

WATER BORNE DISEASE DETECTION SYSTEM USING SUPPORT VECTOR MACHINES

**By:**

**Walter Odhiambo Nyangaga**

**Student No: 97730**

# DECLARATION

I declare that this project has not been submitted to any other university for the award of Bachelor of Science informatics degree

Admission Number:097730

Signature ……………….. Date…………………………

I certify that this work is being submitted for examination with my approval.

Supervisor name: Daniel Machanje

Signature ……………….. Date…………………………

# ABSTRACT

Water-borne diseases are diseases caused by the use of contaminated water or untreated water; this is because the contaminated water contains micro-organisms that cause these illnesses once ingested by humans, they are problem to the country, mainly due to the rapid rate of infection that occurs in cases of an outbreak. The diseases though having adverse conditions are treatable when detected early enough.

The aim of this project is to create a self-diagnosis system that allows patients to be able to diagnose the water borne disease that they are suffering from based on the symptoms that they are exhibiting. This self-diagnosis is meant to reduce outbreaks and spread of these diseases as the patients will be able to seek medical intervention early. The system classifies the symptoms exhibited by the patient based on a trained SVM model. It is meant to provide medical aid to Kenyans in areas that have frequent waterborne disease outbreaks and lack sufficient medical personnel to deal with them.

Contents

[DECLARATION ii](#_Toc53318075)

[ABSTRACT iii](#_Toc53318076)

[TABLE OF FIGURES vi](#_Toc53318077)

[CHAPTER 1: INTRODUCTION 1](#_Toc53318078)

[1.1 Background 1](#_Toc53318079)

[1.2 Problem Statement 2](#_Toc53318080)

[1.3 Objectives 2](#_Toc53318081)

[1.4 Specific Objectives 3](#_Toc53318082)

[1.5 Research Questions 3](#_Toc53318083)

[1.6 Justification of the Study 3](#_Toc53318084)

[1.7 Scope 4](#_Toc53318085)

[CHAPTER 2: LITERATURE REVIEW 5](#_Toc53318086)

[2.1 Introduction 5](#_Toc53318087)

[2.2 Water-Borne Diseases in Kenya 5](#_Toc53318088)

[2.3 Telemedicine 6](#_Toc53318089)

[2.4 Related Applications 6](#_Toc53318090)

[2.4.1 WebMD 6](#_Toc53318091)

[2.4.2 Web-Based System for Malaria Diagnosis 7](#_Toc53318092)

[2.4.3 A Predictive Symptoms-based Classification of Typhoid 9](#_Toc53318093)

[2.5 Gaps in Existing Systems 10](#_Toc53318094)

[2.6 Support Vector Machines 10](#_Toc53318095)

[2.7 Conceptual Framework 11](#_Toc53318096)

[CHAPTER 3: RESEARCH AND DEVELOPMENT METHODOLOGY 13](#_Toc53318097)

[3.1 Introduction 13](#_Toc53318098)

[3.2 Development Model 13](#_Toc53318099)

[3.2.1 Requirements Phase 14](#_Toc53318100)

[3.2.2 The Design Phase 14](#_Toc53318101)

[3.2.3 Implementation and Testing Phase 14](#_Toc53318102)

[3.3 Research Design 14](#_Toc53318103)

[3.3.1 System Architecture 14](#_Toc53318104)

[3.3.2 System Analysis 15](#_Toc53318105)

[3.3.3 System Design 15](#_Toc53318106)

[3.3.4 System Implementation 15](#_Toc53318107)

[3.3.5 System Testing 15](#_Toc53318108)

[3.4 Target Population 15](#_Toc53318109)

[3.5 Data Collection Procedure 16](#_Toc53318110)

[3.6 Data Analysis Procedure 16](#_Toc53318111)

[3.7 Deliverables 16](#_Toc53318112)

[3.7.1 The proposal 16](#_Toc53318113)

[3.7.2 The Application 16](#_Toc53318114)

[3.7.3 The System Documentation 16](#_Toc53318115)

[References 17](#_Toc53318116)

# TABLE OF FIGURES

[Figure 2.1: WebMD online symptoms Checker](#_Toc42762192) 7

[Figure 2.2: Web Based Malaria Diagnosis 8](#_Toc42762192)

[Figure 2.3: Typhoid Diagnosis System……………………………………………………………9](#_Toc42762192)

[Figure 2.4: Support Vector Machines 11](#_Toc42762192)

[Figure 2.5: Conceptual model of Water borne Disease detection 12](#_Toc42762192)

[Figure 3.1: An Iterative Model 13](#_Toc42762192)

# INTRODUCTION

## Background

Water-borne diseases are diseases caused using contaminated water or untreated water; this is because the contaminated water contains micro-organisms that cause these illnesses once ingested by humans (Waterborne Diseases, 2020). The water-borne diseases that have a high infectious rate in Kenya include Typhoid, diarrhea and hepatitis (Kenya Major infectious diseases - Demographics, 2020).

Many of the known water-borne diseases in the world largely affect children because of poor hygiene and weak immunity. Most of these diseases are life-threatening if not treated. The knowledge of the different types of water-borne diseases has increase in the past few years due to increased research.

