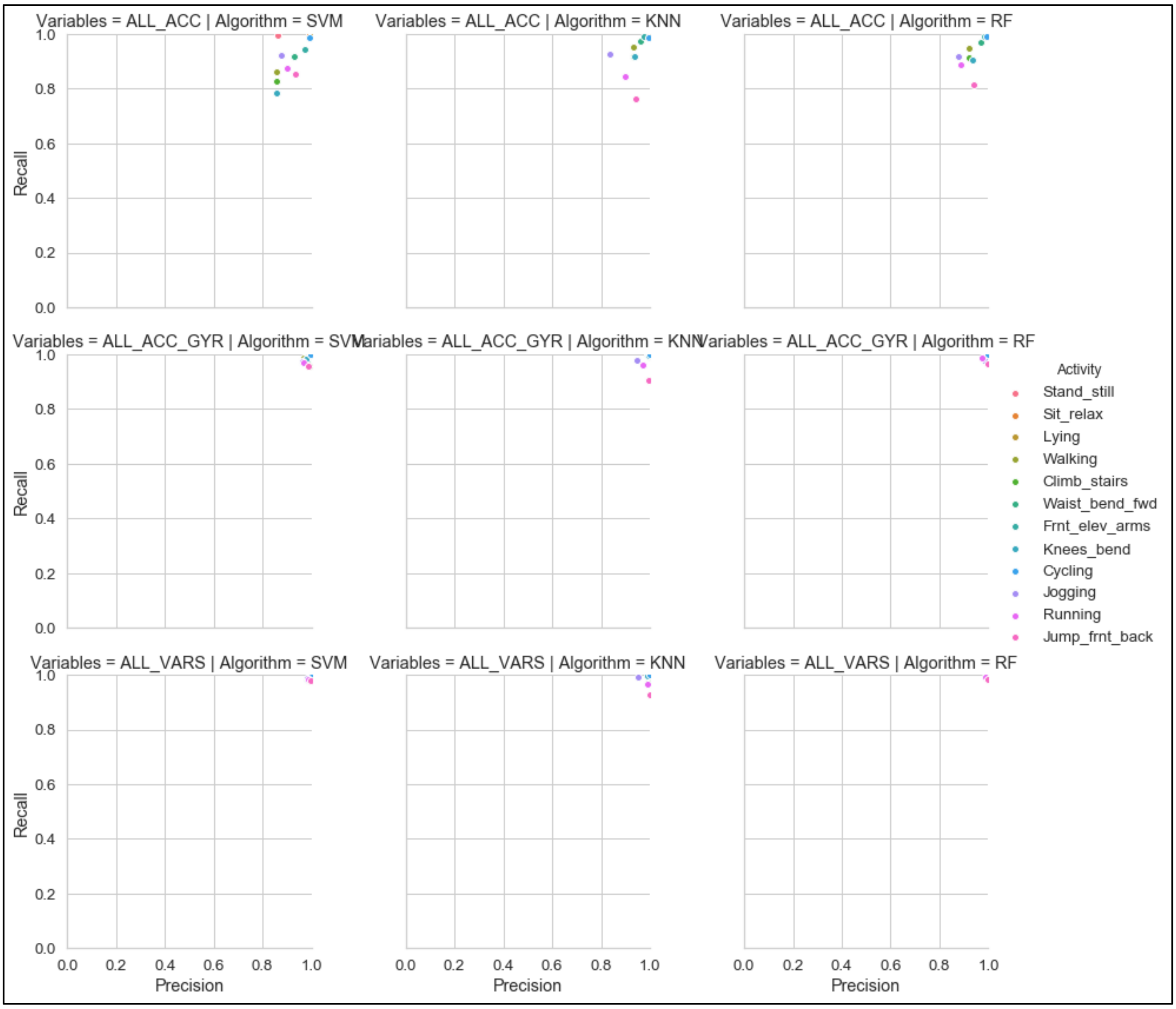


Figure 4.31 Class-wise Precision and Recall Comparison of SVM, KNN and RF Algorithms on Reduced Sensors Dataset

It shows comparable Precision and Accuracy in all three datasets and for all three algorithms. The closely packed group in ACC_GYR and ALL_VARS indicate very high Precision and Recall for all activities. ACC_GYR results indicate the significance of accelerometers and gyros over magnetometers and ECGs.

4.6 Final Note on Overall Accuracy Using All Three Approaches

A close inspection of Overall Accuracy on high performance algorithms and PC14, ACC_GYR and ALL_VARS datasets indicate that 15 columns of sensors data outperform 14 PCA columns explaining 99.9% variance in activity recognition on these algorithms. Moreover, the difference between 15 columns sensors' data results with 23 columns ALL_VARS dataset is very much minor.

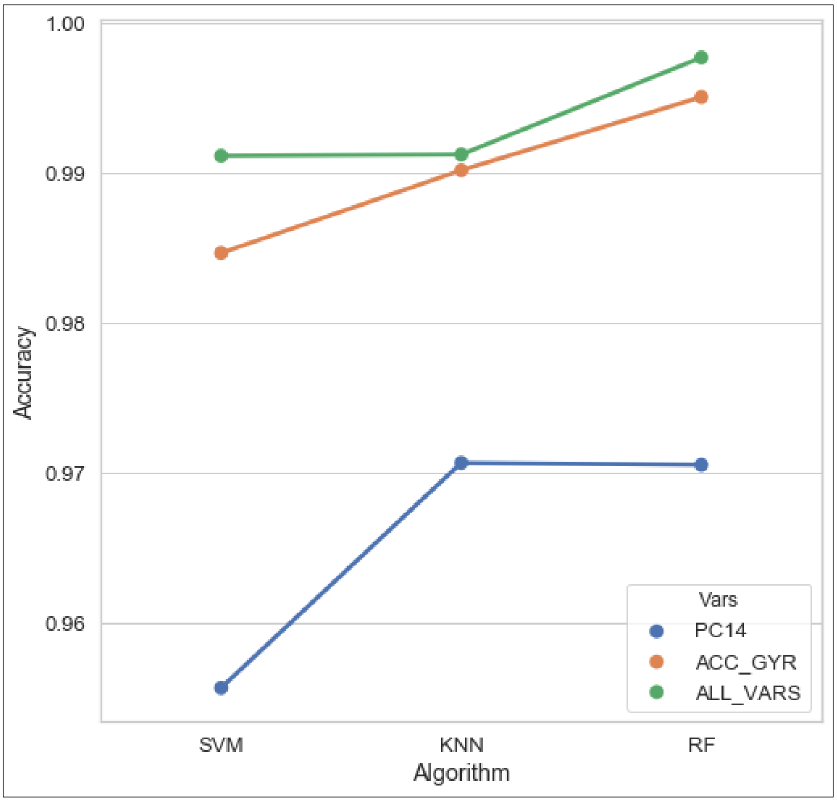


Figure 4.32: Overall Accuracy Comparison of SVM, KNN and RF on three datasets (14 principal component, accelerometer & gyros combined and full MHEALTH data)

4.7 Comparison with Other Studies on Activity Recognition

A number of studies [12][75][76][77] devoted to Human Activity Recognition from Sensors' data have narrated their use of classification algorithms for online or offline recognition. These studies have mainly used supervised algorithms; however, one study [12] has also demonstrated use of unsupervised algorithm for activity recognition. A review of these research reveals that Banos et al. [75] collected the MHealth data, proposed mobile health application framework and used J48 classification algorithm for both online and offline classification. They have reported a very high overall accuracy and have also stated class-wise sensitivity, specificity, predicted values and F-score.

Attal et al. [12] while utilizing the same MHEALTH data have used sliding window technique for input, extracted features in time and frequency domains, used wrapped methods for features selection and applied both supervised and unsupervised algorithms for classification. The details on results are however, given only on overall parameters like accuracy, recall, precision, f-score, specificity, etc. and have also given class-wise confusion matrix for k-NN and HMM.

In contrast to previous two studies Jordao et al. [76] main focus was on establishing a standardized way of analysis on wearable sensors data from multiple available datasets. The have emphasized the requirement of standardized process for both input data sampling and cross-validations. They have used Convolution Neural Network (CNN) for classification and the only metric they have used for reporting results is overall accuracy without any mention of detailed class-wise statistics.

Yogesh and Ghoneim [77] have extracted 15 features like mean, standard deviation, kurtosis etc. from 10 seconds window on MHealth data and compared the overall accuracy on activity recognition on these features with that of reduced dimensionality PCA data. They have reported increase in accuracy with PCA data, even better than that of results from raw data using SVM, Naïve Bayes (NB), J48 and RF algorithms. In a nutshell, these studies have lacked the detailed class-wise analysis and the clarity on application of classification algorithms. Moreover, except for Banos et al. [75], other studies have also used null class (marked as ‘0’ in column#24 of MHealth dataset) for model building. This null class does have some activity however, it is not labelled in the provided data.

In this thesis, the effort has been made to fill in the gaps by devising both a standard mechanism supported by a custom-built data analysis tool, so that experiments can be repeatable. This also provides detailed results on overall statistics and class-wise statistics with visual presentations in form of graphs to infer the results and model performance. This study also provides significance of particular sensor types (accelerometer, gyros, magnetometer, ECG) in activity recognition using classification algorithms. This study also establishes that classification is better with selected sensors information from raw data than with using principal components explaining even 99.9% variance. Finally, all the results are presented using novel visualizations comparing various performance metrics across algorithms and across datasets.

4.8 Discussion & Findings

The thesis describes complete wearable sensors data analysis process; from data loading to classification including features selection and extraction processes. Different ML approaches employed for classification of human activities were discussed in detail in Chapter 2, the research methodology was explained in Chapter 3, while detailed evaluation and results analysis are explained in Chapter 4. The comparative analysis is presented using well-known machine learning classification algorithms (Linear Discriminant analysis, Classification & Regression Trees, k-Nearest Neighbours, Random Forest and Support Vector Machine) applied on a well-established MHEALTH Wearable Sensors’ dataset. Both, raw data and extracted/selected features are used as inputs for the classifiers. The different classification approaches are compared in terms of the recognition of twelve activities in MHEALTH dataset.

During the course of research, following significant outcomes / findings were concluded and are presented in next paragraphs for future guidance:

- CART ML algorithm used for data analysis did not perform well as far as its accuracy is concerned but its feature ranking (variable importance) attribute was quite helpful during feature selection process.
- The results further highlight that the feature selection done through ranking algorithm (CART) was much accurate as compared to feature extraction done through PCA techniques in terms of classifiers' performance. PCA provided a dataset with 99.9% variance when 14 principal components were selected whereas original accelerometer and gyro data comprised 15 features. Later category of data (15 features) outclassed the PC14 data consistently in all high-performance classifiers.
- During analysis, it was observed that out of the five classifiers used, the performance of SVM, KNN and RF on data set was most accurate whereas LDA and CART did not perform well.
- Another very important aspect of the analysis was predictions based upon specific sensors' features or combination of features. During feature selection when features based upon selective sensors were analysed for identifying their relative significance; the Accelerometer & Gyros proved to be most instrumental for activity recognition process.
- Dimensionality reduction achieved through PCA technique is found quite useful and time saving in case of huge sensors data. Comparative analysis of ML classifiers based upon 14 principal components out of 23 components provided a good accuracy closer to the one achieved on actual dataset.
- Based upon their cumulative variance using PCA, three thresholds were selected to compare the classifier performance of the algorithms. After dimensionality reduction through PCA it was found that first 05 principal components explain 95% variance in the dataset, while 07 explain 99% and for 99.9% cumulative variance 14 principal components would be required out of 23.
- Class-wise analysis shows that out of total 12 classes (activities), class-12 (Jump front / back) has the least sensitivity value. It also has the least prevalence and detection rate value as compared to rest of the classes.

