UNIVERSITY of the
WESTERN CAPE

DATA SCIENCE FOR HEALTH-CARE: PATIENT CONDITION RECOGNITION.

by

Munyaradzi Mandava

Masters in Computer Science

A thesis submitted in partial fulfillment
of the requirements for the degree of
Masters of Science

The Department of Computer Science
The University of the Western Cape
Bellville , Western Cape, South Africa

Declaration of Authorship

I, Munyaradzi Mandava, declare that this thesis titled, **Data Science for Health-care: Patient Condition Recognition.** , and the work presented in it is my own. I confirm that:

- This work was done wholly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, the thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is biased on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date: 01 November 2018 Signed: Mandava

Abstract

The emergence of the Internet of Things (IoT) and Artificial Intelligence (AI) have elicited increased interest in many areas of our daily lives. These include health, agriculture, aviation, manufacturing, cities management and many others. In the health sector, portable vital sign monitoring devices are being developed using the IoT technology to collect patients' vital signs in real-time. The vital sign data acquired by wearable devices is quantitative and machine learning techniques can be applied to find hidden patterns in the dataset and help the medical practitioner with decision making. There are about 30000 diseases known to man and no human being can possibly remember all of them, their relations to other diseases, their symptoms and whether the symptoms exhibited by the patients are early warnings of a fatal disease. In light of this, Medical Decision Support Systems (MDSS) can provide assistance in making these crucial assessments. In most decision support systems factors affect each other; they can be contradictory, competitive, and complementary. All these factors contribute to the overall decision and have different degrees of influence [85]. However, while there is more need for automated processes to improve the health-care sector, most of MDSS and the associated devices are still under clinical trials. This thesis revisits cyber physical health systems (CPHS) with the objective of designing and implementing a data analytics platform that provides patient condition monitoring services in terms of patient prioritisation and disease identification [1]. Different machine learning algorithms are investigated by the platform as potential candidate for achieving patient prioritisation. These include multiple linear regression, multiple logistic regression, classification and regression decision trees, single hidden layer neural networks and deep neural networks. Graph theory concepts are used to design and implement disease identification. The data analytics platform analyses data from biomedical sensors and other descriptive data provided by the patients (this can be recent data or historical data) stored in a cloud which can be private local health Information organisation (LHIO) or belonging to a regional health information organisation (RHIO). Users of the data analytics platform consisting of medical practitioners and patients are assumed to interact with the platform through cities' pharmacies , rural E-Health kiosks end user applications. **Keywords:** Medical decision support system, data science, health-care, decision trees, graph theory, artificial intelligence, deep learning and machine learning.

Publications

Mukuzo Fortunat Bagula, Herman Bagula, Munyaradzi Mandava, Claude Kakoko, Antoine Bagula. Cyber-healthcare Kiosks for Healthcare Support in Developing Countries. AFRICOMM 2018: e-Infrastructure and e-Services for Developing Countries, vol. 275 (2019) pp. 185-198 Published by Springer, Cham.

Bagula A, Mandava M, Bagula H. A framework for healthcare support in the rural and low income areas of the developing world. Journal of Network and Computer Applications, vol. 120 (2018) pp. 17-29 Published by Academic Press.

Mandava M, Lubamba C, Ismail A, Bagula A, Bagula H. Cyber-healthcare for public healthcare in the developing world. Proceedings - IEEE Symposium on Computers and Communications, vol. 2016-August (2016) pp. 14-19 Published by Institute of Electrical and Electronics Engineers Inc.

Bagula. A, Lubamba. C, Mandava. M, Bagula. M, Zennaro. M and Pietrosemoli. E. Cloud Based Patient Prioritization As Service in Public Health Care. Conference: ITU-Kaleidoscope 2016, At Bangkok, Thailand.

Acknowledgements

I would like to convey my gratefulness to those people who made this thesis possible. First and foremost I offer my sincerest gratitude to my supervisor, Prof. B. A. Bagula for believing in my abilities, encouraging me and guiding me.

I also owe my gratitude to my family and friends who helped me during the process Fadzai Munyanyi and Michelle Masuku your effort does not go unnoticed. To my wife Thelma Mabambe thank you for putting up with the long sleepless nights of studying.

To my parents and guardian who passed on I would like to sincerely thank you and convey my love for all you made me, my education would not have been possible.

Contents

List of Figures

List of Tables

List of Appendices

Acronyms

α Learning Rate. 79, 83

AI Artificial Intelligence. iii, vi, ix, 13–15

ANN Artificial Neural Network. viii, 79

AVPU Alert, Reacts to Voice, Reacts to Pain, Unresponsive. 28

CART Classification and Regression Decision Tree. vi, 18, 28, 39, 79, 83, 86, 87

DBP Diastolic Blood Pressure. 47

DC Degree Centrality. vii, 56

DNN Deep Neural Network. 28, 29, 39, 46, 47, 79, 83, 86, 103

ECG Electrocardiogram. 29

HSDN Human Symptoms-Disease Network. vii, viii, x, 54–57, 59, 61, 63, 65, 83–85, 87

IoT Internet of Things. iii, 5, 28

KCG Kmeans Clustering and Gausian Estimator. 28

MDSS Medical Decision Support Systems. iii

MeSH Medical Subject Headings. 54, 55

ML Machine Learning. vi, 5, 14, 15, 18, 86

MLiR Multivariate Linear Regression. 28, 39, 79, 83, 86, 87

MLoR Multivariate Logistic Regression. 28, 39, 79, 83, 86, 87

MSD Mean Squared Deviation. 77

MSE Mean Squared Error. viii, 77, 79, 83

NLP Natural Language Processing. vi, 21

R Coefficient of Correlation. viii, 78

SBP Systolic Blood Pressure. 28, 47

SNN Single Hidden Layer Neural Network. 28, 39, 47, 79, 83, 86

SPO2 Peripheral Oxygen Saturation. 47

SVM Support Vector Machine. 28

TRISS Trauma and Injury Severity Score. 19

Chapter 1

Introduction

1.1 Motivation

The incessant pursuit of technological advancements and improvements has significantly increased over the last few decades. One of the fundamental aims of this quest has been to improve human health and well-being through medical technology. The continuous infusion of technology and medicine has led to the conquest of numerous medical barriers as a result of the development of sophisticated high-tech medical apparatus. Whilst technology has always played a paramount role in the medical field, the integration and evolution of technology has been a gradual process. Earlier technological discoveries are the backbone of today's advanced medical technology. Various early scientists have made immense contributions to the field of medical technology, which have spearheaded the improvement of medical apparatus that improve human health and well-being.

One of medicine's pivotal inventions is the invention of the microscope by Zacharias and Hans Jansen. The microscope revolutionized medical engineering. The Jansen's first rendition of the microscope was a multi-lensed tube created in 1595, which had a magnification of 3x to 9x [88]. Thereafter, modifications on this invention began with Galileo Galilei in the early 1600s who included a focusing component to the microscope. Antonie Philips Van Leeuwenhoek and Robert Hooke who both created variations of the microscope between 1630 and 1725 improved on the magnification of this device, increasing it to up to 270x [60]. After centuries of modifications and improvements, microscopes today are used for different purposes, with dissection microscopes having a magnification of up to 70x, compound microscopes having a magnification of 1000x and electron microscopes having a magnification of 10 000 000x. The invention of the microscope was the beginning of the discovery of an endless medical dimension. Its role in increasing knowledge on germs, diseases and life at a cellular level has made it an indispensable apparatus in all areas of medicine. The microscope also opened up doors for the discovery of many other revolutionary medical discoveries such as germ theory, penicillin and genetics.

Another ground-breaking technological invention in medicine was the thermometer. The first functional thermometer - which was a rudimentary water thermoscope that allowed for temperature variations to be taken - was invented by Galileo Galilei in the late 1500s [78]. This invention was later modified by Santorio in 1612 through the addition of a numerical scale which made it the first oral clinical thermometer. However the inaccuracies of these thermometers led Ferdinand II to improve on the design of the device by replacing water with alcohol in the thermometer. The lack of standardization resulted in physicist Daniel Gabriel Fahrenheit modifying the thermometer by replacing alcohol with mercury and introducing a self-named scale - the Fahrenheit Scale - which was able to give more precise temperature readings [50]. Another variation of the same thermometer was introduced by Anders Celsius in 1742 who introduced his own scale - the Celsius Scale. Thereafter, variations of the thermometer were invented for recording temperature in different body parts. Apart from Santorios clinical thermometer, Sir Thomas Allbutt invented the first portable medical thermometer for body temperature in 1867 whilst Theodore Hannes Benzinger invented the first ear thermometer during the Second World War, which was later modified into an infrared ear thermometer by Dr Jacob Farden in 1984 [8]. The invention of the thermometer greatly improved on knowledge of the human body, the nature of diseases and the accurate prescription of medicine as body temperature is used as a necessary measure in many medical diagnoses today.

The invention of the hypodermic needle (which is usually used with a syringe to either extract or inject fluid) changed the way in which medicines were administered to people and increased the rate at which these medicines became effective in the body. The earliest hypodermic needle was invented by Christopher Wren who used it to inject dogs in 1656 [34]. In 1660, Johann Daniel Major and Johann Sigismund Elsholtz were the first to use the hypodermic needle on a human being however there were fatal consequences to this which made the needle a controversial medical tool [22]. The first successful hypodermic needle was invented by Dr Francis Rynd in 1844, and it was later modified by Dr Alexander Wood in 1851 when he created the all glass syringe [34]. The popularity of the hypodermic needle and syringe began in the early 1900s and was used particularly for insulin, morphine and penicillin administration. The invention of the hypodermic needle was one of the most significant innovations in medical technology. Today, the hypodermic needle has paved the way for more modified versions such as microneedles which serve the same purpose however decrease the chance of infection and eliminate pain.

The medical technologies mentioned above show some of the advancements in technology that have vastly improved the ways in which medical practitioners assess and treat diseases today. Technology has, and continues to evolve and the use of technology in the medical field has become indispensable. The countless technological inventions that have emerged over the centuries have progressively led to the development of areas such as biotechnology and information technology which are key elements in the medical field today. Information technology (IT) in particular has been making innovative transformations in the medical field. With an increase in the use of information technologies such as smart phones and tablets in everyday life, the medical field has incorporated these gadgets through the use of telehealth/m-health and electronic medical records.

According to Collen and Ball [23], the rise in the use of computers since the 1950s has led to the assimilation of information systems into medicine and healthcare. As a result, there is a vast spectrum of applications of information technology in healthcare today. One area that has been significantly transformed by information technology is medical imaging. Images used in medicine are of extremely high richness of detail and analysis of such images cannot accurately be done without the use of information technology. Advances in information technology have enhanced the quality of images which assist with the virtual construction/deconstruction of the body for clinical analysis, the detection of abnormalities and testing out medical interventions. Currently, there has been a gradual increase (particularly in developed countries) in the use of 3D medical imaging systems. These IT innovations have allowed surgeons to assess and solve medical anomalies prior to operating on patients and this has reduced risky exploratory medical procedures.

Through IT, healthcare practitioners have been able to bridge the physical distance between themselves and patients through telehealth. M-Health (also known as telehealth and e-health) is "the use of digital information and communication technologies, such as computers and mobile devices, to access health care services remotely and manage your health care" [19]. Such technologies include personal devices such as cellphones and computers, or remote patient monitoring devices used by doctors. M-health has been particularly useful in remote patient treatment. It has allowed for medical services to be available to people who may have transportation challenges or restricted mobility. Additionally, m-health enables people to make virtual appointments with doctors and have web-based doctor's appointments. The infusion of IT into medicine has not only eased access to healthcare, it has also broadened the scope of information available for research purposes.

The use of technology in the medical field has increased the effectiveness, efficiency and progression of medical research. Medical research and innovation play a vital role in the healthcare sector. Health informatics give medical researchers the necessary tools to assess and adjust medical treatments and procedures, and to uncover possible side-effects of drugs. Additionally, technology has aided medical research through innovations which allow for the cellular analysis of various organisms. This can be attributed to the micro-scaling of medical research technologies which make it possible to manipulate biological matter. Moreover, with the rapid digitalization of the world, there has also been a subsequent increase in the amount of data available to medical researchers. Today there are numerous online medical databases that are at the disposal of researchers which help with data comparisons.

With information technology expanding the amounts and types of data available to researchers in the medical field, it has become increasingly important to have a data science strategy which will effectively channel and organize data. Data collected in the medical field is highly personal and extensive, and requires effective data aggregation and handling. According to a study an estimated thirty percent of the world's stored data is produced in the medical field [3]. Data science allows for the medical practitioners to have a vast amount of medical knowledge at their disposal which can greatly improve their efficiency and the quality of their services. Additionally, a functional data science strategy handles data quality paving the way for effective data analysis and storage methods. Without an effective data science strategy, medical institutes risk deriving inaccurate and jeopardous conclusions. Moreover, this can also lead to the unlawful access of information by outsiders. A well-designed data science strategy will encourage precision medicine and facilitate learning health systems.

Effective medical data science management requires a data repository which will keep an inventory of all data assets in an organized fashion. The creation of a repository also involves the reconciliation of vast amounts of data. Such a repository will allow for healthcare practitioners to access electronic healthcare data which can help improve their services. Due to the highly personal information that can be found in a data science repository, security is of high importance. The protection of patient information through respecting privacy and honoring anonymity is maintained through data governance frameworks which warrant that these measures are secured by cyber security. Due to the large volume of data, it is necessary to have a team of dedicated data scientists who are responsible for the "processing, cleaning, statistics, visualization, operational research, management, and archiving/curation" of data [9]. Lastly, a reliable feedback system such as predictive analytics is valuable in such a broad system. The use of a data science strategy plays a great role in facilitating improved patient care and simplifying data access for medical research purposes. With the increased assimilation of technology into medicine and healthcare, improvements in data care systems will ensure the safeguarding of medical data and knowledge. An overview of a generic healthcare system is depicted in the Figure 1.1 which shows its ecosystem include many entities and stakeholders from the medical field. These include i) radiotherapy ii) imaging iii) diagnostic laboratories iv) medical emergencies v) medical home care vi) primary health care and specialist forming a multi-disciplinary team vii) pharmacies and viii) surgeries.

Healthcare evolution

The IoT and artificial intelligence have played a key role in the evolution of modern healthcare. Healthcare has indeed evolved from the use of invasive methods to evaluate blood samples to the use of IoT devices that make use of ultraviolet light to get body vital signs such as: peripheral oxygen saturation, heart rate, pulse, diastolic blood pressure, systolic blood pressure and many more. An example of such a device is the Samsung Simband. The devices are being developed and are available for research.

On the other hand, artificial intelligence is being massively deployed in the health sector to mitigate the fact that humans have limitation when it comes to processing massive sets of clinical data. While data emerging from genomics, medical imaging, scientific research and daily care seems too complex for humans to comprehend especially when the aim is to find patterns in huge datasets, artificial intelligence is currently used to find meaningful patterns in such data. Artificial intelligence combined with Big Data technologies are also helping to increase the speed and accuracy required in the medical field when performing surgical operations, identifying diseases (for example cancer through image processing), and patient prioritisation from real time data.

1.2 Research Objective

Research plays a fundamental role in healthcare. Through research, medical practitioners are able to provide people with effective and equitable healthcare. Over the years, technological advancements have led to the growth of medical databases and electronic medical records and

Figure 1.1: Healthcare overview.

this has opened up numerous opportunities for researchers in the health sector. The computerization of medical information has made it easier for researchers to access information which contributes to the understanding and discovery of disease patterns, the effectiveness of treatment and other factors in the medical field. However, there is a great need for researchers to have access to information such as anonymized patient data and database access in order to make research more effective. In an address to the American Medical Association, Barack Obama stated that too many doctors and patients are making decisions without the benefit of the latest research [96]. Researchers and medical practitioners are in need of increased access to the most recent, high quality medical data in order for them to make improvements in the healthcare sector.

This research aims to motivate research institutes to make anonymised medical data available to research for the advancement of medical science through machine learning algorithms. Which can only be effectively used to solve problems only if they have been trained using reliable data. The Machine Learning (ML) algorithms can be used together with IoT devices to provide cheap and reliable health-care. The application and importance of medical data availability for diagnostic purposes will be shown.

1.3 Research Gap

The first triage systems used in the medical field were used in battlefield situations where there were mass casualties. Today, triage systems are still used to assess mass casualties in hospitals through the categorization of patients who are in need of immediate, urgent or non-urgent medical attention. However, current triage systems are based on pseudoscience and hence cannot manage a wide range of medical presentations. Triage systems that merely sort patients into simple categories without specificity are inadequate, especially with the technological innovations taking place in the medical field. Triage systems should be able to manage critical illness and injury; giving advice based on the patient's age. Additionally, certainty can be added to the triage process through the integration of physiological measured parameters. This will improve the care given to patients and allow for the system to be used more especially for emergency treatment purposes. The use of pseudo-science based triage systems has inspired the need for the development of a scientific based triage system which can help address the shortcomings of some of the triage systems currently in use. The development and use of a scientific based triage system will assist medical staff in the patient prioritization process with increased in specificity; and allow for the applicability of this system in emergency hospital use. This research will be making use of a scientific based triage system in the patient prioritization process.

As stated in the research objective, access to medical data plays a fundamental role in research. Ventola [96] estimated that about 30 percent of the world's electronic data storage was used by the health sector. This data can be used to improve medical diagnosis through the integration of machine learning algorithms in healthcare. However, the use of machine learning algorithms in healthcare has been limited due to barriers to data access. The key issue today of integrating machine learning into healthcare has been collecting/accessing and using diverse data for the analysis and treatment of people. In order for machine learning to be effective, a vast amount of data needs to be available. Machine learning algorithms in healthcare have proven to be greatly beneficial in a number of ways. Machine learning has been used in the identification/-diagnosis of various diseases. One notable machine learning innovation in this area has been carried out by IBM Watson Health which has incorporated cognitive computing and genomic tumour sequencing in its diagnostic procedures [102]. Machine learning has also been used in the development of personalized treatments for patients. IBM Watson Oncology has been making strides in this regard through its integration of machine learning algorithms to improve diagnosis by using patient medical information and history [61]. Additionally, the use of machine learning in drug discovery has also become an area of interest in the medical field, with the MIT Clinical Machine Learning Group making use of and developing more algorithms to improve the treatment of diseases such as diabetes [36]. Although various big companies have started making use of machine learning in the medical field, a gap still exists in patient prioritization machine algorithms. This research aims to contribute on machine learning integration in healthcare by showing the benefits of increased medical data access to improved patient prioritization and personalized treatment development.

In summary, three main research gaps have been identified in this work. These include:

1. The lack of more scientific based triage systems in the patient prioritisation process in healthcare.
2. Inadequate integration of machine learning algorithms in patient prioritisation processes.
3. The lack of intelligent tools to support medical practitioners with disease identification.

1.4 Research Questions

The main aim of this research is to address the research gaps described above by assessing the effective use of machine learning algorithms in patient diagnosis and prioritisation. This will be achieved through the use of a science-based solutions. In order to achieve this, four research questions have been considered to be addressed as core subject of this research work. These include:

1. Are scientific medical decision support systems more accurate than pseudo-science health systems?
2. What is the best way to implement a scientific medical decision support systems?
3. What is the most effective way to analyse quantitative and qualitative medical data for decision making?
4. Is it possible to automate medical knowledge in medical decision support systems?

1.5 Thesis contribution and outlines

This thesis revisits cyber physical health systems (CPHS) with the objective of designing and implementing a data analytics platform that provides patient condition monitoring services in terms of patient prioritisation and disease identification [1]. Different machine learning algorithms are investigated by the platform as potential candidate for achieving patient prioritisation. These include multiple linear regression, multiple logistic regression, classification and regression decision trees, single hidden layer neural networks and deep neural networks. Graph theory concepts are used to design and implement disease identification. The data analytics platform analyses data from biomedical sensors and other descriptive data provided by the patients (this can be recent data or historical data) stored in a cloud which can be private local health Information organisation (LHIO) or belonging to a regional health information organisation (RHIO). The users of the data analytics platform consisting of medical practitioners and patients are assumed to interact with the platform through cities' pharmacies , rural

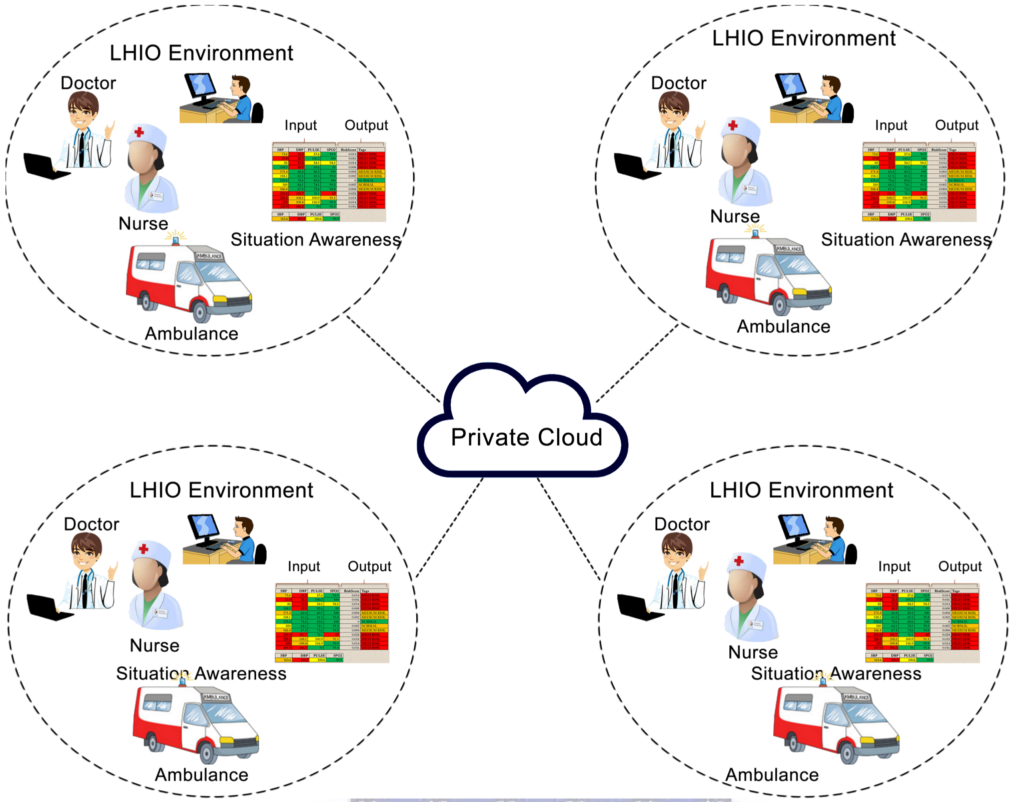

Figure 1.2: A Regional Health Information Organization Infrastructure.

E-Health kiosks end user applications. This data analytics platform is provided as a platform-as-service in a cloud ecosystem depicted by Figure 1.3 and Figure 1.2. As revealed by Figure 1.2 , a number of local health information organisation (LHIO) infrastructures can be federated into a cloud platform owned by a regional health information organisation (RHIO) to provide services to citizen. These include i) patient condition recognition ii) public health monitoring to detect the outbreaks of epidemics such as Cholera and Ebola iii) public health management by taking advantage of the massive data sets generated by the platform to get useful insights concerning the public health iv) remote access to medical expertise and v) many other rich services including "over-the-cloud" access to shared medical resources such as scanners, M-Rays, etc.

Figure 1.3 shows different entities that may benefit from the services provided by the RHIO ecosystem. These include

Home Care: Home care patients for patients with reduced mobility and elderly people could be improved through the cloud services provided.

Emergency: Emergency services will benefit from the cloud in terms of quicker and better information dissemination leading to faster response to casualties and other medical emergencies.

Surgery: Through cloud services, additional expertise and resources can be provided to a

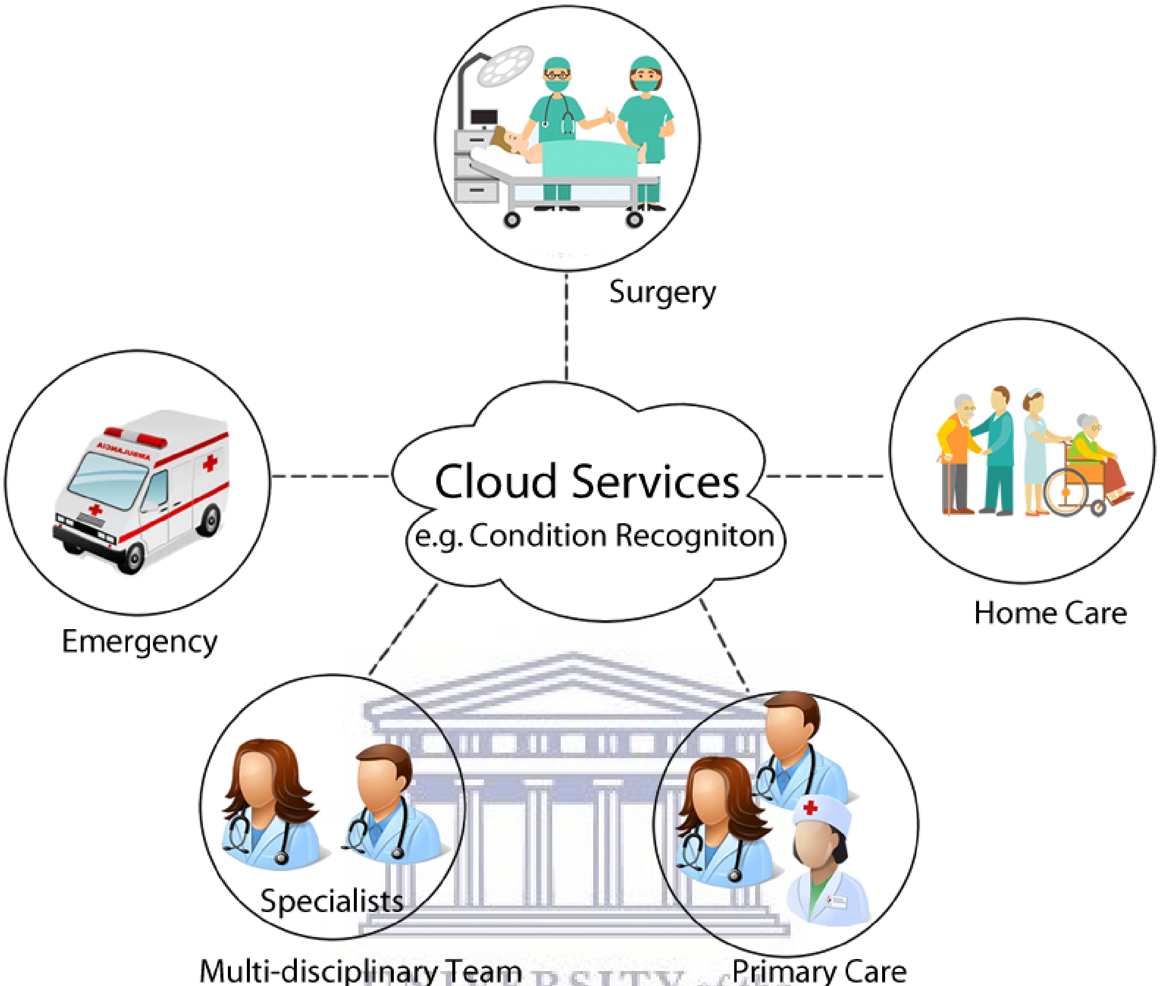

Figure 1.3: Ehealth cloud services end users.

medical team operating in a hospital theater.

Multi-disciplinary Team: These are specialists in one or more fields in medical sciences. An example of such a team is illustrated in Figure 1.4.

Primary Care Physician: They are also known as primary care providers. These are health care professionals who practice general medicine. They are also the first contact for a person with a health concern. See examples of areas of focus in Figure 1.5.

The E-health kiosks can also be used by the community to collect patients' vital signs and other descriptive symptoms. The patient data is then relayed into a network and stored in a cloud. The cloud services then perform situation recognition (disease identification and patient prioritization). The cloud supports information sharing among a number of entities or organizations to enable institutions, medical professionals, emergency workers, researchers and public health planners to access an integrated healthcare system for utilization, planning and advanced research.

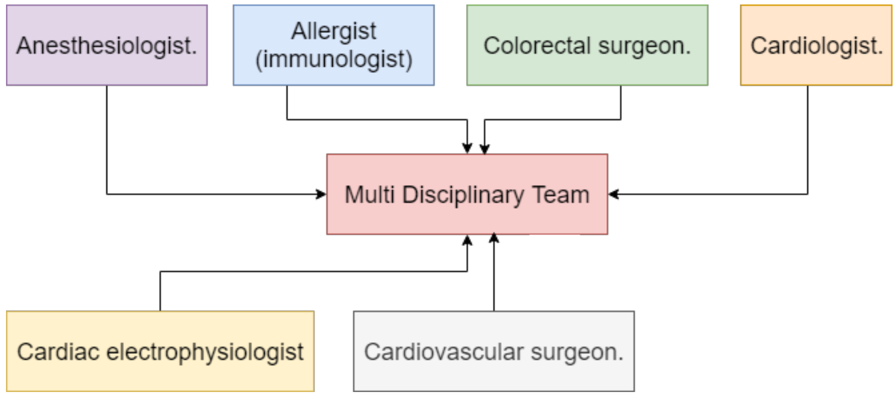

Figure 1.4: Medical multi-disciplinary team.

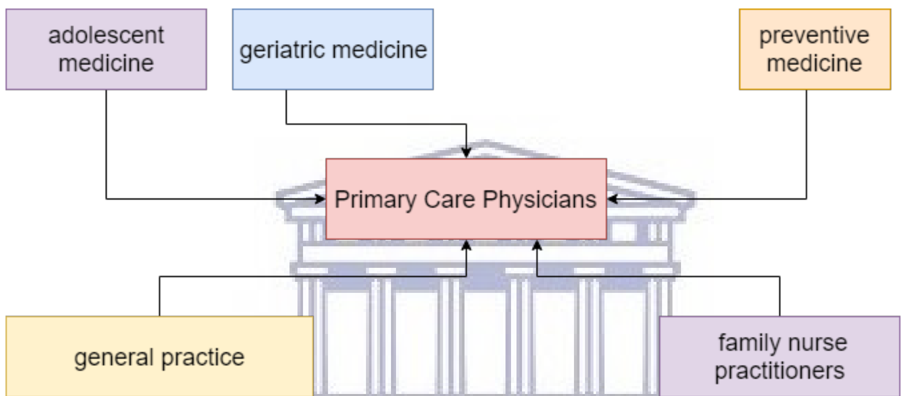

Figure 1.5: Primary care physicians areas of focus.

Some of the advantages of this system are [51]:

- Cost effective; unnecessary hospital admissions are avoided and thereby saving costs.
- Time saving; when deployed in well funded or private hospitals. The nurses on duty do not have to go around taking every patient's readings since each patient can have a low costly E-health kit or other smart devices which continuously take readings and update the medical record database hosted in the private cloud.
- Instant processing; the cloud services can be equipped with an intelligent data analysis algorithm capable of performing patient prioritization and situation recognition. Once new readings are sent to the database the doctors can monitor and access the patients' decision enabling visualizations remotely from their smart devices, tables and smart phones with no real time constraint.
- Reliable decision making; the machine learning algorithms implemented in this study for example the deep neural network is known to reach an accuracy close to 100 percent hence human errors can be eliminated. If given more time to train and more data the

algorithm's accuracy increases, training is achieved in a just a few seconds or just a few minutes depending on the size of the data and number of times given to adjust the biases. Advances in hardware has made it possible to make use of such algorithms. Personal computers now posses high computational power. Twenty years ago the fastest personal computer would take months or years to train such an algorithm.