Time and again, cases of spread of water borne disease in Kenya have been reported. This has been the case because access to safe quality drinking water is a problem for a lot of Kenyans, a recent article by the water project website sited Kenya as one of the countries in Africa being affected by water scarcity (Water In Crisis - Spotlight Kenya, 2020).

The situation has made these Kenyans without access to water have to use water that they find available, this puts them at risk as they end up using contaminated water leading to them being infected by these waterborne diseases.

A recent article by agencia revealed that close to half of all Kenyans are at risk of contracting water borne diseases. The article stated that people in Nairobi’s slum areas start the day by collecting water for use, this being a problem as stated in the earlier paragraph they do not have access to safe water for use and hence are forced to make do with what is readily available for them. It also stated that the level of hygiene and sanitation in these slums is low there by increasing the risk of contracting these waterborne diseases (Agencia, 2018).

Currently only hospitals are acquainted with the facilities to test for and treat the water borne diseases, this causes a problem as not all Kenyans have access to these facilities, and more to the point the people at a higher risk of being affected by these water borne diseases have less access to proper medical care.

The situation has led to rise in cases of outbreaks of water borne diseases due to their high infectious rate. There have been measures to try and counter these shortfalls. The Kenyan government has been trying to provide safe water for Kenyans and the introduction of mobile clinics that can diagnose and help treat some of these diseases.

Apart from these online solutions to detecting diseases such as WebMD an online medical system that can diagnose patient diseases based on their symptoms are also available. The problem with WebMD is that it is too generalized and does not try to solve the waterborne diseases problem.

The proposed system will take a targeted approach as it will be meant to target and identify these waterborne diseases and try to help reduce the number of outbreak cases that are reported in the country. This is meant to be done by training a model to identify cases of waterborne disease infection based on the symptoms that a patient manifest. The system is also expected to be more used in places that have had outbreaks of these water borne diseases there by helping to prevent future outbreaks by early detection.

## Problem Statement

A recent study by the Central Intelligence Agency has revealed a group infectious disease that affect a lot of people in our country. The diseases have been classified as having a high infection rate (Kenya Major infectious diseases - Demographics, 2020). It has cited these waterborne diseases as some of these diseases that have a high infectious rate.

With the recurrent cases of flooding that our country experiences, the waterborne diseases are a major problem as the flooding increases the risk of infection for the diseases with already high rates of infection. Because of this, there needs to be ways of coming up with quick and accurate diagnosis of these waterborne diseases, this will in turn help reduce spread of the diseases since with early diagnosis and intervention the diseases will not allowed to spread as is currently the situation with the outbreaks of these waterborne diseases.

## Objectives

To develop a system that can detect and diagnose Waterborne diseases based on the patient symptoms with a high degree of accuracy.

## Specific Objectives

1. To review the current waterborne disease detection system in the different parts of our country.
2. To identify the challenges with the current means of waterborne disease diagnosis and detection.
3. To design a waterborne disease detection system.
4. To develop a working waterborne disease detection system.
5. To test the water borne disease detection system.

## Research Questions

1. Are there problems with the current methods used to detect water borne disease?
2. What solutions are can be applied to solve the challenges with the current methods to detect water-borne diseases?
3. Which machine learning model can best be used for water borne disease detection?
4. What data is necessary to train a machine learning model for water-borne disease detection?
5. How will the water-borne detection system be tested?

## Justification of the Study

The study is important as tries to identify methods of early detection and diagnosis of waterborne diseases. The system is meant to provide ability for patients to self-diagnose accurately since the waterborne diseases sited are expected to be prevalent within the country.

With alternatives methods to diagnosis the waterborne diseases rather than having to go visit doctors which may not always be viable to the residents, they can be able to faster and more easily deal with cases of outbreaks due to early detection, early diagnosis and easy medication once the Waterborne diseases have been correctly diagnosed.

## Scope

The system should be able to diagnose if a patient has contracted one of the three water borne diseases that have been highlighted in the Kenya Major infectious diseases – Demographics. The disease are under the group highly infectious water borne diseases. They are Typhoid, Diarrhea and Hepatitis.

The waterborne disease group was chosen because of the countries recurrent state of flooding due to heavy rains (Kenya: Floods Flash Update, 2020). The heavy rains make the water borne disease more prevalent as most people are exposed to the water during the times of these floods.

The Geographic scope of the system will be limited to our country since the diseases that were chosen were sighted as being most prevalent in our country. The area around Ahero district will be the focus since the area is prevalent to flooding, these stagnant watery conditions aid in the outbreaks of water-borne diseases in the areas.

# LITERATURE REVIEW

## Introduction

The information to be reviewed includes documentations of previous similar projects but with a wider scope. The review also focuses on the successes and failures of the previous works on the topic and narrows down to the improvement areas of such a system, where potential for growth and bigger developments exists.

## Water-Borne Diseases in Kenya

Water-borne diseases are diseases that are caused by using contaminated water or untreated water. This contaminated water contains micro-organisms that cause these illnesses once ingested by humans (Waterborne Diseases, 2020).