4.9 Summary

The analysis of MHEALTH sensors' data shows that SVM, KNN and RF performs with very high accuracy and other class-wise statistics while LDA and CART are not suitable predictive ML algorithms for activity recognition from MHealth data. The significant sensors are accelerometers and gyroscopes as indicated through variable importance from CART algorithm and confirmed via results from high performance classifier algorithms. In dimensionality reduction, the application of classifiers on reduced dataset obtained from features selection through ranking algorithm gives better classification than reduced dimensionality dataset obtained from PCA technique. It also shows the ease and usefulness of the developed tool and visualization scripts which helps in quick analysis of the algorithms on classification data.

Chapter 5
Conclusion & Future Work

5.1 Thesis Work Summary

In contrast to other studies as mentioned in Chapter 1 Section 1.3 an effort was made to fill in the gaps by devising a standard process supported by a custom-built Data Analysis Tool so that the analysis can be repeatable with different configurations. The thesis describes complete wearable sensors data analysis process; from data loading to classification including features selection and extraction processes. The comparative analysis with detailed results on overall statistics and class-wise statistics is presented using well-known machine learning classifier algorithms (Linear Discriminant analysis, Classification & Regression Trees, k-Nearest Neighbours, Random Forest and Support Vector Machine) applied on a MHEALTH dataset. Furthermore, it has also provided significance of particular sensor types i.e. accelerometers, gyros, magnetometer and ECG in activity recognition process. It has also been established in the study that classification is slightly better with selected sensors from raw data than after PCA even selecting 14 columns out of 23 explaining 99.9% cumulative variance. Finally, all the results are presented using graphs comparing various performance statistical metrics across algorithms and datasets.

5.2 Contributions

The major contributions of this research work are briefly given below:

- The existing research on wearable sensors data for activity recognition encompassed a process which is mostly non-repeatable or non-reproducible. The work done in this thesis is fully reproducible as complete details regarding experimental setup and process have been mentioned. The custom-built Data Analysis Tool used has also been made using open source languages i.e. R-Shiny and Python.
- Although, ML algorithms being used during 'classification/ activity recognition' stage have been mentioned along with its results in different wearable sensors' data research works but very few have discussed

statistical comparative analysis of these approaches and that too in cursory manner. In this thesis, a detailed comparative analysis is presented for five multi-class classifier algorithms namely; Linear Discriminant Analysis, Classification & Regression Trees, k-Nearest Neighbours, Random Forest and Support Vector Machines as applied on a MHEALTH dataset. The comparison is made on overall accuracies, class-wise sensitivity and specificity of each algorithm, class-wise detection rate and detection prevalence in comparison to prevalence of each class, positive and negative predicted values etc. Furthermore, detailed confusion matrix of each algorithm on full dataset as well as class-wise precision and recall statistics on different levels of reduced datasets are presented.

- In contrast to previous studies, the thesis has explored the activity recognition results of MHEALTH sensors data in three different ways of inputs to classifier algorithms. These are original dataset; extracted principal components with three different cumulative variance thresholds; and reduced information datasets from selective sensors. All the resultant statistics have been compared through novel visualization techniques for ease of understanding and inference.
- Critical assessment was provided on relative significance of sensor(s) i.e. accelerometers, accelerometers along with gyros and all the sensors viz-a-viz multiple physical activities by exploring and analysing multiple ML algorithms.
- The development of a custom-made Data Analysis Tool used for wearable sensors data for the purpose of activity recognition is also a novel approach adopted to support this study and to assist further researchers doing similar studies. The Analysis Tool developed is generic and can be utilised in the study of other datasets. This analysis tool has greatly reduced the time to conduct analyses and would help future researchers to focus on the original problem being faced instead of the developmental effort. Low code, no code paradigm was kept in focus while implementation.

5.3 Research Limitations

During the research and analysis phases certain issues and limitations were faced which are discussed below:

- Activity Recognition through wearable sensors data analysis requires more research to reach its full potential. Comparison between different approaches for activity recognition is hindered and becomes unquantifiable as each researcher uses a different dataset for activity recognition.
- It is evident that the lack of large and realistic sensors datasets for AR is a significant challenge that needs to be addressed. An ideal dataset should cover several topics, including diversity in human poses for the same activity, a wide range of ground truth labels, and variations in data collection and quality.
- Most studies including this use classification methods that are trained offline, thereby making the training process static. These systems may not adapt to new users. The performance of such statically trained classifiers is dependent on the type of users on which they are being tested and can affect the recognition performance. This problem can be solved by providing an online training option in these implementations. With the help of such an option, users can adapt these classifiers according to their own needs, making it more personalized.
- The data loading and analysis requires huge processing power especially in case of big sensors data. The analysis process can be enhanced with the help of parallel processing techniques and inclusion of big data frameworks like Hadoop and Spark.
- The research should be reproducible. This means that the implementation details of the employed classification algorithms and pre-processing steps should be explicitly described. Different implementations of the same algorithm can lead to different evaluation results, which is important for comparison purposes.
- Wearable sensors will only become even more ingrained in the world. A future where healthcare is revolutionized by sensors that not only detect vitals but potentially predict medical conditions is entirely conceivable.

However, that future will only happen if it is built on a solid foundation where wearable sensor data is accurate, valid and made useful for everyone.

5.4 Future Work

There is lot of potential for future research in the field of human activity recognition using different kinds of wearable sensors data.

- The focus of this research is to address sensor data analysis issues by developing such an analysis tool that loads the data, runs different ML algorithm on the data, outputs the result in the form of different statistical performance matrices and provides visualizations in the form of graphs. The tool can be enhanced with more features to help pre-processing of data, addition of more ML algorithms using different settings for each algorithm and graphical depiction based upon different set of features and attributes.
- Activities recognized in existing systems have been simple and atomic, which could be a part of more complex composite behaviours. Recognition of composite activities can enrich context awareness. There is also a great research opportunity to recognize overlapping and concurrent activities.
- There is a need for publicly available and widely acceptable, benchmark dataset from wearable sensors covering many activities. Currently, the datasets are developed and labelled for single activity at a time. The actual human activities pattern is complex and composite. There is a need for research into covering these complex patterns into datasets. The activity recognition may also be studied for overlapping and concurrent activities.
- Deep learning techniques like convolution neural networks (CNNs) have been used for activity recognition problems in studies, however, recurrent neural networks (RNNs) also seem have a great potential for online activity recognition. The concept of ‘word embeddings’ may also be explored to have similar kind of implementations in context of activity recognitions and this may ultimately result into recognition systems able to predict actions before they take place. Such a development would be revolutionary in many fields.

List of References

[1] Sazanow, E., and M. R. Neuman. "Wearable Sensors: Fundamentals Implementation and Applications." *Academic, Cambridge, MA, USA* (2014).

[2] Bagot, K. S., Stephen Augustus Matthews, M. Mason, Lindsay M. Squeglia, J. Fowler, K. Gray, M. Herting et al. "Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health." *Developmental cognitive neuroscience* 32 (2018): 121-129.

[3] http://archive.ics.uci.edu/ml/datasets/mhealth+dataset

[4] Chertchom, Prajak. "A comparison study between data mining tools over regression methods: Recommendation for SMEs." In *2018 5th International Conference on Business and Industrial Research (ICBIR)*, pp. 46-50. IEEE, 2018.

[5] S. Seneviratne et al., "A Survey of Wearable Devices and Challenges," in IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2573-2620, Fourthquarter 2017, doi: 10.1109/COMST.2017.2731979.

[5a] Jovic, Alan, Karla Brkic, and Nikola Bogunovic. "An overview of free software tools for general data mining." In *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1112-1117. IEEE, 2014.

[6] Rodgers, Mary M., Vinay M. Pai, and Richard S. Conroy. "Recent advances in wearable sensors for health monitoring." *IEEE Sensors Journal* 15, no. 6 (2014): 3119-3126.

[7] Edwards, John. "Wireless sensors relay medical insight to patients and caregivers [special reports]." *IEEE Signal Processing Magazine* 29, no. 3 (2012): 8-12.

[8] Montgomery, Kathryn, Jeff Chester, and Katharina Kopp. "Health wearables: ensuring fairness, preventing discrimination, and promoting equity in an emerging Internet-of-Things environment." Journal of Information Policy 8 (2018): 34-77.

[9] Kress-Rogers, E. "Biosensors and electronic noses for practical applications." *Handbook of Biosensors and Electronic Noses, Medicine, Food, and the Environment* (1997): pp. 3-39.

[10] Ortiz, Jorge Luis Reyes. *Smartphone-based human activity recognition*. Springer, 2015.

[11] Jovanov, Emil, Aleksandar Milenkovic, Chris Otto, and Piet C. De Groen. "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation." *Journal of NeuroEngineering and rehabilitation* 2, no. 1 (2005): 6.

[12] Attal, Ferhat, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. "Physical human activity recognition using wearable sensors." *Sensors* 15, no. 12 (2015): 31314-31338.

[13] Patel, Shyamal, Hyung Park, Paolo Bonato, Leighton Chan, and Mary Rodgers. "A review of wearable sensors and systems with application in rehabilitation." *Journal of NeuroEngineering and Rehabilitation* (2012).

[14] Aroganam, Gobinath, Nadarajah Manivannan, and David Harrison. "Review on wearable technology sensors used in consumer sport applications." *Sensors* 19, no. 9 (2019): 1983.

[15] Hildeman, Anders. "Classification of epileptic seizures using accelerometers." (2011).

[16] Salazar, António José, Ana Sofia Silva, C. M. Borges, and Miguel Velhote Correia. "An initial experience in wearable monitoring sport systems." In *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*, pp. 1-4. IEEE, 2010.

[17] Passaro, Vittorio, Antonello Cuccovillo, Lorenzo Vaiani, Martino De Carlo, and Carlo Edoardo Campanella. "Gyroscope technology and applications: A review in the industrial perspective." *Sensors* 17, no. 10 (2017): 2284.

[18] Brunner, Thomas, Jean-Philippe Lauffenburger, Sébastien Changey, and Michel Basset. "Magnetometer-augmented IMU simulator: In-depth elaboration." *Sensors* 15, no. 3 (2015): 5293-5310.

[19] Aroganam, Gobinath, Nadarajah Manivannan, and David Harrison. "Review on wearable technology sensors used in consumer sport applications." *Sensors* 19, no. 9 (2019): 1983.

[20] Anuva. A Resource Guide to Wearable Device Sensors. Available online: https://anuva.com/blog/a-resourceguide-to-wearable-device-sensors/ (accessed on 12 October 2018).

[21] Aroganam, Gobinath, Nadarajah Manivannan, and David Harrison. "Review on wearable technology sensors used in consumer sport applications." *Sensors* 19, no. 9 (2019): 1983.