The main purpose of this research can be summarised as providing decision support through situation recognition using the cyber domain. Figure 1.6 shows patient's condition recognition that includes two main components which form the core of this thesis work:

- Patient prioritisation using the triage system
- Disease identification though symptom check

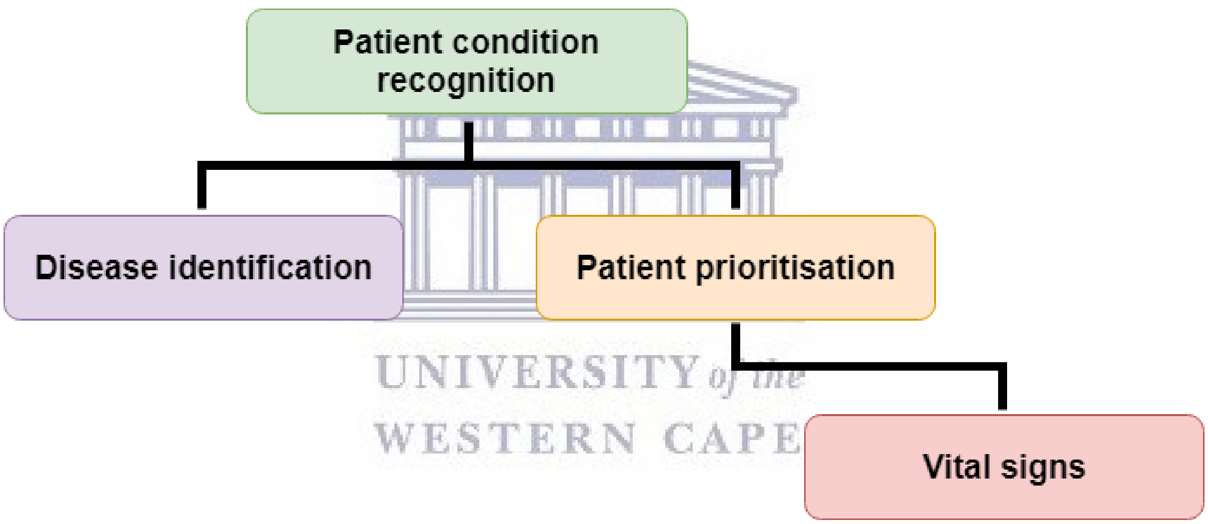

Figure 1.6: Patient's Condition Recognition.

The remainder of this thesis is as follows:

Chapter 2 defines and give a brief background of decision support systems, artificial intelligence, machine learning techniques, deep learning, graph theory, natural language processing, and fuzzy cognitive systems. It also discusses the application of AI in healthcare.

Chapter 3 recalls the history evolution, issues, motivation of the use of "Patient's Condition Recognition" and other related areas of technology in healthcare.

Chapter 4 presents the implementation and applications of machine learning techniques: "Logistic Regression, Linear Regression, Classification and Regression Decision Tree, Single Hidden Layer Neural Network (Shallow Neural Network) and Multiple Layer Neural Networks (Deep Neural Network)" to calculate the risk score.

Chapter 5 discusses the use of graph theory and social network analysis techniques in disease identification (patients symptoms are matched to symptoms in the database or knowledge repository).

Chapter 6 discusses the architecture and implementation of the proposed situation recognition system.

Chapter 7 evaluates the proposed patient prioritisation and disease identification algorithms.

Chapter 8 concludes and outlines future work of this study. The limitations and drawbacks encountered during the study are also discusses.

Chapter 2

Background

2.1 Decision Support Systems

Decision support systems in healthcare – known as clinical decision support systems – are applications used to analyse data which will assist medical practitioners with timely, clinical decision making. In the medical field, decision support systems are used to prepare and review medical diagnosis in order to ensure accuracy of the final result. Data mining is also used to assess the patients' medical history and any research that may have been done on similar cases. An analysis such as this can make predictions which include symptoms and potential cures and other possible. Decision support systems are mainly used to manage routine tasks such as warning of potential problems or diagnosis suggestions. The use of a decision support system in healthcare centres such as hospitals can greatly improve efficiency, reduce patient waiting time and lower costs to both the hospital and the patient. The use of decision support systems in healthcare has been increasing in countries such as the USA which have called for the implementation of such systems. The Health Information Technology for Economic and Clinical Health (HITECH) Act states that healthcare centres should "demonstrate the meaningful use of health IT... [healthcare] providers must implement one clinical decision support rule..." [49]. This shows the importance of incorporating such systems in today's healthcare system. As this study makes use of machine learning techniques and graph theory to solve medical service delivery problems. The following sections will give an introduction to machine learning, artificial intelligence and graph theory.

2.2 Artificial Intelligence (AI)

Contrary to natural intelligence (NI), which is the intelligence exhibited by human-beings and animals, Artificial Intelligence (AI) (which is also referred to as Machine Intelligence) is intelligence shown by machines. In 1956, Artificial Intelligence (AI) was established as an academic discipline which experienced a number of optimistic discovery's and creations. However, the discipline also went through numerous failures which led to decreased funding - this period was referred to as the “AI winter.” Thereafter, the invention of new approaches in Artificial Intelligence (AI) revitalized the discipline. Traditionally, Artificial Intelligence (AI) research targets areas such as perception, reasoning and object manipulation. One of the long term goals of the discipline is general intelligence. The AI discipline incorporates information from a number of fields which include computer science, neuroscience and psychology. The belief that human intelligence “can be so precisely described that a machine can be made to simulate it,” was the foundation of the AI field [35]. The twenty-first century has seen a significant rise in the use of AI, with AI fast becoming the backbone of the technology industry and resolving a number of issues in computer science.

In the field of Computer Science, research on Artificial Intelligence is known as ‘the study of intelligent agents.’ An intelligent agent would be any device that is able to identify and react to the environment it is in, in order to attain a goal. From a colloquial standpoint, AI is implemented when a machine imitates human cognitive functions which include problem solving [44]. There has been debate regarding the scope of AI. With the capabilities of machines today broadening at a fast pace, the ‘AI effect’ has resulted in some previously recognized AI tasks being removed from the definition of AI. Hence Tesler’s Theorem states that “AI is whatever hasn’t been done yet,” [39]. An example would be optical character recognition which is often omitted from AI as it has become a regular technology. As of 2016, some of the recognized AI capabilities include military simulations, analysis of complex data that includes images and videos, and autonomous cars [76].

Artificial Intelligence has also become an integral part of the healthcare sector, with AI being used to assist with patient treatments, medication recommendations and patient monitoring. Bloom Technology states that Microsoft has developed AI to assist with the discovery of cancer treatments [14]. Generally, research on cancer is extensive with a vast amount of drugs having been developed to fight the disease. There are well over 800 different medicines and vaccines all related to cancer treatment. This has made it difficult for doctors to find the right drugs for patients as there are too many possibilities and combinations available. In light of this problem, Microsoft has set out to develop “Hanover” - an AI machine which will memorize all relevant cancer related information and make predictions on the drug combinations best suited to a patient. AI has also been found to be useful in assisting doctors identify various skin cancers. In yet another study, AI has been used to monitor multiple high risk patients which has been done through data acquired from live doctor-patient interactions [79]. One of the most well-known and advanced AI inventions was created by IBM which created the IBM Watson. This AI machine has been able - at some level - to surpass human intelligence. Watson was able to successfully perform a leukaemia diagnosis.

Figure 2.1 is an illustration of AI in a Venn Diagram.

2.3 Machine Learning (ML)

Machine learning is a sub-field or subset of AI and is in fact an approach to achieve Artificial Intelligence. It refers to any computer program that can ”learn” by itself without having to be explicitly programmed by a human. The underlying idea and the phrase originated in Alan

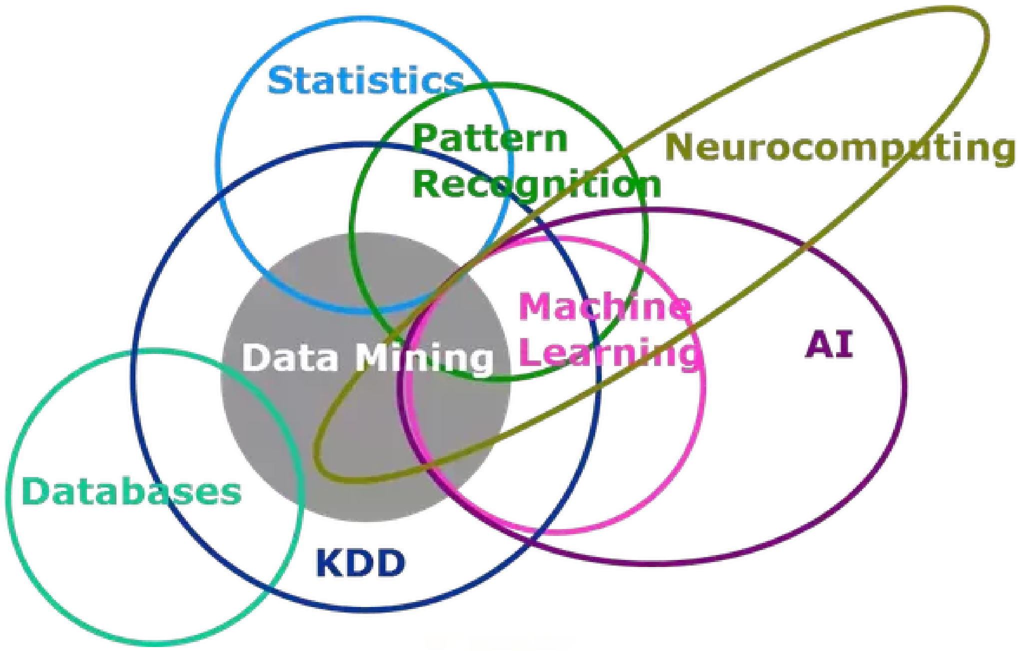

Figure 2.1: Understanding AI in a Venn-diagram.

Turing's seminal paper "Computing Machinery and Intelligence", which featured a section on his "Learning Machine" that could fool a human into believing that it's real [90]. The machine takes a small or huge volumes of data, learns from it and then makes a prediction about something that is related to the data.

Example of machine learning algorithms:

- Linear Regression
- Logistic Regression
- Artificial Neural Networks
- Random Forests
- Naïve Bayes Classifier
- K Means Clustering
- Decision Trees
- Support Vector Machine
- Apriori
- Nearest Neighbour

2.4 Deep Learning

Deep learning is also known as hierarchical learning or deep structured learning. It is a technique for implementing machine learning and to be more specific a class or subset of methods based on data representations. It can make use of supervised or unsupervised algorithms. Deep learning characteristics are as follows:

- use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input
- use some form of gradient descent (stochastic, mini batch or batch) for training via back-propagation
- use supervised learning (e.g. classification) and/or unsupervised learning (e.g. pattern analysis)
- learn multiple levels of representations that correspond to different levels of abstraction; the levels form a hierarchy of concepts.

Figure 2.2 is an illustration of deep learning in a Venn Diagram. The strong relationship among the three: deep learning, machine learning and artificial intelligence is also shown. There are all related to big data and other enabling technology for example pattern recognition and neuro computing as shown in Figure 2.1.

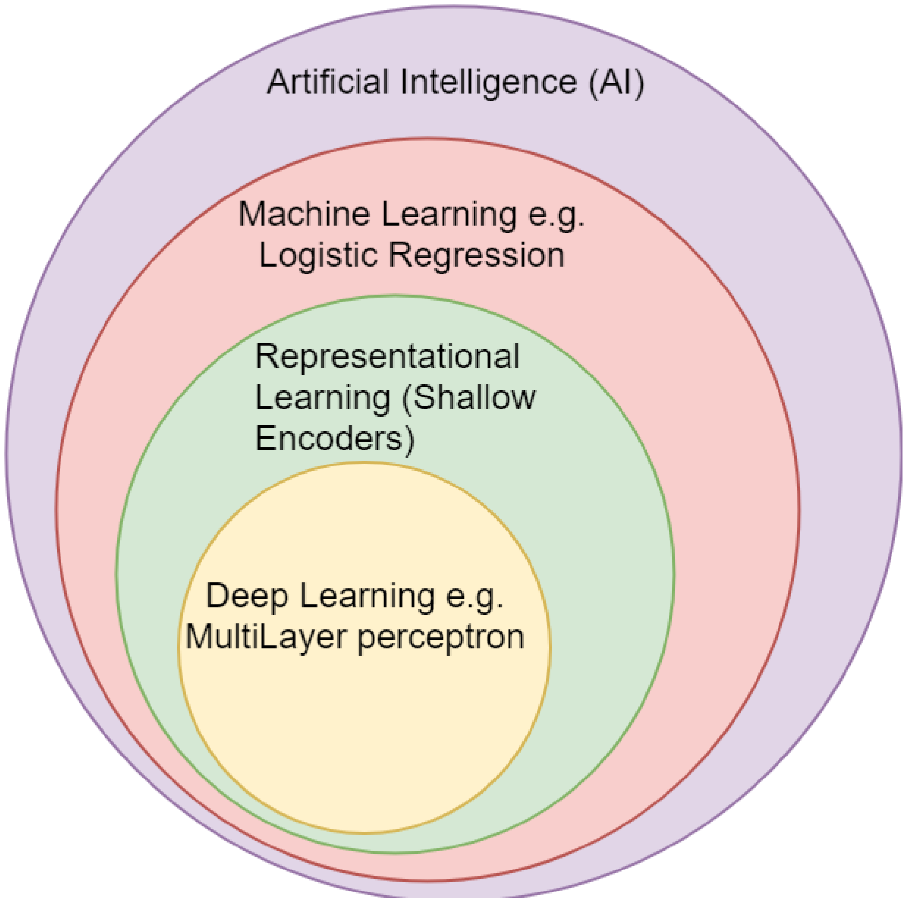

Figure 2.2: Understanding Deep Learning in a Venn-diagram.

2.4.1 Deep Neural Network

Deep neural network is known to be the development of machine learning theory. Deep neural learning was introduced as a way to counter the limitations of the perceptron learning algorithm which could not fully understand the ambiguity of 'or' within natural language. In light of this, Marvin Minsky - a well renowned Artificial Intelligence expert - emphasised the need for advancements in learning algorithms. Numerous deep learning methods make use of neural networking structures hence deep learning is often known as deep neural networks.

A deep neural network can be described as a complex neural network with over two layers. Scholars such as Rifkin (2014) and Danzer (2007) have stated that neural networks normally have "an input layer, an output layer and at least one hidden layer in between." Therefore deep neural networks can be described as 'feed-forward' networks. These networks utilize intricate mathematical modelling for data processing. Generally, neural networks are algorithms which are used to mimic human brain activity; particularly pattern recognition through the multiple layers of replicated neural connection. These layers use the 'feature hierarchy' process which entails using machine perception for the categorisation and ordering raw data. The main use of neural networks is management of unstructured data. Neural networks identify numeric data in vector. The phrase "deep learning" describes the level at which information is sorted, going beyond mere input-output practices. This is a form of machine learning which utilizes Artificial Intelligence. Hence deep neural networks are part of bigger machine learning applications which make use of algorithms for fortified learning, regression and cataloguing.

2.5 Decision Trees

Decision trees are basic machine learning models which are used to analyse regression data. Decision trees create regression models which are structured like a tree. It categorizes a dataset into multiple, small subsets whilst also accumulatively developing a decision tree. This tree has decision nodes and leaf nodes. A decision node has at least two branches which each represents a value for the test in question. A leaf node represents a decision on the numerical target. The root node is the ultimate decision node as it correlates with the best predictor. Decision trees are able to operate with both categorical and numerical data. The regression decision tree is used when the target variable is continuous in nature, including for predictive purposes. Decision tree learning makes use of decision trees for predictive purposes. This approach is mainly used in fields such as machine learning and statistics.

Decision trees make use of algorithms and one of these is the Classification and Regression Decision Tree (CART) algorithm which was introduced by Leo Breiman [63]. CART can be used to evaluate and resolve classification or regression predictive modelling problems. This algorithm provides an important base for other essential algorithms such as random forest and bagged decision trees. According to [74], CART tree is a binary decision tree that is constructed by splitting a node into two child nodes repeatedly, beginning with the root node that contains the whole learning sample. CART is amongst the most popular decision tree algorithms used today.

2.6 Classification and Regression Decision Tree (CART)

A decision tree is an ML algorithm that uses a set of instances in-order to make decisions. The decision node and the leaf node are the only two types of nodes. Every leaf node is a terminal node of the tree and specifies the ultimate decision of the tree. The decision nodes are used for testing a particular attribute.

The unlabeled instance is classified by routing the test case down the tree according to the values of the attributes tested in successive decision nodes until a leaf is reached. The unlabeled instance is then classified according to the probability distribution over all classification possibilities. The 'divide and conquer' approach is used to construct the decision tree. Initially an attribute is selected and placed at the root node of the tree. The root node then splits up and divides the dataset into different subsets, one for every value of the attribute. The value is specified by the edge between the parent node and its child node. The process is recursively repeated for every branch of the tree [31].

2.7 Logistic Regression

The logistic regression model used in fields such as statistics, machine learning and medical fields is a regression model which makes use of a categorical dependent variable. This model was developed by 1950s statistician David Cox. In instances where the dependent variable elicits multiple results, analysis is carried out using multivariate logistic regression. Logistic regression is also suitable for analysing dichotomous dependent variables. Moreover, it is a predictive analytic tool. The logistic regression model is used for descriptive and explanatory purposes particularly where there is a relationship between a dependent binary variable and numerous independent variables. According to [28], logistic regression generates the coefficients of a formula to predict a logit transformation of the probability of presence of the characteristic of interest.

Within the medical field, logistic regression has been used in machine learning to assist with predictive diagnosis in cases where there is trauma or severe injury for example. This approach was originated from Boyd, Tolson and Copes 1987 Trauma and Injury Severity Score (TRISS) evaluation method which used logistic regression. Logistic regression has been used to predict whether or not a person has a particular disease based on various attributes such as age and body mass index. For the purposes of this research, multivariate logistic regression will be used in a similar manner. Variables such as systolic blood pressure, diastolic blood pressure, heart rate and peripheral oxygen saturation will be used to predict a person's risk score. The risk score will reflect an individual's degree of risk in light of various risk factors.

2.8 Graph Theory

The medical research field has a lot of literature and graph based algorithms have been used for information retrieval and natural language processing. There are high chances of disease and symptom relationships being mixed up and as a result patients being misdiagnosed and given the wrong medication. For example the symptoms: Fatigue, Fever, Headache and Sore throat are related to about 172 medical conditions. To mention but a few: viral pharyngitis, acute sinusitis, mononucleosis, mumps, medication reaction or side-effect, aseptic meningitis, thalassemia, multiple sclerosis, sunburn, lyme disease and Diabetes type 2 [48].

One of the most universal models of both natural and human-made structures in computer science are graphs. Graphs have been used to model various relations and process dynamics in this field such as biological systems. Graph theory is a division of mathematics which has been used in computer science. Graph theory "is the study of graphs which are mathematical structures used to model pair wise relations between objects,". Graphs have been used to develop problem solving systems and have been applied in various areas of computer science such as database networking and data mining. The use of graph theory in computer science has led to the development algorithms which have been used in various applications. This theory was developed in 1913 after having built up on and improved the Eulerian graph concept created by Leonhard Euler in 1735.

2.8.1 Graph Theory Algorithms

In computer science, graph theory has mainly been used to develop graph algorithms. A number of algorithms have been used to resolve problems displayed in graph form. Such algorithms are applied to the resolution of graph theoretical concepts which are then used to resolve the parallel computer science application problems. Examples of such algorithms include:

- Finding a minimum spanning tree.
- Algorithms to find cycles in a graph.
- Finding graph planarity.

2.8.2 Graph Applications

Graph are used to solve social, physical, information and biological systems.

- Semantic networks - in computational linguistics and natural language processing.
- Travel eCommerce Websites - determining the cheapest route.
- Global Positioning System (GPS) - finding the fastest route or shortest path.
- Social networking - e.g. network of family and friends on social media applications such as Facebook.

2.8.3 Graph theory and Machine Learning

The recursive neural network model is able to process directed acyclic graphs with labelled edges. There a a number of fields that make use of graphs in machine learning examples are computer vision and image processing.

Other graph applications in the medical sciences are as follows[98]:

- Genomics Profiles
- Molecular networks
- Cancer Sub-networks and Pathways
- Drug and Disease Phenotype Network
- Network-based Drug Re-positioning
- Biomedical and Molecular Networks
- TCGA Studies

2.8.4 Natural Language Processing (NLP)

Natural Language Processing (NLP) is a field of machine learning that studies how computers process and analyze natural language data. The two categories are rule based and statistical.

2.9 Fuzzy Cognitive Systems and Machine Learning

The fuzzy cognitive maps (FCMs) described in this study are used for disease identification and pre-diagnosis qualitative predictions. Machine learning algorithms such as Multivariate Linear Regression (MLR), Neural Networks and Support Vector Machines (SVM) are used to perform regressions, predict and score patients. The results from both the FCMs and quantitative machine learning algorithms are then used to give a much more reliable patient prioritization list. Algorithms that perform quantitative computations such as the MLR are less reliable on their own. Hence the use of an hybrid system which makes use of qualitative FCMs to extract relevant information and provide some quantitative and qualitative output. The output from FCMs together with quantitative inputs such as body vital signs can be feed into a machine learning algorithm and compute one score which can be used to prioritize patients. Other more descriptive results are also given as output and can be used by medical practitioners to make decisions.

Table 2.1: Machine learning algorithms characteristics and features.

Learning Algorithm	**Type**	**Function**	**Output**
DNN	supervised, semi-supervised or unsupervised	linear and non linear	real values
SNN	supervised learning	linear and non linear	real values
MLR	supervised	linear	real values
MLoR	supervised	linear	real values
CART	supervised	non linear	real values

2.10 AI in Healthcare

DermaCompare is one of the revolutionary AI inventions that have transformed medical diagnosis in the healthcare industry. DermaCompare is a skin cancer detection application which runs on "cloud-based artificial intelligence technology," [83]. Early detection of skin cancer is crucial for patient treatment, and increases the chances of survival. Dermatologists have been using Total Body Photography (TBP) – an innovation created 15 years ago – to detect skin cancer. However, using TBP has proved to be time-consuming and inaccurate at times as it is a manual diagnostic procedure. Additionally it is expensive and this has made it inaccessible to many people. DermaCompare was created to overcome these challenges by providing a cost-effective, accurate and speedy way of detecting skin cancer. The application mainly uses image processing software (which is inclusive of knowledge from both military image software), big data analytics and patented comparing algorithms to aid the analysis of medical images. DermaCompare works closely with dermatology for the early detection of skin cancer. The DermaCompare application works by using AI algorithms to compare over 50 million known moles on their database to a picture taken by a person on their smartphone. A summary of relevant information is then generated by the app and a doctor is notified. This enables doctors to identify melanoma in a speedier and more accurate manner and has vastly improves telemedicine of skin cancer diagnostics. Babylon – like DermaCompare – is another AI application used in the healthcare industry that has vastly improved the diagnosis process. "Babylon is a subscription health service provider that enables users to have virtual consultations with doctors and health care professionals via text and video messaging through its mobile application," (www.babylonhealth.com). Babylon has combined the latest AI technology with the experience, knowledge and expertise of healthcare practitioners to make access to medical advice and diagnosis simpler, faster, affordable and accessible. Babylon works through an AI-powered chatbot triage service which gives patients access to consultations, health monitoring, tests and kits and AI. The consultation feature gives patients' access to a GP who can answer questions on common medical topics. Prescriptions are emailed to the patient. There is also access to therapists to discuss topics such as depression. The health monitor feature allows users to sync their activity tracking devices (e.g. smart watches) to the Babylon app. The delivery service – test and kits – provides users with the option to request diagnostic kits which can test for sugar levels, cholesterol, etc. The AI feature pre-screens a user's symptoms or health condition, increasing the accuracy of triaging patients. Data-driven techniques such as machine learning have been known to reinforce AI and subsequently improve healthcare services and systems. One company harnessing the power of machine learning and AI in the healthcare sector is Google DeepMind. Google DeepMind is a division of Google's AI company has been applying machine learning to medical data for diagnostic purposes on its healthcare platform. It has been applied to the diagnosis of eyes, cancer screening and electronic patient management. Google DeepMind is an app which aims to assist (and not replace) medical practitioners; allowing them to work more efficiently at the touch of a button. The early detection of various ailments which warrant rapid response will increase patient treatment. Google DeepMind has collaborated with various hospitals and organizations where it has rendered its services. The tools mentioned above are just some of the innovations in healthcare AI that have made significant impacts. The table below gives a more extensive list of AI tools currently available in healthcare:

Table 2.2: Machine Learning Algorithms.

AI Tool	Company	User	Use	Specialisation	Algorithm
Google DeepMind Health	Alphabet Inc. (Google conglomerate)	Practitioner	Data mining Electronic patient record management	Kidney data Eye diseases	Machine learning
Verily	Alphabet Inc. (Google conglomerate)	Practitioner	Data mining Disease monitoring	Genetic data	PageRank
WatsonPaths	IMB Watson	Practitioner	Electronic medical records management	Decision making	Bowyer-Watson
CareSkore	CareSkore	Practitioner Patient	Predictive analysis Patient communication service platform Personalized population management	Patient readmission prediction Value based care	Zeus
Butterfly Network	Butterfly Network, Inc.	Practitioner	Medical imaging	MRI and ultrasound	Deep learning
Zephyr Health	Zephyr Health	Research	Data mining	Data visualization	Machine learning
CloudMedX	CloudMedX Health	Practitioner	Data mining Generate real-time clinical insights	Natural language processing	Evidence-based Machine learning
Dragon Medical Virtual Assistant	Nuance	Patient	Enhancing clinician-patient interactions	Communication	Arbitration

The application of AI tools in healthcare has been briefly introduced. The next chapter provides more history about the evolution of healthcare and then discuss some few related studies in more details.

Chapter 3

Related Work

Over the last decade, wearable devices have been receiving an increased attention in both the academic and industrial sectors. Constantly evolving medical devices have resulted in the development of portable wearable devices which have been used to empower both health care professionals and their patients. Portable wearable devices in the medical field are used by individuals in order to support their general health or recovery through monitor aspects of their health over a period of time. These devices are defined as “devices that can be worn or mated with human skin to continuously and closely monitor an individual’s activities, without interrupting or limiting the user’s motions [10].” Portable wearable devices are unique health support tools as they can be worn as an accessory or implanted into clothes. The use of non-intrusive physiological detectors, wireless data dispatch systems and medical data processing algorithms in these devices allows them to constantly monitor and record patient health progress. Additionally, both patients and doctors can access real-time feedback and the alerting system can warn users of any potential health risks. The simple design makes it suitable for unskilled persons to understand and use. The emergence of portable wearable medical devices has vastly improved the delivery of point-of-care health services and distance medical management, monitoring and support. This has been particularly helpful for chronically ill individuals, the elderly and people with disabilities. In order to understand the progress that has been made in the field of wearable technology, it is important to look at the history of these devices.

The market and demand for wearable biomedical sensor devices has been catapulted by the fitness and health industries. According to Tillmann [89] demand has grown by 25 percent from 2014 to 2015 and it has been predicted that this rapid growth rate will continue over the coming 5 years. Synchronously, wearable biomedical sensor devices have started being utilized more to keep track of stress and human performance during strenuous activities, as well as to enhance home-based care specifically made for patients with noted health problems. At the moment, wearable biomedical sensors have been developed to provide ongoing supervision of human vital signs in a comfortable and noninvasive manner. The consistent monitoring of an individual’s vital signs enables the establishment of a health baseline which enables both users and health care practitioners to be signaled when there are any abnormal variations that may require medical attention.

The shift towards more flexible monitoring devices that amalgamate well with an individual’s skin and body can aid the increased acceptance of these devices as well as refine monitoring mechanisms. The authors in [103] states that significant efforts are being made to incorporate advanced fabrication methods and materials, and apply them to sensors and electronics with the aim of establishing pliable and conformal electronic devices. The large and inflexible nature of traditional silicon-based devices can obstruct their "applications in epidermal and implantable medical sensing" [43]. Materials such as elastomeric substrates and plastic which have been used as alternates in flexible devices are viewed as being conformal by nature and weightless, hence providing a more desirable interface with an individual's' skin and soft tissue.There are other studies in conductive and stretchable material used in sensors [65, 87]. In the same light, Kim et al look into alternative approaches which explore the transmission of thin, minute traditional devices onto flexible substrates. It has been found that silicon-based electronics offer unprecedented performance in the processing of data and in terms of general functionality. Hence the most desirable design for a wearable medical device which is both comfortable and flexible would require a fusion of progressive, tractile materials and silicon integrated circuits. A compound approach such as this one takes into account the multiplicity of body silhouettes and sizes that exist in the populace, easing the production of customized medical devices [42]. Flexibility, adjustability and adaptability boost the quality of signals recorded, and also reduce inaccuracies [82].

3.1 History

Portable wearable technologies come in various forms which include clothing and accessories such as watches and glasses. All these technologies make use of progressive computer and electronic systems. Most designs make use of simple, user-friendly features however each has its own aesthetic agenda. Currently, there has been a surge in portable wearable devices on the market. The history of portable wearable devices dates back to the early 1600s, when the first abacus necklace was invented [40]. Additionally, in China during the Qing Dynasty (1644-1911) the first abacus ring was invented to make calculations easier for traders. The 1800's saw the introduction of the wristwatch which was first worn by the Queen of Naples [13]. Computerized timing devices were created by mathematicians Edward Thorp and Claude Shannon in 1961 as cheating devices during gambling games [17]. The device was improved on by various people in the 1970s making it more sophisticated. In 1977, C.C Collins introduced a camera-to-textile wearable system for the blind. The popularity of wearable computers increased in the 1980s [75]. A study states that in 1975, Pulsar released the first calculator watch which became a trend setter during that time [80]. The watch had a small numeric keypad and came with a stylus for users. However these elaborate watches - which were made of solid gold - became popularized when companies such as Casio and Seiko created their own less expensive version of these watches with some more complex calculators in the early 1980's. The need for mobility in the media field led Steve Mann to create "a backpack-mounted multimedia computer in 1981" [40]. In 1994, Mann went on to create a wearable, wireless webcam. Also in 1994, Edgar Matias and Mike Riucci invented the first wrist computer. The 21st century saw a significant rise in portable wearable technology advancements and developments. The introduction of augmented-reality support wearable devices started in 2002 with Kevin Warwick's Project

Cyborg which electronically linked his nervous system to an electrode embedded in his wife's necklace [99]. The late 2000s saw the beginning of the production of wrist-phones. In 2011, Google presented the first model of smart glasses in the Google Glass Project. Thereafter, a number of big companies such as Samsung and Apple introduced their own portable wearable technology. Today wearable portable technology comes in various forms from rings and headbands to glasses and watches. These devices have been used for a multitude of purposes which include fitness tracking and health monitoring. Portable wearable devices for medical purposes have become increasingly popular, particularly the ones in the form of a wrist watch.

3.2 Patient's Condition Recognition

Healthcare is a complex industry generating massive volumes of data emerging from medical imaging, genomics, sensors, daily care and scientific research which may have hidden patterns which are difficult to analyse by human beings, both medical and non-medical professionals. Artificial intelligence is able to provide more insights into this data as it has the potential to process a huge number of combinations that the human brain cannot mathematically handle, thus enabling medical professionals to handle the treatment and medical care more effectively and with more time for empathy. Artificial intelligence has been used in medicine for many different specialized applications. For example, a reliable epileptic seizure detection model using an improved wavelet neural network was proposed in [56] while an acute ischaemic stroke prediction model from physiological time series patterns was proposed in [81] and believed to be useful in optimizing stroke recovery by manipulating physiological variables. Artificial intelligence prediction was also proposed in [104] to improve elective surgery scheduling.

3.3 Patient Prioritization

Patient prioritization is an instrumental process in the treatment of patients. It has become increasingly important for hospitals to integrate a patient prioritization system into their everyday system as it has proven to be greatly beneficial in many aspects which include time management, rate of service delivery and diagnosis accuracy. Public hospitals have struggled with funding and this has impaired the quality of health care provided. Issues which include a lack of adequate resources; insufficient and outdated medical equipment, limited health care professionals have all contributed to service provision problems. A study conducted in 2004 showed that approximately there are are at least 6.7 doctors per 10 000 people [12]. As increasing the number of medical practitioners available would not be feasible due to financial constraints, the most effective solution would be to create more efficient hospital systems through the incorporation of patient prioritization systems. This research aims to reduce mortality through the use of patient prioritization. This will ensure patients with time-sensitive illnesses, diseases or injuries are catered to on time thus reducing the rate of unnecessary deaths. This research will achieve this by identifying the most suitable patient prioritization and disease identification algorithm.