Access to clean drinking water is not easy in Kenya. People who live in slum areas in Nairobi find it particularly difficult to access this commodity as it is not readily available for them, a study showed that for most people living in this areas their first activity of the day is to try and find water to use. The study revealed that only 56 percent of all Kenyans have access to clean water and that half the residents in Nairobi live in slum areas which have inadequate access to clean water and proper hygiene and sanitation facilities (Namale, 2020).

The study also revealed that close to 80 percent of people who go to hospitals have preventable diseases and out of these more than 50 percent have water-borne diseases. What is more worrying is that infants are more vulnerable, a study revealed that the situation in the country resulted in approximately 10,000 infant deaths annually (Namale, 2020).

The main Water borne diseases that affect Kenyans include Cholera, Typhoid, Diarrhea, Hepatitis, and Amoeba. Out of those common in Kenya the Kenya Major infectious diseases – Demographics cited cholera, Hepatitis and Diarrhea as having the highest infectious rate which is why they have been selected as the subject of this study (Kenya Major infectious diseases - Demographics, 2020).

In 2015 it was revealed that after the heavy rains and flooding, over 200 schools were forced to close due to an increase of water-borne disease infection. This was caused by the flood water mixing with poorly built toilets in the slum areas and causing contamination of the water and subsequently infecting school children (Nyamai and omulo, 2020).

## Telemedicine

Telehealth is the delivery of medical care over telecommunication means such as phones and the internet. Telehealth has been able to overcome challenges to treatment delivery caused by distance between patient and doctor which can be because of inadequate access to transportation or fragmentation of medical treatment due to gaps in time between appointments, and lack of available doctors (What are the Advantages of Telemedicine? | Doxy.me, 2020).

Telemedicine has a variety of advantages that include, increased access to care, distance traveled for patients to get access to medical care can limit access to medical care, since telemedicine makes this distance a non-factor, more and more people can now get access to medical care, this is particularly important for patients in communities without access to proper medical care (What are the Advantages of Telemedicine? | Doxy.me, 2020).

Telemedicine improves the type of medical care, reduced number of patients in hospitals and clinics, the resources at these places can be directed to provide care for the more critical cases that arise that need more specialized care (What are the Advantages of Telemedicine? | Doxy.me, 2020).

Telemedicine reduces healthcare costs, this is done by increase efficiency of patient treatment, reduce expenses for the treatment of the patients or transporting to other hospitals, and can even keep patients out of the hospitals completely. A study showed that telemedicine treatment had 19 percent savings on inpatient treatment cost (What are the Advantages of Telemedicine? | Doxy.me, 2020).

## Related Applications

### WebMD

Though the organization was initially and still is largely recognized as an online publisher of information pertaining to human health and well-being, the company developed and implemented a symptoms tracker on their website. The primary goal of the symptom tracker is to identify the most likely disease that a patient has contracted based on the symptoms that they presented on the website.

Their system tracker also works on other parameters such as location, age, and gender to be able to determine the likely disease that the patient is suffering from. This information is used to get an idea of the diseases likely affecting people in the area, it does not completely rely on the symptoms that the patient is suffering from (WebMD, 2020).

A further flaw with this system is that it does not narrow down your disease, when used it gives a long list of diseases that one might be suffering from. This gives a large margin for error as a patient will not be sure what diseases to get treated for due to the lack of specialization

(WebMD, 2020).