[22] Tamura, Toshiyo, Yuka Maeda, Masaki Sekine, and Masaki Yoshida. "Wearable photoplethysmographic sensors—past and present." *Electronics* 3, no. 2 (2014): 282-302.

[23] Avila, Lisa, and Mike Bailey. "The wearable revolution." *IEEE Computer Graphics and Applications* 35, no. 2 (2015): 104-104.

[24] Zhao, Neil. "Full-featured pedometer design realized with 3-axis digital accelerometer." *Analog Dialogue* 44, no. 06 (2010): 1-5.

[25] Aroganam, Gobinath, Nadarajah Manivannan, and David Harrison. "Review on wearable technology sensors used in consumer sport applications." *Sensors* 19, no. 9 (2019): 1983

[26] Anindya Nag, S.C. Mukhopadhyay. “Wearable Electronics Sensors: Current Status and Future Opportunities”. Wearable Electronics Sensors. Smart Sensors, Measurement and Instrumentation, Vol 15. Springer. May, 2015. DOI: 10.1007/978-3-319-18191-2_1.

[27] Ching Yee Yong, Rubita Sauirman, Ahmad Hazwan Ab Rahim, Nasrul Humaimi Mahmood, Kim Mey Chew. “Jogging and Walking Analysis Using

Wearable Sensors". Published in Engineering 5b (05): 20-24. July, 2013. DOI: 10.4236/eng.2013.55B005.

[28] Akin Avci, Stephan Bosch, Mihai Marin-Perianu, Raluca Marin-Perianu, Paul Havinga. "Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey". 23rd International Conference on Architecture of Computing Systems, ARCS 2010, Hannover, Germany. pp 167-176.

[29] Ermes M, Pärkka J, Mantyjarvi J, Korhonen I. "Detection of Daily Activities and Sports with Wearable Sensors in Controlled and Uncontrolled Conditions". IEEE Transactions on Information Technology in Bio Medicine. Jan, 2008, 12(1):20-6. DOI: 10.1109/TITB.2007.899496.

[30] Nagender Kumar Suryadevara, Subhas Chandra Mukhopadhyay. "Wireless Sensor Network Based Home Monitoring System for Wellness Determination of Elderly". IEEE Sensors Journal, Vol. 12, No. 6, June 2012. pp. 1965-1972.

[31] Yun C. Zhang, James M. Rehg. "Watching the TV Watchers". Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. Volume 2, Issue 2, June 2018. Article No. 88. DOI:10.1145/3214291.

[32] Hussein, Aziza I. "Wearable computing: Challenges of implementation and its future." In *2015 12th Learning and Technology Conference*, pp. 14-19. IEEE, 2015.

[33] Parvarandeh, Pirooz, and Anthony Stephen Doy. "Wearable device for using human body as input mechanism." U.S. Patent Application 14/616,865, filed January 21, 2016.

[34] Fitzgerald, Alissa M., Ely K. Tsern, David J. Mooring, and James A. Gasbarro. "Gesture-based power management of a wearable portable electronic device with display." U.S. Patent 8,344,998, issued January 1, 2013.

[35] González, José Luis, Antonio Rubio, and Francesc Moll. "Human powered piezoelectric batteries to supply power to wearable electronic devices. *"International journal of the Society of Materials Engineering for Resources* 10, no. 1 (2002): 34-40.

[36] Ullah, Sana, Henry Higgins, Bart Braem, Benoit Latre, Chris Blondia, Ingrid Moerman, Shahnaz Saleem, Ziaur Rahman, and Kyung Sup Kwak. "A comprehensive survey of wireless body area networks." *Journal of medical systems* 36, no. 3 (2012): 1065-1094.

[37] Nag, Anindya, Subhas Chandra Mukhopadhyay, and Jürgen Kosel. "Wearable flexible sensors: A review." IEEE Sensors Journal 17, no. 13 (2017): 3949-3960.

[38] "Wearable Medical Devices Market to Surpass US$ 29 Bn by 2026; Rise in Prevalence of Chronic Diseases, Increase in the Incidence of Diabetic Population, and Surge in Adoption of Technologically Advanced Wearable Medical Devices to Drive Market: Transparency Market Research." Wearable Medical Devices Market to Expand CAGR of more than 17.0% from 2018 to 2026 | TMR. Accessed November 22, 2019. https://www.transparencymarketresearch.com/pressrelease/wearable-medical-devices.htm. Accessed 22 November, 2019.

[39] Rodgers, Mary M., Vinay M. Pai, and Richard S. Conroy. "Recent advances in wearable sensors for health monitoring. "*IEEE Sensors Journal*15, no. 6 (2014): 3119-3126.

[40] McAdams, Eric, Asta Krupaviciute, Claudine Gehin, Andre Dittmar, Georges Delhomme, Paul Rubel, Jocelyne Fayn, and Jad McLaughlin. "Wearable electronic systems: Applications to medical diagnostics/monitoring." In *Wearable monitoring systems*, pp. 179-203. Springer, Boston, MA, 2011.

[41] Tobii Tech - What is eye tracking?, September 17, 2015. https://www.tobii.com/tech/technology/what-is-eye-tracking/. Accessed 22 November, 2019.

[42] Velikova, Marina, Josien Terwisscha van Scheltinga, Peter JF Lucas, and Marc Spaanderman. "Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare." *International Journal of Approximate Reasoning* 55, no. 1 (2014): 59-73.

[43] Portela, Filipe, Manuel Filipe Santos, and Marta Vilas-Boas. "A pervasive approach to a real-time intelligent decision support system in intensive medicine." In *International Joint Conference on Knowledge Discovery, Knowledge*

Engineering, and Knowledge Management, pp. 368-381. Springer, Berlin, Heidelberg, 2010.

[44] Sun, Jimeng, Candace D. McNaughton, Ping Zhang, Adam Perer, Aris Gkoulalas-Divanis, Joshua C. Denny, Jacqueline Kirby, Thomas Lasko, Alexander Saip, and Bradley A. Malin. "Predicting changes in hypertension control using electronic health records from a chronic disease management program." *Journal of the American Medical Informatics Association* 21, no. 2 (2013): 337-344.

[45] Zheng, Ya-Li, Xiao-Rong Ding, Carmen Chung Yan Poon, Benny Ping Lai Lo, Heye Zhang, Xiao-Lin Zhou, Guang-Zhong Yang, Ni Zhao, and Yuan-Ting Zhang. "Unobtrusive sensing and wearable devices for health informatics." *IEEE Transactions on Biomedical Engineering* 61, no. 5 (2014): 1538-1554.

[46] Andreu-Perez, Javier, Daniel R. Leff, Henry MD Ip, and Guang-Zhong Yang. "From wearable sensors to smart implants-–toward pervasive and personalized healthcare." *IEEE Transactions on Biomedical Engineering* 62, no. 12 (2015): 2750-2762.

[47] Wei, Joseph. "How Wearables Intersect with the Cloud and the Internet of Things: Considerations for the developers of wearables." *IEEE Consumer Electronics Magazine* 3, no. 3 (2014): 53-56.

[48] Brooks, Frederik P., and No Silver Bullet. "Essence and accidents of software engineering." *IEEE computer* 20, no. 4 (1987): 10-19..

[49] Breiman, Leo. *Classification and regression trees*. Routledge, 2017.

[50] Wentworth, Tom. “RapidMiner Pricing.” RapidMiner. RapidMiner, September 26, 2019. https://rapidminer.com/pricing/. Accessed 20 Nov, 2019.

[51] Meyer, Daniel Z., and Leanne M. Avery. "Excel as a qualitative data analysis tool." *Field methods* 21, no. 1 (2009): 91-112.

[52] Wallace, Byron C., Issa J. Dahabreh, Thomas A. Trikalinos, Joseph Lau, Paul Trow, and Christopher H. Schmid. "Closing the gap between methodologists and end-users: R as a computational back-end." *J Stat Softw* 49, no. 5 (2012): 1-15.

[53] Ozgur, Ceyhun, Taylor Colliau, Grace Rogers, Zachariah Hughes, and B. Myer-Tyson. "MatLab vs. Python vs. R." *Journal of Data Science* 15, no. 3 (2017): 355-372.

[54] Preece, Stephen J., John Y. Goulermas, Laurence PJ Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. "Activity identification using body-mounted sensors—a review of classification techniques." *Physiological measurement* 30, no. 4 (2009): R1.

[55] Attal, Ferhat, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. "Physical human activity recognition using wearable sensors." *Sensors* 15, no. 12 (2015): 31314-31338.

[56] Fida, Benish, Ivan Bernabucci, Daniele Bibbo, Silvia Conforto, and Maurizio Schmid. "Pre-processing effect on the accuracy of event-based activity segmentation and classification through inertial sensors." *Sensors* 15, no. 9 (2015): 23095-23109.

[57] Banos, Oresti, Juan-Manuel Galvez, Miguel Damas, Hector Pomares, and Ignacio Rojas. "Window size impact in human activity recognition." *Sensors* 14, no. 4 (2014): 6474-6499.

[58] de Quadros, Thiago, Andre Eugenio Lazzaretti, and Fábio Kürt Schneider. "A movement decomposition and machine learning-based fall detection system using wrist wearable device." *IEEE Sensors Journal* 18, no. 12 (2018): 5082-5089.

[59] Culhane, K. M., G. M. Lyons, D. Hilton, P. A. Grace, and D. Lyons. "Long-term mobility monitoring of older adults using accelerometers in a clinical environment." *Clinical rehabilitation* 18, no. 3 (2004): 335-343.

[60] Boyle, Justin, Mohan Karunanithi, Tim Wark, Wilbur Chan, and Christine Colavitti. "Quantifying functional mobility progress for chronic disease management." In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5916-5919. IEEE, 2006.

[61] Coley, Brian, Bijan Najafi, Anisoara Paraschiv-Ionescu, and Kamiar Aminian. "Stair climbing detection during daily physical activity using a miniature gyroscope." *Gait & posture* 22, no. 4 (2005): 287-294.

[62] Nyan, M. N., F. E. H. Tay, K. H. W. Seah, and Y. Y. Sitoh. "Classification of gait patterns in the time–frequency domain." *Journal of biomechanics* 39, no. 14 (2006): 2647-2656.

[63] Bourke, Alan K., and Gerald M. Lyons. "A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor." *Medical engineering & physics* 30, no. 1 (2008): 84-90.

[64] Godfrey, A. C. R. M. D. O. G., Richard Conway, David Meagher, and Gearoid ÓLaighin. "Direct measurement of human movement by accelerometry." *Medical engineering & physics* 30, no. 10 (2008): 1364-1386.

[65] Parkka, Juha, Miikka Ermes, Panu Korpipaa, Jani Mantyjarvi, Johannes Peltola, and Ilkka Korhonen. "Activity classification using realistic data from wearable sensors." *IEEE Transactions on information technology in biomedicine* 10, no. 1 (2006): 119-128.

[66] Ermes, Miikka, Juha Pärkkä, Jani Mäntyjärvi, and Ilkka Korhonen. "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions." *IEEE transactions on information technology in biomedicine* 12, no. 1 (2008): 20-26.