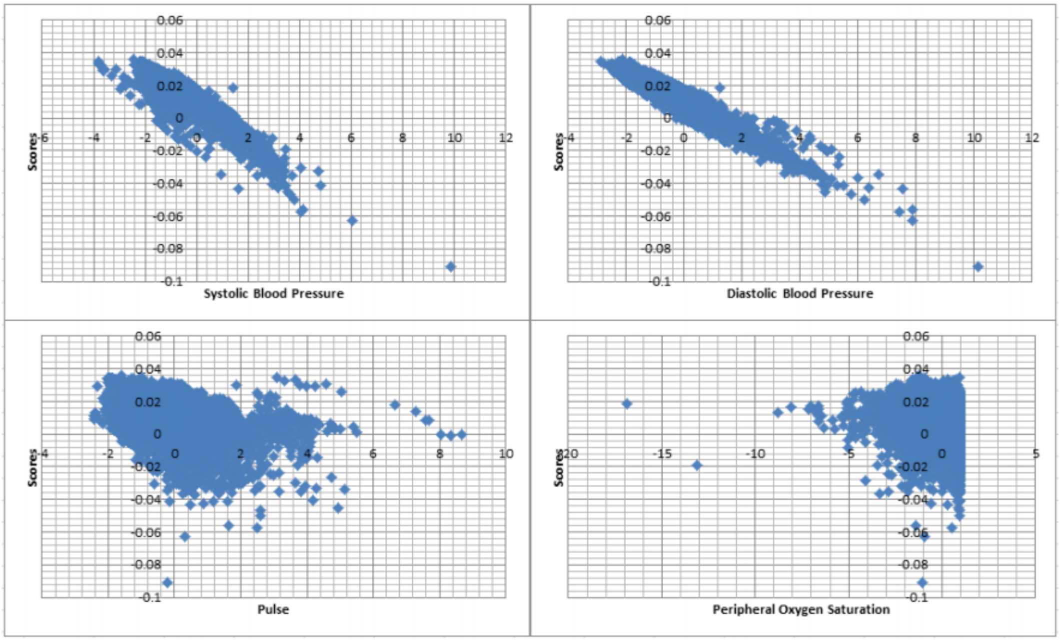

Figure 3.1: Physiological data exhibiting characteristics.

The availability of medical data helps with the investigation of important parameters that greatly affect the value to be predicted. In a previous study [62], data patterns were identified as shown in Figure 3.1. These patterns are a key player in discovering new algorithms and ground breaking implementations and inventions. The Multivariate Linear Regression (MLiR) algorithm played a huge role in identifying these patterns. A comparison to other algorithms such as the Support Vector Machine (SVM) and Kmeans Clustering and Gausian Estimator (KCG) was made. MLiR was found to be the best algorithm. The algorithm was successfully implemented to solve multiple classification for patient prioritization.

In this study other algorithms such as Multivariate Logistic Regression (MLoR) and Classification and Regression Decision Tree (CART) are implemented and the performance of each is compared to the Multivariate Linear Regression (MLiR). The other algorithms for disease identifications are also introduced. A web application, a cloud service applications and visualizations are designed and implemented.

The Multivariate Logistic Regression (MLoR), Classification and Regression Decision Tree (CART), Single Hidden Layer Neural Network (SNN), and a Deep Neural Network (DNN), were implemented and compared to the Multivariate Linear Regression (MLiR), from the previous study. The DNN is famous for its ability to handle very difficult prediction problems. Patient Prioritization makes use of qualitative and quantitative variables. An example is the South African Triage system makes use of qualitative parameters such as Mobility, Trauma, and (Alert, Reacts to Voice, Reacts to Pain, Unresponsive also known as AVPU). The numeric parameters are Respiratory Rate, Temperature, Systolic Blood Pressure (SBP) and Heart rate. The advancement of IoT has seen a lot of vital sign monitoring devices penetrating clinical trials and research. An example is the Samsung Simband wearable device which measures parameters such as Electrocardiogram (ECG), Blood Flow, Blood Pressure, Heart Rate, C02 and Oxygen Level and Skin Temperature (see Figure 3.2). Most of these can be useful and can be used together with the South African Triage parameters. The DNN can take many parameters and still achieve great performance.

3.4 Vital signs monitoring using biomedical wearable sensors

Continuous vital sign monitoring plays an important role in and contributes immensely to the treatment and prevention of various health issues. Generally, when there is a need to assess a patient's vital signs it usually takes place either at a hospital or at the patient's residence with the use of considerably large, expensive medical equipment which requires health care personnel to be available to operate it. However, contemporary society has developed and introduced wearable biomedical sensors as a way to reduce costs, increase patient mobility and possibly improve the quality of data health care practitioners use in patient diagnoses. This notion is concurred by Hao and Foster [38] who state that the wireless nature of these devices enables better mobility for patients and improves efficiency levels. In addition to this, the study looks into the use of wireless body sensor network (WBSNs) in health related applications and found that the chief driving forces behind the shift towards the use of WBSNs and vital sign monitoring were the increased obesity levels and the aging of the populace. This has led to the need for significant medical interventions with incidentally significant cost. Similar research focused on diabetes and cardiovascular disease and concluded that constant monitoring of blood glucose levels is essential to prevent any further complications related to these illnesses [67]. There are a various vital signs that are monitored by of wearable vital sign monitoring devices which include blood pressure, elecotrocardiograph (ECG) patterns and blood glucose. These will be discussed in greater detail below with specific reference to various studies that have been conducted regarding the effectiveness of wearable vital sign monitoring devices.

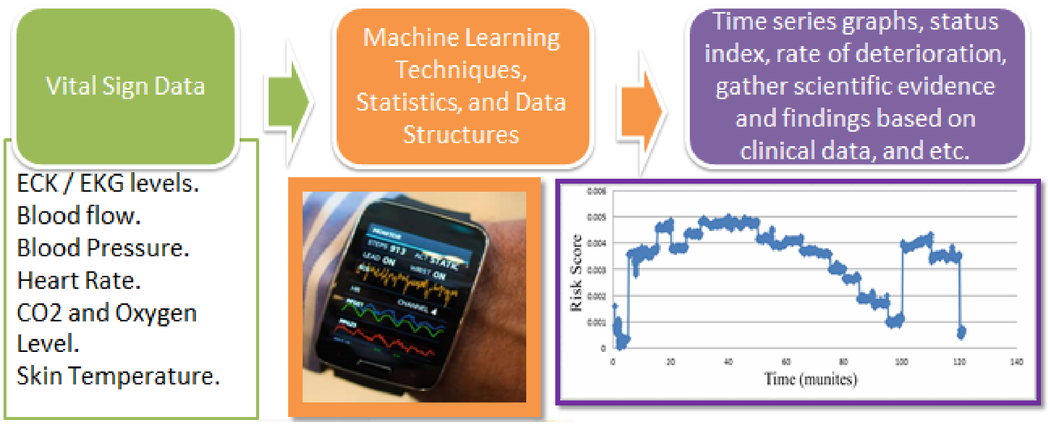

Figure 3.2: Vital Sign Monitoring using the Samsung Simband.

3.5 Continuous assessment of heart rate and cardiac activity

Traditionally, an individual's cardiac activity is assessed and monitored in a clinical setting through the recording of electrocardiograph (ECG) signals. Holter monitors have typically been used for these purposes, especially after a patient has undergone heart surgery [105]. Even though Holter monitors can effectively monitor cardiac activity, the machine itself is bulky and has numerous wires which connect to the central unit of the monitor. Hence the use of such large equipment impedes a patient's ability to go about their daily activities, is time consuming and is not ideal for constant cardiac activity monitoring.

Technological breakthroughs in the wearable biomedical sensor domain have resulted in the evolution and miniaturization of Holter monitors making them more flexible, effective and portable. Such devices have commonly come in the form of watches which provide real time information regarding an individual's cardiac activity as well as allow a person to store and analyze the recorded information at a later stage on a computer [73]. In recent years, advancements in technology have allowed for such information to be linked directly to a person's cellphone using bluetooth which has increased the accessibility of information. HealthGear is one such device which allows for the monitoring of heart rate through a person's pulse. Data collected using this device is accessible on a cellphone where it can be stored and analyzed in real time [70]. Other innovations regarding cardiac activity monitoring include more invasive but comfortable devices. Park et al [45] developed a successful wearable piezoelectric sensor which is inserted in the ear canal and is able to detect heart rate through monitoring the pressure variances and waveforms of the ear canal area. The success of this innovation was attributed to the use of an embedded algorithm and the hardware circuit model which allowed for the refining and detection of ear pulse waves. The last decade has seen the transformation of the devices used for the monitoring of heart rate and cardiac activity. A study looks into the development of a t-shirt that measures EKG signals using textile based computing or 'e-broidery' [86]. Biomedical sensors in the form of minute EKG electrodes which were embroidered with conductive yarn were incorporated into the material of a stretch t-shirt. The results of this innovation were positive and showed room for further improvement through the incorporation of more biomedical sensors in the fabric of the t-shirt. Similarly, another study also look into the use of biomedical sensors in garments through using the WEALTHY system as a means to attain ECG signals. The WEALTHY system is a patented interface used in some wearable biomedical sensor devices. The incorporation of the WEALTHY system into a knitted garment enabled users to concurrently record vital signs and parameters extrapolation and inter-signal elaboration' that aid in the generation of alert messages and the production of synoptic patient tables. This system was geared towards monitoring cardiac patients during their rehabilitation [66].

3.6 Continuous assessment of blood glucose levels

A majority of the devices used to measure, monitor and assess an individual's blood glucose levels make use of invasive techniques which usually require the drawing of a small blood sample. This blood sample is then used together with a digital monitor to determine the blood glucose level. Individuals with diabetes are required to repeat this process at least 5 times daily in order to closely monitor their metabolic levels. In spite of this, some individuals forgo this process as it has been reported to be painful which puts them at risk of developing further health problems linked to inadequate blood glucose level monitoring and management.

In order to improve people's management of diabetes, it is imperative to have a non-invasive system that allows for the constant monitoring of blood glucose levels. Modern technological advancements have resulted in the development of devices which are minimally invasive. A study gives an example of the GlucoWatch Biographer which made use of transdermal fluid to determine the blood glucose level [6]. However the production of this product was discontinued as it resulted in skin irritation amongst some users. Another study also reviews a similar product - the European produced Pendra - which is a watch like blood glucose monitoring device [2]. This device utilized "impedence spectroscopy requiring no extraction of tissue fluid,". Moreover, the Pendra could provide minutely real time readings and hyperglycemia and hypoglycemia alerts. However, home use of the Pendra revealed that readings were imprecise which resulted in it being pulled from the market for further enhancements.

3.7 Continuous assessment of blood pressure

Although it is usual overlooked, blood pressure (BP) is one of the most important vital signs that require monitoring. A significant number of people who have hypertensions do not exhibit any symptoms hence they have no need to seek medical attention which leads to them paying no heed to this ailment. The significance of BP has been highlighted by scholars such as Kikuya et al [55] and Pringle et al [29] who both purport that BP can be linked to morbidity and mortality due to its close association with cardiovascular illness. Over the years wearable biomedical sensors for BP monitoring and assessment have been developed on the basis of traditional techniques like the oscillometric technique. Such devices usually measure BP through the brachial artery in the upper arm. Modifications have been made to BP measuring biomedical devices to incorporate sensors in a watch like device which uses the radial artery in the wrist for measurement. The first watch based BP monitoring system was developed by Casio in 1992. However this venture was discontinued however since then modifications have continuously been made to perfect and improve this innovative device [100]. Devices which have built on Casio's BP watch system include the MediWatch which utilizes arterial tonometry to record radial pulse waveforms and supply BP measurements [47]. Similarly, another study conducted earlier research which looked at Vasotrac which is also a wearable wrist device for BP monitoring [46]. Vasotrac also measures BP after 12 - 15 beats through waveforms acquired from the radial artery. A study looked into the development of "a novel adaptive algorithm for calibrating non-invasive pulse transit time (PTT) measurements to arterial blood pressure (BP) [25]." This pioneering algorithm - which was incorporated into a wearable sensor - aimed to allow for the comprehensive calibration of PTT to BP with no need for an oscillometric blood pressure cuff.

3.8 Continuous assessment of individual health and well-being

The use of remote vital sign monitoring systems has become more and more important as the shift towards 'at home' care - especially for elderly patients and for rehabilitation purposes - has increased. In light of this, extensive research has been conducted to evaluate the efficacy of wearable sensors in keeping track of the daily activities of elderly people, individuals who suffer from chronic illnesses and people undergoing rehabilitation in the home environment. A research conducted by Lovell et al. [54] showed that the use of accelerometers in determining the level of performance of active daily living activities by elderly people being monitored in their homes was effective and expedient. Various innovations have been developed in order to assist with this monitoring system. A study looked into the monitoring of activities such as sitting, walking and standing through developing a pressure and acceleration sensor system which was in-built in shoes. This system was also able to concurrently sense if the individual was also engaging in arm movements [30]. Similarly, another study devised a step counting device based on an accelerometer for people with Parkinson's disease [27].

Apart from monitoring activity, wearable sensors have also been implemented in patient recovery tracking. A study illustrated how wearable sensors could be used to monitor the recovery process of people who had undergone abdominal surgery [69]. Additionally, a number of research projects focusing on the monitoring of patient activity have shown that such applications can significantly increase an individual's compliance to exercise.Two studies look at the use of wearable devices in monitoring the physical activities of individuals who suffer from obesity and concur that such devices are useful especially when used in clinical interventions which aim to promote an active and healthy lifestyle amongst patients [7, 52]. In line with this, there are a number of projects that are investigating the usefulness of incorporating wearable devices in clinical assessments. A study gives an example of LiveNet which is a monitoring system that measures electrocardiogram (ECG), electromyography and 3D acceleration [59]. This system developed at MIT system has been considered for use in observing the symptoms of Parkinson's disease and detecting epilepsy induced fits. Similarly, another study highlighted the effectiveness of LifeGuard which is a monitoring system designed to keep track of the health status of people residing in areas with extreme conditions such as space [16]. Wearable monitoring devices have also been developed to monitor the vital signs of high risk individuals with illnesses such as cardio-respiratory diseases. With regards to this, another study speaks of how the FP5 program under the European Commision led to the development of a device worn on the wrist that can detect and record an individual's blood pressure, temperature and ECG [91]. In addition to this, a number of similar projects under the European Commission that have also developed devices of a similar nature which include the MyHeart project which was put in place to prevent and diagnose cardiovascular diseases through monitoring an individual's vital body signs [37].

3.9 Fuzzy Cognitive Mappings as a tools for medical decision making

The application of Fuzzy Cognitive Mapping (FCM) techniques in the generation of various intelligent systems has been vastly successful in a number of applications. The intricate construction process and infrastructure of Fuzzy Cognitive Maps is reflective of the immense knowledge and experience required in creating FCMs that accurately mimic the decision making and logicizing process [85]. In the area of healthcare and medicine, such decisions tend to be climacteric and need to be executed with promptness. In light of this, FCMs have been incorporated in the development of Decision Support Systems (FCM-DSS) which have been used for differential diagnosis purposes [93]. A study states that FCMs are especially suitable for healthcare related applications because of the complex nature of medical systems and their use of imprecise, irresolute and equivocal data that is constantly changing [58].

3.10 Fuzzy Cognitive Mappings as a diagnostic tool

The underutilization and excessive use of diagnosis tests can have significant impacts on health and cost of health care. Hence decision support technology aims to maximize diagnostic test use in the healthcare sector through assisting healthcare practitioners with the establishment of diagnosis [71]. The use of Clinical Decision Support Systems (CDSS) helps to enhance the quality of patient safety and care. A number of studies assessing the implementation of CDSSs in health care have found that they are highly effective particularly with regards to patient diagnosis and therapy [26]. In light of this, Fuzzy Cognitive Maps (FCMs) have been used as tools to model and control intricate systems with innumerate variables and factors in medical decision support. As a result, FCMs have been utilized in the diagnosis of a variety of illnesses.

Fuzzy Cognitive Mapping (FCM) has proved to be a useful tool in assisting healthcare practitioners to make accurate patient diagnosis of various diseases. A study investigated the use of a customized FCM model in the diagnosis of infant and child meningitis [95]. The FCM in this case - which had a 95 percent accuracy rate - exhibited results with 83.3 percent sensitivity and

80 percent specificity showing that FCM can provide practitioners with a dependable, front-end decision-making tool. Research has also been conducted on the use of FCM in predicting the severity of diabetes mellitus. Bhatia and Kumar developed a FCM system for decision support to help healthcare practitioners evaluate a patient's health status before diagnosis [11]. This FCM system was specifically tailored for the diagnosis of diabetes mellitus using 50 cases. With an accuracy of 96 percent, the system was able to efficiently predict a patient's diabetes severity. Positive results were also found by another study which looked into the integration of FCM in the differential diagnosis of specific language impairment (SLI) from dyslexia and autism [94]. Although the research was in it's initial stage, there was promise of it being successful in complimenting other diagnostic methods to aide speech pathologists in the discernment of of language and communication disorders. According to Mulik and Jadhav FCM can be useful in the diagnosis and treatment of multiple kidney diseases [64]. Their FCM system looked into the diagnosis and treatment of "nephrotic syndrome, renal tubular acidosis, hyponatremia, kidney stone, urinary tract infection, kidney cancer, renal anemia, calculation of kidney transplant allocation score and cytomegalovirus" [64]. Through the development of the 'Fuzzy System for Diagnosis and Treatment of Kidney Disease (FSDTKD),' which was used to assist doctors and their assistants in decision making, a study found that such a FCM system can be useful given that it provided an acceptable accuracy level [64]. However, the FSDTKD called for further enhancement through having additional decision variables.

The prediction of pulmonary infections using FCM is an area that has also been researched. Papageorgiou and Froelich [72], pioneered the investigation of this area as a way to enhance the predictive and diagnostic capabilities of medical practitioners. The FCM system utilized in this research intended to revolutionize the FCM algorithms through a multi-step enhancement process. The incorporation of this evolved FCM system exhibited promising results which point towards the success and advantageous nature of using this in predicting pulmonary infections such as pneumonia. However, the study expressed that further work is still needed to further enhance the system [72]. Yet another pioneering study was undertaken by Beena and Ganguli who looked at the development of a new algorithmic approach - which was based on a FCM together with Hebbian learning [10]. The incorporation of Hebbian learning was to enhance the FCM and the results it produced. The levels of damage detection from the system were 100 percent at a 50 percent noise level; whilst at a 70 percent noise level, damage detection was at 98.8 percent.

These tools provide useful outputs in form of monitoring displays, graphs representing a person's health over a period, priority lists and warnings generated when threshold values are reached. These outputs are archived by performing functionality such as: time series analysis, patient prioritization, computing the rate of deterioration and monitoring threshold values. The outputs should help medical professional to make decisions.

3.11 Continuous Glucose Monitoring Systems: Categories and Features

There are a number of continuous glucose monitoring systems that were approved by the Food and Drug Administration (FDA). The FreeStyle Libre Pro system (Abbott, Alameda, CA) is an example. It consists of a sensor and a single reader device. The sensor is applied to the back of a patient's upper arm and requires a two minute activation period. Glucose levels are recorded in 15 minutes intervals for up to 14 days [41].

Another successful professional Continuous Glucose Monitoring (CGM) is the Medtronic iPro2 system (Medtronic, Northridge, CA). Is consists of the Enlite glucose sensor and the iPro2 digital recorder. The Enlite glucose sensor is wearable for up to 6 days and glucose readings are recorded every 5 minutes. At least one blood glucose entry be aqcuired in every 12 for system uploading [41].

Several personal CGM systems are also available. The Abbott FreeStyle Libre flash CGM system was FDA approved in September 2017 for stand-alone use. It consists of the FreeStyle Libre sensor and the the FreeStyle Libre reader. The FreeStyle Libre sensor is wearable for up to 10 days [41].

Another personal CGM is the Medtronic Enlite sensor which is used with and 630G insulin pumps and Medtronic MiniMed 530G. Rate-of-change trend arrows and Real-time glucose data are available every 5 minutes on the pump screen. Both pumps use Medtronic SmartGuard technology. When glucose levels fall below a preset threshold insulin delivery is suspended for up to 2 hours . The Enlite sensor is calibrated every 12 hours [41].

3.12 Conclusion

There are many technological solutions being implemented to improve healthcare. The following chapter discusses the low level software solutions that were proposed to solve the main objectives of this study.

An in-depth understanding of the machine learning and graph algorithm implemented during this study will be discussed in the following chapter 4.

Chapter 4

Patient Prioritization: Calculating the Risk-score

The aim of a triage system is to determine a quantitative measure of patients' medical conditions and then give priorities to the most urgent cases. Some of the requirements to be met by the algorithm behind such as system include: Interpretability, Speed, Simplicity, Scalability and Accuracy. The algorithm should be scalable that is it should be portable enough to run on small devices e.g. smart phones, tables, iPad, biomedical sensors and smart watches without any problems. The solution must be efficient, accurate and easy to interpret. Figure 4.1 depicts the patient condition recognition system proposed in this thesis. It has been implemented as a data analytics platform that includes the following key components:

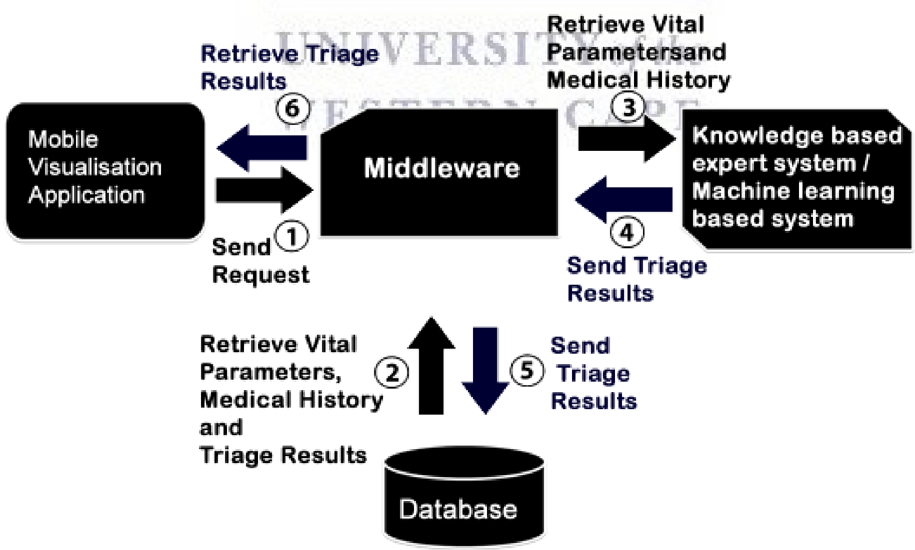

Figure 4.1: High-level overview of the situation recognition [1] [56].

Database: The database stores medical record history for the patients, time stamped patient physiological parameter readings from the bio-medical sensors and also stores the triage results for every patient record [56].

Knowledge Based System: The scoring system uses domain knowledge to score the un-scored patient history data from the database. The knowledge based system calculates the score using weights provided by the experts (physicians) no learning is involved. The final triage scores for every record are then fed directly into a machine learning model as training data [56].

Machine Learning Based System: A machine learning technique can be used to perform intelligent data analysis in order to perform situation recognition and patient treatment prioritization [56].

Mobile Visualization: This application interfaces with the middleware to retrieve the patient records and other data from the Database. The application presents the doctor with prioritized patient records, real time heart rate monitoring, risk score time series monitoring and is also able to visualize patterns discovered by a machine learning technique based system in the form of graphical representations [56].

Middleware: The middleware interfaces with the database and the mobile visualization application. The middleware interfaces with a) the database to acquire triage results obtained by the scoring system and 2) the mobile visualization application to provide the prioritized patient records and graphical representations [56].

4.1 Predictive Models

4.1.1 Gradient Descent Algorithm

Gradient descent can be defined as a "first-order iterative optimization algorithm for finding the minimum of an objective function," [32]. Gradient descent is also referred to as 'steepest descent.' This algorithm is generally used in machine learning to find coefficients of ML algorithms like logistic regression and neural networks. The gradient descent algorithm works by using a model to make predictions on data and prediction error to create an updated model that has a reduction in error. This algorithm aims to find model parameters which have minimized model error through making changes along the gradient toward a minimum value in the model. Hence the algorithms' name 'gradient descent.' Based on the different ways of calculating data, there are various types of gradient descent. The type of gradient descent to use will depend on:

1. Data volume
2. Time complexity
3. Algorithm accuracy

The three gradient descent variations are:

1. Batch Gradient Descent
2. Stochastic Gradient Descent
3. Mini-batch Gradient Descent

4.1.2 Batch Gradient Descent

Batch gradient descent - which is also known as vanilla gradient descent – is a version of the gradient descent algorithm which enumerates the gradient of a whole dataset. This works well with fairly smooth error manifolds. Thereafter, an optimum solution can be attained, be it local or global. Because batch gradient descent utilizes a complete dataset for a single update, it tends to be somewhat slow. However, this version of gradient descent is computationally efficient as it requires significantly less updates when compared to stochastic gradient descent. Additionally, a more stable error gradient can be obtained with this algorithm because of the infrequency of updates.

4.1.3 Stochastic Gradient Descent

According to Nicholas [68], stochastic gradient descent can be described as a simple, effectual "approach to discriminative learning of linear classifiers under convex loss functions such as (linear) Support Vector Machines and Logistic Regression." It is another version of the gradient descent algorithm which computes the error and updates the model one example at a time. This gradient descent algorithm is often referred to as an 'online machine learning algorithm.' In order to use stochastic gradient descent, the data set needs to be randomized first. Thereafter, to update the parameters only one example in each iteration is used to calculate the gradient of cost function. Hence this gradient descent algorithm tends to be faster than batch gradient descent. However, accuracy may not be achieved using this approach. Stochastic gradient descent is useful for error manifolds with numerous local maxima/minima.

4.1.4 Mini-batch Gradient Descent

The mini-batch algorithm is widely used as it generates precise results at a fast pace. This approach makes use of a prescribed set of 'm' training data to calculate the gradient of the cost function. Generally, mini-batch sizes are between 50–256 however this can fluctuate depending on the application. As a result, the mini-batch gradient descent algorithm has reduced discrepancies in the parameter updates and this leads to convergence stability. Additionally, gradient computation can be done efficiently through the use of highly optimized matrix. Table 4.1 compares the three forms of gradient descent.

Algorithm	Batch	Mini Batch	Stochastic
Accuracy	High	Moderate	Low
Time Consuming	More	Moderate	Less

Table 4.1: Comparing the types of gradient descents.

4.2 Machine Learning Techniques

This chapter discusses the following models in detail MLiR, MLoR, CART, SNN and DNN. Each algorithm performs 10 steps in order to prioritize patients. The steps are shown in Figures 4.2, 4.3, 4.4, 4.5 and 4.7.

The un-scored history data is scored and then fed into the machine learning model as training data. If the history data is scored then steps 1 to 4 can be skipped and the scored history data can be fed directly into the machine learning model. The most distinguishing characteristics of the five machine learning algorithms discussed in this dissertation are step 5 and step 8.

4.3 Logistic Regression

Logistic regression is used for modelling event probabilities which can either be dichotomous (binary) or multiple. The algorithm can either have one variable affecting the outcome (univariate) or multiple dependent or independent variables affecting the outcome (multivariate).

There are three types of logistic regression that is ordinal, multinomial and binomial. The binomial regression can have only one of the two possible outcomes. For example the possible outcomes from playing a game of chess with someone at the same level as you is either a win or a lose and can be represented by 1 and 0 respectively.

The multinomial logistic regression has many possible outcomes for example a sick patient can either be diagnosed with disease A, disease B or disease C. The ordinal logistic regression deals with dependent ordered variables.

The probability of an event can be measured in three ways that are regarded to be equivalent in some way:

- Probability of the event, π is a number between 0 and 1. The value one means it is certain the event will occur and a zero value means it is certain the event will not occur. The probability π that Y = y is represented as

$$\pi = P(Y = y) \tag{4.1}$$

- Odds in favour of the event, ODDS(Y = y) is a number between 0 and ∞. The odds in favour of Y = y is represented by the equation 4.2. The value zero means it is certain the event will not occur, the value one corresponds to the probability one half and ∞ corresponds to certainty the event occurs.

$$ODDS\,(Y = y) = \frac{P(Y = y)}{1 - P(Y = y)} = \frac{\pi}{1 - \pi} \tag{4.2}$$

- Log-odds in favour of the event, $\log(ODDS(Y = y))$ is a number between -∞ and ∞. A value of -∞ means we are certain the event will not occur and -∞ means the event will occur.

$$\log ODDS(Y = y) = \log \frac{P(Y = y)}{1 - P(Y = y)} = \log \frac{\pi}{1 - \pi} \tag{4.3}$$

The probability of an event was measured using the Log-odds in favour of the event which is an number between -∞ and ∞. The weights are calculated in step 5.

Figure 4.2 shows that the derived estimator in step 8 is defined as

$$logit(\pi(x)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \tag{4.4}$$

Where $logit(\pi(x))$ is the log-odds in favour of the event, β_0 to β_p are weights, and x_1 to x_p are input variables [1].

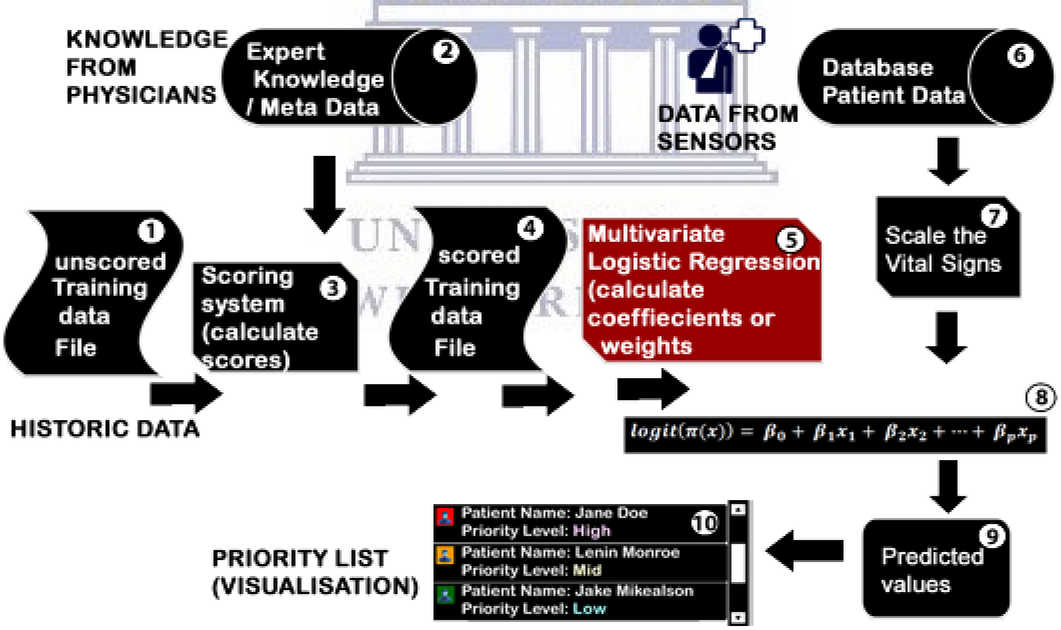

Figure 4.2: Classification using the Multivariate Logistic Regression algorithm [1].