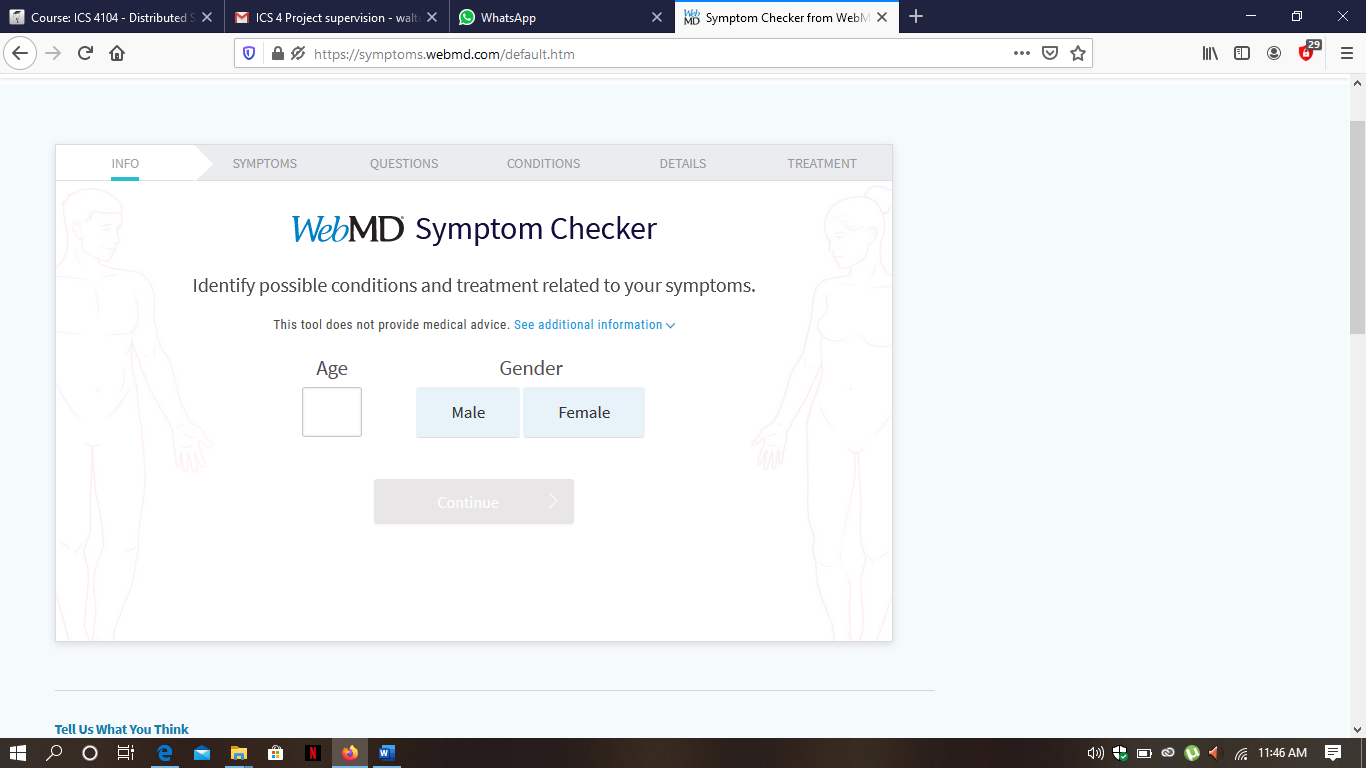


Figure 2.1: WebMD online Symptoms Checker (WebMD, 2020).

### Web-Based System for Malaria Diagnosis

Malaria is a diseases caused when a female anopheles mosquito bites an infected person and then subsequently bites someone healthy and in this way transmitting the malaria parasite from an infected person to a healthy person.

Malaria is widespread in the tropical and subtropical regions of the world as these provide suitable conditions for the survival of the mosquitos that are responsible for the spread of the parasite from one person to another(Adetunmbi, Oguntimilehin, Falaki, 2012.).

Malaria deaths are frequently due to late diagnosis and treatment of the disease. This is due to miss diagnosis of malaria symptoms. Another problem is that severe symptoms may sometimes be misdiagnosed as being malaria such as fever and septic shock which may lead to other more life-threatening illnesses not being treated and again also leading to fatalities (Adetunmbi, 2012).

Yearly there are between 300-500 million reported cases of infection of malaria with a fatality rate of between 1-3 million of those infected, with close to 90% of these fatalities occurring in sub-Saharan Africa (Adetunmbi, 2012).

Because of these worrying statistical figures and the adverse effects that malaria has on the people and the economy, a study was done with the goal of trying to develop a classification model for malaria diagnosis using rough set and develop a web-based medical diagnosis and treatment system for malaria based on classification model generated.

The system uses data collected from a hospital in Nigeria, the data mainly comprised of symptoms that were presented by the hospitals malaria patients that were recorded by an doctor at the hospital and those that the patients said they were experiencing. The data was collected over 2 separate periods and then they were grouped into 5 different categories based on the severity of the cases (Adetunmbi, 2012).

The rough set was used because it is a mathematical tool that can be used to overcome the insufficient data and knowledge to identify patterns in the dataset. It again helps with easy interpretation of collected data. The system uses LEM2 algorithm to classify the patient’s systems based on a set of classification rules.