[67] Mathie, M. J., Branko G. Celler, Nigel H. Lovell, and A. C. F. Coster. "Classification of basic daily movements using a triaxial accelerometer." *Medical and Biological Engineering and Computing* 42, no. 5 (2004): 679-687.

[68] Zhang, Tong, Jue Wang, Liang Xu, and Ping Liu. "Using wearable sensor and NMF algorithm to realize ambulatory fall detection." In *International conference on natural computation*, pp. 488-491. Springer, Berlin, Heidelberg, 2006.

[69] Bao, Ling, and Stephen S. Intille. "Activity recognition from user-annotated acceleration data." In *International conference on pervasive computing*, pp. 1-17. Springer, Berlin, Heidelberg, 2004.

[70] Bedogni, Luca, Marco Di Felice, and Luciano Bononi. "By train or by car? Detecting the user's motion type through smartphone sensors data." In *2012 IFIP Wireless Days*, pp. 1-6. IEEE, 2012.

[71] Huynh, Tâm, and Bernt Schiele. "Towards less supervision in activity recognition from wearable sensors." In *2006 10th IEEE International Symposium on Wearable Computers*, pp. 3-10. IEEE, 2006.

[72] Mannini, Andrea, and Angelo Maria Sabatini. "Machine learning methods for classifying human physical activity from on-body accelerometers." *Sensors* 10, no. 2 (2010): 1154-1175.

[73] Brownlee, Jason. *Master Machine Learning Algorithms: discover how they work and implement them from scratch*. Machine Learning Mastery, 2016.

[74] Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." (2011).

[75] Banos, Oresti, Claudia Villalonga, Rafael Garcia, Alejandro Saez, Miguel Damas, Juan A. Holgado-Terriza, Sungyong Lee, Hector Pomares, and Ignacio Rojas. "Design, implementation and validation of a novel open framework for agile development of mobile health applications." Biomedical engineering online 14, no. 2 (2015): S6.

[76] Jordao, Artur, Antonio C. Nazare Jr, Jessica Sena, and William Robson Schwartz. "Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art." arXiv preprint arXiv:1806.05226 (2018).

[77] Yogesh, K. M., and Osama A. Ghoneim. "Performance Evaluation of Advanced Classification Models on Spatial Location Based Shimmer2 Sensor Data Sets." In 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1925-1931. IEEE, 2018.

[78] Brownlee, Jason. “A Gentle Introduction to a Standard Human Activity Recognition Problem.” Machine Learning Mastery, August 5, 2019. Accessed December 19, 2019. https://machinelearningmastery.com/how-to-load-and-explore-a-standard-human-activity-recognition-problem/.

[79] “What Is RAD Model? Advantages & Disadvantages.” Guru99. Accessed December 15, 2019. https://www.guru99.com/what-is-rad-rapid-software-development-model-advantages-disadvantages.html.

[80] Avci, Akin, Stephan Bosch, Mihai Marin-Perianu, Raluca Marin-Perianu, and Paul Havinga. "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey." In *23th International conference on architecture of computing systems 2010*, pp. 1-10. VDE, 2010.

[81] Ermes, Miikka, Juha Pärkkä, Jani Mäntyjärvi, and Ilkka Korhonen. "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions." *IEEE transactions on information technology in biomedicine* 12, no. 1 (2008): 20-26.

[82] Baker, Chris R., Kenneth Armijo, Simon Belka, Merwan Benhabib, Vikas Bhargava, Nathan Burkhart, Artin Der Minassians et al. "Wireless sensor networks for home health care." In *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, vol. 2, pp. 832-837. IEEE, 2007.

[83] Liu, Luyang, Cagdas Karatas, Hongyu Li, Sheng Tan, Marco Gruteser, Jie Yang, Yingying Chen, and Richard P. Martin. "Toward detection of unsafe driving with wearables." In *Proceedings of the 2015 workshop on Wearable Systems and Applications*, pp. 27-32. ACM, 2015.

[84] A. Ayman, O. Attalah and H. Shaban, "Smart System for Recognizing Daily Human Activities Based on Wrist IMU Sensors," 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), Al Madinah Al Munawwarah, Saudi Arabia, 2020, pp. 1-6, doi: 10.1109/AECT47998.2020.9194154.

[85] A. A. Badawi, A. Al-Kabbany and H. Shaban, "Daily Activity Recognition using Wearable Sensors via Machine Learning and Feature Selection," 2018 13th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 2018, pp. 75-79, doi: 10.1109/ICCES.2018.8639309.

[86] I. Khokhlov, L. Reznik, J. Cappos and R. Bhaskar, "Design of activity recognition systems with wearable sensors," 2018 IEEE Sensors Applications Symposium (SAS), Seoul, 2018, pp. 1-6, doi: 10.1109/SAS.2018.8336752.

[87] A. Nandy, J. Saha, C. Chowdhury and K. P. D. Singh, "Detailed Human Activity Recognition using Wearable Sensor and Smartphones," 2019 International

Conference on Opto-Electronics and Applied Optics (Optronix), Kolkata, India, 2019, pp. 1-6, doi: 10.1109/OPTRONIX.2019.8862427.

[88] L. Cheng, Y. Guan, Kecheng Zhu and Yiyang Li, "Recognition of human activities using machine learning methods with wearable sensors," 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 2017, pp. 1-7, doi: 10.1109/CCWC.2017.7868369.

[89] Yuqing Chen and Yang Xue, "A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer," in IEEE International Conference on Systems, Man, and Cybernetics, 2015, pp. 1488– 1492.

[90] Sojeong Ha and Seungjin Choi, "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," in International Joint Conference on Neural Networks, 2016, pp. 381–388.

[91] Cagatay Catal, Selin Tufekci, Elif Pirmit, and Guner Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," Applied Soft Computing, vol. 37, pp. 1018–1022, 2015.

[92] Wenchao Jiang and Zhaozheng Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," in 23rd Annual Conference on Multimedia Conference, 2015, pp. 1307–1310.

[93] Rasha M. Al-Eidan, Hend Al-Khalifa, Abdul Malik Al-Salman, "A Review of Wrist-Worn Wearable: Sensors, Models, and Challenges", Journal of Sensors, vol. 2018, Article ID 5853917, 20 pages, 2018. https://doi.org/10.1155/2018/5853917

[94] Qi, Jun, Po Yang, Atif Waraich, Zhikun Deng, Youbing Zhao, and Yun Yang. "Examining sensor-based physical activity recognition and monitoring for healthcare using Internet of Things: A systematic review." Journal of biomedical informatics 87 (2018): 138-153.

[95] H. Yu, S. Cang and Y. Wang, "A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems," 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), Chengdu, 2016, pp. 250-257, doi: 10.1109/SKIMA.2016.7916228.

[96] Ahola, T.M. (2010). Pedometer for running activity using accelerometer sensors on the wrist. Medical Equipment Insights 2010 (3): 1–8.

[97] N. Statt. (2017). Apple Reportedly Developing a Dedicated AI Chip for the iPhone. [Online]. Available:https://www.theverge.com/2017/5/26/15702248/apple-neural-engineai-chip-iphone-ipad

[98] Y. Kang et al., “Neurosurgeon: Collaborative intelligence between the cloud and mobile edge,” in Proc. 22nd Int. Conf. Archit. Support Program. Lang. Oper. Syst., 2017, pp. 615–629.

[99] Wearable Technology 2016–2026. (2016). [Online]. Available: http://www.idtechex.com/research/reports/wearable-technology-2016-2026-000483.asp

[100] Wearable Technology Market Is Quickly Diversifying Into New Device Categories, Application Markets, and Services. (2016). [Online]. Available: https://www.tractica.com

[101] The WT—Wearable Technologies 2015 Market Assessment n Smart Patches. (2016). [Online]. Available: https://www.wearable-technologies.com

[102] J. Chauhan, S. Seneviratne, M. A. Kaafar, A. Mahanti, and A. Seneviratne, “Characterization of early smartwatch apps,” in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops), Sydney, NSW, Australia, Mar. 2016,pp. 1–6.

[103] IDC Forecasts Wearables Shipments to Reach 213.6 Million Units Worldwide in 2020 With Watches and Wristbands Driving Volume While Clothing and Eyewear Gain Traction. (2016). [Online]. Available: http://www.idc.com

[104] Shipments of Cellular M2M Modules Reached 96.0 Million Units in 2015. (2016). [Online]. Available: http://www.berginsight.com/ news.aspx

APPENDIX ‘A’: FUNCTIONAL SPECIFICATIONS

Functional specifications present the system’s perspective and indicate the minimum product functionality, required to satisfy requirements identified in requirement analysis phase. Functional requirements as derived from requirement analysis are described in the following tables.

FR01: Data Input Mechanism

FR01-01	System shall allow user to load data from the desired hard drive locations.
FR01-02	System shall open the input dialog to navigate and search the user desired data from the desired location.
FR01-03	System should allow user to load the data in csv file format.
FR01-04	System shall display the loading bar to display progress of data loading.

FR02: Data Visualisation

FR02-01	System shall allow user to view data structure and data dimensions.
FR02-02	System shall allow user to view the Head, Tail or Complete data records loaded from the file.
FR02-03	System shall allow user to select the desired data display options from the view selection panel.
FR02-04	System shall allow user to view the desired data record entries through the selectable options of record display.
FR02-05	System shall provide the option to search a specific record from the loaded data.
FR02-06	System shall display the summary statistics of all the variables in the loaded data file.

FR03: Variable Selection for the Model

FR03-01	System shall allow user to see all the independent variables in the data file with option to select the desired independent variables for running an algorithm model.
FR03-02	System shall allow user to select the dependent variable/variables from the desired variable list.

FR04: Running the Desired Algorithm on Data

FR04-01	System shall provide an option to select the desired type of Algorithm model available in a dropdown list.
FR04-02	System shall allow user to change the desired training and test ratios for running the Algorithm models.
FR04-03	System shall allow user to change algorithm parameters for running different compositions on data.
FR04-04	System shall display the results of algorithm model applied on the loaded data in the form of model accuracy and Confusion matrix.

FR05: Comparing the performance of chosen Algorithm on Data

FR05-01	System shall provide an option to select the desired Algorithms' model for running on the data and available in a dropdown list.
FR05-02	System shall allow user to change the desired training and test ratios for running the Algorithm models.
FR05-03	System shall display the results of comparison applied on the loaded data in the form of RMSE and Rsquared values.
FR05-04	System shall allow user to display the results of comparison applied on the loaded data in the form of different graph types. 1. Box & Whisker plot 2. Density plot 3. Dot plot 4. Scatter plot 5. XY plot

APPENDIX 'B': USER INTERFACE

The main functionalities of software as articulated in the user requirements are implemented in GUI that enable the user to:

- Load the dataset from local drive.
- View the dataset in a tabular form.
- Apply selected classification ML algorithms on data.
- Compare results of algorithms.
- Visualise the results in the form of summary and graphs.
- Save the experimental results.
- Repeat the experiments if required.

B.1 Main GUI

The main GUI is divided into three main parts for data loading, selections to view the loaded data and also to select desired algorithm(s) on data and the third part is for viewing the results, "Output", as annotated in the following figure.