4.3.1 Logistic Regression Gradient Descent

Algorithm 1 $GradientDescent(trainList[][], alpha, maxEpochs)$, Logistic Batch Gradient Descent

```
epoch = 0
hypothesis = []
weights = []
while epoch < maxEpochs do
   epoch = epoch + 1
   for integer index = 0 <trainList.Count do
      double computed = ComputeSigmoid(trainList[index], theta)
      hypothesis[index] = computed
      integer targetIndex = trainList[index].Count − 1
      double target = trainList[index][targetIndex]
      weights[0]+ = alpha ∗ (target − computed) ∗ 1
      for integer i = 1 <weights.Count do
         weights[i]+ = alpha ∗ (target − computed) ∗ trainList[index][i − 1]
         i = i + 1
      end for
      index = index + 1
   end for
end while
return weights
```


Algorithm 2 $ComputeSigmoid(dataItem[], weights)$, Batch Gradient Descent: Sigmoid

```
z = 0.0
z+ = weights[0]
integer i = 0
for i <weights.Count do
   z+ = (weights[i + 1] ∗ dataItem[i])
   i = i + 1
end for
return 1.0/(1.0 + Math.Exp(−z))
```

4.4 Linear Regression Gradient Descent

The Figure 4.3 shows that the derived estimator in step 8 is defined as follows:

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + ... + \theta_p x_p \tag{4.5}$$

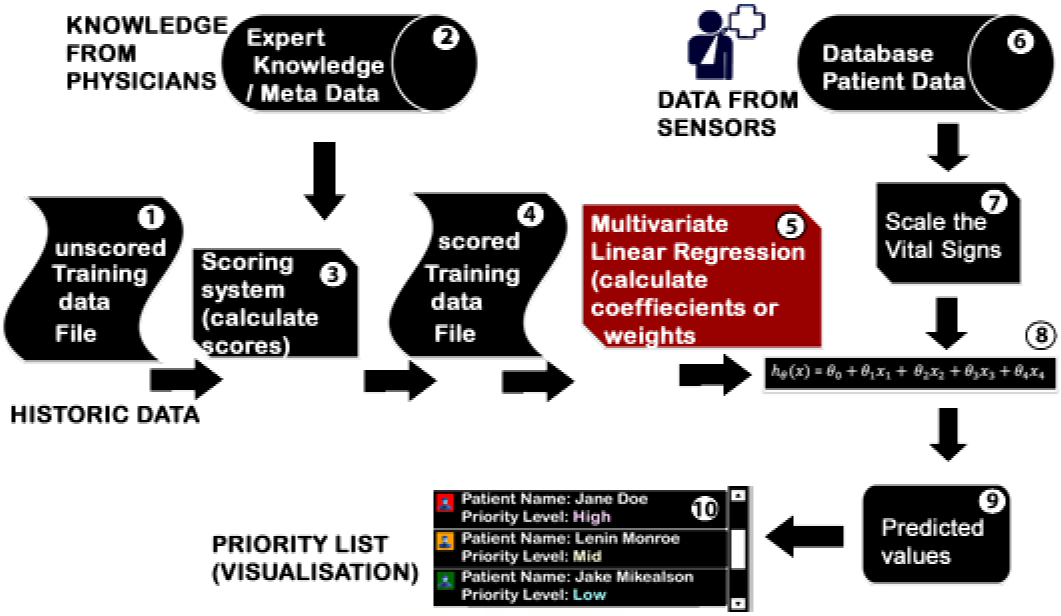

Figure 4.3: Classification using the Multivariate Linear Regression algorithm [1] [57].

Where $\hat{y}$ is the estimated y output value, θ_0 to θ_p are weights, and x_1 to x_p are input variables. The estimate $\hat{y}$ is calculate in step 8 and the weights θ_0 to θ_p are calculated in step 5 as shown in Figure 4.3.

Algorithm 3 *GradientDescent*(*trainList*[][], *alpha*, *maxEpochs*), Linear Batch Gradient Descent

```
integer epoch = 0
integer col = 0
integer row = 0
double columnTotal = 0
hypothesis = []
weights = []
while epoch < maxEpochs do
    integer index = 0
    for index < trainList.Count do
        columnTotal = 0
        col = 0
        for col < numberOfColumns do
            columnTotal = columnTotal + weights[col] *trainList[row][col]
            hypothesis[row] = columnTotal
            index = index + 1
        end for
    end for
    col = 0
    for col < numberOfColumns do
        sum_HminusYxX[col] = 0
        col = col + 1
    end for
    row = 0
    for row < numberOfRows do
        col = 0
        for col < numberOfColumns do
            sum_HminusYxX[col] += (hypothesis[row] −trainList[row][numberOfColumns]) *
            trainList[row][col]
        end for
    end for
    col = 0
    for col < numberOfColumns do
        theta[col] = weights[col] −(alpha * sum_HminusYxX[col])/(double)(numberOfRows)
    end for
    epoch = epoch + 1
end while
return weights
```

4.5 Classification and Regression Decision Tree

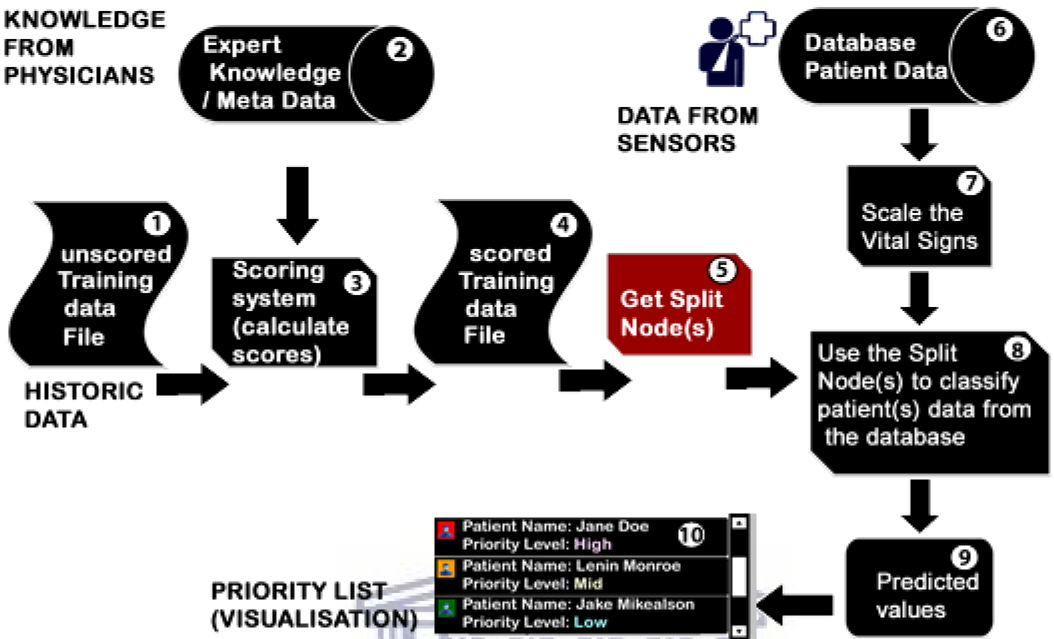

Figure 4.4: Classification using the Classification and Regression Decision Tree algorithm.

Algorithm description: Unlike the other machine learning algorithms discussed in this dissertation. The CART algorithm derives the split node(s) in step 5. The split node(s) is(are) then used to split the patient(s) data into classes in step 8 as shown in the Figure 4.4.

Algorithm 4 $GrowTree(D, F)$: grow a feature tree from training data [33].

```
Input: data D; set of features F.
Output: feature tree T with labelled leaves.
if Homogeneous(D) then
   return Label(D)
end if
S = BestSplit(D, F)
split D into subsets Di according to the literals in S
for each i do
   if Di is not empty then
      Ti = GrowTree(Di, F)
   else
      Ti is a leaf labelled with Label(D)
   end if
end for
return a tree whose root is labelled with S and whose children are Ti
```


Algorithm 5 $BestSplit(D, F)$: find the best split for a decision tree [33].

```
Input: data D; set of features F.
Output: feature f to split on.
Imin = 1
for each f in F do
   split D into subsets D1, ..., Dl according to the values Vj of f
   if Impurity(D1, ..., Dl) < Imin then
      Imin = Impurity(D1, ..., Dl)
      fbest = f
   end if
end for
return fbest
```

4.6 Classification and Regression Decision Trees using Directed Graphs

Step 1: Building the Decision Tree using a Directed Graph

Step 1.1: Expert Knowledge Input Format:

Parent Node, Child Node, Value 1 (e.g. Descriptive Text) , Value 2 (e.g. Quantitative Weight such as Degrees of Freedom and Probabilities)

Step 1.2: Build the directed graph

Step 2: Classifying Attributes, Searching and Matching

Step 2.1: Attributes Input Format

Parameter 1 Value 1 (e.g. Descriptive Text) , Parameter 1 Value 2 (e.g. Quantitative Weight such as Degrees of Freedom)

Parameter 2 Value 1 (e.g. Descriptive Text) , Parameter 2 Value 2 (e.g. Quantitative Weight such as Degrees of Freedom)

Parameter 3 Value 1 (e.g. Descriptive Text) , Parameter 3 Value 2 (e.g. Quantitative Weight such as Degrees of Freedom)

...

Parameter N Value 1 (e.g. Descriptive Text) , Parameter N Value 2 (e.g. Quantitative Weight such as Degrees of Freedom)

Where N is the number of parameters available for classification

Step 2.2: Use Depth First Search / Breath First Search algorithm to Traverse the Graph

If the tree has costs (quantitative values) for each decision node And node not leaf then

use probabilities or degrees of freedom to compute a value that will be used for classification

else

match the attribute to the decision node and return the decision node

4.7 Deep Neural Network

The DNN calculates the weights which are used to predict the scores for the patient vital parameters. The Figure 4.5 shows a software architecture for training and classification using a DNN.

The DNN shown in Figure 4.6 is a 8-(10,10,10)-4 neural network. There are 8 input nodes, 10 by 3 hidden nodes and 4 output nodes. This might correspond to a problem where the goal is

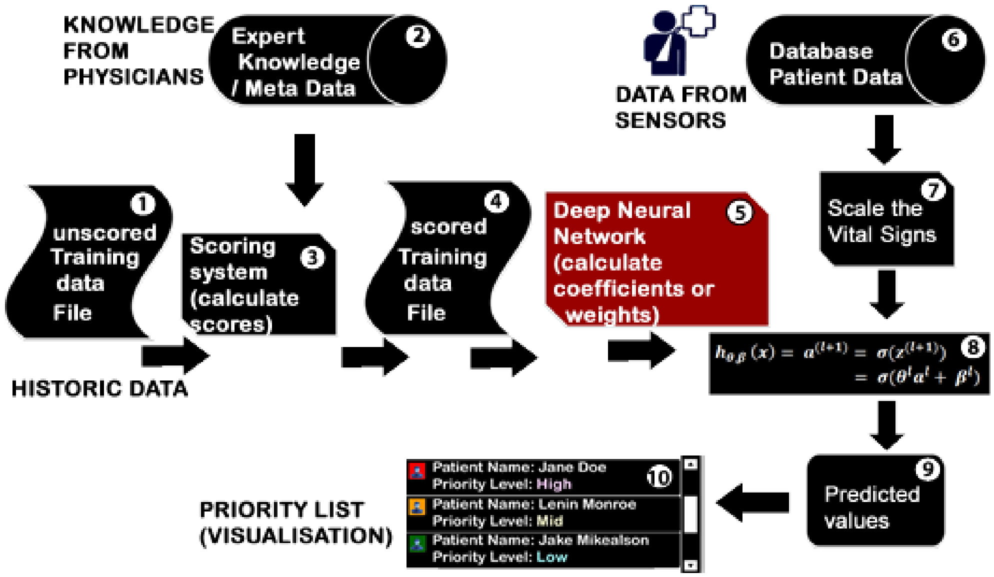

Figure 4.5: Classification using the Deep Neural Network algorithm [1].

to predict the priority level (Normal, Low, Medium and High) of a patient based on Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Heart Rate and Peripheral Oxygen Saturation (SPO2), where the input values are scaled using the gaussian normalizer. The SBP = 231.8, DBP = 111.7, Heart Rate = 76.1 and SPO2 = 67 are scaled to 3.459, 5.075, 0.357 and -3.166 respectively [1].

If Normal is encoded as (1,0,0,0) Low is encoded as (0,1,0,0) and Medium is encoded as (0,0,1,0) and High is encoded as (0,0,0,1), then the DNN predicts a priority level of Low for someone with SBP = 231.8, DBP = 111.7, Heart Rate = 76.1 and SPO2 = 67 because the last output value 0.9453 is the largest.

Every connection between two nodes represents a weights and its bias. The weight is defined as the strength or amplitude of a connection between two nodes from two successive layers.

4.8 Single Hidden Layer Neural Network

Similarly; the SNN in Figure 4.8 is a 8-(10)-4 neural network. There are a 8 input nodes, 10 hidden nodes and 4 output nodes. The goal is to predict the priority level (Normal, Low, Medium and High) of a patient based on SBP = 231.8, DBP = 111.7, Heart Rate = 76.1 and SPO2 = 67 [1].

If the parameters are encoded as in the DNN then the SNN predicts a priority level of Low because the last output value of 0.999 is the largest and corresponds to Low as shown in Figure 4.8.

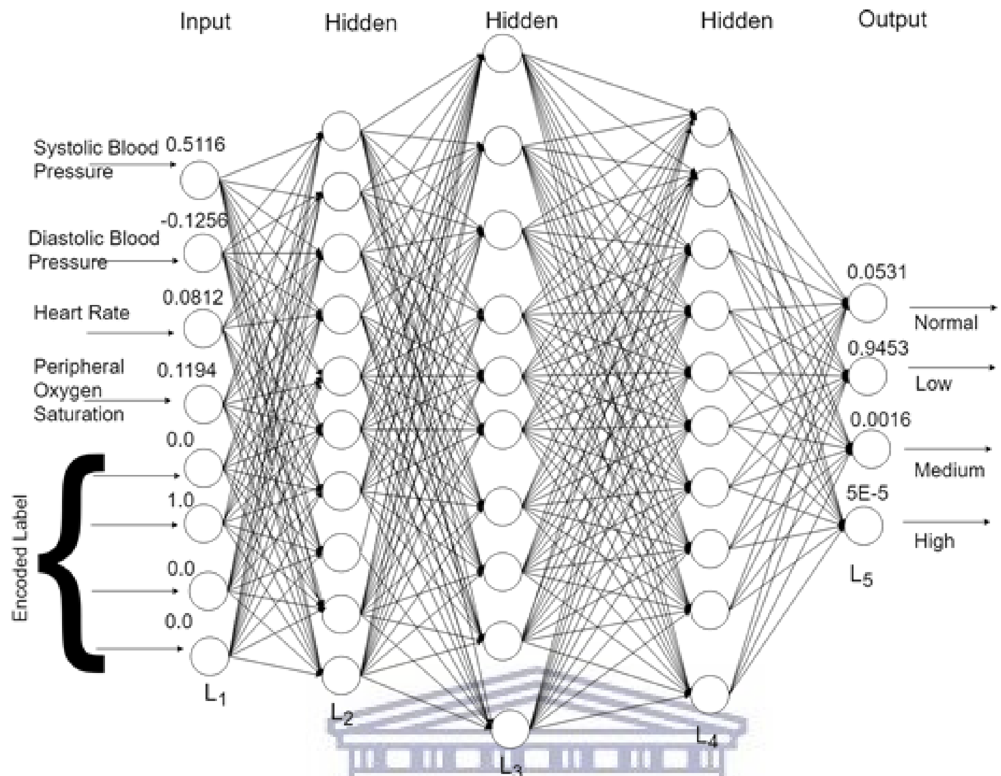

Figure 4.6: Deep Neural Network 8-(10-10-10)-4 [1].

4.9 Multiple Layer Neural Network

The weights for both the DNN and the SNN are calculated in step 5 as shown in Figure 4.5 and Figure 4.7. The outputs in a neural network are calculated using the equation; A.4 which can be derived as shown in appendix A.

The machine learning algorithms implementation have been discussed at a very low level. The following chapter will look at a cloud based solution at a high level.

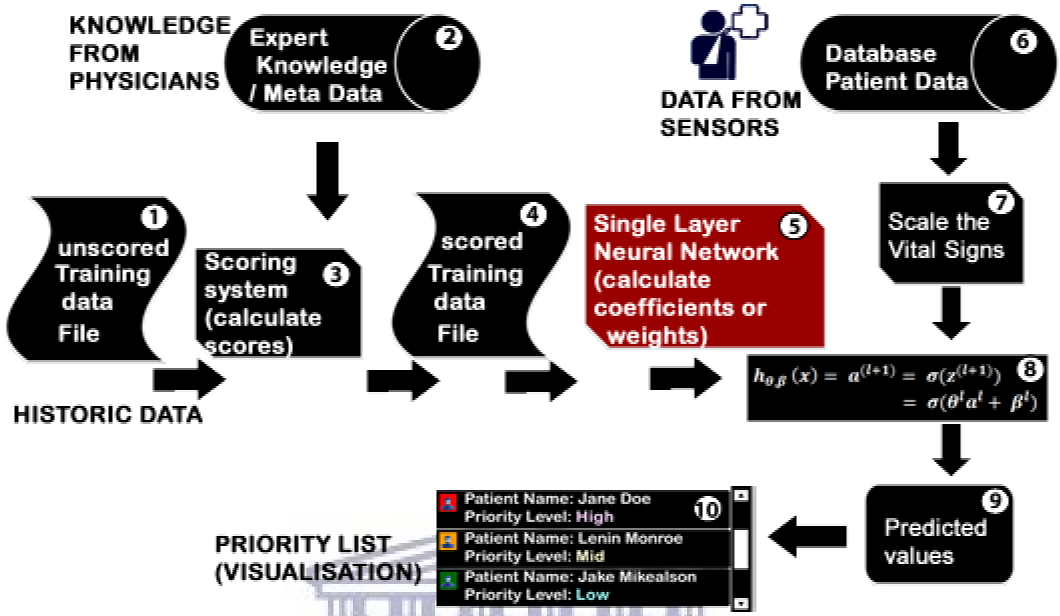

Figure 4.7: Classification using the Single Hidden Layer Neural Network algorithm [1].

Algorithm 6 *GradientDescent*(*trainData*, *maxEpochs*, *learnRate*, *momentum*) Back Propagation: High Level Pseudo Code

---

Input: *trainData* the matrix of features,
maxEpochs the number iterations it takes to adjust the weights,
learnRate, determines how fast the algorithm learns or how fast the weights change
momentum, a value between 0 and 1 that is increased to speed convergence by increasing the size of steps taken towards the minimum by jumping from a local minima.
Output: *weights*
while some condition is met **do**
  compute gradients of each output node
  compute gradients of all hidden layer nodes
  update *weights* and *biases*
  return *weights*
end while

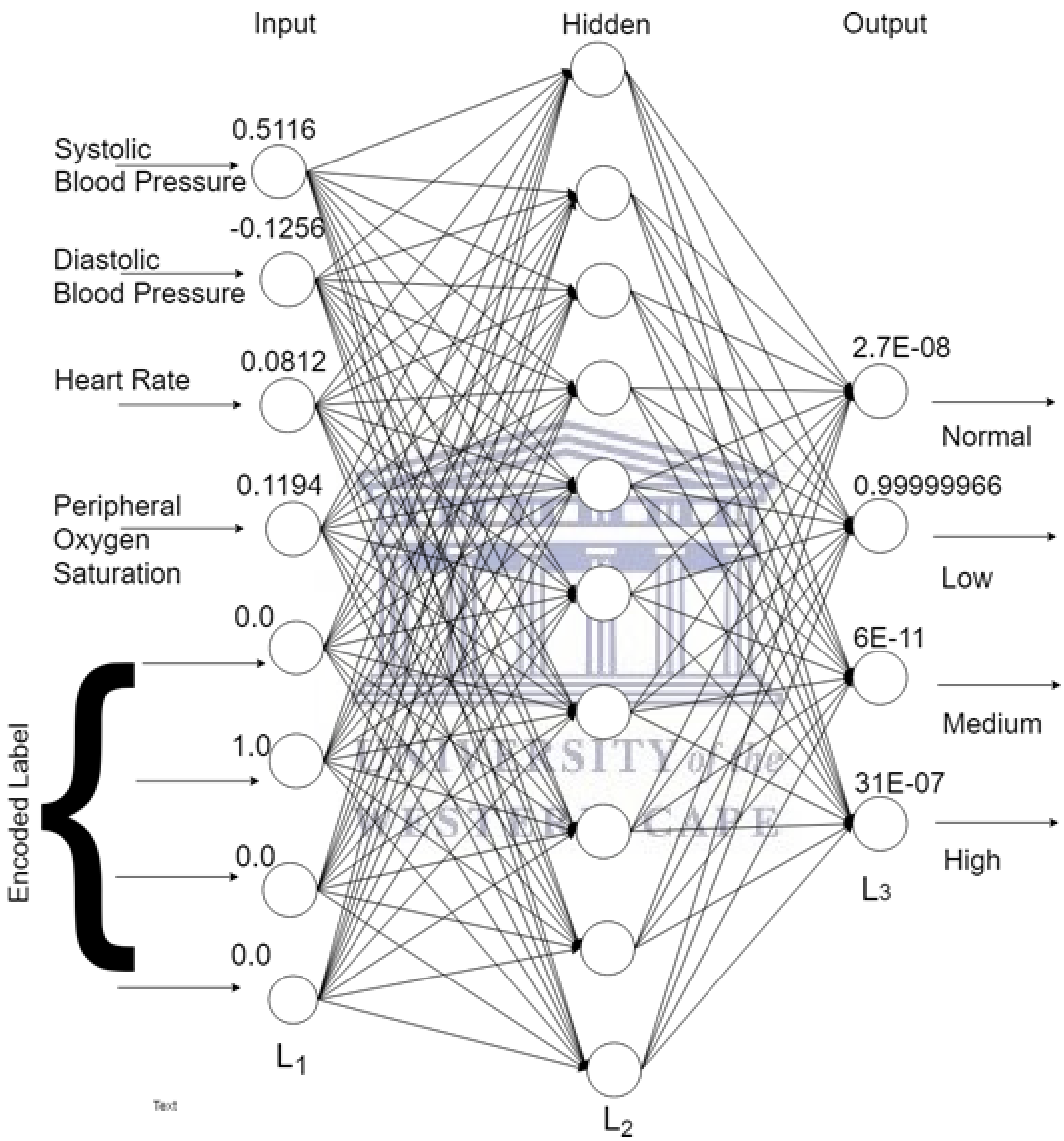

Figure 4.8: Neural Network 8-(7)-4 [1].

Chapter 5

Disease Identification

5.1 Introduction

Two studies [4][5] use graph theory and social network analysis techniques to ease the burden caused by chronic conditions on hospital by preventing unnecessary hospital admissions and associated high costs. In this study graph theory is used to identify a patient's illness by matching the patient's symptoms to the symptoms in the database or knowledge repository. This study however uses machine learning techniques to calculate risk-scores. Patient prioritization is then used to give priority to high risk patients thereby preventing unnecessary hospital admissions.

A study described a decision tree as a structure that represents procedure for classifying objects based on their attributes. And each object is represented as a set of attribute/value pairs and a classification [84].

Assuming a set of medical symptoms might be represented as follows:

Table 5.1: Classified patient medical symptoms [84].

Patient	Cough	Fever	Weight	Pain	Classification
Mary	no	yes	normal	throat	flu
Fred	no	yes	normal	abdomen	appendicitis
Julie	yes	yes	skinny	none	flu
Elvis	yes	no	obese	chest	heart disease

The input (training data) and output (correct classifications) are provided therefore algorithms such as the CART or the Decision Trees ID3 can be used to classify a new set of input data [84]. A decision tree such as the one illustrated in Figure 5.1 can be used to represent the Table 5.1.

The ID3 and CART algorithms are machine learning techniques. Machine learning techniques are used when learning from experience is required before making a decision. However if a knowledge base of disease symptoms relations exists then this study also proposes considering use of data structures such as graphs and trees together with natural language processing to implement a symptom checker. A related study also makes use of natural language processing [97]. Algorithms such as depth first search (Appendix B) and breath first search (Appendix C) can be used to traverse graphs.

5.1.1 Directed graph for decision making

The Figure 5.1 is a disease identification tree or graph. Each blue rectangle represents a condition to be tested. The red oval represents a match (positive result) for a disease and the green oval represents a no match (negative result) meaning no disease or illness. A single test case can have many possible outcomes. The data structure can be constructed in such a way that it can be used as a test for all the diseases and their symptoms. Each test case leads to another path and terminates when a leaf node with the final outcome is reached. Tests are in the form of a conversation and the algorithm is in control of the conversation. This algorithm is more ideal in virtual nursing.

An example of a patient and virtual nurse conversation used to test the algorithm:

Question 1: Where is your pain [abdomen, throat, chest or none]?

Answer: none

Question 2: If previous answer is no pain. Are you coughing [yes or no]?

Answer: yes

Question 3: If previous answer is yes: Do you have a fever [yes or no]?

Answer: yes

Decision: you have flu

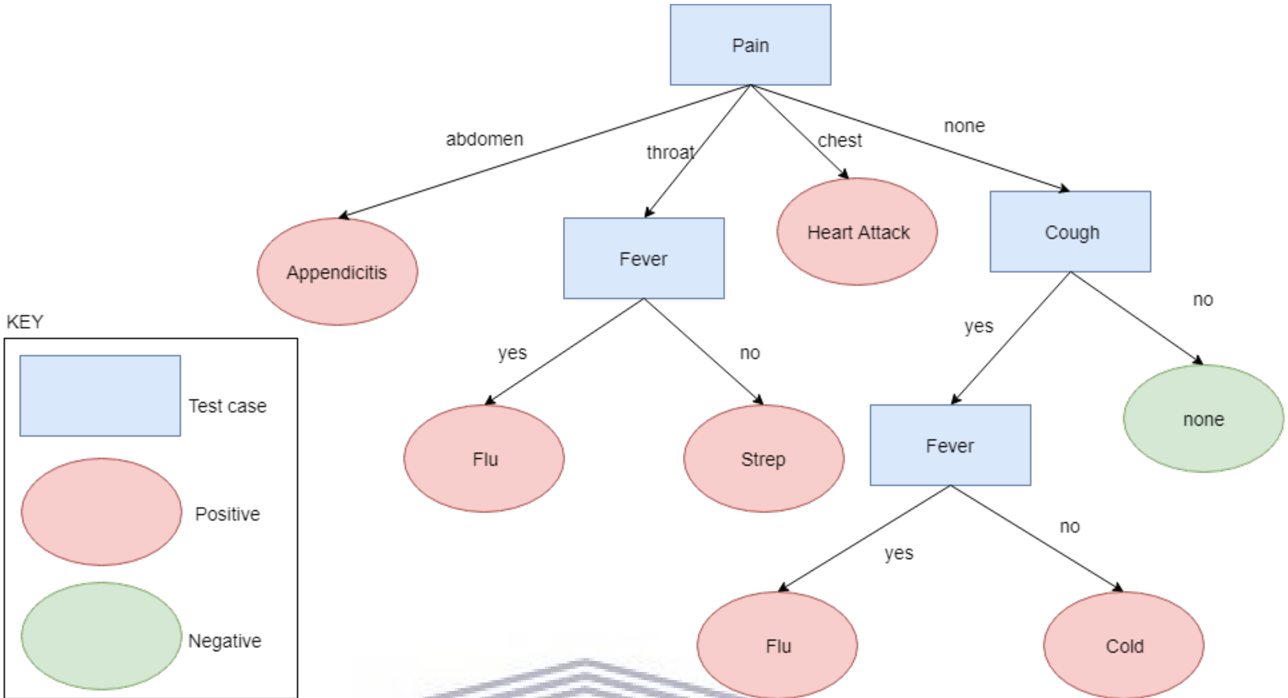

Figure 5.1: Disease identification using a graph data algorithm [84].

A practical chest tree or graph used by medical practitioners is show in the following Figure 5.2.

Another example of a disease identification graph is the abdominal pain, chronic graph shown in the following Figure 5.3.

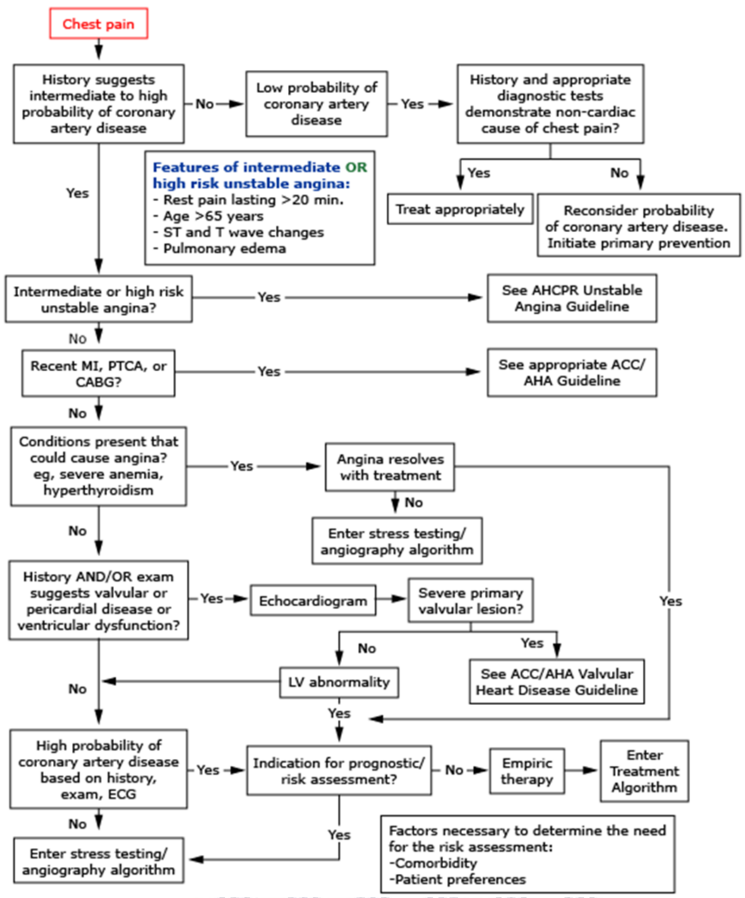

Figure 5.2: A more practical disease identification tree graph for a chest algorithm [24].

5.2 Human Symptoms-Disease Network (HSDN)

In [101], medical bibliography data is used to generate a human symptom disease network. The connection between shared phenotypes (symptoms) and common genotypes (protein-protein interactions) of two diseases are also illuminated. The study extracted data from the extensive medical bibliographic literature records and the related Medical Subject Headings (MeSH) metadata from PubMed [101]. There are seven different files that were created. The following is a listing of the columns available in the first four files relevant to this study [101]:

1. MeSH Disease Term and PubMed occurrence
2. MeSH Symptom Term and PubMed occurrence
3. MeSH Symptom Term, MeSH Disease Term, PubMed occurrence and TFIDF score
4. MeSH Disease Term, MeSH Disease Term and symptom similarity score

The following networks were established from the data [101]:

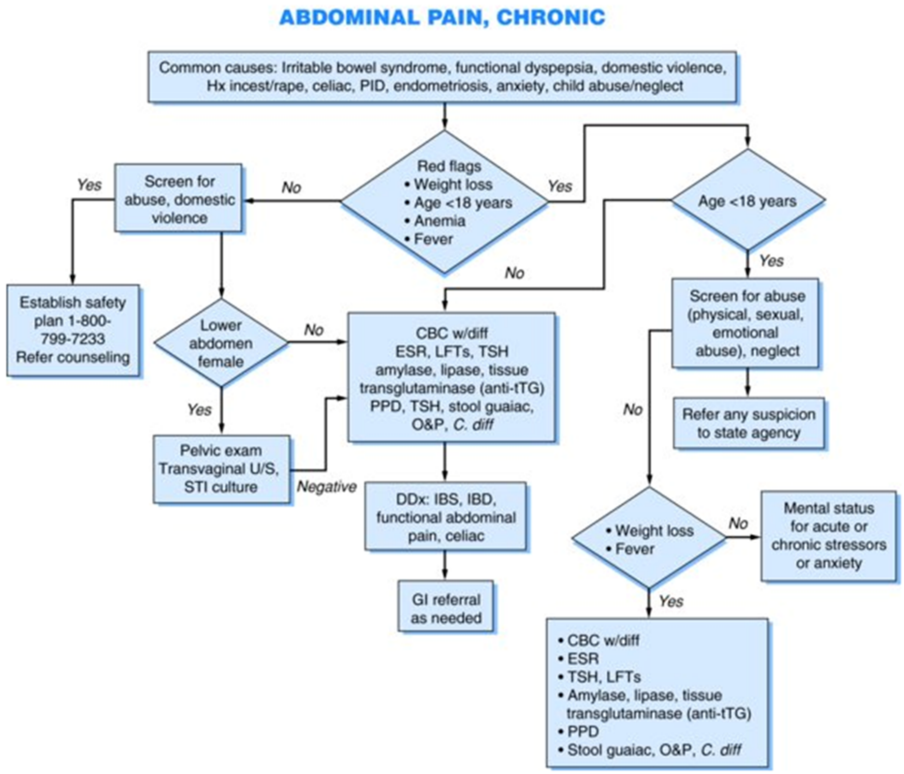

Figure 5.3: Abdominal Pain or Chronic Graph [20].