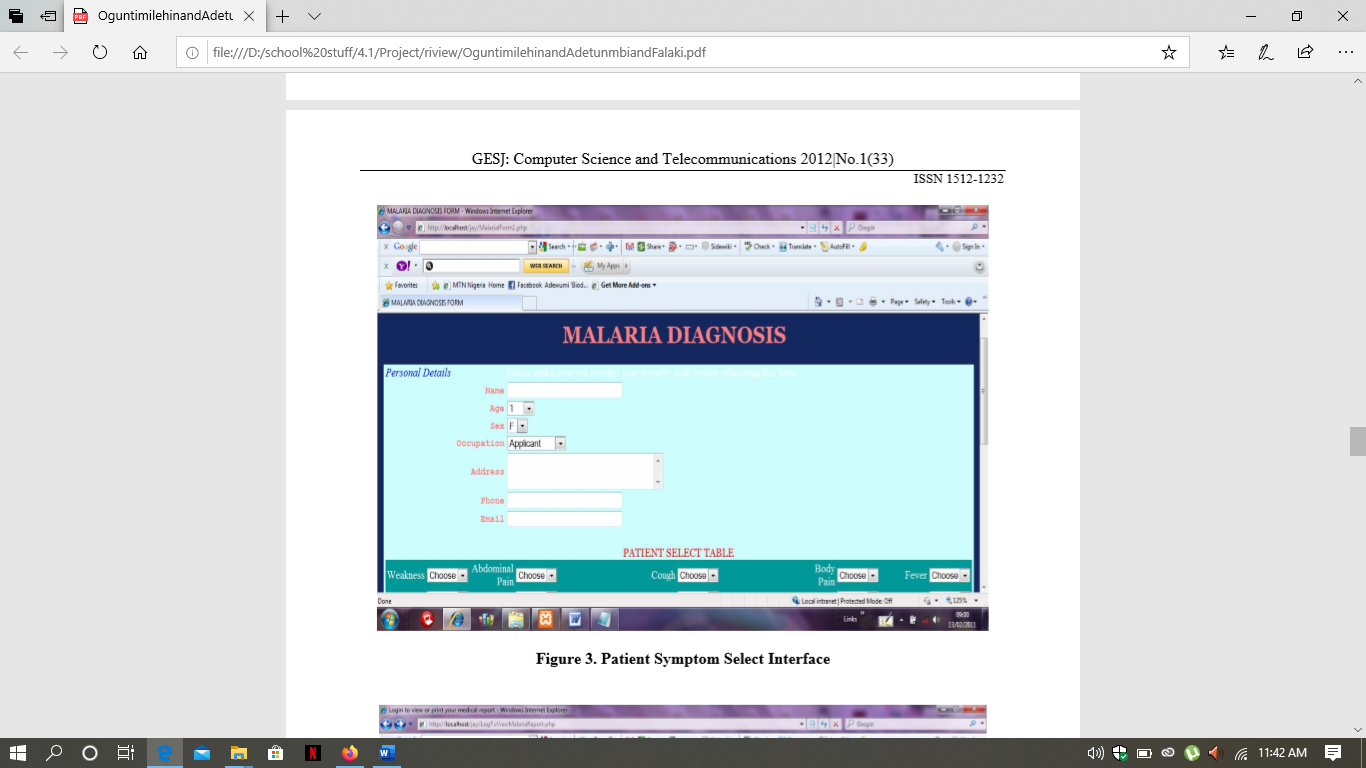


Figure 2.2: Web based Malaria diagnosis (Adetunmbi, 2012).

### A Predictive Symptoms-based Classification of Typhoid

This study was done to create a classifier system that accurately determines which disease a patient has between Malaria and Typhoid based on the symptoms that they are experiencing (Oguntimilehin, Adetunmbi and Abiola, 2013).

The data set was collected from a hospital, they include symptoms that were observed in patients that were suffering from these two diseases. These were then grouped and classified based on the severity of the cases of the diseases from each patient. The main aim of the system was to build on a pervious system and enhance the classificational accuracy between the two diseases as they had similar symptoms.

The System uses Support Vector Machine models. The SVM uses maximum margin hyperplane in identifying the Support Vectors (SV) which it uses to make classification decisions based on the symptoms that a patient has (Oguntimilehin, 2013).