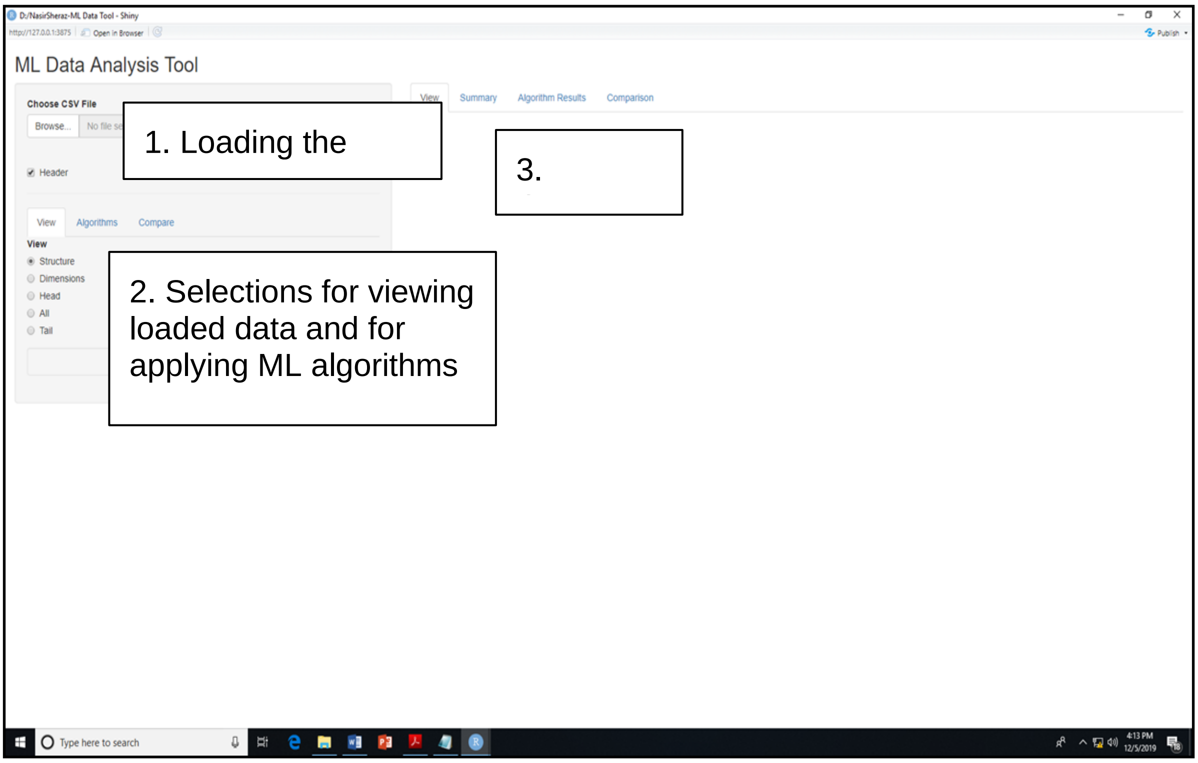


Figure 3.3 : Main Graphical User Interface

B.2 Data Loading

The data is loaded with a simple “Browse” button which opens a “Choose File Dialog Box” to select a CSV type of file having data in rectangular format. As the file is loaded into the software, a progress bar indicates its completion. There is a checkbox to let the user select or deselect if the loaded data file has first row as indicating headers of the columns.

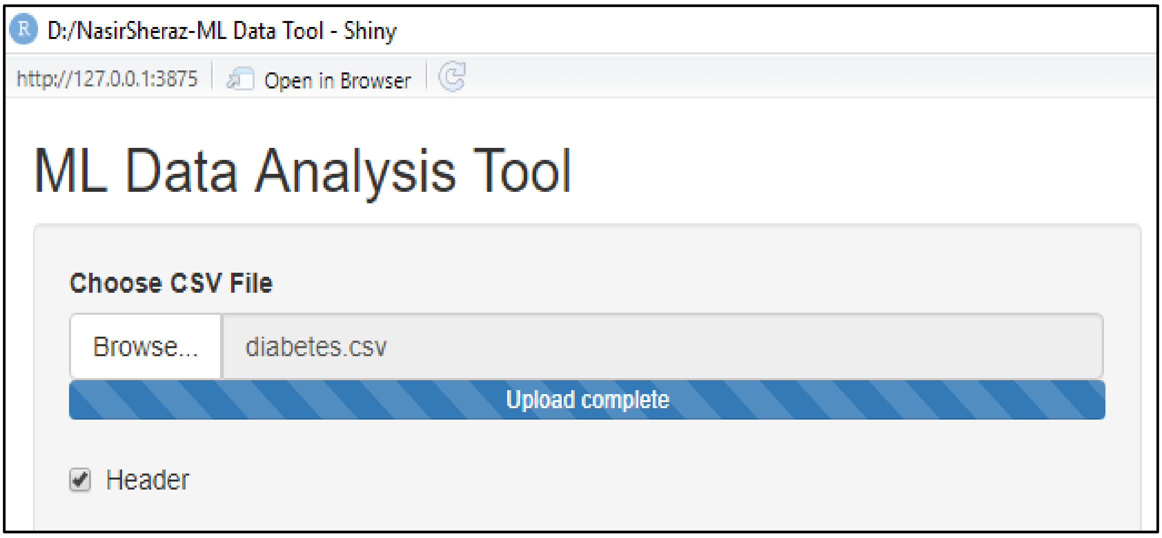
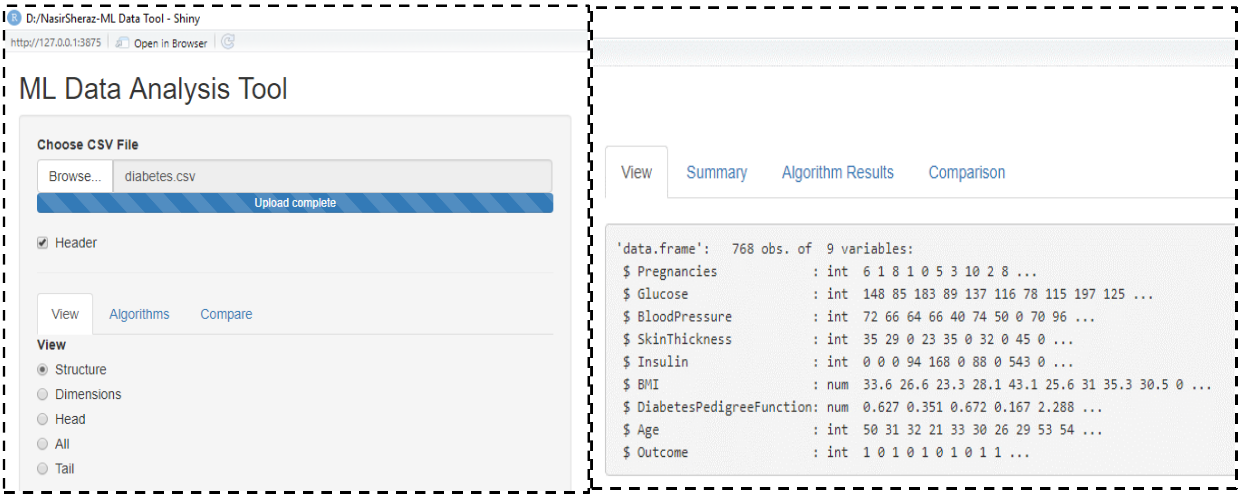


Figure 3.4 : Load CSV File

B.3 View Loaded Data

The loaded data can be viewed for its general properties for selecting options from the View Tab in “Selections” area of GUI as described above.

Figure 3.5 : Dataset Views / Data Structure



B.4 Data Dimension

To view the dimensions of rectangular data, select the ‘Dimensions’ radio button in the View Tab. The dimensions are displayed in the output area.

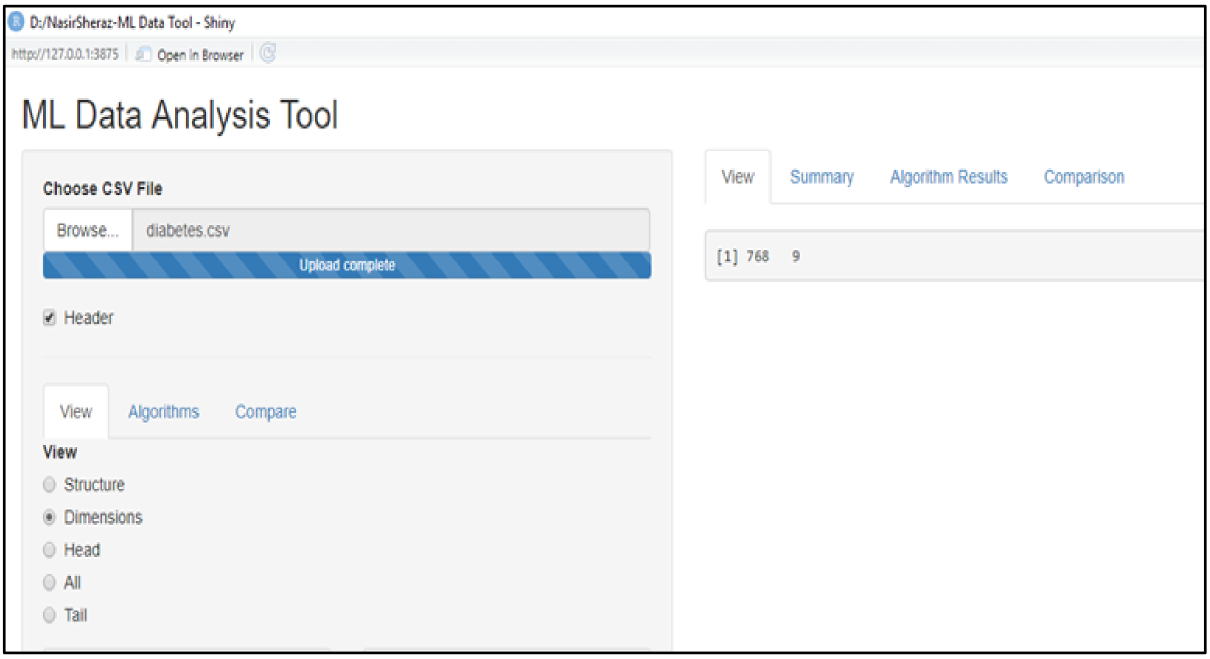


Figure B: Dataset Views / Data Dimensions

B.5 Data Top Rows

To view the top rows of the data, select the ‘Head’ radio button. The top six rows in the data are displayed in the Output area with all the columns.