1. disease-symptoms relationships
2. disease-gene relationships
3. disease-disease networks based on symptom similarity
4. disease-disease networks based on genes (Protein-Protein Interactions)

5.2.1 Processing pipeline

This current study contributes by proposing an architecture that makes use of the networks established by the study [101]. The acquisition of symptom and disease relationships was covered by the authors of the study [101].

Medical Data knowledge base - This study uses disease-symptoms relationships data and MeSH list of diseases. The diseases-symptoms data is used to construct the data analysis system and the list of diseases is used to verify the suggested disease from the literature.

Data Network Analysis - The relationships between diseases, symptoms and genes are established by building a HSDN. The HSDN can be queried and output is given for example symptoms can be provided as input and the degree centrality of each matching disease-symptoms relationship is given as output.

Data Visualization - This presentation of the output (diseases with at-least one matched searched symptoms) from the data network analysis.

5.2.2 Why use data structures to perform disease identification

A medical data knowledge base containing disease-symptoms, disease-disease, and disease-gene relationships was provided and data structures such as graphs and trees can easily represent these relationships. Machine learning algorithms such as CART or the Decision Trees ID3 can be used when the classification (i.e. the disease) related symptoms is not known.

5.2.3 Graph

A graph G is defined as an ordered pair $G = (N, E)$ where N is a set of nodes or vertices or points and the set E of edges connects the vertices N. $E_{i,j}$ is an edge for the nodes N_i and N_j, where i and j is an index for vertex N_i and N_j respectively [53].

5.2.4 Degree Centrality (DC)

The degree centrality DC of any node N_i in a graph is a network theoretic measure and is defined as the number of edges n attached to the node N_i. Where $i \in \{0, 1, 2...i...n_1\}$ is the node's index [15].

The number of symptoms in disease D_i exhibited by the patient is n. The DC of any disease D_i measures the number of symptoms n.The HSDN is a graph G and its output is a set L of DC, such that $DC \in L$.

5.2.5 Building and Searching the Symptoms Disease Network

The types of HSDN networks were established; binary tree based, *zero*-Edge graph based and *n*-Edge graph based. The three types of HSDN were named based on the state of the disease-symptoms sub graphs S, $S \in G$. A tree is an undirected graph in which any two nodes are connected by exactly one edge(path).

zero-Edge graph based HSDN

The disease-symptoms relationship is represented by a graph with no edges. The disease is represented by the node on index zero of the graph node list and the other nodes represent the symptoms. The order of the symptoms in the node list is not important. This is illustrated by Figures 5.7 and 5.8.

n-Edge graph based HSDN

The disease-symptoms relationship is represented by a graph with n edges and n symptom nodes. The disease is represented by the node on index zero of the graph node list and the other nodes represent the symptoms. The order of the symptoms in the node list is not important and all symptoms nodes are connected to the disease node. The order of the symptoms is not important. This is illustrated by Figures 5.9 and 5.10.

Algorithm 7 $Search(G, symptoms)$, Symptom Search Algorithm: Graph Based

```
Input: symptoms, the symptoms to be searched
Input: G, the human symptoms disease network (HSDN graph)
Output: degree centrality list L, DC ∈ L
for each node ∈ G do
  Initialize degree centrality graph G, G ∈ L
  for each symptom ∈ symptoms do
    Initialize sub graph, S, S ∈ G
    if symptom ∈ S then
      if G node count is zero then
        add node at index zero of S to G
      end if
      add symptom to L
    end if
  end for
  if G has nodes then
    add G to L
  end if
end for
return L
```


Binary tree based HSDN

The disease-symptoms relationship is represented by a binary tree. The disease is represented by the root node of the tree structure and the other nodes represent the symptoms. The order of the symptoms is not important. This is illustrated by Figures 5.11 and 5.12.

The binary tree based network has better performance since the binary tree algorithm has a better search time complexity than that of the graph algorithm as shown by Table 5.2.

Pattern Matching

The Boyer Moore pattern matching algorithm was used to match the symptom terms in the tree and graph based networks.

Algorithm	Time Complexity Search	Time Complexity Insert	Responses
Binary Tree	$O(\log n)$ where n is the number of nodes in a tree.	$O(\log n)$	Two
Graph	$O(\log \lvert E \rvert + \log \lvert V \rvert)$ where E is the number of Edges in a graph and V is the number of vertices in a graph.	$O(1)$	Many

Table 5.2: Comparing the binary search tree and a graph.

Algorithm 8 $Search(G, symptoms)$, Symptom Search Algorithm: Tree Based

Input: $symptoms$, the symptoms to be searched
Input: G, the human symptoms disease network (HSDN graph)
Output: degree centrality list, $L, DC \in L$
for $node \in G$ **do**
  Initialize degree centrality graph, G, $G \in L$
  for $symptom \in symptoms$ **do**
    Initialize sub tree, $S, S \in G$
    if $symptom \in S$ **then**
      if G node count is zero **then**
        add the S root node to G
      end if
      add $symptom$ to G
    end if
  end for
  if G has nodes **then**
    add G to L
  end if
end for
return L

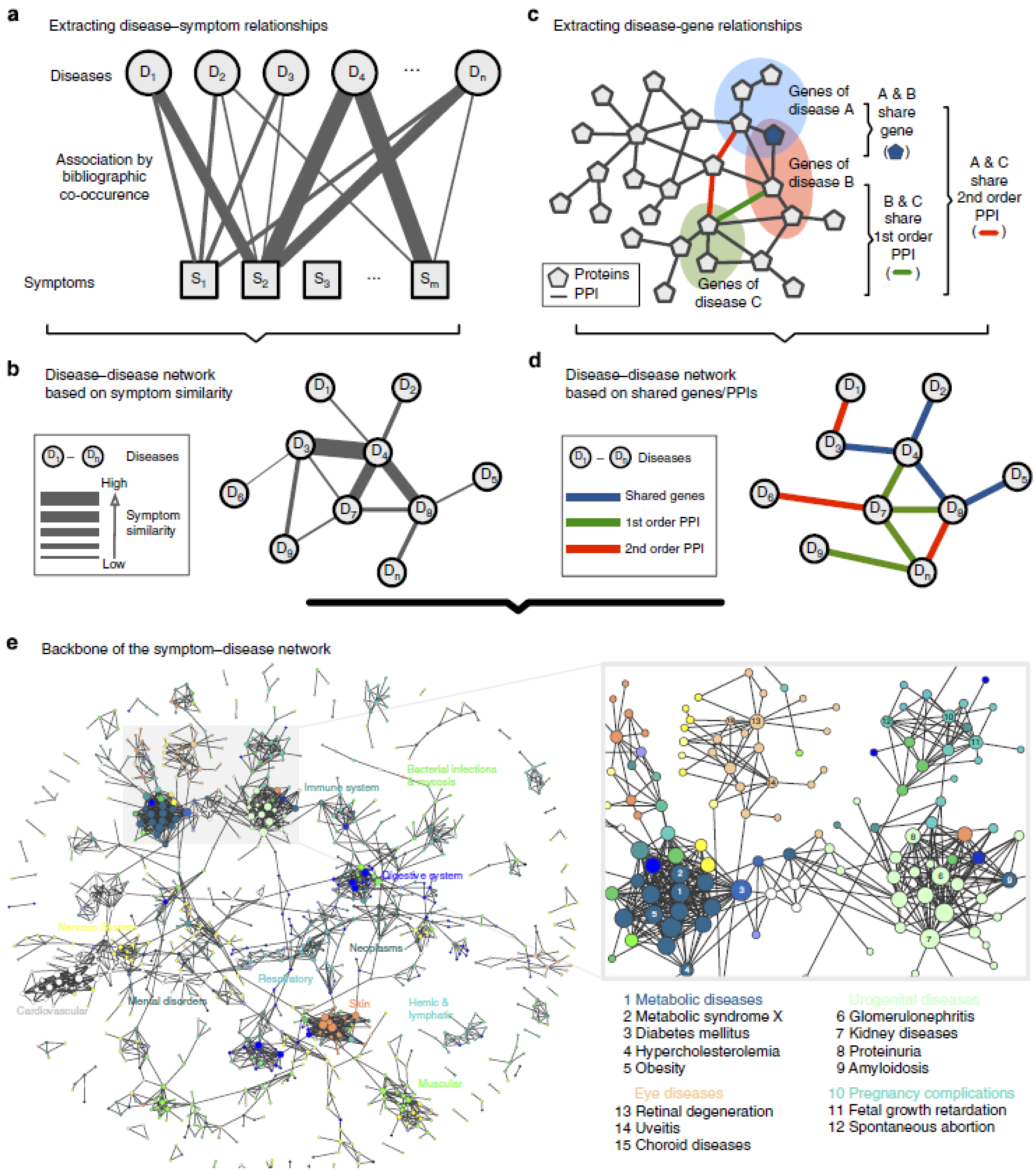

Figure 5.4: Construction of the HSDN [101].

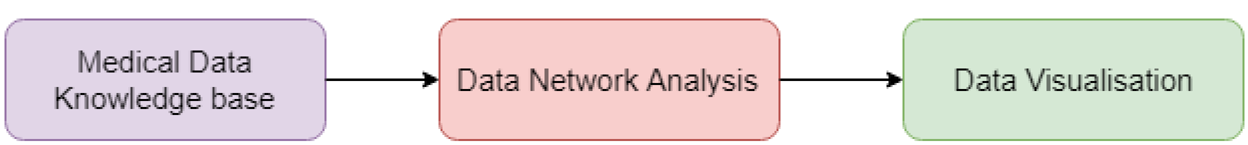

Figure 5.5: Processing Pipeline.

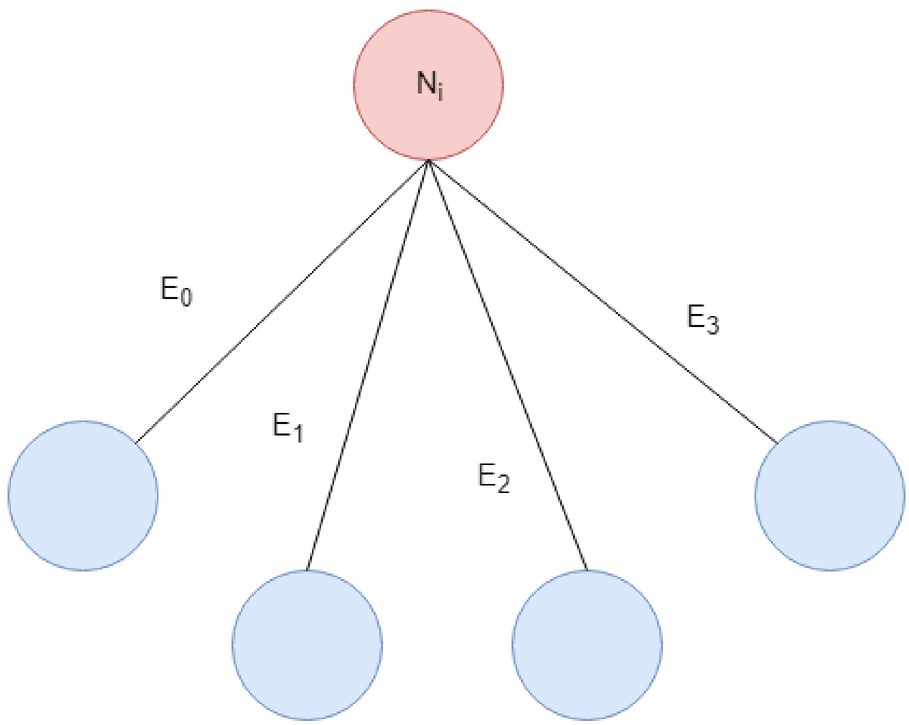

Figure 5.6: Example node N_i with a DC value of 4.

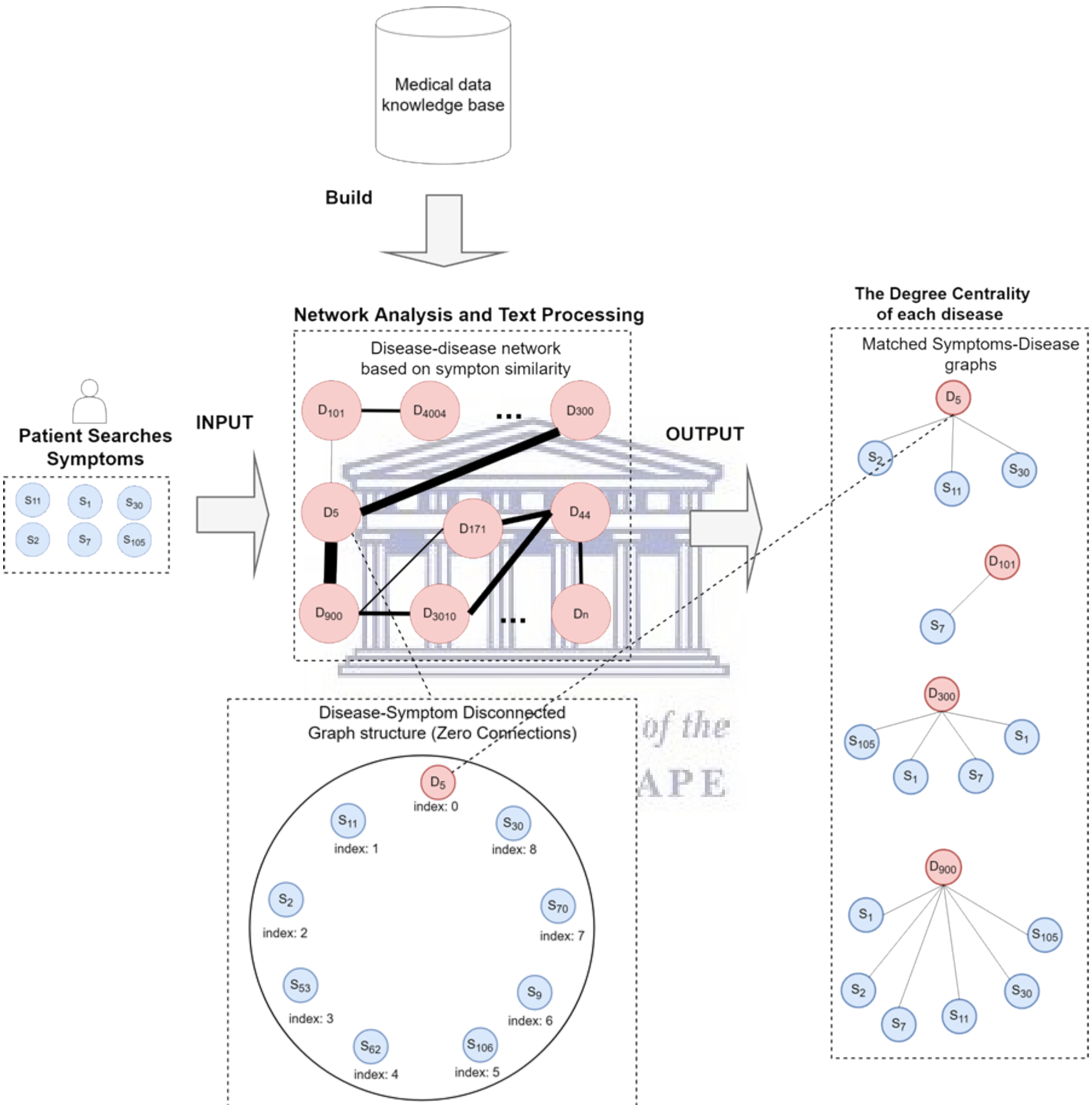

Figure 5.7: Disease identification through symptom search (*zero*-Edge Graph).

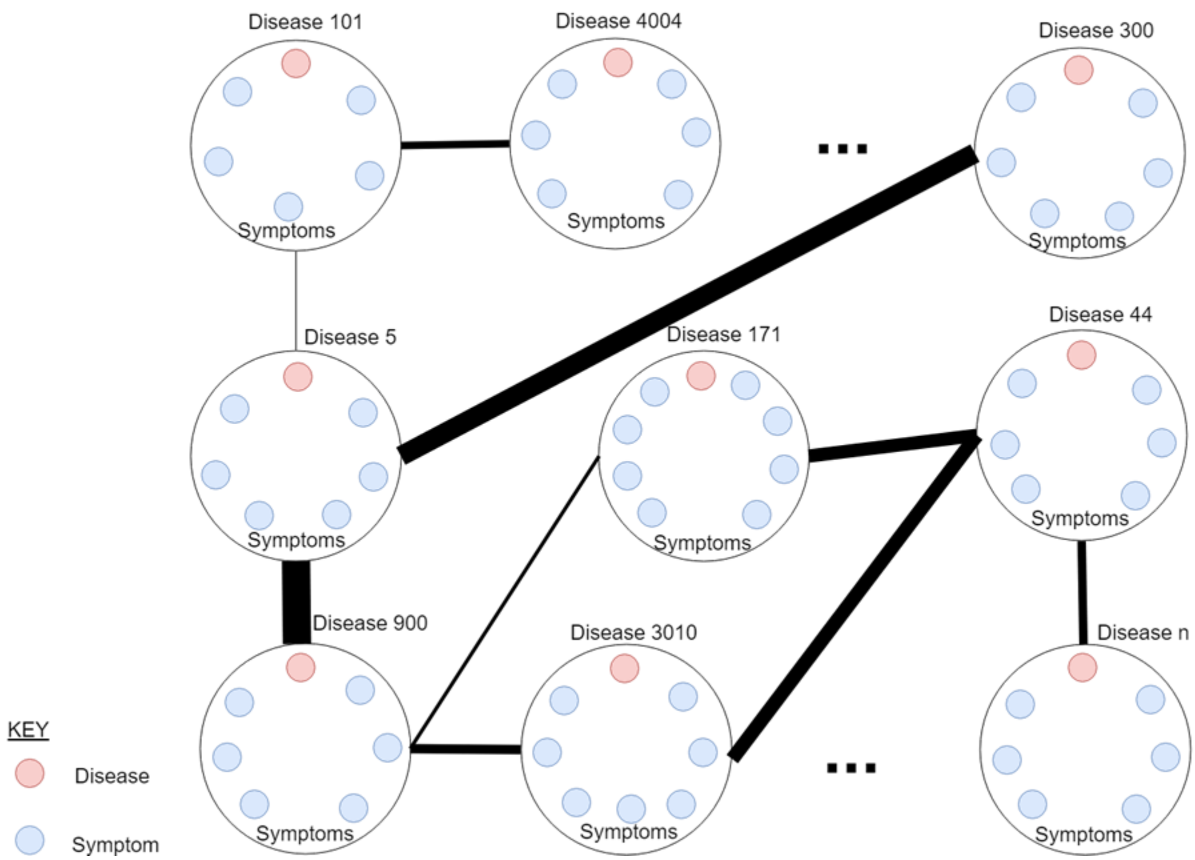

Figure 5.8: Disease-disease network based on symptom similarity (*zero*-Edge Graph).

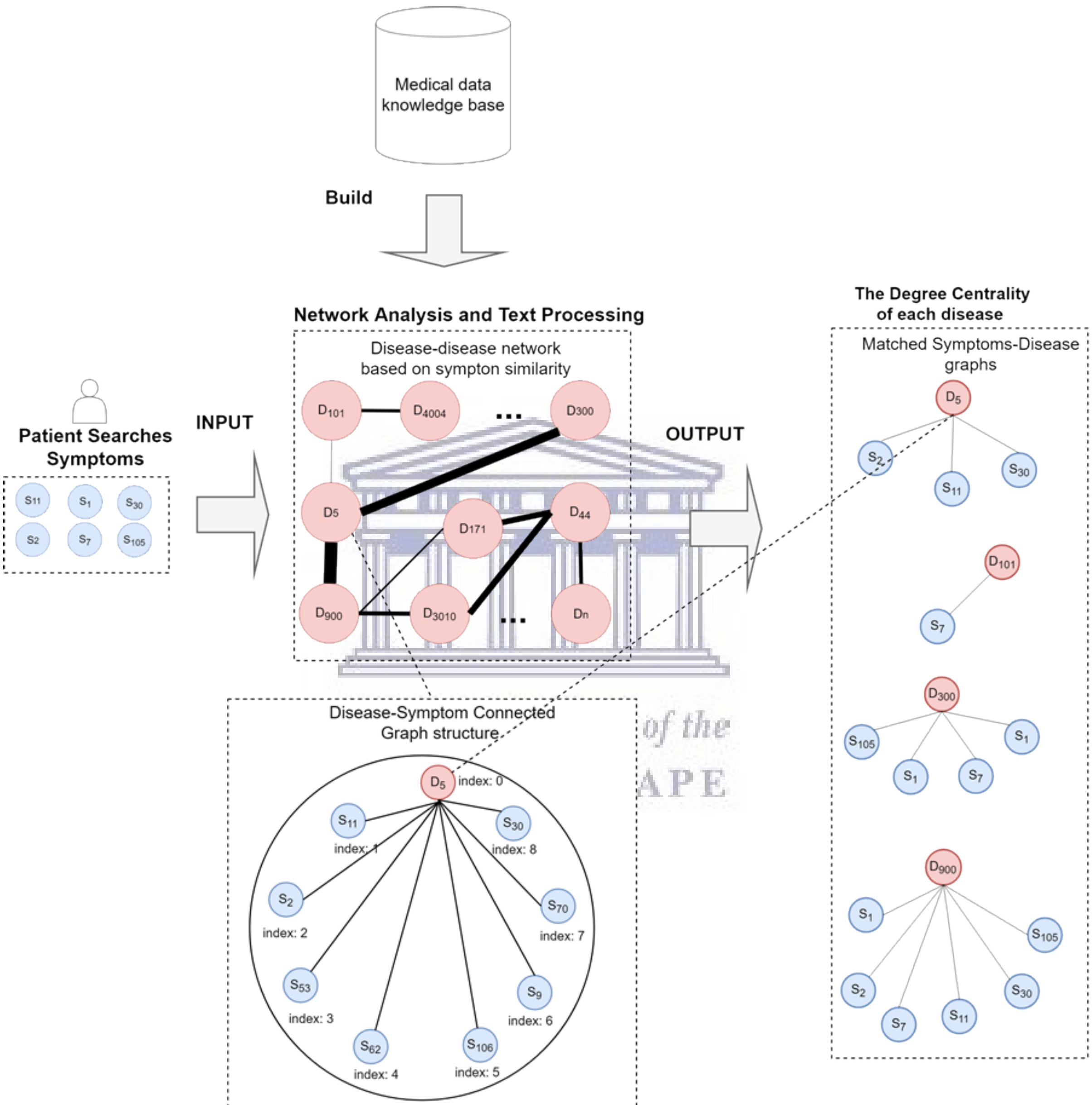

Figure 5.9: Disease identification through symptom search (n-Edge Graph).

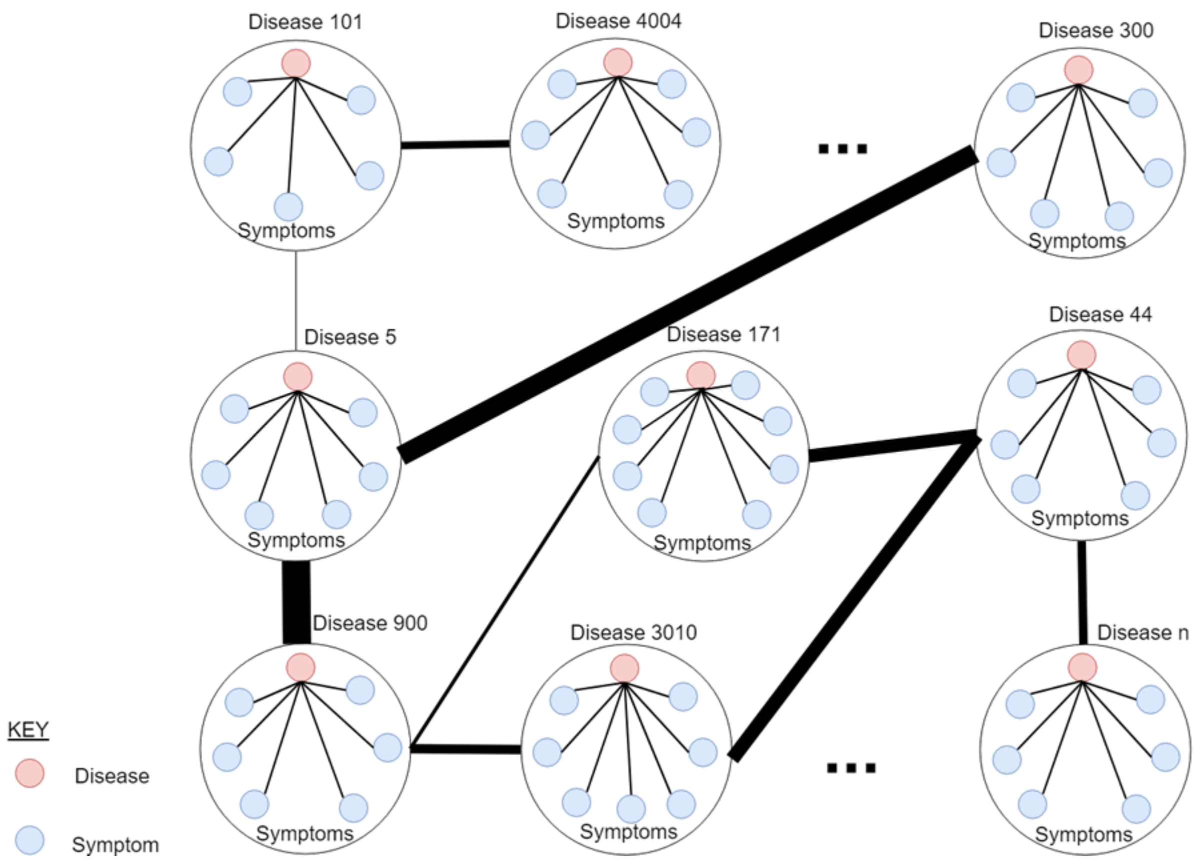

Figure 5.10: Disease-disease network based on symptom similarity (*n* Edge Graph).

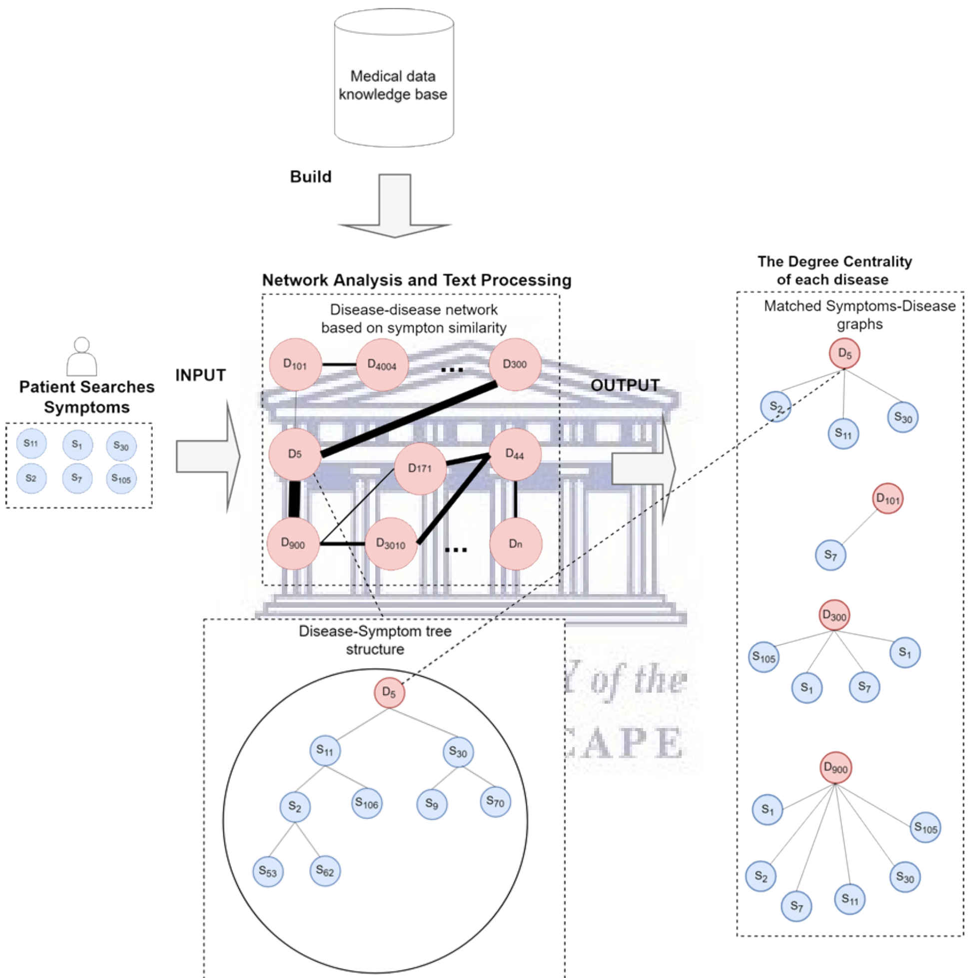

Figure 5.11: Disease identification through symptom search (Binary Tree).

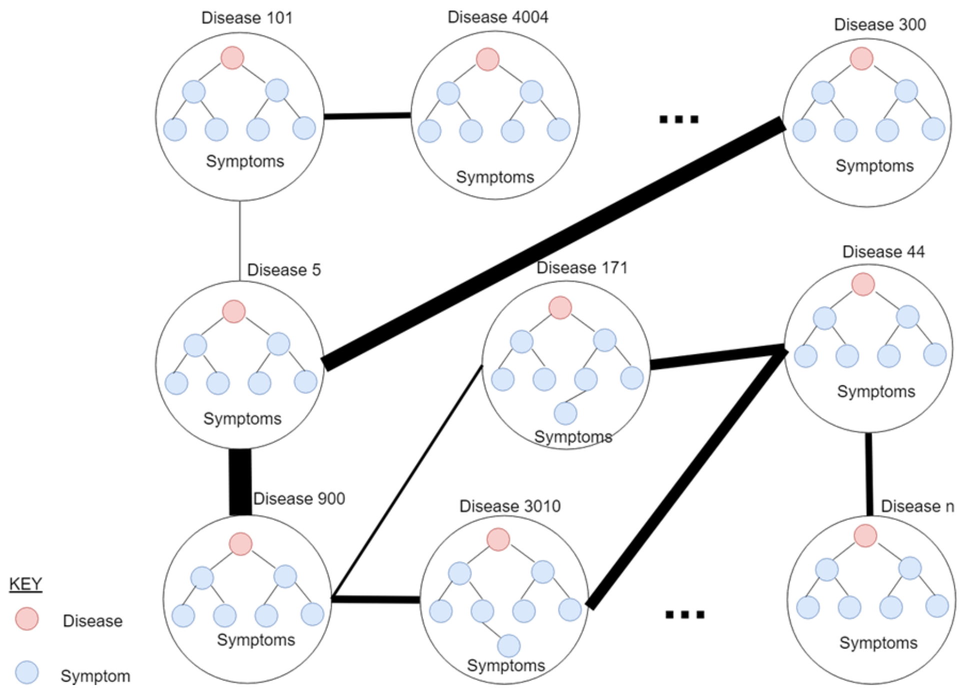

Figure 5.12: Disease-disease network based on symptom similarity (Binary Tree).

Chapter 6

Software Analysis and Design

6.1 Introduction

In this section the problem and requirements are identified and the system is decomposed into components.