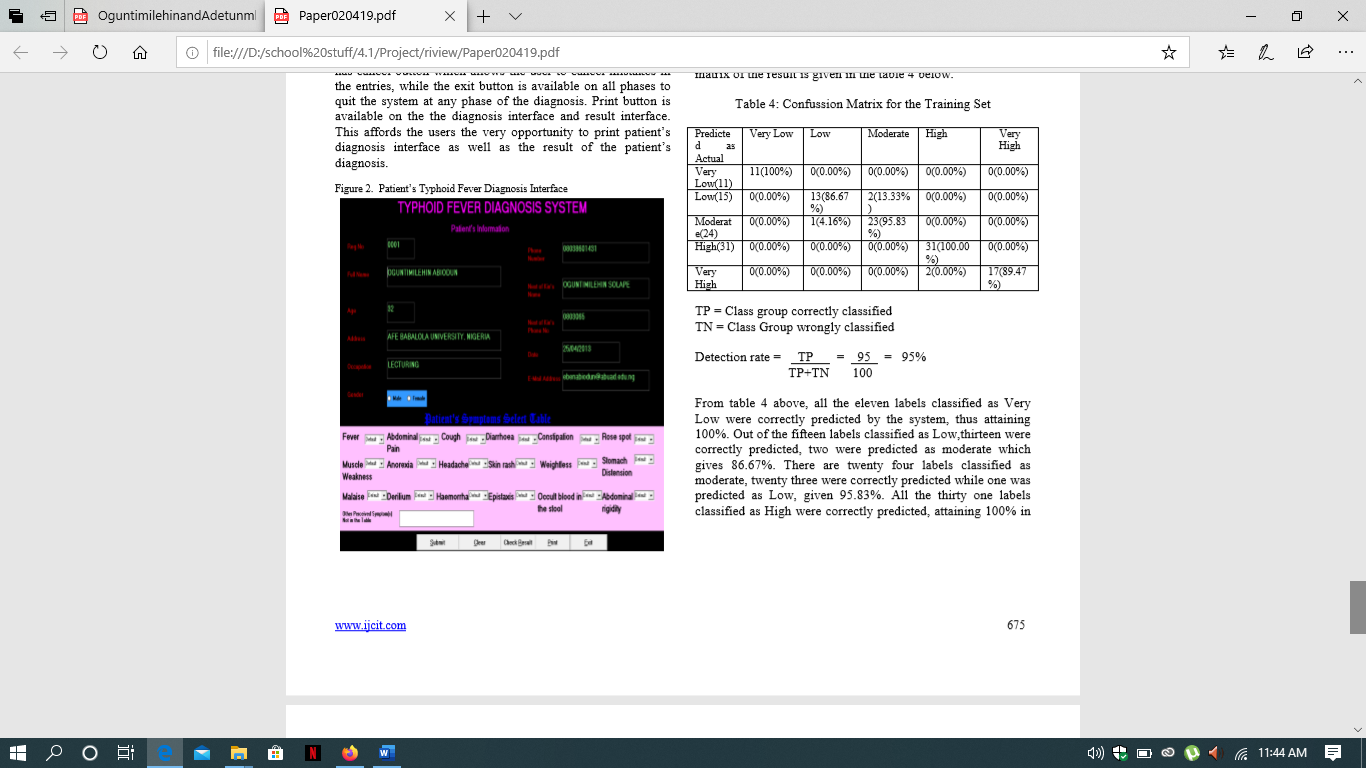


Figure 2.3: Typhoid Diagnosis System (Oguntimilehin, 2013).

## Gaps in Existing Systems

The current systems are all web based and so work on the assumption of easy internet access to all its patients. This problem can be overcome by implementing the system over other less internet intensive platforms such as WhatsApp or telegram since access to smart phones has become easier.

The current systems might also have problems as patients might not know what information to give about the symptoms that they are experiencing, to solve this I believe having the system implemented over a chatbot will allow the systems to get better structured information to be able to more accurately classify the waterborne diseases.

The System can be improved by providing a better interface for use as the current interface may confuse patients and thus, they may not be able to get the best out of the system. This also can be improved by implementing the systems using a chatbot that will be able to guide the patients as they give their symptoms and hence providing them with a more interactive experience.

## Support Vector Machines

The new system will use SVMs to classify the disease that a patient has based on the symptoms that they are exhibiting.

SVMs work by identifying a hyperplane that is used to separate the different classes, the hyper plane is like a line with the 2 different classes falling on either side of it. The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the number of features) that distinctly classifies the data points (Gandhi, 2020).

once the dataset is entered into system, the model will decide to which class the data belongs. The plane works by identifying the maximum distance of margin between the two data classes. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line. If the number of input features is 3, then the hyperplane becomes a two-dimensional plane (Gandhi, 2020).

Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, we maximize the margin of the classifier (Gandhi, 2020).