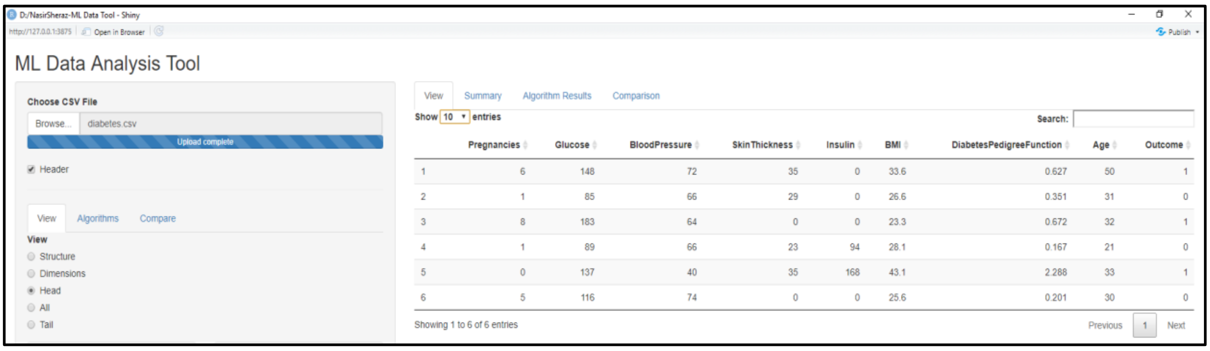


Figure 3.7: Dataset Views / Data Head ValuesB.6 Include Variables

The “Column Headers” of the input dataset are displayed as ‘Included Variables’ with ‘Checkboxes’. These checkboxes are selected by default. Deselecting any of the checkbox would exclude the variable from model during algorithmic analysis.

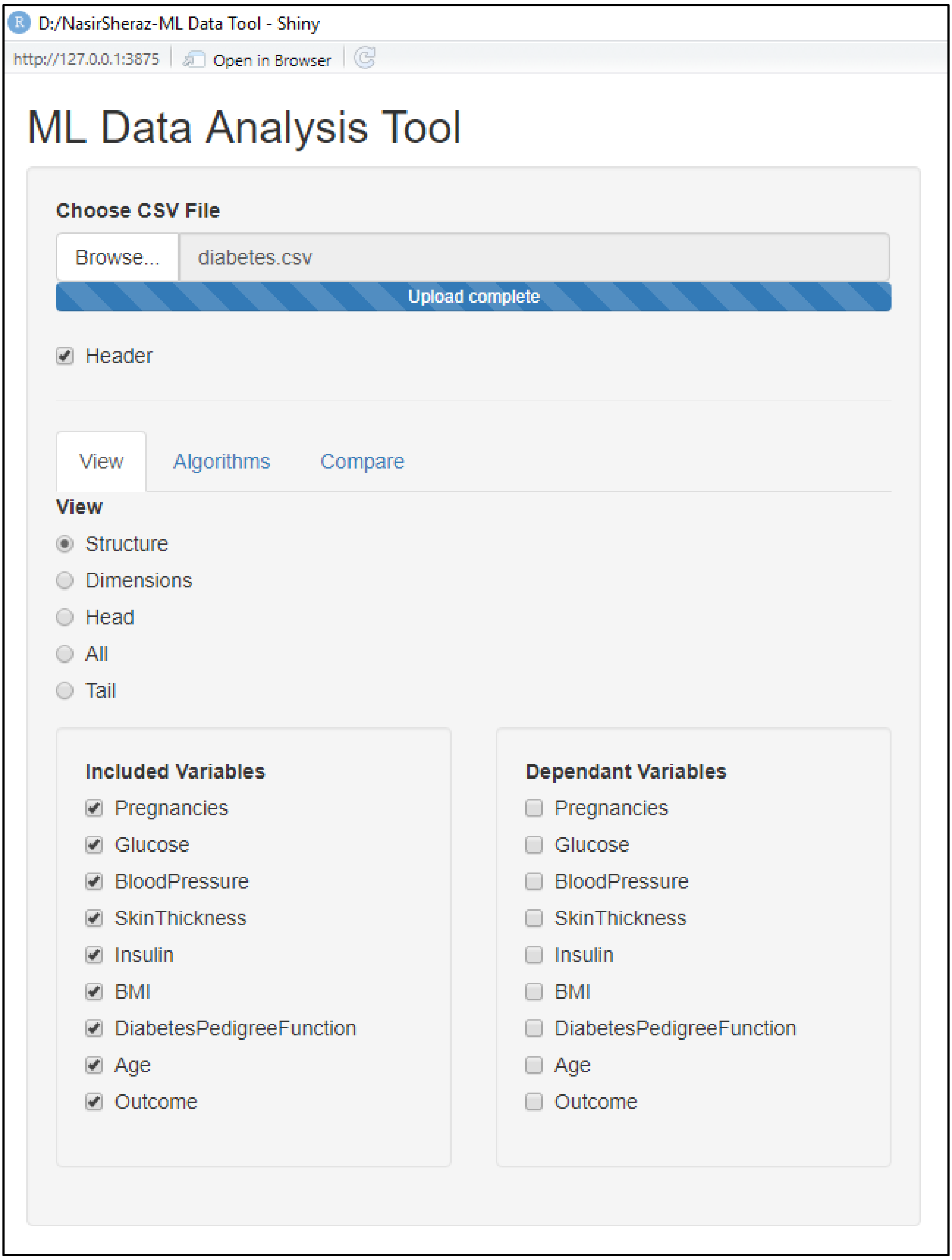


Figure 3.8 : Include Variables from Dataset

B.6 Dependent Variable Selection

The dependent variable in this tool contains classification information of the training / test data. One dependent variable can be selected among all the selected variables.

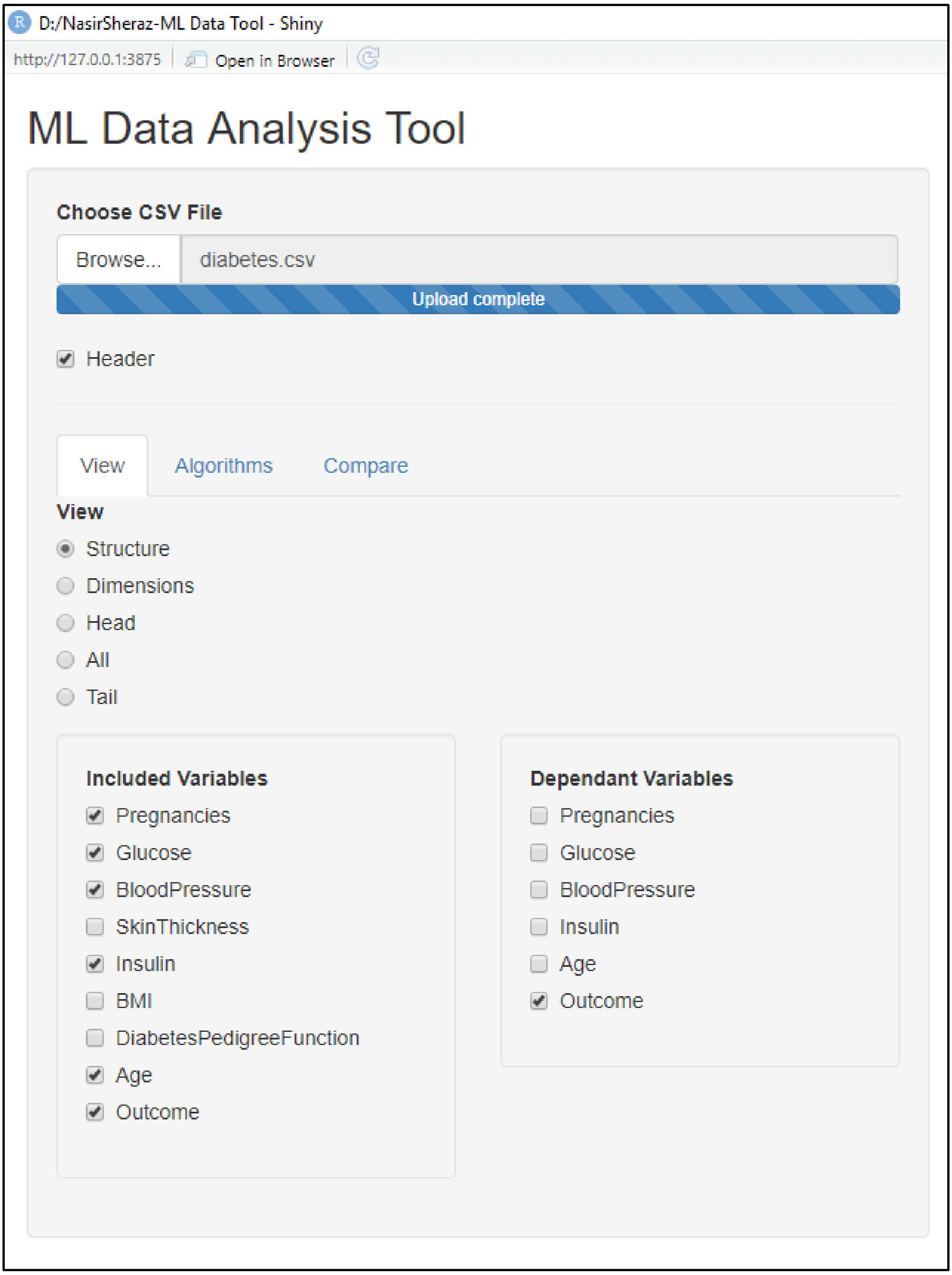
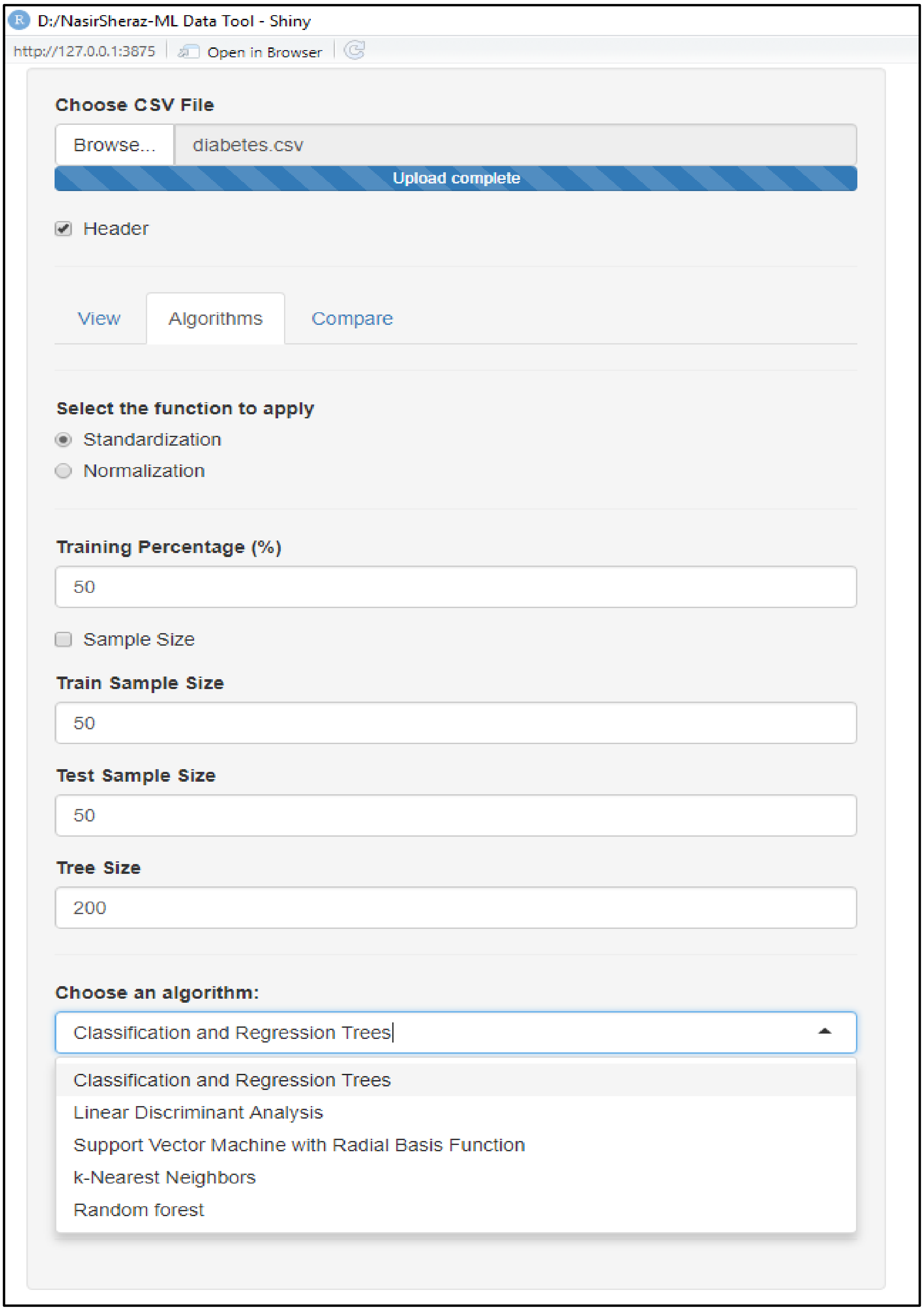