6.1.1 Requirements

There are multiple Ehealth business capabilities that can be implemented to help automate some manual processes in public hospitals and clinics. Ehealth solutions can be large and complex and the proposed solution should enable continuous delivery and deployment. Two business capabilities are proposed; patient prioritization and condition recognition. Some parts of the application should be accessible through authentication and also depending on the user roles. The application should be deploy-able on any cloud platform and should be able to run locally on a mobile device when there is no connectivity. Should be able to connect and update the database once there is internet connectivity.

6.1.2 Monolithic Architecture

The figure 6.1 below shows the relationship between the hexagonal architecture and the clean architecture.

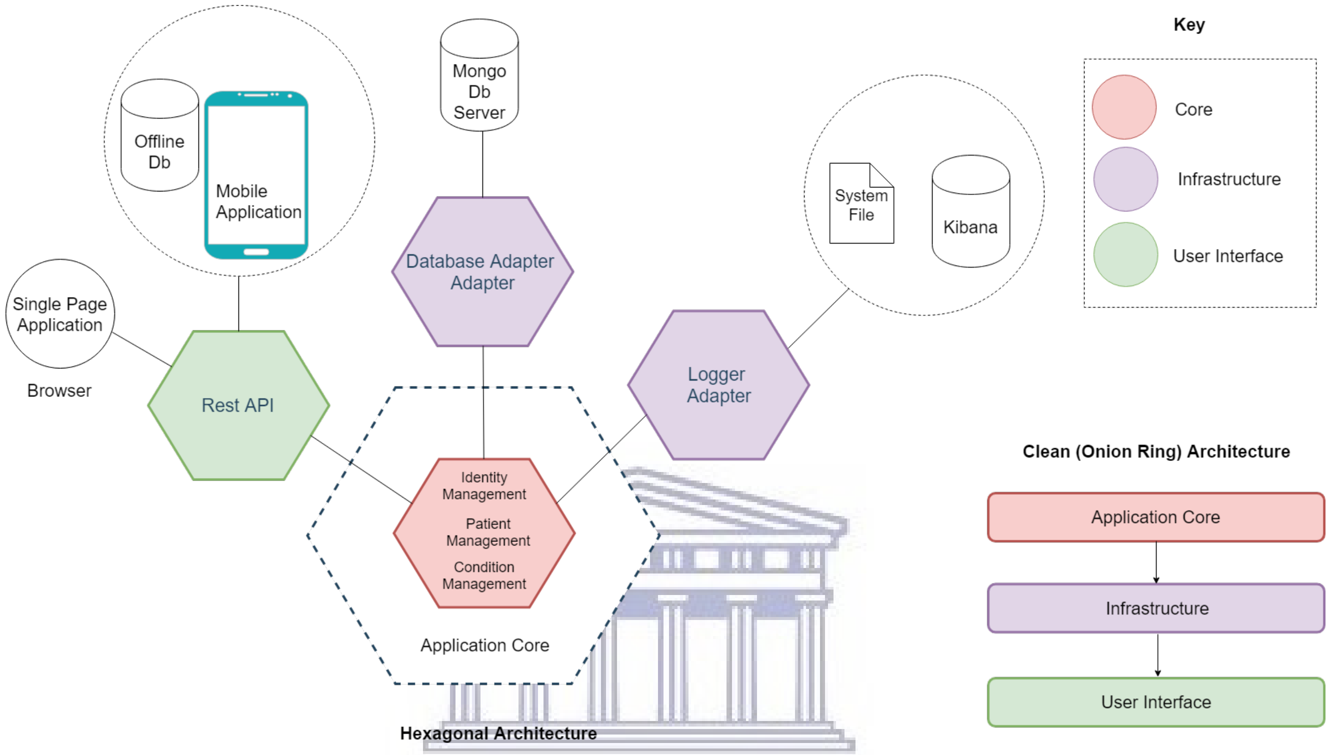

Figure 6.1: Monolithic Hexagonal and Clean Architecture.

The hexagonal architecture is an example of another architecture design that is used to develop monolithic application. As more functionality is developed a monolithic application becomes big and monstrous which leads to undesirable characteristics such as) [77]:

- difficult to understand and modify
- unable to support continuous delivery and deployment
- not easy to scale
- does not support polygloting

6.1.3 Proposed Architecture

The microservice architecture was proposed since it enables continuous delivery and deployment.

Microservice architecture is an a architectural style that is used to implement business capabilities as loosely coupled services (microservices) [77]. The problem is decomposed into different business capabilities and each one of them is represented by a service.

In software development containerization is an approach in which a service or an application, its configuration and its dependencies are packaged together as a container image. The containerized application can be deployed as a container image instance to the host operating system (OS) and the unit can also be tested as a unit. Containers are lightweight, easily packed and designed to run anywhere. A microservice is an application has a single business capability [18].

Figure 6.2 shows how three services i.e. identity, patient and condition were derived from their business capabilities identity management, patient management and condition management respectively.

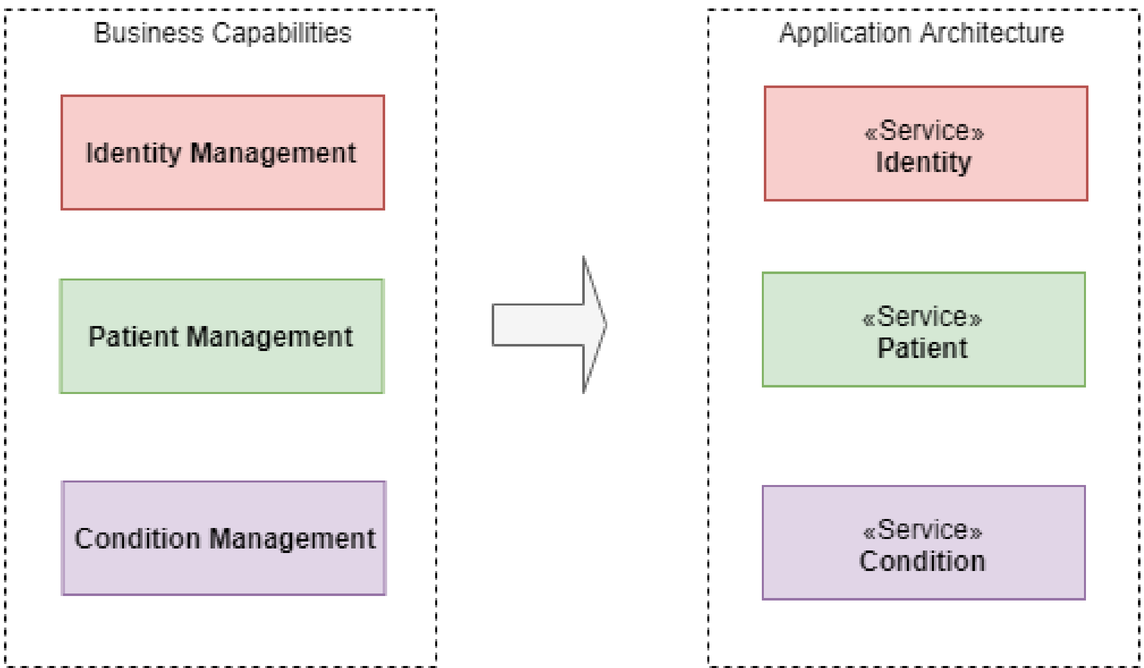

Figure 6.2: Decomposition by business capabilities.

The solution follows a microservices architecture and each service has it own context. This means each service will have its own database see Figure 6.4. The three main microservices are as follows:

- Patient Service: This microservice is responsible for patient management and its main purpose is to calculate the risk score and perform patient prioritization. It consists of the following algorithms: deep neural networks, single hidden layer neural networks, multiple variate linear regression, multiple variate logistic regression and classification regression decision trees.

- Identity Service: The identity microservice is responsible for user management or authentication.

- Condition Service: This condition microservice is responsible for disease identification or condition recognition.

Each microservice has its own context, database, unique port and can be deployed into its own container.

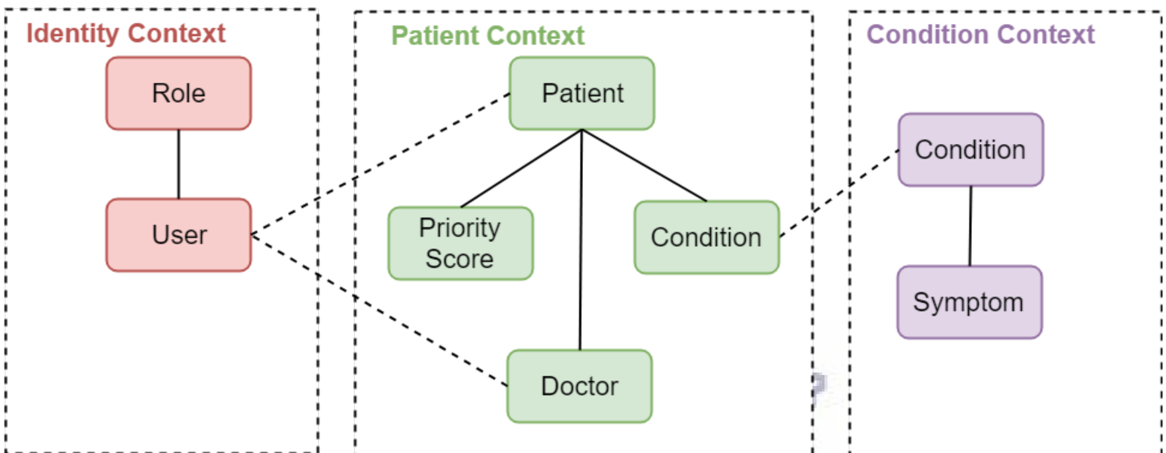

Figure 6.3: Bounded Context.

Bounded context defines boundaries of a complex domain into business context. It allows a ubiquitous language that is shared and valid within a boundary to be defined) [77]. The domain expert of the condition management system may not understand anything about the patient management system) [77]. The bounded context supports high cohesion. Figure 6.4 shows how the services in proposed solution are bounded to each other.

The microservice architecture is a collection of different architectural patterns. The event driven architecture, command pattern and the asynchronous request response pattern are some of the patterns used to implement the solution. See the illustration shown in Figure 6.4. Each micro service is hosted in its container hence why the different ports illustrated in Figure 6.5. The ip address is the same for all the containers showing that they are all hosted on the same machine.

6.2 Modeling Tools

Figure 6.6 shows the tools used to prototype the algorithms and the programming language used to develop the solution.

The prototyping of the proposed system was done using octave and then implemented using c sharp. Scikit a python library was used from classification and regression.

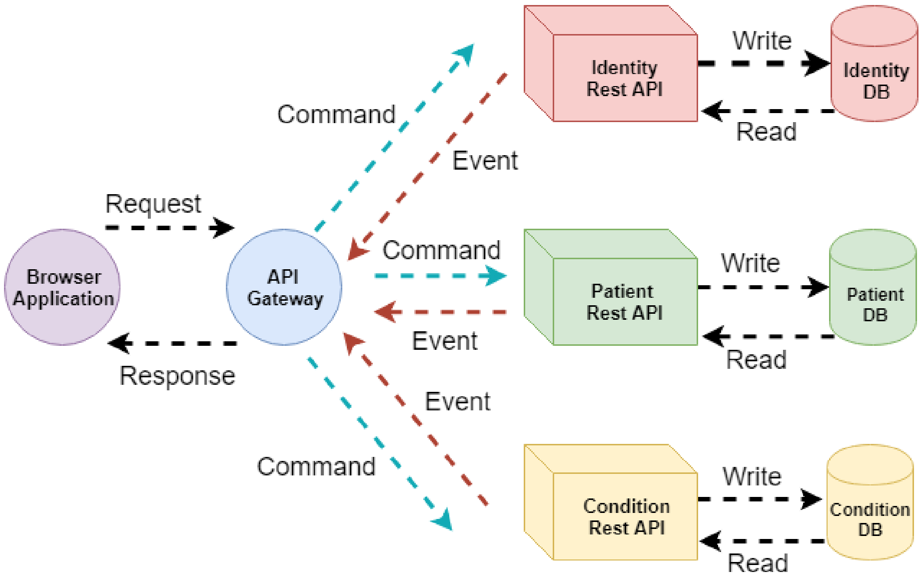

Figure 6.4: Microservices Architecture: Requests, Responses, Commands and Events.

The advantages of using programming languages like c and c-sharp is that they are supported by most electronic devices. The program can be run on small portable wearable devices.

6.3 Development Tools, Languages, Libraries and Frameworks

The following languages were used to build the client application:

- MsSQL and Mongo database: used to store historical patient data and other data like diseases and symptoms
- Firebase database: is a local storage database used by the client application when there is no internet or intranet connection.
- C sharp language: developing the client application back-end and micro-services. It is type safe and this comes with its advantages. Its syntax is very well designed and code is easily architected into readable code.
- Type Script language: used to integrated external libraries like D3 with frameworks like Angular 2+. Type Script supports strongly typed programming, access modifiers which allows encapsulation, supports generic programming. It makes it easy to follow SOLID principles and Object Oriented Programming to build your front-end application.
- Angular 2+: used to build the front-end client application. Angular 2+ was chosen because it makes use of typescript. Angular is a very rich front-end framework.

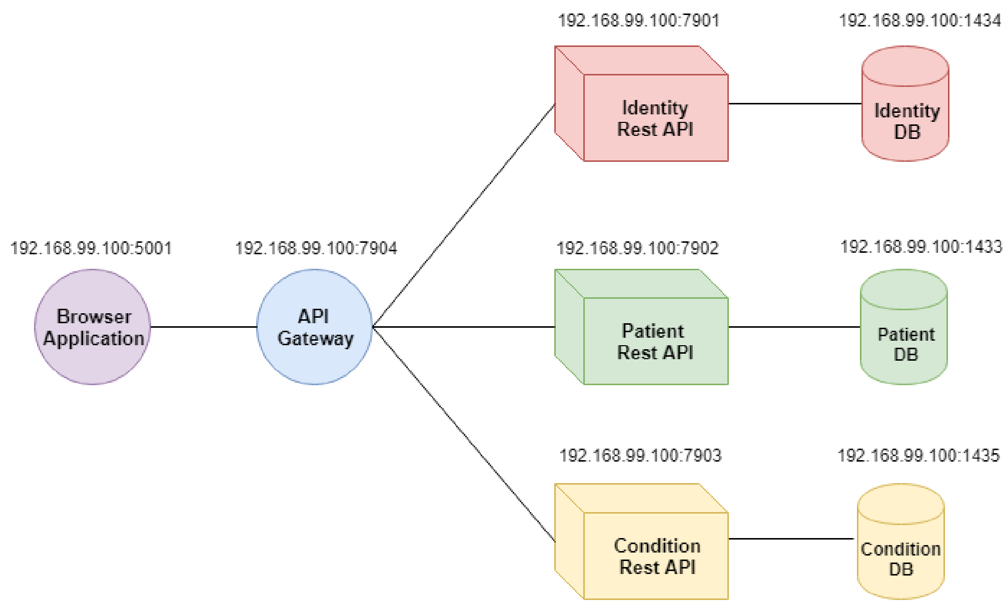

Figure 6.5: Microservices Architecture IP:PORT.

- D3.js visualization library: used for graphical visualizations. D3 can be used to develop advanced visualizations.

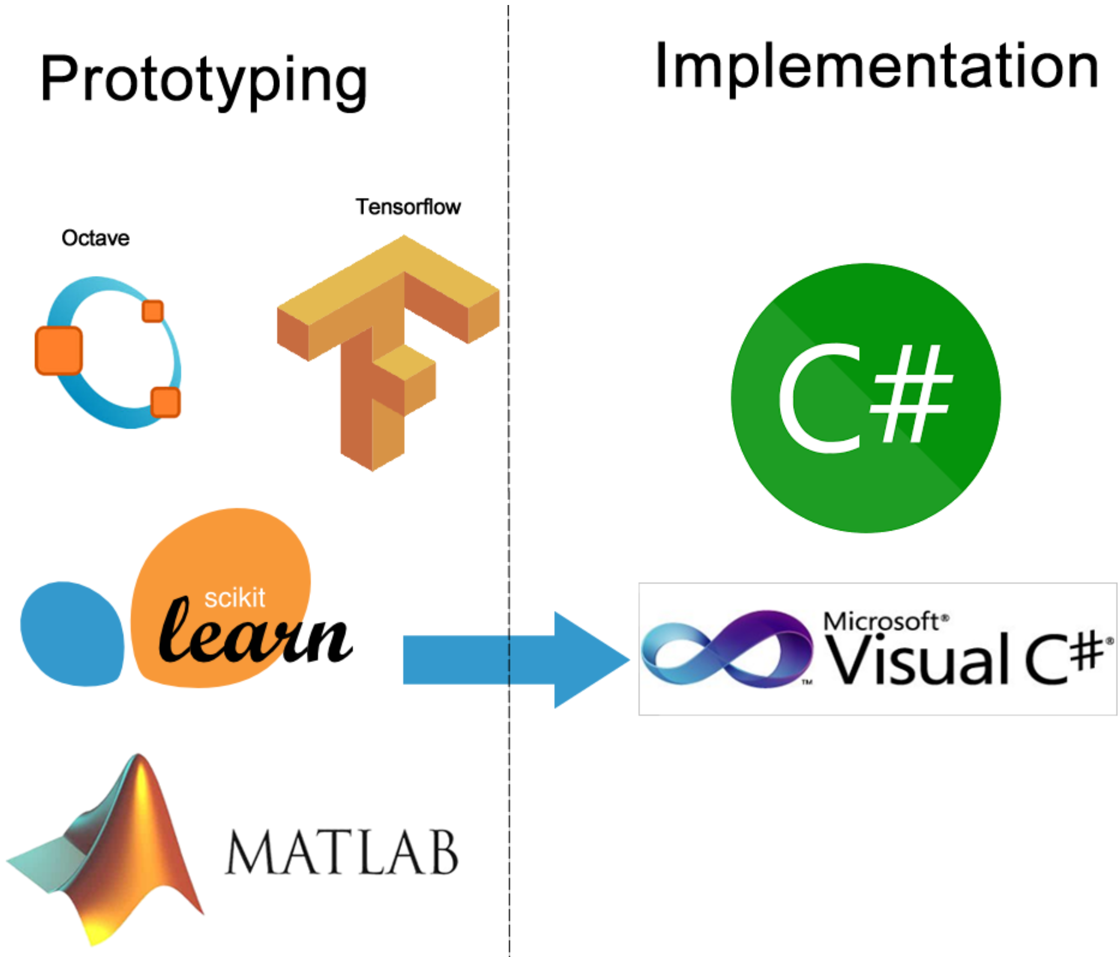

Figure 6.6: Prototyping, Programming Languages and Tools.

6.4 Visualizations

Visualizations were implemented in Type Script using a Java Script library D3 a data visualization library. There are five main visualizations that were designed and developed see the following Figures 6.7, 6.8, 6.9 and 6.10. The visualizations presented in this chapter are the main features for the proposed solution and more information is explained in Appendix D to Appendix H.

The patient vital signs are used to classify each patient and priority list as one shown in Figure 6.8 is generated.

The gauge on the left illustrated in Figure 6.7 shows the patient's priority score which can either be 1, 2 or 3 representing normal, low, medium or high priority respectively. The priority scoring system follows the South African Triage system and this can be useful in hospitals and clinics.

The gauge on the right shows the patient's overall health which can be any value between 0 and

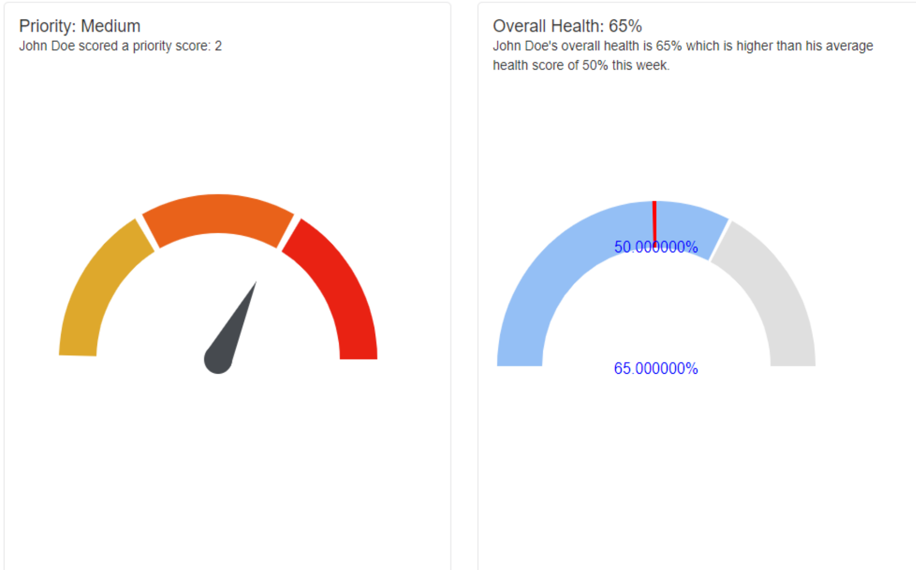

Figure 6.7: Priority and overall health visuals.

100 percent. The score is then compared to other individuals scores saved in the database. This can be connected to a wearable device such as a watch that also measures a person's heart rate. This gauge basically measures an individual's activities and basically the aim is to promote a health living. An individual can also track their diet and a score is calculated.

Patient Priority List

Firstname	Lastname	Priority
Edd	Eddie	High
Jane	Doe	High
Test	Test	Medium
John	Tulip	Low
John	Doe	Normal

Figure 6.8: Prioritization.

Figure 6.8 shows the patient's priority list which is colour coded as red, orange, yellow, and green representing a priority score of 3, 2, 1 and 0 respectively.

The risk score for each individual can be plotted on a line graph as shown on Figure 6.9.

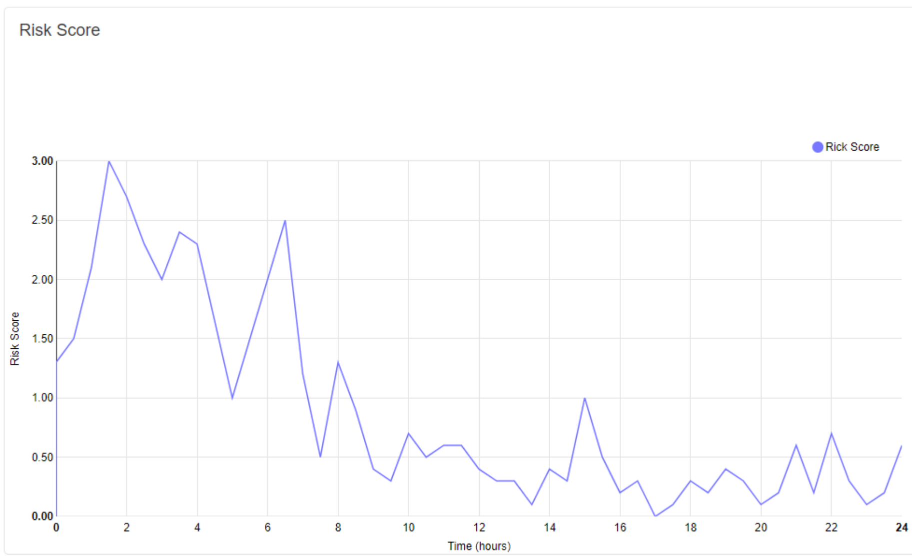

Figure 6.9: Risk score time series monitoring.

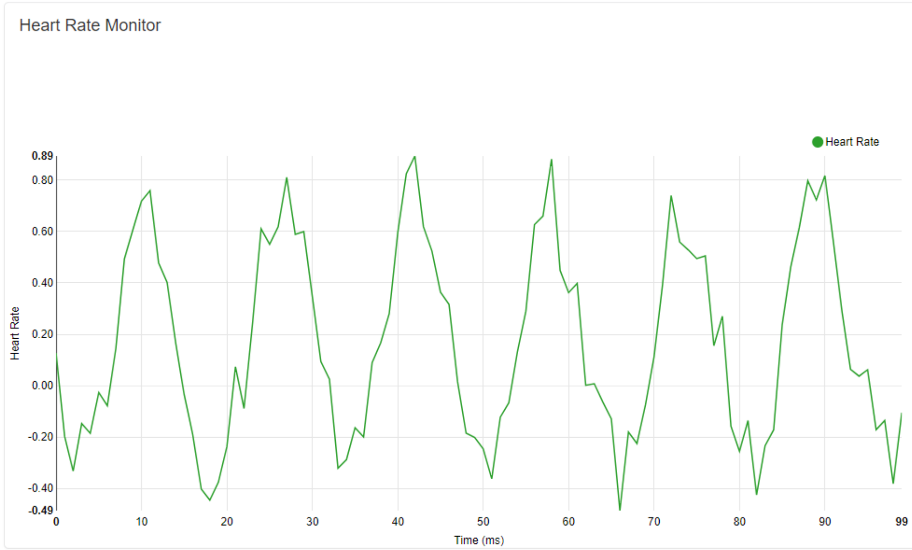

Figure 6.10: Heart rate time series monitoring.

Figure 6.10 shows that the heart rate can also be tracked using line graph visualisation. The plot does not show actual patient data. The illustration shows future implementations that can be included in the solution.

The user stories are described in Appendix D, Appendix E, Appendix F, Appendix G and Appendix H.

There are five roles medical practitioner, medical trainee, developer, administrator and patient. The developer role has access to the advanced settings or DNN optimization settings shown in (Figure F.1 and Figure F.2) The developer role and the administrator have access to the manage accounts screen (Figure G.1), the manage account screen (Figure G.2) and role assigning (Figure G.3). Every role has access to the manage personal profile screen (Figure H.1). The medical practitioner, medical trainee, and administrator role have access to the patient priority list screen (Figure D.1). All the roles have access to the priority score and overall health gauges (Figure E.1), heart rate monitor (Figure E.2) and risk score monitor (Figure E.3).

Chapter 7

Experimental Evaluation

This chapter presents the experimental evaluation for both proposed patient prioritisation and disease identification systems.

7.1 Patient Prioritisation

7.1.1 Methods of Evaluation

The regression algorithms were evaluated using a data set containing 17255 records and training to test ratio of 40:60. At the beginning of the study hospitals were contacted in-order to acquire data to train the algorithms. However the acquisition was unsuccessful therefore an alternative source of data was used. The continuous vital sign data for thirty two individuals undergoing surgical cases who underwent anaesthesia was acquired from The University of Queensland [92]. The records were combined to make the 17255 records data set.

Mean Squared Error (MSE)

The MSE also known as Mean Squared deviation (MSD) of an estimator measures the average of squares of deviations of the estimated from the estimator. An MSE value of zero means there was no variance between the estimated and the estimator. Therefore values close to zero indicate a very small difference between the estimator and the estimated.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \tag{7.1}$$

Where $\hat{Y}$ is a vector of estimated or predicted values and Y is a vector of the actual output values that are used to train the algorithm.

7.1.2 Coefficient of Correlation (R)

Is a measure of correlation and dependence, that is the statistical relationship between two or more random or bi-variate variables. Also defined as the degree to which two or more variable are associated with each other.

The values range from -1 to 1, the value -1 represents a strong negative goodness of fit, the value 1 represents a strong positive goodness of fit and a value close to zero represent a weak relationship.

Accuracy

The model is trained using 40 percent of the data and the other 60 percent of the data is used for cross validation.

In this study machine learning techniques are used to score the data. The triage score is then verified by the standardized South African triage system and its values assigned a status true positive (TP), false positive (FP),true negative (TN) and false negative (FN). These statuses are used to construct a confusion matrix whose values are used to calculate the *Precision* as defined by equation 7.4. The accuracy is then computed from the precision as depicted in the equation 7.6.

The recall and precision are calculated inorder to find the overall performance of the algorithm. Patients records are classified in four classes high, medium,low and normal. A record classified as normal means the patient is not sick and if classified as low, medium and high then the patient is sick. For an easy of interpretation and computation of the overall performance of the algorithm, the precision and recall can be taken as a binary classification. The class normal is represented by the class "not-patient" and the classes (low, medium and high) are represented by the class "patient". This is possible since during testing every predicted value is always compared to the original target value regardless of the class (normal, low, medium and high). Therefore this assumption does not affect the overall performance results.

The accuracy is also calculated by comparing the estimator and the estimated and at the same time counting the number of matches.

$$Accuracy(\%) = \frac{NoM}{NoR} \times 100 \tag{7.2}$$

Where NoM is the number of matches and NoR is the number of records in the test sets.

Epochs

The number of iterations the gradient descent algorithm takes to adjust weights and biases in-order to attain the highest accuracy.

7.1.3 Predictive Models

The Multivariate Logistic Regression (MLoR), Multivariate Linear Regression (MLiR), Single Hidden Layer Neural Network (SNN), Classification and Regression Decision Tree (CART) and Deep Neural Network (DNN) algorithms are evaluated and compared to determine the best algorithm to perform multiple classification.

Artificial Neural Network: Optimization

The Table 7.1 shows an attempt to increase the accuracy of the SNN algorithm by increasing epochs. Increasing the epochs had no effect on the accuracy of the algorithm. Also the best optimum α and momentum have a positive effect on time it takes to train the algorithm.

Algorithm	α	MSE	Accuracy (%)	Epochs	Time (min)
SNN 8-(7)-4	0.3	0.12	93.60	1000	0.80
SNN 8-(10)-4	0.3	0.05	97.88	1000	0.08
SNN 8-(10)-4	0.3	0.05	97.88	10000	0.12
SNN 8-(10)-4	0.3	0.05	97.88	100000	0.11

Table 7.1: Optimizing the Single Hidden Layer Neural Network.

In Table 7.2 the best Learning Rate (α) value was found to be 0.02 and increasing epochs also improved the accuracy.

Algorithm	α	MSE	Accuracy (%)	Epochs	Time (min)
DNN 8-(10-10-10)-4	0.01	0.026	94.97	1000	0.74
DNN 8-(9-11-9)-4	0.01	0.0010	97.86	1000	7.17
DNN 8-(9-11-9)-4	0.02	0.0040	99.16	1000	0.74
DNN 8-(9-11-9)-4	0.02	0.0033	99.16	2500	1.92
DNN 8-(9-11-9)-4	0.02	0.0013	99.68	5000	3.75
DNN 8-(9-11-9)-4	0.02	0.0008	99.68	7500	5.56
DNN 8-(9-11-9)-4	0.02	0.0007	99.86	10000	7.42
DNN 8-(9-11-9)-4	0.02	0.0007	99.86	12500	11.38
DNN 8-(9-11-9)-4	0.02	0.0007	99.86	15000	11.49
DNN 8-(9-11-9)-4	0.02	0.0007	99.83	17500	13.88
DNN 8-(9-11-9)-4	0.02	0.0006	99.88	20000	15.20
DNN 8-(9-11-9)-4	0.02	0.0006	99.88	22500	20.11
DNN 8-(9-11-9)-4	0.02	0.0006	99.88	25000	19.42

Table 7.2: Optimizing the Deep Neural Network.

The Figure 7.1 shows that the time it takes to train the DNN increases as the number of epochs increases. In Figure 7.2 the DNN's accuracy also increases as the number of epochs and time to train increases.

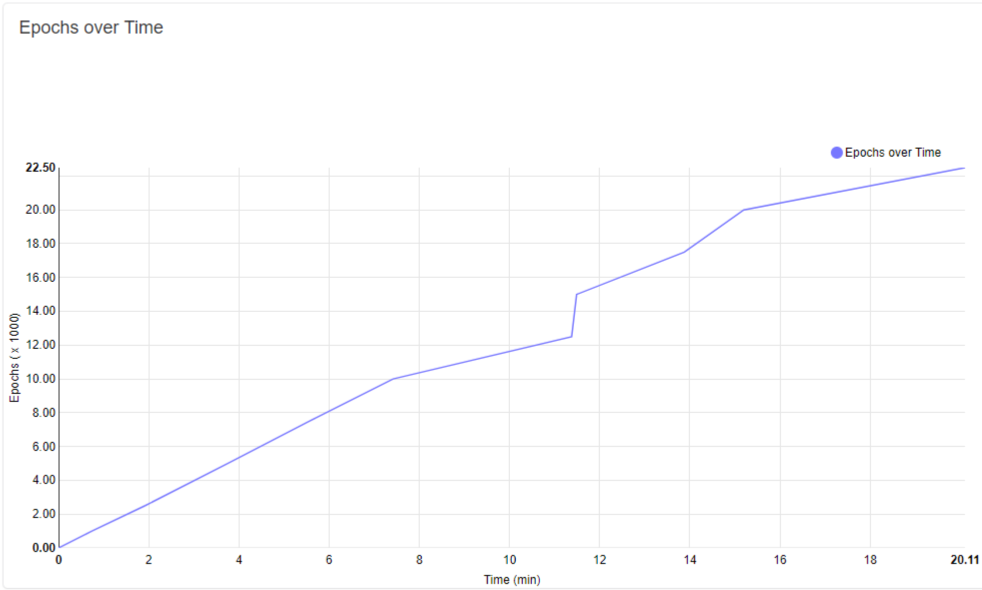

Figure 7.1: Epochs over Training Time.