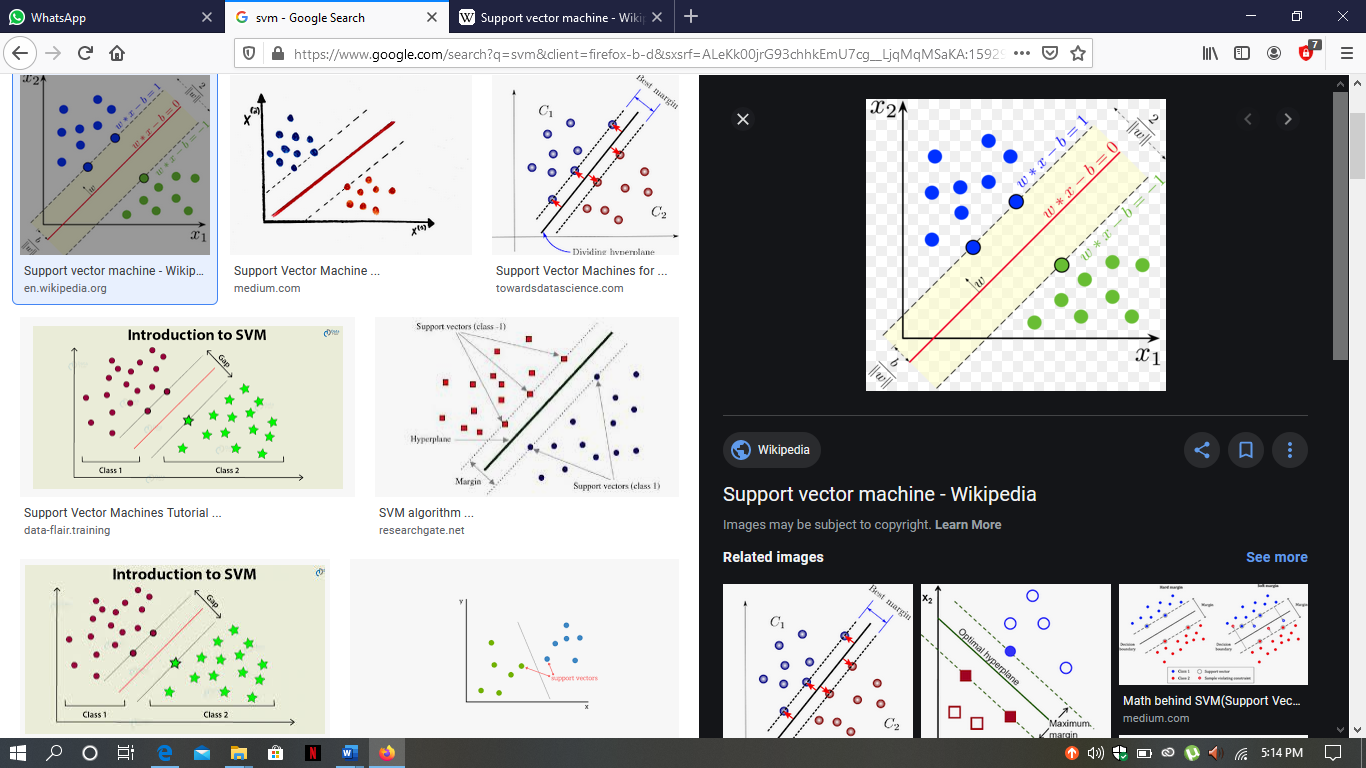


Figure 2.4: Representation of a Support Vector Machine (Gandhi, 2020).

## Conceptual Framework

The conceptual design tries to explain the interaction between most of the components of the waterborne disease diagnosis system.

The system works by a patient imputing the symptoms that they are experiencing, and the system identifies the waterborne disease that they are suffering from based on the classifier model.

The classifier model to be used will be Support vector machines, these will be trained to classify the water borne diseases. The system model will be trained to classify the 3 of water borne diseases, once the SVM is trained it will create hyperplanes to distinguish between them, on the that is done the 3 classes will be the 3 waterborne diseases.

To identify which waterborne disease a patient is suffering from, the patients symptoms will be input into the system model, based on the classes that the model will have created during model training. Once the symptoms are input, the SVM model will work at matching the symptoms presented to the disease class that it belongs to.