Figure 3.9 : Dataset Variables / Mark dependent

B.7 ML Algorithm Selection

The ML algorithms can be selected from the 'Algorithms' Tab next to the 'View' tab. Currently, five classification algorithms are available in the selection as shown



in the

Figure 3.10 : Algorithm selection

figure. In the same tab, fraction of data to be used as training can be given as input in a text box. Alternately, sample sizes for both training and test data can be mentioned. The samples are randomly drawn from the data without replacement. Some parameters specific to applied algorithm are also given in the same tab.

B.8 Algorithm Comparison

The algorithmic comparison is done by selecting desired subset of algorithms in the 'Compare' tab.

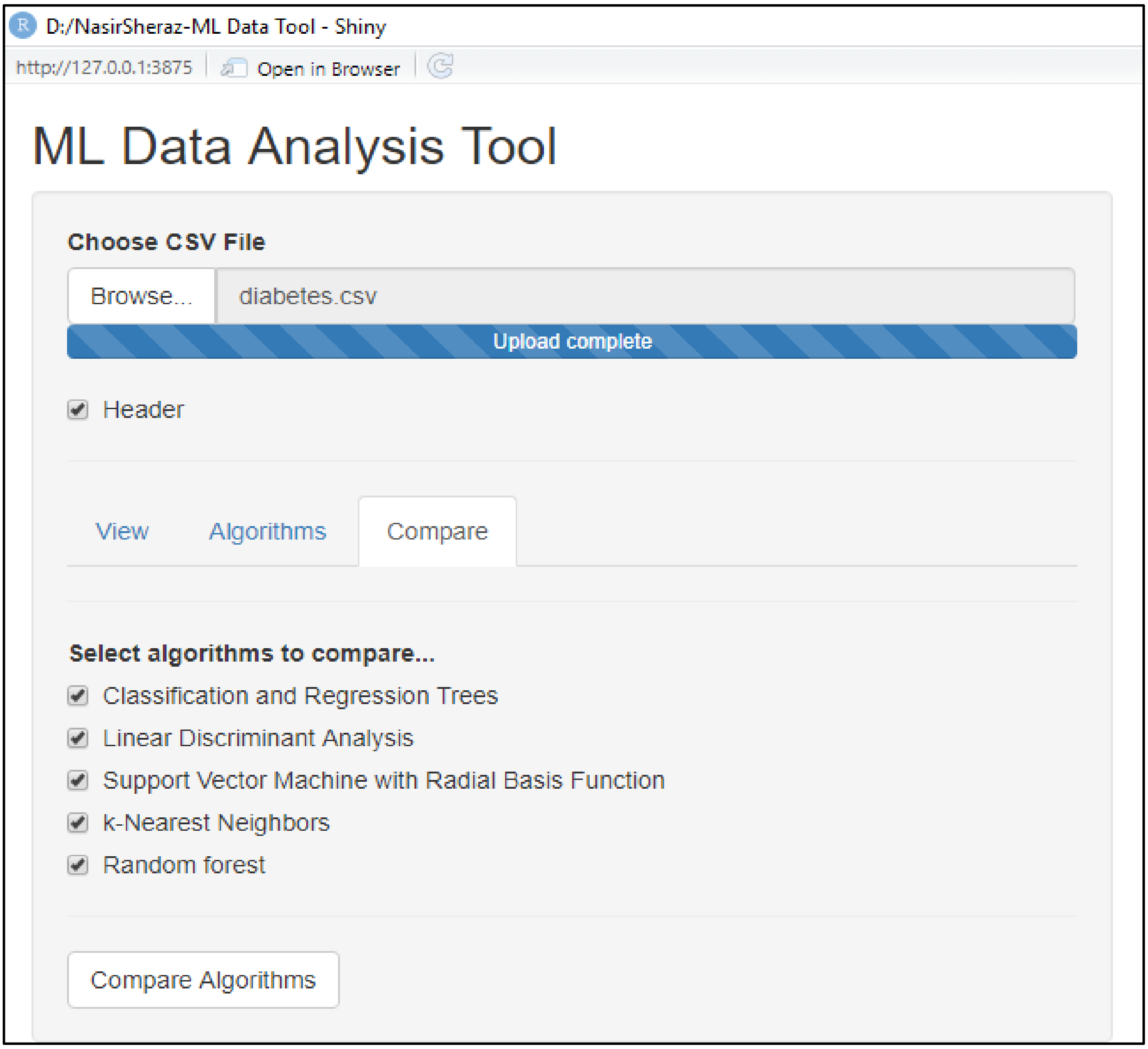


Figure 3.11 : Algorithm comparison

B.9 Algorithm Results

The results of algorithms are shown in the output area under the ‘Algorithm Results’ tab. It further has two sub-tabs ‘Summary’ and ‘Confusion Matrix’.

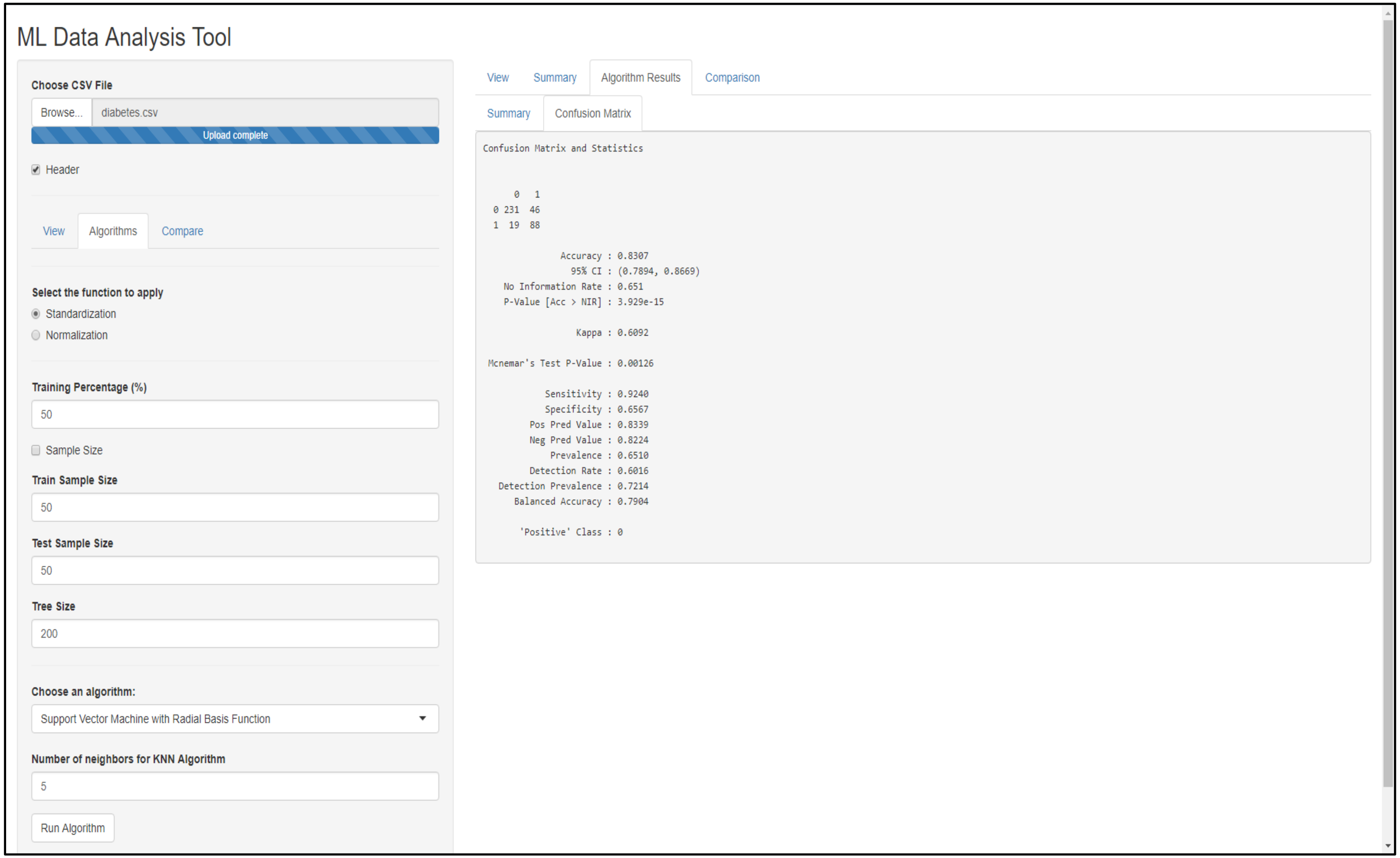


Figure 3.12 : Algorithm Results

B.10 Graphical Results – Box & Whisker Plots

The graphical results show Box and Whisker Plots which are quite useful to look at the spread of the estimated accuracies for different methods and how they relate. The boxes are ordered from highest to lowest mean accuracy. The mean values (dots) and the overlaps of the boxes (middle 50% of results) are depicted here.