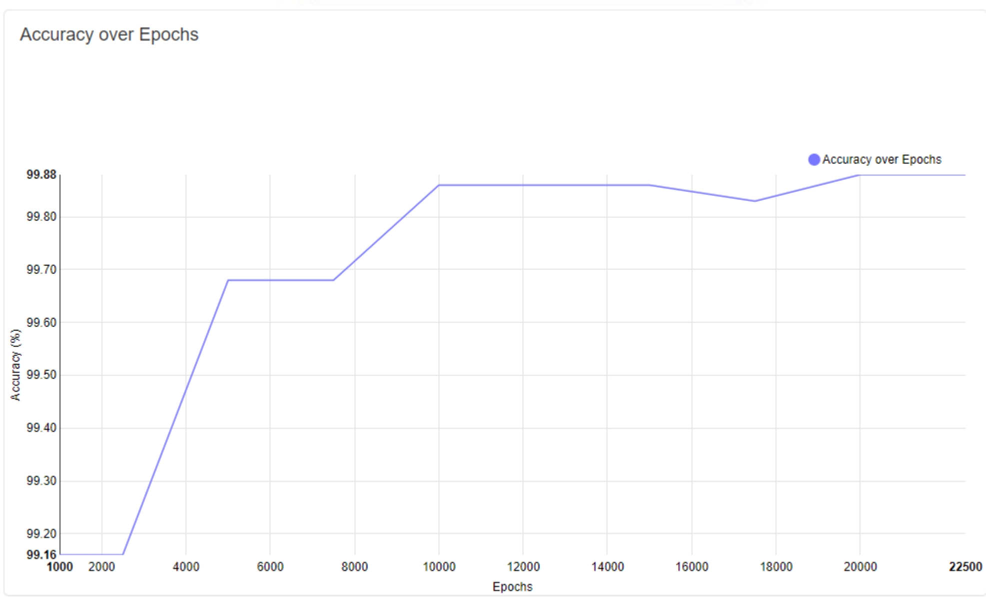

Figure 7.2: Accuracy over Epochs.

Confusion Matrix

Table 7.4 shows data for patients who have been given anesthesia. If the data is scored using a triage system and the triage score is verified by the doctor and assigned a status false positive (FP), true positive (TP), false negative (FN) and true negative (TN). The confusion matrix shown in Table 7.3 is calculated using equations 7.3, 7.4 and 7.5.

Recall or Detection or Hit rate or True Positive Rate

$$Recall = \frac{TP}{TP + FN} \tag{7.3}$$

$$Precision = \frac{TP}{TP + FP} \tag{7.4}$$

$$Falserate = \frac{FP}{TN + FP} \tag{7.5}$$

		Prediction outcome		
		p	**n**	**total**
actual value	**p′**	True Positive	False Negative	P′
	n′	False Positive	True Negative	N′
	total	P	N	

Table 7.3: Confusion Matrix.

Case	SBP	DBP	HR	SPO2	PULSE	Score	Status
1	95	48	55.9	98.9	55.6	0.07348	FN
2	98.5	54	78.5	95.7	78.3	-4.3819	TP
3	100.5	55.4	67.7	99.5	67.4	-3.3693	TP
4	114.1	63.3	113.1	98.55	113	-9.798391	TP
5	198.2	89.1	62.1	99.9	61.1	-22.8152	TP
6	104.7	56.9	67.2	98.1	66.7	-4.6763	TP
7	96.7	46.23	55.11	99.99	55.06	0.97306	FN
8	104.99	65.95	62.18	99.92	61.45	-7.0776	FP
9	106.72	46.33	70.45	99.53	70.17	-0.78477	TN
10	103.95	54.14	81.22	97.89	80.91	-4.38861	TP
11	111.69	72.51	88.34	99.99	87.89	-11.1947	FP
12	98.187	55.99	69.93	99.1549	69.7179	-3.611	TP
13	94.39	63.95	77.4	99.87	76.99	-6.26558	TP
14	141.56	86.78	99.21	99.04	81.22	-20.22591	TP
15	96.4656	44.84	55.39	99.43	57.28	1.369666	FN

Table 7.4: An example patient data set.

		Actual		
		Normal	Patient	Total
Predicted	Normal	$TN = 1$	$FN = 3$	4
	Patient	$FP = 2$	$TP = 9$	11
	Total	3	12	1

Table 7.5: Confusion matrix for the example data set.

The confusion matrix for the data shown in Table 7.4 can be represented by Table 7.5.

The confusion matrix precision parameter can be used to calculate the accuracy as shown in equation 7.6. The accuracy calculated using equations 7.2 and 7.6 are equivalent.

$$Accuracy = Precision \times 100 \tag{7.6}$$

7.1.4 Best Performance

The MLoR, MLiR and SNN algorithms were trained using a learning rate alpha value of 0.3 in 1000 iterations or epochs in order to achieve their best performance. The DNN algorithm was trained using a learning rate value of 0.02 in 10000 epochs. The DNN's accuracy increases as the number of epochs increases see Figure F.2 however this has a negative effect on the time it takes to train the algorithm.

The DNN has the best accuracy however it has the worst running/execution time. The MLiR has the best running/execution time but not the best accuracy. Both accuracy and time complexity of the algorithms are important. Both algorithms calculate weights which are used to calculate the output /estimate when given input parameters. These weights can be cached, saved in a file or database hence no need to recalculate the weights every time when calculating the output/estimate of new input values. The weights are only recalculated when given a new set of training data therefore the running/execution time is not as important as the accuracy.

Figure F.1 in the Appendix D shows a visualization that can be used to optimize the algorithm. Three parameters i.e. learning rate, momentum and max epochs can be adjusted until the best performance is achieved.

Algorithm	α	MSE	Accuracy (%)	Classification	Time (s)
MLoR	0.3	0.06	85.00	Multiple	0.22
MLiR	0.3	0.05	90.60	Multiple	0.12
CART	N.A	0.00	100.00	Binary	14.07
SNN	0.3	0.05	97.88	Multiple	4.81
DNN	0.02	0.001	99.86	Multiple	430.03

Table 7.6: Machine Learning Algorithms Performance Comparison.

7.2 Disease Identification

7.2.1 Method of Evaluation

The HSDN analysis algorithms is evaluated using a data set containing 147978 symptoms-disease records. The data structures are known to consume a lot of resources (memory and cpu). Performance testing is carried out by measuring the execution time and CPU usage of the algorithm as the data increases. When two or more programs are running simultaneously they tend to compete over resources. A constant environment is established by making sure only the HSDN program is running whilst monitoring the CPU usage and memory consumption.

7.2.2 Building and Searching the HSDN

Four symptoms (Fever, Pain, Vomiting and Diarrhea) are searched during the evaluation of the three networks (n-Edge graph based, $zero$-Edge graph based and the binary tree based). About 2548 diseases with at-least one of the four mentioned symptoms are found. The illustrations Figure 7.3 and Figure 7.4 show that the execution time of the three networks increases as the amount of data increases.

The binary tree based network performs better than the other two networks.

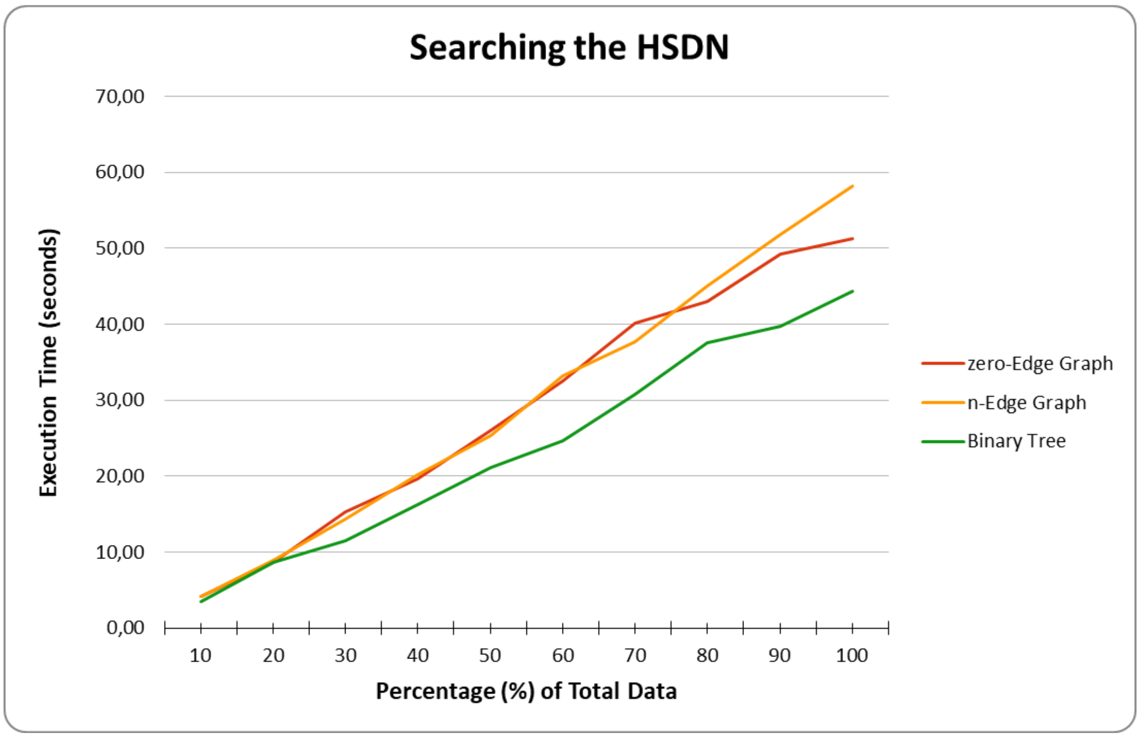

Figure 7.3: Execution time over amount of data (Searching the HSDN).

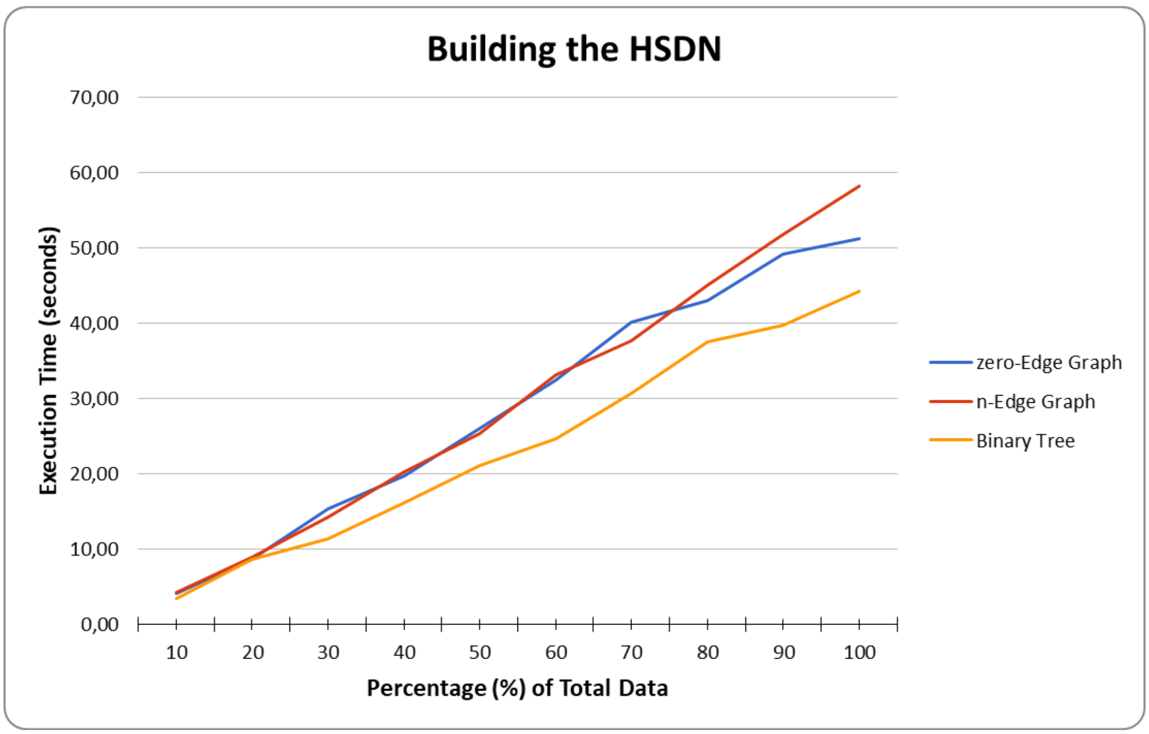

Figure 7.4: Execution time over amount of data (Building the HSDN) .

Chapter 8

Conclusion and Future Work

8.1 Conclusion

Medical knowledge can be accessed electronically in a more effective way using directed graphs which can easily show the relationships between diseases and their symptoms. Medical conditions can share symptoms and a condition can lead to another. Most pseudo science methods fail to eliminate human errors and other systematic errors. In machine learning scientific methods have ways to reduce biases during training. The abnormal values tend to move away from the model's best line of fit and hence can be ignored Open source technology was chosen and hence there are no subscription fees paid to make use of libraries, frameworks or services.

This study supports the use of vital parameters to measure patient risk scores. Pre-diagnosis and virtual nursing was successfully achieved through disease identification from symptoms using a key value binary tree and a directed graph. A heart attack pre-diagnosis graph was provided as an example of a early warning or pre-diagnosis test. The same graph data structure can be used to implement pre-diagnosis tests for other diseases such as diabetes. The graph data structure can be used to represent disease-symptoms and disease-disease relationships and there by automating medical knowledge. Visualizations were designed to help medical experts interpret the results from ML algorithms.

The MLoR, MLiR, SNN and DNN were found to be useful in multiple classification. The CART algorithm produced excellent binary classification results.

Some ML algorithms can achieve a very high accuracy close to 100%. As the problem becomes complex ML algorithms such as DNNs can be used. The DNN can take many variables and still achieve high accuracy by learning over time through the increase of epochs. DNN can also process qualitative data by encoding the values to numeric values. The use of automated systems eliminates human error altogether. The algorithms are trained using a small sample and can accurately classify any other set of parameters. In this study forty percent of the data was used to train the algorithm and the other sixty percent was used to test the accuracy of the ML algorithms. This proves that scientific methods are more useful than pseudo-science health systems.

This study has proved to be a very good start to build an ehealth solution that is cost effective, reliable and the results can be accessed in real time.

The disease-symptoms relationships analysis network was successfully implemented using graphs and binary trees. The binary tree based HSDN was found to be the best performing network.

The relationship between the main objectives of this study i.e. patient prioritization and disease identification was not shown. Patient prioritization should be also be dependent on disease identification and not only vital parameters. This relationship depends most on real patient data and availability of medical professionals to validate correctness of the proposed architectures.

8.2 Future Work

The CART algorithm has an excellent performance and can be further configured to support multiple classification. Multiple classification using CART will most likely produce better results than the other two algorithms MLoR and MLiR.

Future work will include implementing applications of other disease identification networks that is (1) disease-gene relationships and (2) disease-disease networks based on genes (Protein-Protein Interactions). Processing of grammatical search phrases using machine learning will be implemented in future.

8.3 Limitations

Access to medical data in South African hospitals still remains a problem. An attempt was made to get real patient data from hospitals however the attempts were not fruitful. This study is a motivates the importance of real data in research. There are rules and regulations that are created to govern the use of medical technology within the health-care industry. There are Privacy and Security Acts that play an important role in making sure that there is security and safety with regards to patient information [21]. However that data can be anonymous and made available for research.

Bibliography

[1] H. Bagula A. Bagula, M. Mandava. A framework for healthcare support in the rural and low income areas of the developing world. *Journal of Network and Computer Applications*, July 2018.

[2] M. Mueller F. Dewarrat J Klisic M. Donath L. Heinemann W.A. Stahel A. Caduff, M. S. Talary. Non-invasive glucose monitoring in patients with type 1 diabetes: a multisensor system combining sensors for dielectric and optical characterisation of skin. *Biosens Bioelectron*, 24:2778–2784, February 2009.

[3] E. Menasalvas E. Xu H. F. González A. R. Vasiljevs F. Sasaki C. Garattini F. M. Aarestrup G. Roesems-Kerremans C. Costa-Soria D. Hamelinck J. Thoms M. M. Sanchez P. Casla M. Despenic P. Czech M. Romao R. Wuyts M. Melideo S. Consoli M. A. Mayer N. S. Rüping M. Petkovic N Stojanovic S. Campbell S. Chatterjea I. W. Kraaji A. Heinrich, A. Lojo and N. Boujemaa. Big data technologies in health care: Needs, opportunities and challenges. *Big Data Value Association: Bruxelles*, December 2016.

[4] U. Srinivasan A. Khan, S.Uddin. Adapting graph theory and social network measures on healthcare data: a new framework to understand chronic disease progression. *ACSW 2016 Proceedings of the Australasian Computer Science Week*, 2016(66), February 2016.

[5] U. Srinivasan A. Khan, S.Uddin. Comorbidity network for chronic disease: A novel approach to understand type 2 diabetes progression. *International Journal of Medical Informatics*, 115:1–9, February 2018.

[6] R. H. Guy A. Sieg and M. B. Delgado-Charro. Noninvasive and minimally invasive methods for transdermal glucose monitoring. *Advanced Technologies & Treatments for Diabetes (ATTD)*, 7(1):174–197, February 2005.

[7] O. Amft and G. Troster. Recognition of dietary activity events using on-body sensors: Artificial intelligence in medicine. *Elsevier Science Publishers Ltd. Essex, UK*, 42(2):121–136, February 2008.

[8] G. D. Baura. *Medical Device Technologies: A Systems Based Overview Using Engineering Standards.* Academic Press, October 2011.

[9] S. Baškarada and A. Koronios. Unicorn data scientist: The rarest of breeds. *Program*, 51(1):65–74, 2017.

[10] P. Beena and R. Ganguli. Structural damage detection using fuzzy cognitive maps and hebbian learning. *Applied Soft Computing*, 11(1):1014–1020, December 2011.

[11] N. Bhatia and S. Kumar. Prediction of severity of diabetes mellitus using fuzzy cognitive maps. *Advances in Life Science and Technology*, 29:71–78, 2015.

[12] M. Breier. The shortage of medical doctors in south africa. *Depart of Labour South Africa*, 2008.

[13] F. J. Britten. *Old Clocks and Watches and Their Makers: Being an Historical and Descriptive Account of the Different Styles of Clocks and Watches of the Past, in England and Abroad; To Which Is Added a List of Ten Thousand Makers*. London, B. T. Batsford, 2016.

[14] E. Brynjolfsson and A. McAfee. The second machine age: Work, progress, and prosperity in a time of brilliant technologies. *New York: W.W. Norton & Company*, 2014.

[15] K. Kabino C. Spreckelsen. How is my field evolving? - network based analysis of biomedical scientific discourse. *Studies in Health Technology and Informatics*, 223:199–206, 2016.

[16] U. E. Udoh V. N. Barker G. C. Thonier A. M. Tellier R. D. Ricks R. B. Darling Y. D. Cagle N. A. Cabrol C. W. Mundt, K. N. Montgomery. A multiparameter wearable physiologic monitoring system for space and terrestrial applications. *IEEE Transactions in Information Technology in Biomedicine*, 9(4):382–391, september 2005.

[17] A. Calvo. *Beginning Android Wearables: With Android Wear and Google Glass SDKs*. New York: Apress L.P, 2015.

[18] M. Rousos Cesar de la Torre, B. Wagner. *.NET Microservices: Architecture for Containerised .NET Applications*. Microsoft Corporation, 2018.

[19] S. Chatterjee and S. Laxminarayan. 5th international workshop on enterprise networking and computing in healthcare industry. *IEEE*, August 2003.

[20] J Chery. Abdominal pain, chonic. *MD Neurogastroenterol Motil*, 18(7):499–506, February 2016.

[21] J. J. Chin. The use of information technology in medicine defining its role and limitations. *Springer Computational Linguistics and Intelligent Text Processing*, 44(3):149–151, 2005.

[22] L. Clendening. *Source Book of Medical History*, volume 32. American Journal of Public Health Nations Health., September 1960.

[23] F. M. Collen and M. J. Ball. *The History of Medical Informatics in the United States*. New York: Springer, 2015.

[24] Cursoenarm.net. Clinical assessment of patients suspected to have ischemic chest pain.

[25] A. T. Reisner D. B. Combie and H. H. Asada. Motion based adaptive calibration of pulse transit time measurements to arterial blood pressure for an autonomous, wearable blood pressure monitor. *IEEE EMBS*, pages 989–992, August 2008.

[26] J. A. Osheroff B. Middleton J. M. Teich J. S. Ash E. Campbell D. F. Sittig, A. Wright and D. W. Bates. Grand challenges in clinical decision support. *Journal of Biomedical Informatics*, 41(2):387–392, April 2008.

[27] G. Maccioni D. Giansanti and S. Morelli. An experience of health technology assessment in new models of care for subjects with parkinson's disease by means of a new wearable device. *Telemed J E Health*, 14(5):467–472, June 2008.

[28] J. Dzau and C. Liew. Cardiovascular genetics and genomics for the cardiologist. *Wiley-Blackwell*, 2016.

[29] L. Thijs C. Davidson J. A. Staessen P. W. D. Leeuw M. Jaaskivi C. Nachev G. Parati E. T. O'Brien J. Tuomilehto J. Webster C. J. Bulpitt E. Pringle, C. Phillips and R. H. Fagard. Systolic blood pressure variability as a risk factor for stroke and cardiovascular mortality in the elderly hypertensive population. *Journal of Hypertension*, 21(12):2251–2257, December 2003.

[30] N. Sazonova E. S. Sazonov, G. Fulk and S. Schuckers. Automatic recognition of postures and activities in stroke patients. *Conference Proceedings in the IEEE Engineering in Medicine and Biology Society*, pages 2200–2203, 2009.

[31] C.D. Stylios2 E.I. Papageorgiou1 and P.P. Groumpos. Hybrid model based on decision trees and fuzzy cognitive maps for medical decision support system. *World Congress on Medical Physics and Biomedical Engineering 2006*, 6(24):3689–3692, 2006.

[32] O. Erdinc. Optimization in renewable energy systems: Recent perspectives. *Cambridge: Elsevier Ltd*, 2017.

[33] P. Flach. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press, 2012.

[34] A. Gedeon. Science and technology in medicine: An illustrated account based on ninety-nine landmark publications from five centuries. *Springer Science and Business Media Inc*, 2006.

[35] M. Gleiser. The island of knowledge: The limits of science and the search for meaning. *New York: Basic Books*, 2014.

[36] P. Gupta. Machine learning: The future of healthcare. *Harvard Science Review*, 2017.

[37] J. Habetha. The myheart project–fighting cardiovascular diseases by prevention and early diagnosis. *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 6746–6749, September 2006.

[38] Y. Hao and R. Foster. Wireless body sensor networks for health-monitoring applications. *Institute of Physics and Engineering in Medicine*, 29(11):27–56, October 2008.

[39] J. Hendler and A. M. Mulvehill. Social machines: The coming collision of artificial intelligence, social networking, and humanity. *New York: Apress*, 2016.

[40] J. Holland. *Wearable Technology and Mobile Innovations for Next-Generation Education*. Hershey PA: IGI Global, April 2016.

[41] MD. Chamberlain J. James. *Role of Continuous Glucose Monitoring in Diabetes Treatment.* American Diabetes Association, 2018.

[42] H. Cheng Z. Xie S. Xu K-I. Jang J. W. Lee Z. Liu P. Gutruf X. Huang P. Wei F. Liu K. Li M. Dalal R. Ghaffari X. Feng Y. Huang S. Gupta U. Paik J. Kim, A. Banks and J. A. Rogers. Epidermal electronics with advanced capabilities in near-field communication. *WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim*, February 2015.

[43] H. J. Shim R. Ghaffari H. R. Cho D. Son Y. H. Jung M. Soh C. Choi S. Jung K. Chu D. Jeon S. T. Lee J. H. Kim S. H. Choi T. Hyeon J. Kim, M. Lee and D. H. Kim. Stretchable silicon nanoribbon electronics for skin prosthetics. *Nature Communications*, 5(5747), December 2014.

[44] N. Rochester J. McCarthy, M. L. Minsky and C. E. Shannon. A proposal for the darthmouth summer research project on artificial intelligence. *Dartmouth Summer Research Project on Artificial Intelligence*, August 1956.

[45] J. W. Park J. Park, D. Jang and S. Youm. Wearable sensing of in-ear pressure for heart rate monitoring with a piezoelectric sensor. *Sensors*, 15(9):23402–23417, 2015.

[46] J. Hynson T. Hartmann H. Reyford J. M. Martino M. Poliac K. Belani, M. Ozaki and R. Miller. A new noninvasive method to measure blood pressure: Results of a multicenter trial. *Anesthesiology*, 91(3):686–692, 1999.

[47] J. H. Yeo K. W. Sim W. L. Peh N. H. Chua N. K. Chua K. G. Ng, C. M. Ting and F. Kwong. Progress on the development of the mediwatch^{TM} ambulatory blood pressure monitor and related devices. *Blood Pressure Monitoring*, 9(3):149–165, 2004.

[48] D. Peikes L. Moreno and A. Krilla. Using directed graph based bdmm algorithm for chinese word segmentation. *Springer Computational Linguistics and Intelligent Text Processing*, pages 214–217, 2005.

[49] D. Peikes L. Moreno and A. Krilla. Necessary but not sufficient: The hitech act and health information technology's potential to build medical homes. *Rockville, MD: Agency for Healthcare Research and Quality*, (10), 2010.

[50] B. G. Lipták and K. Venczel. *Measurement and Safety*, volume 1. CRC Press, October 2016.

[51] M. Mandava C. K. Lubamba A. B. Bagula M. F. Bagula, H. Bagula. Cyber-healthcare kiosks for healthcare support in developing countries. *AFRICOMM 2018: e-Infrastructure and e-Services for Developing Countries*, 275:185–198, March 2019.

[52] A. Di Gioia L. Conti L. Berti L. D. Esposti G. Tarrini N. Melchionda M. G., Benedetti and S. Giannini. Physical activity monitoring in obese people in the real life environment. *Journal of NeuroEngineering and Rehabilitation*, 6(47), December 2009.

[53] Y. Toussaint M. Hassan, A. Coulet. Learning subgraph patterns from text for extracting disease-symptom relationships. *CEUR Workshop Proceedings*, 1202:81–96, 2014.

[54] N. H. Lovell B. G. Celler S. R. Lord M. J. Mathie, A. C. Coster and A. Tiedemann. A pilot study of long-term monitoring of human movements in the home using accelerometry. *Journal of Telemedicine and Telecare*, 10(3):144–151, January 2004.

[55] T. Ohokubo I. Tsuji M. Michimata M. Matsubara M. Ota K. Nagai T. Araki H. Satoh S. Ito S. Hisamichi M. Kikuya, A. Hozawa and Y. Imai. Prognostic significance of blood pressure and heart rate variabilities. *Hypertension*, 36(5):901–906, November 2000.

[56] A. Ismail A. Bagula H. Bagula M. Mandava, C. Lubamba. Cyber-healthcare for public healthcare in the developing world. *Proceedings IEEE Symposium on Computers and Communications*, 2016 August:14–19, 2016.

[57] C. K. Lubamba M. Zennaro E. Pietrosemoli M. Mandava, A. B. Bagula and H. Bagula. Cloud based patient prioritization as service in public health care. *ITU-Kaleidoscope*, November 2016.

[58] M. Lenic M. Sprogar and S. Alayon. Evolution in medical decision making. *Journal of Medical Systems*, 26(5):479–489, 2002.

[59] C. Marci M. Sung and A. Pentland. Wearable feedback systems for rehabilitation,' journal of neuroengineering and rehabilitation. *licensee BioMed Central Ltd*, 2(17), June 2005.

[60] L. N. Magner. A history of the life sciences, revised and expanded. *CRC Press*, August 2002.

[61] J. L. Malin. Envisioning watson as a rapid-learning system for oncology. *Journal of Oncology Practice*, 9(3):155–157, 2013.

[62] M. Mandava. *Participatory Healthcare System (Situation recognition and Patient Prioritization)*. University of the Western Cape - Computer Science Department, 2015.

[63] J. J. McArdle and G. Ritschard. Contemporary issues in exploratory data mining in the behavioural sciences. *New York: Routledge*, 2013.

[64] S. R. Mulik and B. T. Jadhav. Fuzzy system for diagnosis and treatment of kidney diseases. *International Journal of Scientific Engineering and Research (IJSER)*, 3(12):56–59, 2014.

[65] T. Yokota H. Jinno K. Kuribara T. Sekitani N. Matsuhisa, M. Kaltenbrunner and T. Someya. Printable elastic conductors with a high conductivity for electronic textile applications. *Nature Communications*, 6(7461), June 2015.

[66] A. Dittmar N. Taccini, G. Loriga and R. Paradiso. Knitted bio-clothes for health monitoring. *IEEE-EMBS*, pages 2165–2168, September 2004.

[67] J. D. Newman and A. P. F. Turner. Home blood glucose biosensors: A commercial perspective. *Biosensors and Bioelectronics*, 20(12):2435–2453, 2005.

[68] P. R. Nicholas. Scala for machine learning. *Birmingham: Packt Publishing*, 2015.

[69] B. Lo M. Elhelw L. Wang G. Z. Yang O. Aziz, L. Atallah and A. Darzi. A pervasive body sensor network for measuring postoperative recovery at home. *Surgical Innovation*, 14(2):83–90, June 2007.

[70] N. Oliver and F. Flores-Mangas. Healthgear: A real-time wearable system for monitoring and analyzing physiological signals.' paper presented at international workshop on wearable and implantable body sensor networks. *IEEE*, April 2006.

[71] J. Dhaliwal D. Koff J. A. Mackay L. Weise-Kelly D. W. Navarro N. L. Wilczynski P. S. Roshanov, J. J. You and R. B. Haynes. Can computerized clinical decision support systems improve practitioners' diagnostic test ordering behavior? a decision-maker-researcher partnership systematic review. *Implementation Science*, 6(88):1–12, August 2011.

[72] E. I. Papageorgiou and W. Froelich. Multi-step prediction of pulmonary infection with the use of evolutionary fuzzy cognitive maps. *Neurocomputing*, 92:28–35, September 2012.

[73] S. Park and S. Jayaraman. Enhancing the quality of life through wearable technology. *IEEE Engineering in Medicine and Biology Magazine*, 22(3):41–48, 2003.

[74] P. Perner. Machine learning and data mining in pattern recognition: 12th international conference. *Switzerland: Springer International Publishing*, 2016.

[75] K. A. Popat and P. Sharma. Wearable computer applications a future perspective. *International Journal of Engineering and Innovative Technology (IJEIT)*, 3(1):213–217, 2013.

[76] G. O' Regan. Introduction to the history of computing: A computing history primer. *International Publishing*, 2016.

[77] C. Richardson. *Microservices patterns*. Manning, 2018.

[78] A. Rooney. *The Story of Physics: From Natural Philosophy to the Enigma of Dark Matter (The Story of Series)*. Arcturus Publishing Limited, 2011.

[79] R. S. M. Soroushmehr F. Navidi D. A. Beard S. Belle, R. Thiagarajan and K. Najarian. Big data analytics in healthcare. *BioMed Research International*, 2015.

[80] M. Gannon S. D. Guler and K. Sicchio. *Crafting Wearables: Blending Technology with Fashion*. New York: Springer, 2016.

[81] Hansen S. Khanna, A. Sattar. Artificial intelligence in health the three big challenges. *Australasian Medical Journal*, 6(5), 2013.

[82] A. Liao E. J. Leeflang Y. Khan F. J. Pavinatto K. Mann A. Naujokas D. Young S. Roy M. R. Harrison A. C. Arias V. Subramanian S. L. Swisher, M. C. Lin and M. M. Maharbiz. Impedance sensing device enables early detection of pressure ulcers in vivo. *Nature Communications*, 6(6575), March 2015.

[83] T. Shevchenko. Most awesome examples of using artificial intelligence in medicine. *Australasian Medical Journal*, August 2016.

[84] L. Sinapova. Decision trees id3 algorithm.

[85] C. S. Stylios and V. C. Georgopoulos. Fuzzy cognitive maps for medical decision support – a paradigm from obstetrics. *the 32nd Annual International Conference of the IEEE EMBS*, pages 1174–1177, September 2010.

[86] R. Aschenbrenner T. Linz, C. Kallmayer and H. Reichl. Fully integrated ekg shirt based on embroidered electrical interconnections with conductive yarn and miniaturized flexible electronics. *IEEE*, April 2006.

[87] H. Maeda T. Fukushima T. Aida K. Hata T. Sekitani, H. Nakajima and T. Someya. Stretchable active-matrix organic light-emitting diode display using printable elastic conductors. *Nature Materials*, pages 494–499, May 2009.

[88] R. B. Thompson and B. F. Thompson. *Astronomy Hacks: Tips and Tools for Observing the Night Sky*. Sedastopol, CA: O'Reilly Media, Inc, June 2005.

[89] K. Tillmann. Understanding the connected health and wellness market. consumer electronics association, 2014.

[90] A. M. Turing. Computing machinery and intelligence. *Oxford University Press on behalf of the Mind Association*, 56(236):433–460, October 1950.

[91] P.Lukowicz G. Tröster F. Dolveck M. Baer F. Keita E. Schenker F. Catarsi L. Coluccini A. Belardinelli D. Shklarski M. Alon E. Hirt R. Schmid M. Vuskovic U. Anliker, J. A. Ward. A wearable multiparameter medical monitoring and alert system. *IEEE Transactions on Information Technology in Biomedicine*, 8(4):415–427, January 2004.

[92] University of Queensland. The University of Queensland Vital Signs Dataset. https://outbox.eait.uq.edu.au/uqdliu3/uqvitalsignsdataset/browse.html, 2018. Online; accessed 17 October 2018.

[93] G. A. Malandraki V. C. Georgopoulos and C. D. Stylios. A fuzzy cognitive map approach to differential diagnosis of specific language impairment. *Artificial Intelligence in Medicine*, 29(3):261–278, November 2003.

[94] G. A. Malandraki V. C. Georgopoulos and C. D. Stylios. A fuzzy cognitive map approach to differential diagnosis of specific language impairment. *Journal of Artificial Intelligence in Medicine*, 29(3):261–278, November 2003.

[95] R. Woolrych V. K. Mago, R. Mehta and I. E. Papageorgiou. Supporting meningitis diagnosis amongst infants and children through the use of fuzzy cognitive mapping. *BMC Medical Informatics and Decision Making*, 12(98), September 2012.

[96] C. L. Ventola. President obama's health care reform policies. *A Peer-Reviewed Journal for Managed Care and Hospital Formulatory Management*, 34(6), June 2009.

[97] A. S. Kanter W. Ogallo. Using natural language processing and network analysis to develop a conceptual framework for medication therapy management research. *AMIA Annual Symposium Proceedings Archive*, 2016(PMC5333323):984—-993, January 2016.

[98] J. Yong R. Kuang W. Zhang, J. Chien. Network-based machine learning and graph theory algorithms for precision oncology. *npj Precision Oncology*, 1(1):25, August 2017.

[99] K. Warwick. I, cyborg. *University of Illinois Press*, 2004.

[100] C. C. Y. Poon X. Teng, Y. Zhang and P. Bonato. Wearable medical systems for p-health. *IEEE Reviews in Biomedical Engineering*, 1:62–74, 2008.

[101] A. Barabási A. Sharma X. Zhou, J. Menche. Human symptoms-disease network. *Nature Communications*, 5, 2014.

[102] J. D. Elenee Argentinis Y. Chen and G. Weber. Ibm watson: How cognitive computing can be applied to big data challenges in life sciences research. *Clinical Therapeutics*, 38(4):688–701, April 2015.

[103] C. M. Lochner A. Pierre Y. Khan, A. E. Ostfeld and A. C. Arias. *Monitoring of Vital Signs with Flexible and Wearable Medical Devices*, volume 28. Advanced Materials, June 2016.

[104] O. Pauline Z. Zainuddin, L.K. Huong. Reliable epileptic seizure detection using an improved wavelet neural network. *Australasian Medical Journal*, 6(5):308–314, May 2013.

[105] P. J. Zimetbaum and M. E. Josephson. The evolving role of ambulatory arrhythmia monitoring in general clinical practice. *Annals of Internal Medicine*, 130(10):848–856, May 1999.

Appendix A

Neural Network Equations

Let n denote the number of layers in a neural network and m the number of nodes in a layer. The layer l is labeled as L_l while the input layer is labelled L_1 and the output layer as L_n, $l \in [1 \dots n]$. Let θ^l_{ij} denote the weight associated with the connection between node i in layer l, and node j in layer $l+1$ with nodes $i, j \in [1 \dots m]$. Each layer can have a different number of nodes M. Therefore each layer $L_l \in [L_1 \dots L_n]$ has $m_l \in [m_1, \dots m_n]$ number of nodes. The activation or output value is denoted by $a^{(l)}_i$, where the $i - th$ input i is the node number in the layer l and β_i is the bias of the weight. Consider that x_1 to x_m are input variables fed into the neural network and m is the number of input variables. For $l = 1$, $a^{(1)}_i = x_i$ can be used to denote the $i - th$ input.

The first layer's output for the neural network is a total weighted sum of inputs x_i to node i in layer l computed as follows:

$$h_{\theta,\beta}(x) = a^n_1 = \sigma(\theta^{(n-1)}_{11} x_1 + \theta^{(n-1)}_{12} x_2 + \dots + \theta^{(n-1)}_{1m} x_m + \beta^{(n-1)}_1) \quad \text{(A.1)}$$

The sum $\sum^m_{j=1} \theta_{i,j} a^l_j$ can be written in short form as $\theta^l a^l$ where σ is the sigmoid function represented by equation A.2.

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad \text{(A.2)}$$

Therefore, a general expression of the estimator at a layer is an output value $h_{\theta,\beta}(x)$ from the previous layer which is expressed by the equation A.3

$$h_{\theta,\beta}(x) = a^{(l+1)} = \sigma(\theta^l a^l + \beta^l) \quad \text{(A.3)}$$

Where the $\theta^l \in [\theta^1 \dots \theta^m]$ are weights, and x is the vector $[x_1 \dots x_m]$ of the input variables which are fed to the neural network and $a^l \in [a^1 \dots a^m]$ are input to a given node i in layer l while β^l is the bias at that layer. The equation A.3 can be explicitly defined by equation A.4.

$$a^{(l+1)}_i = \sigma(\sum^m_{j=1} \theta_{i,j} a^l_j + \beta^l) \quad \text{(A.4)}$$

Appendix B

Depth First Search

```
public IEnumerator<GraphNode<T>> DepthFirstSearch(Graph<T> graph, T start)
{
    HashSet<GraphNode<T>> visited = new HashSet<GraphNode<T>>();
    Stack<GraphNode<T>> stack = new Stack<GraphNode<T>>();

    GraphNode<T> startNode = (GraphNode<T>)graph.Find(start);
    stack.Push(startNode);

    while (stack.Count > 0)
    {
        GraphNode<T> current = stack.Pop();
        visited.Add(current);
        yield return current;
        foreach (GraphNode<T> neighbour in current.Neighbors.Reverse())
            stack.Push(neighbour);
    }
}
```

Appendix C

Breath First Search

```
public IEnumerator<GraphNode<T>>
BreathFirstSearch (Graph<T> graph, T start)
{
        Queue<GraphNode<T>> queue = new Queue<GraphNode<T>>();
        GraphNode<T> startNode = graph.Find(start);

        queue.Enqueue(startNode);
        while (queue.Count > 0)
        {
                GraphNode<T> current = queue.Dequeue();
                yield return current;
                foreach (GraphNode<T> neighbour in current.Neighbors)
                        queue.Enqueue(neighbour);
        }
}
```

Appendix D

Priority List

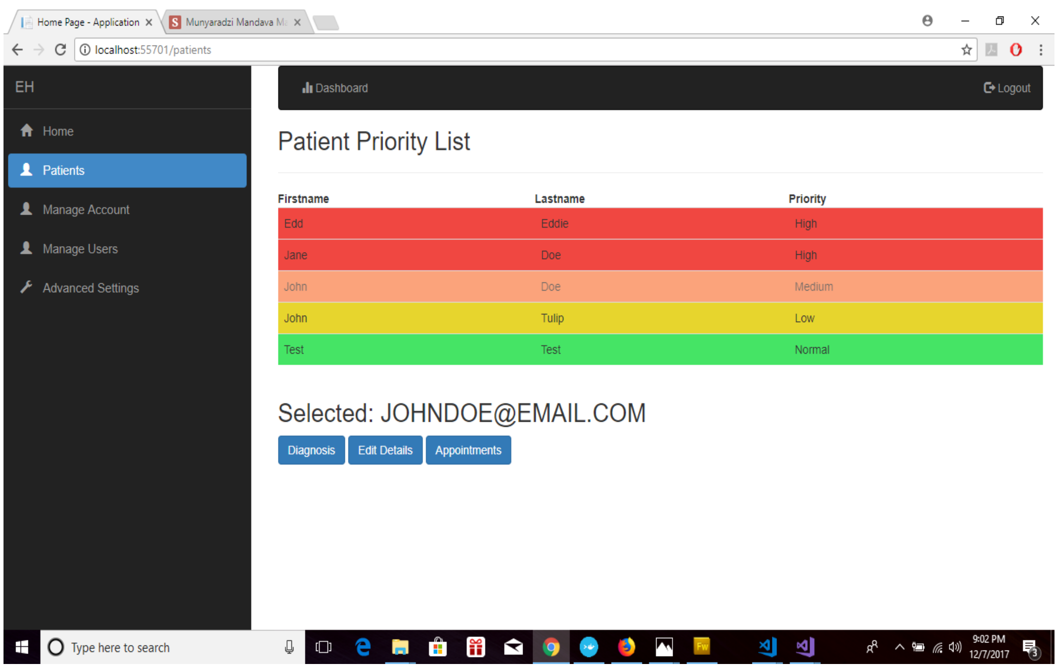

Figure D.1: Patient Priority List.

To view the diagnosis results of a patient, select a record and click the button diagnosis. This will open a dashboard with five visualizations mainly: the priority score gauge, overall health gauge, heart rate monitor and risk score monitor see Appendix E.

Appendix E

Dashboard

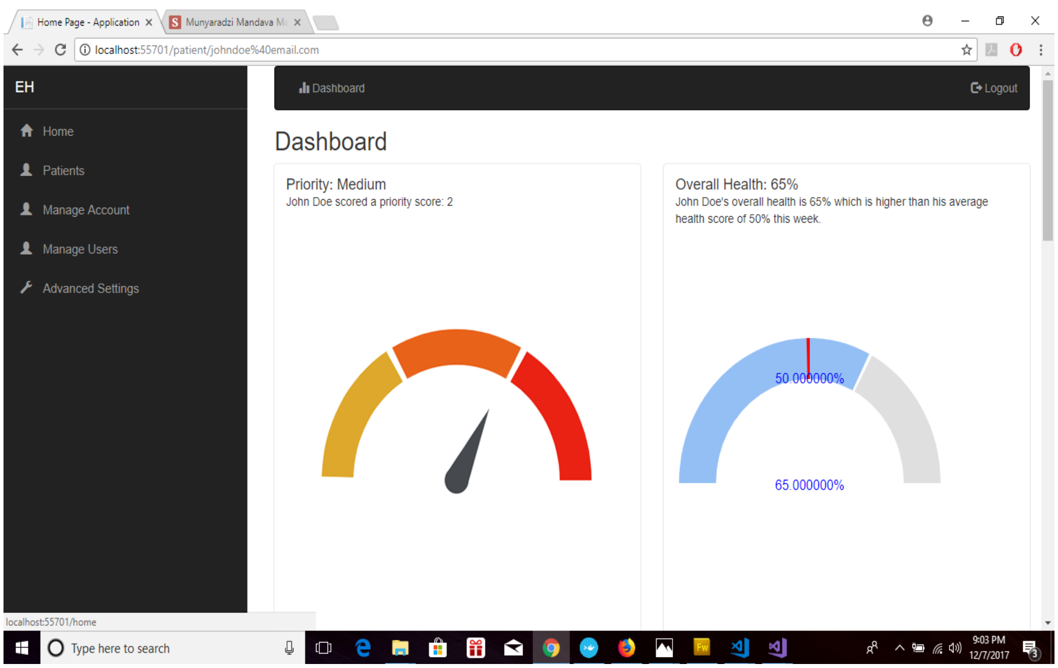

Figure E.1: Priority score and Overall health gauges.

The two gauges in Figure E.1 represent John Doe's overall's score from the patient priority list.

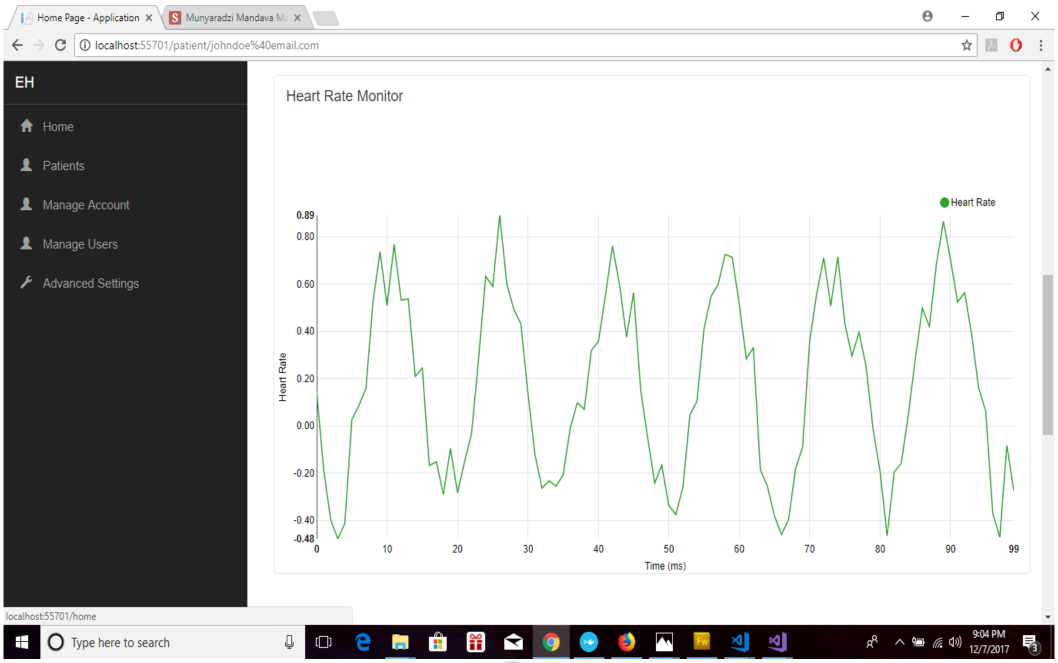

Figure E.2: Heart rate monitor.

The Figure E.2 represents John Doe's heart rate monitor. Real time monitoring is achieved using SingalR, angular 4 and D3 visualizations.

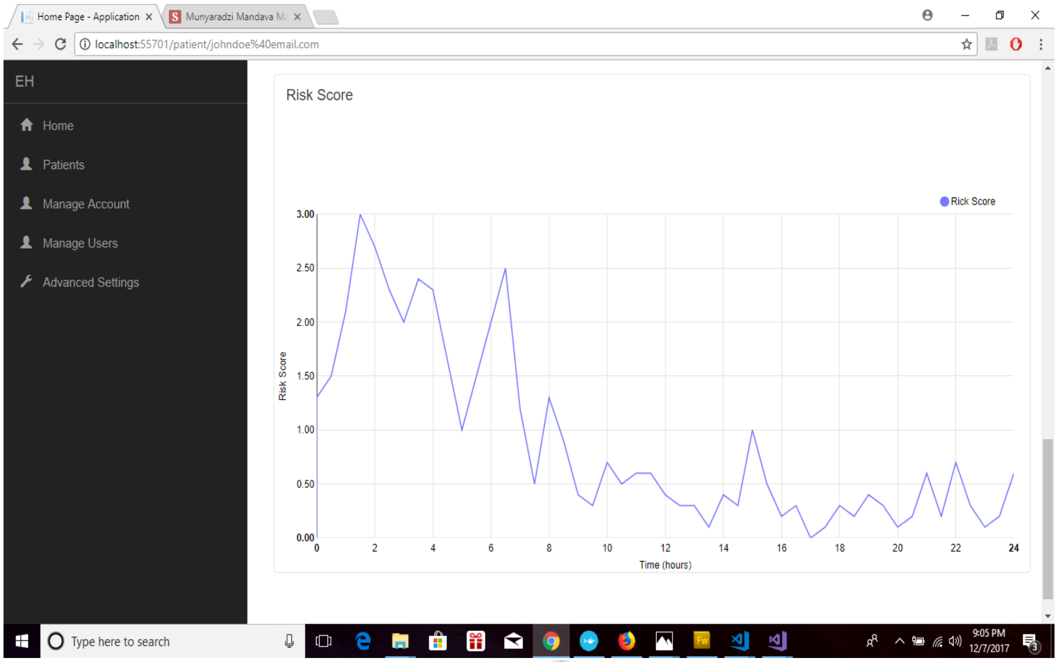

Figure E.3: Risk score monitor.

The Figure E.3 is a visualization showing John Doe's hourly risk score.

Appendix F

Advanced Settings and Visualizations

The Figure F.1 and Figure F.2 show the DNN optimization settings and visualizations.

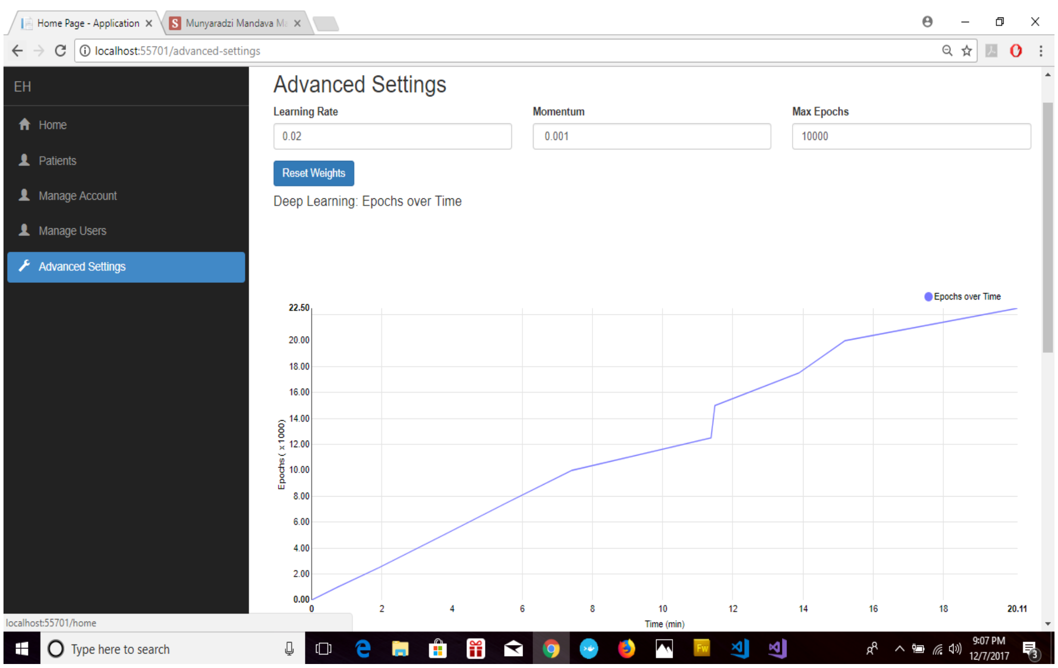

Figure F.1: DNN optimization settings.

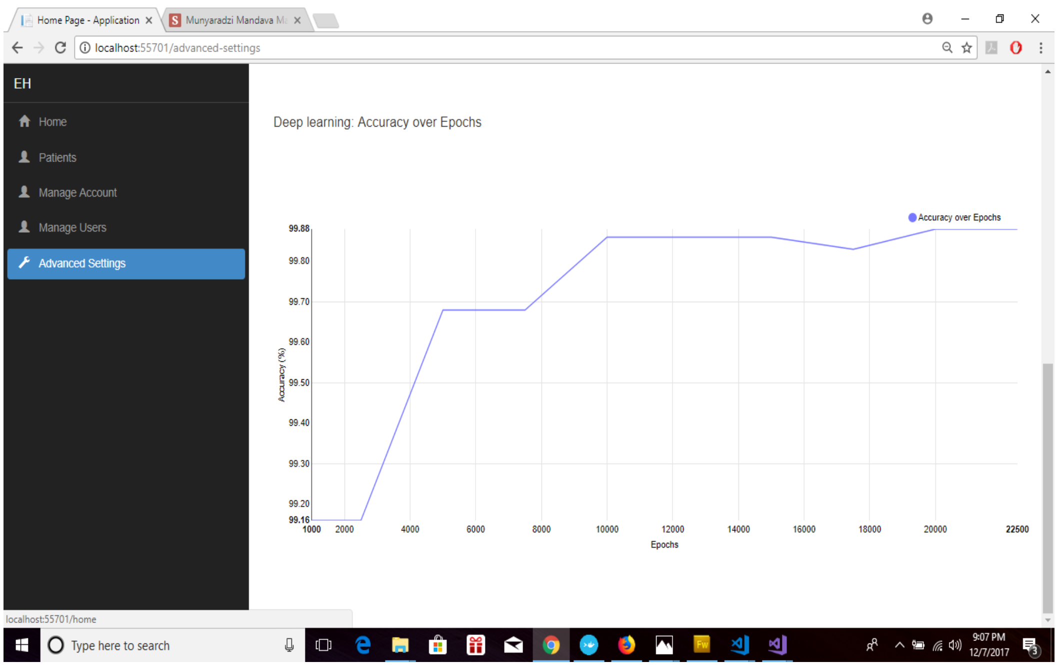

Figure F.2: DNN optimization settings.

Appendix G

Manage System Users

The Figure G.1, Figure G.2 and Figure G.3 illustrate role management. There are five possible roles: medical practitioner, medical trainee, developer, administrator and patient. Access to different sections of the system depend on the role assigned to the system user.

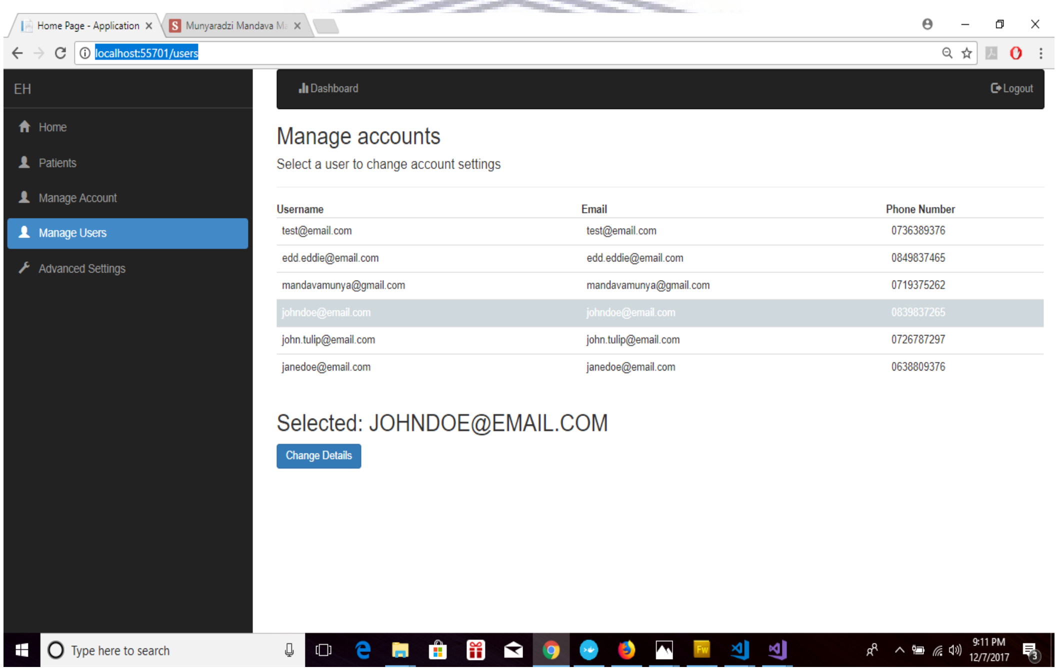

Figure G.1: View all users.

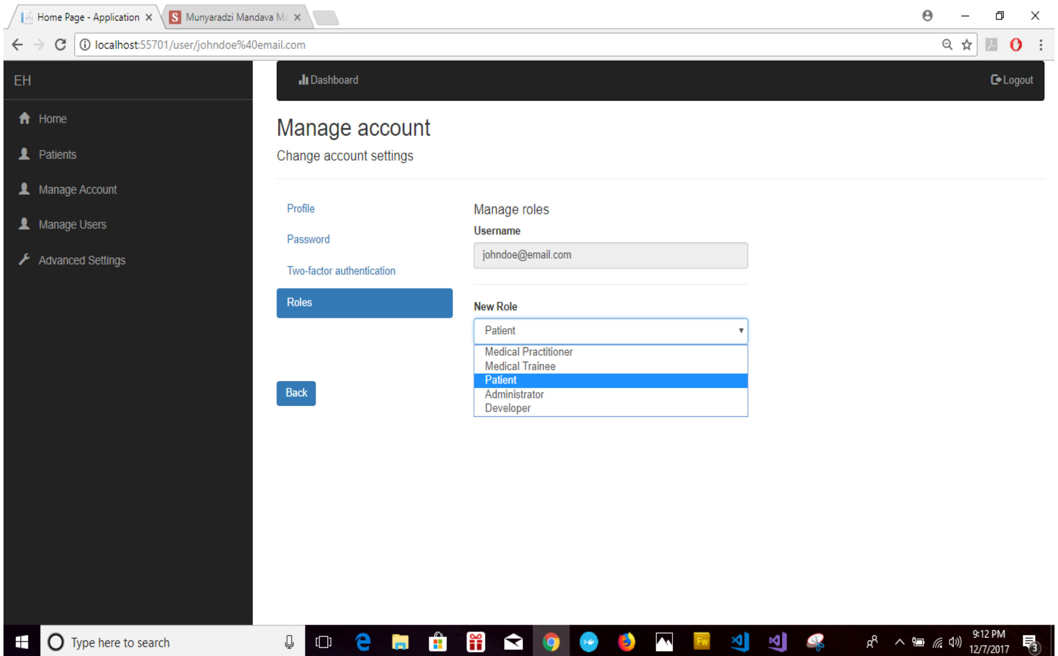

Figure G.2: Choose appropriate role and assign.

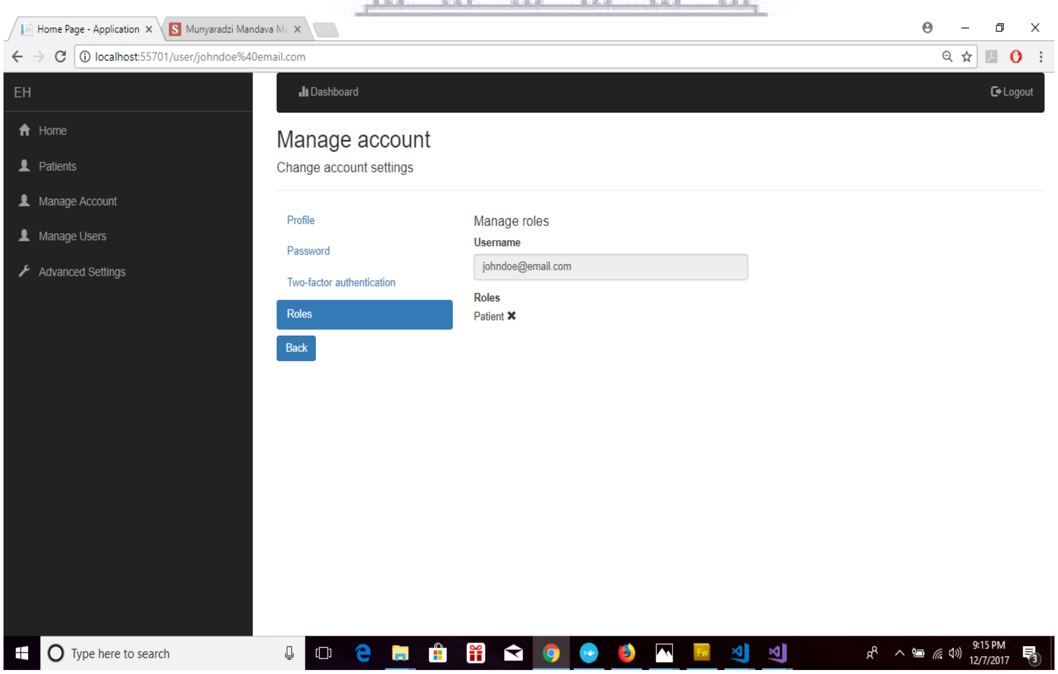

Figure G.3: Role assigned.

Appendix H

Personal Profile

The manage account page allows each system user to edit his or her own details.

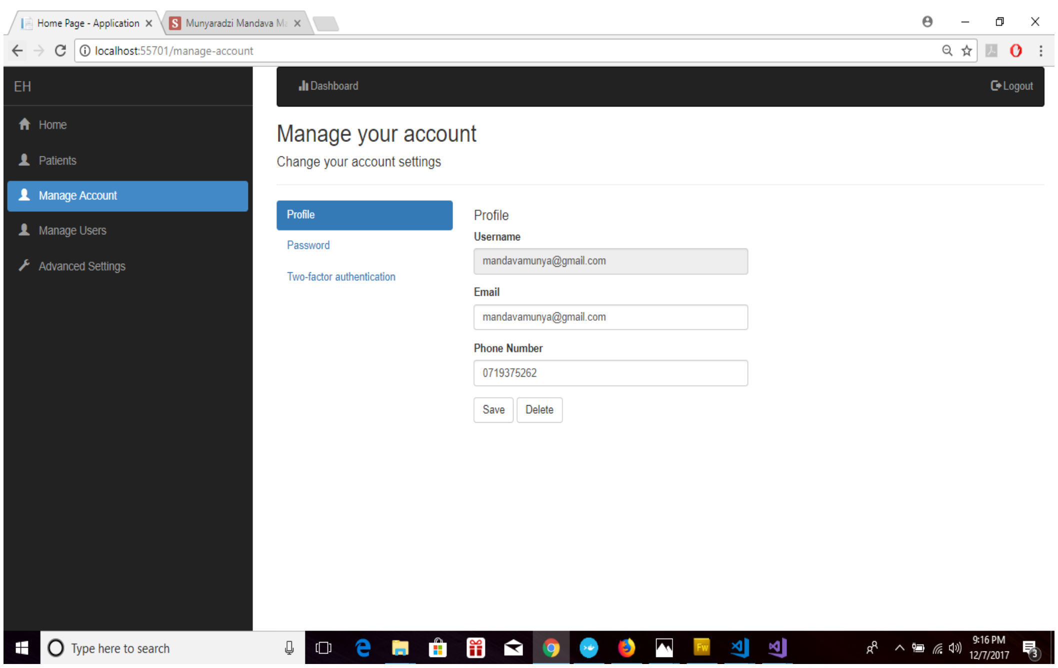

Figure H.1: Manage personal profile.