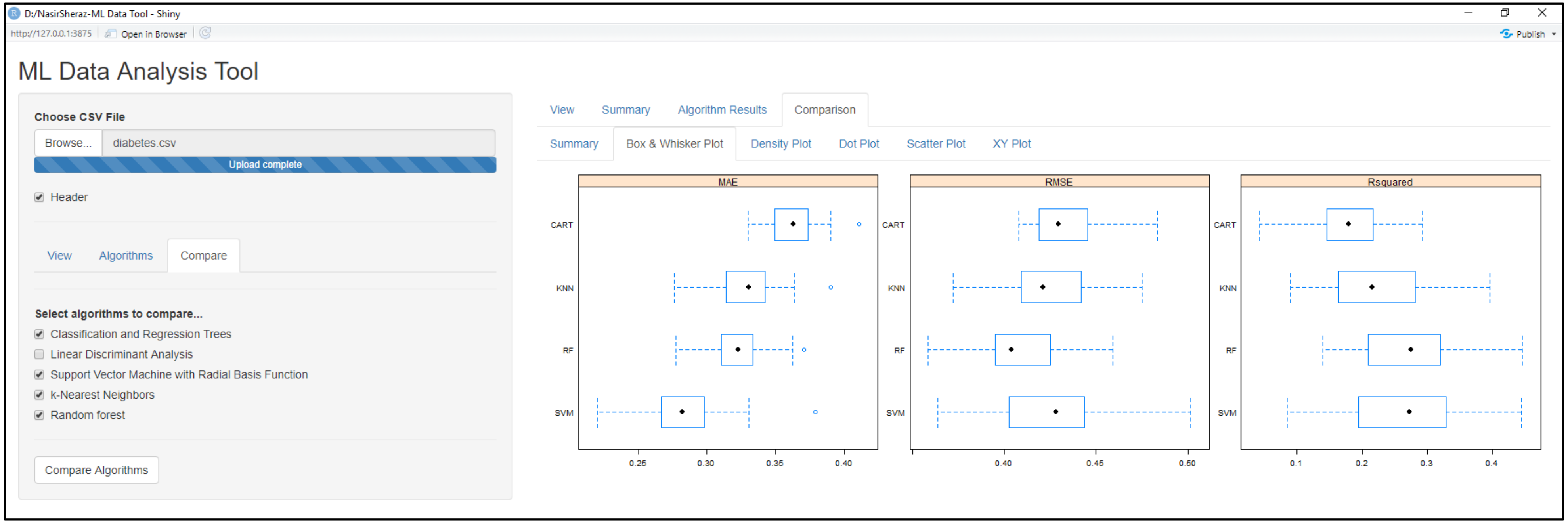


Figure 3.13 : Algorithms Comparison / Box & Whisker plots

B.11 Graphical Results – Density Plots

The distribution of model accuracy is depicted as density plots. This is a useful way to evaluate the overlap in the estimated behaviour of algorithms

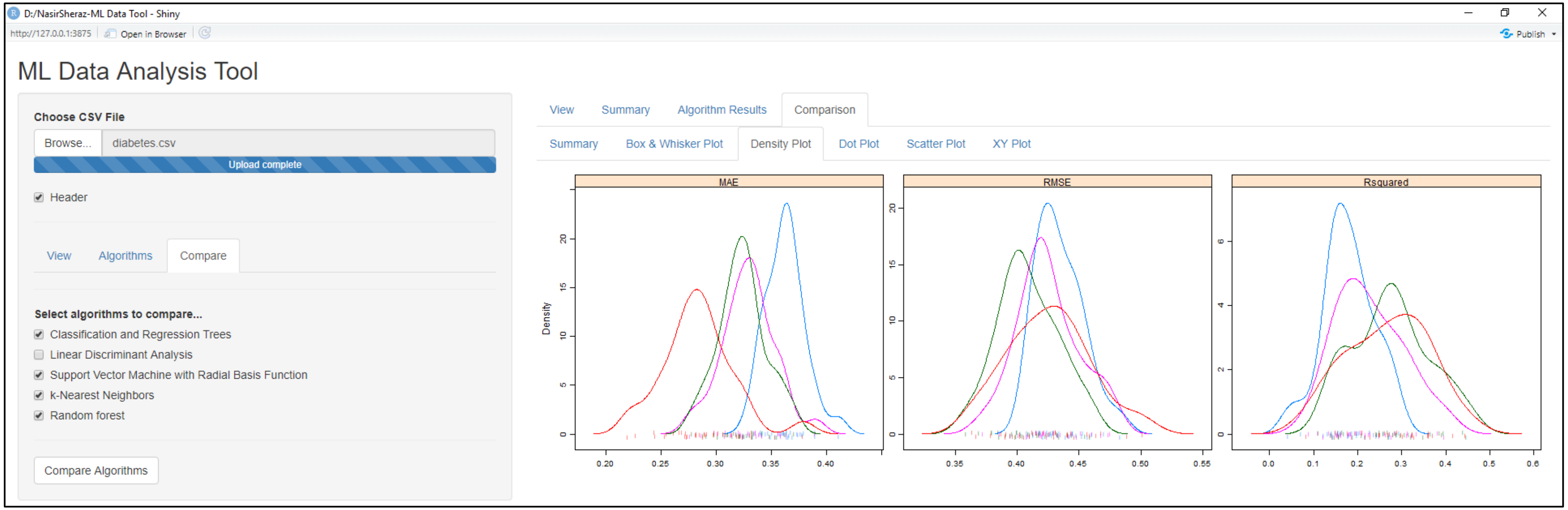


Figure 3.14 : Algorithms comparison - density plots

B.12 Graphical Results – Dot Plots

These are useful plots as they show both the mean estimated accuracy as well as the 95% confidence interval (e.g. the range in which 95% of observed scores fell). It is good to compare the means and eye-ball the overlap of the spreads between algorithms.

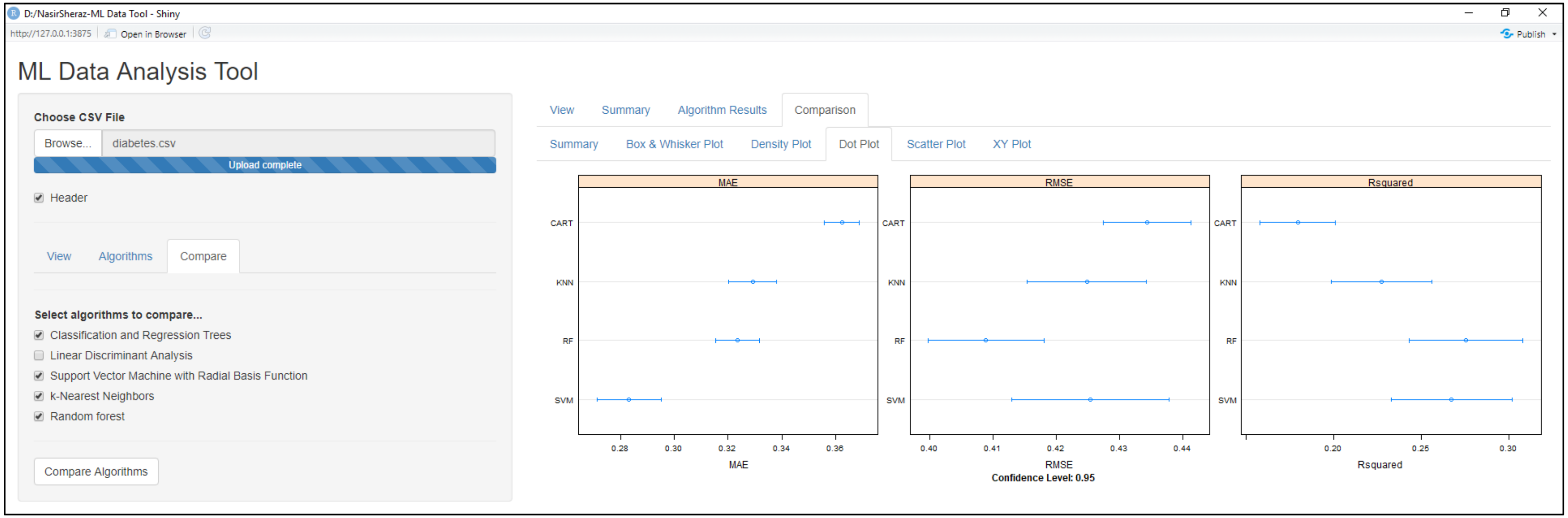


Figure 3.15: Algorithms Comparison – Dot Plots

B.13 Graphical Results – Scatter Plots

This is invaluable when considering whether the predictions from two different algorithms are correlated. If weakly correlated, they are good candidates for being combined in an ensemble prediction. Looking at the graphs it looks like RF and KNN look strongly correlated, whereas SVM and CART look weakly correlated.

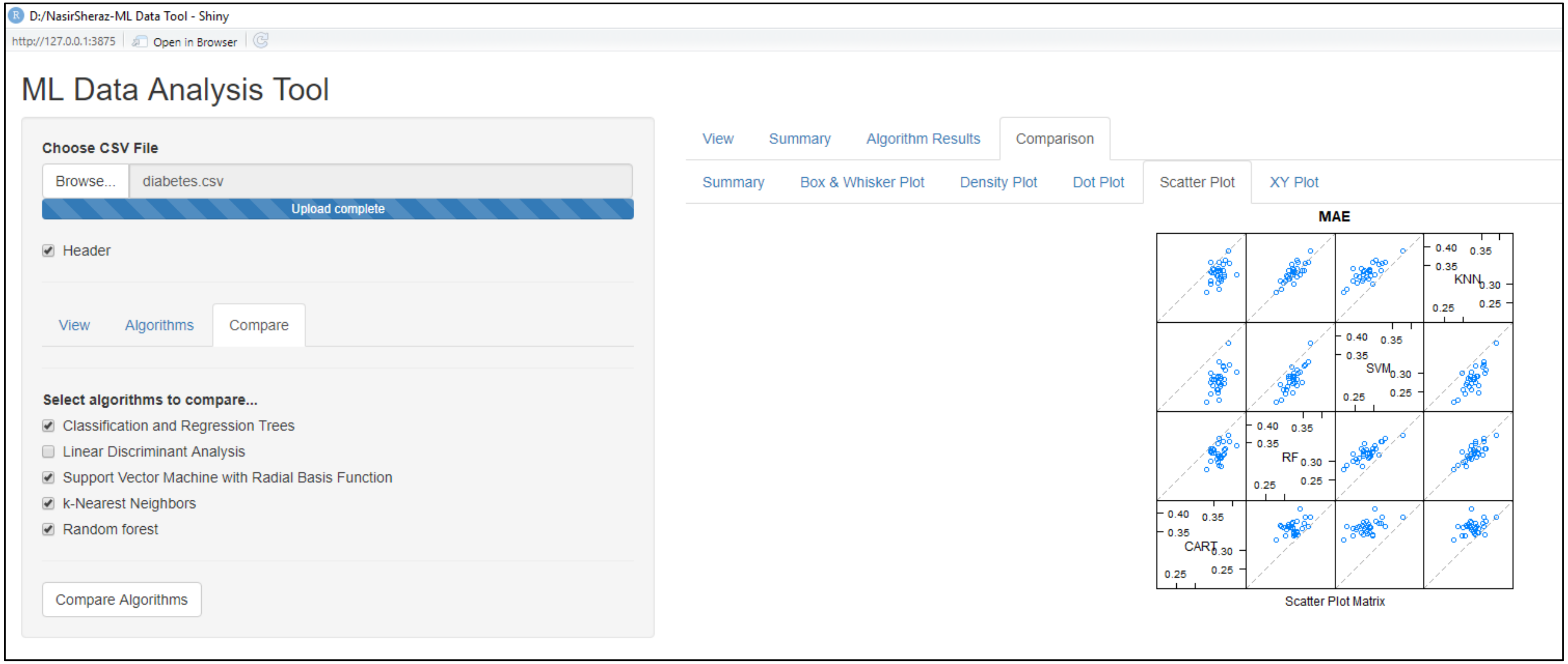


Figure 3.16: Algorithms Comparison - Scatter Plot