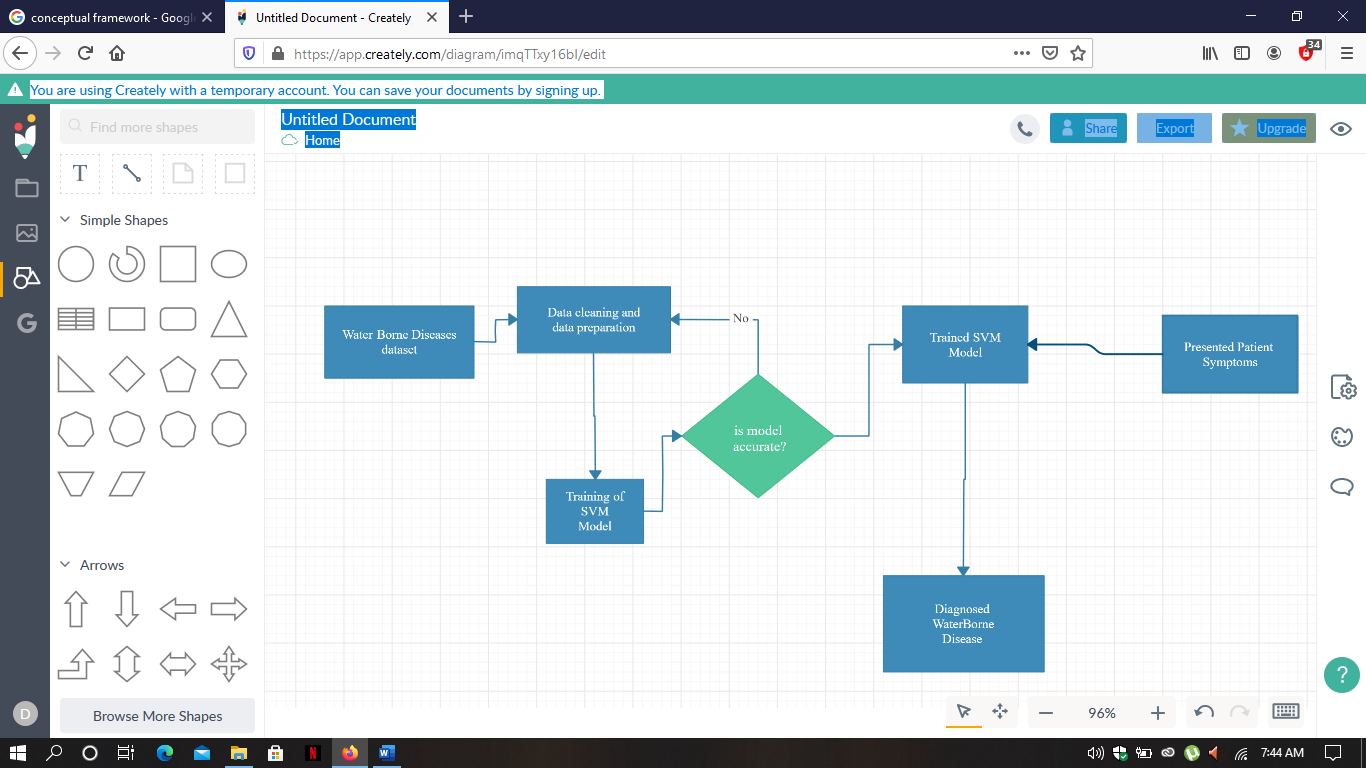


Figure 2.5: Conceptual model of Water borne disease diagnosis system

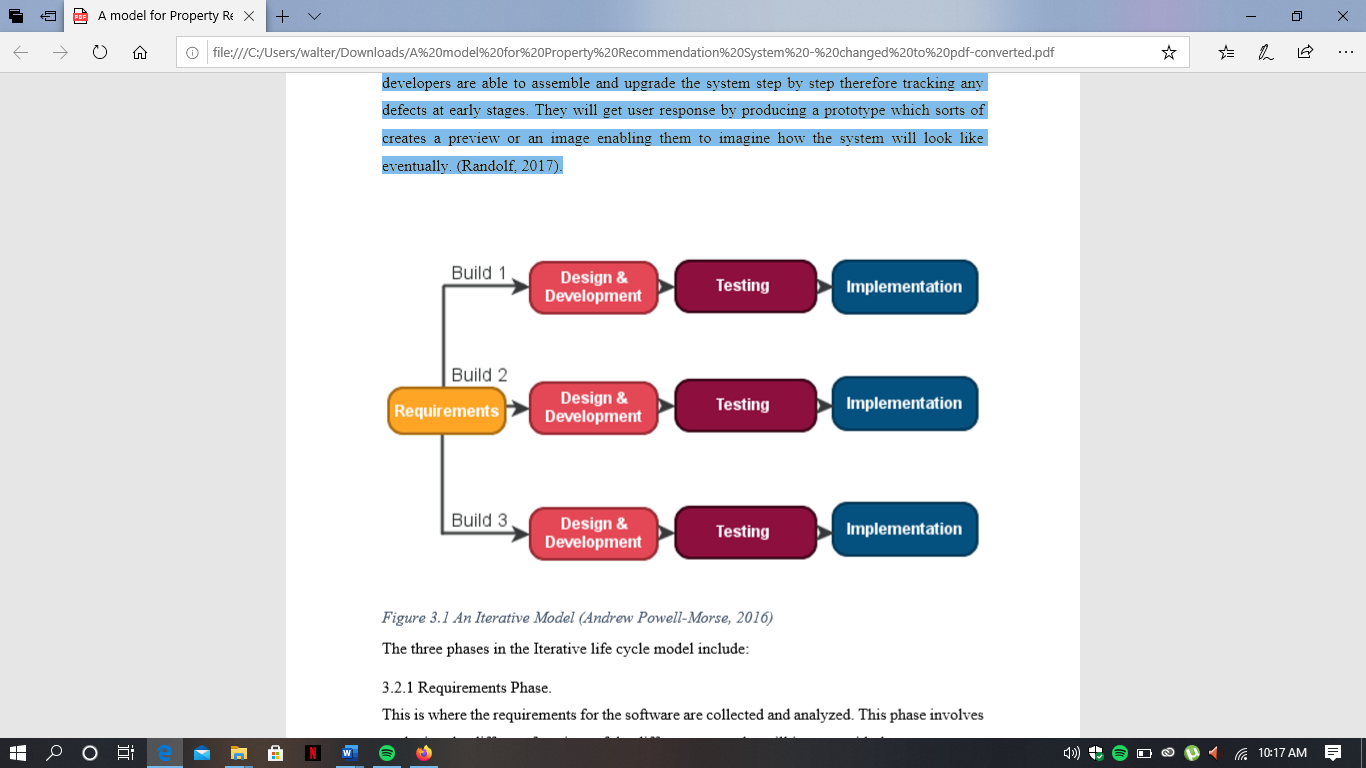
# RESEARCH AND DEVELOPMENT METHODOLOGY

## Introduction

This chapter describes some of the methods used for conducting the research and viability. The chapter then analyses the methodologies to be used in a system architecture, system analysis, system design, system development, system implementation, system testing and evaluation of the system.

## Development Model

A system development methodology is essential in structuring and formulating the development process of an information system (Kashif, 2014). For an iterative model, development commences by identifying the software requirements early and then the developers are able to assemble and upgrade the system step by step therefore tracking any defects at early stages. They will get user response by producing a prototype which sorts of creates a preview or an image enabling them to imagine how the system will look like eventually (Randolf, 2017).



**Figure 3.1 An Iterative Model (Powell-Morse, 2016)**

The three phases in the Iterative life cycle model include:

### Requirements Phase

This is where the requirements for the software are collected and analyzed. This phase involves analyzing the different functions of the different users who will interact with the system.

### The Design Phase

In the design phase, a software solution to meet the requirements is designed. This involves designing a friendly user interface where the users can navigate through the system easily.

### Implementation and Testing Phase

In the Implementation and Testing phase, the software is coded, integrated, and tested. The major functionalities that will allow the users to interact with the system will are developed here.

## Research Design

The research will use a quantitative and a qualitative research methods where qualitative research will aim at gaining deep understanding of specific organizations or events and how its deriver meaning from the surroundings and how the meaning influences the behaviour. This will be used to enhance good processing factor of the beverage ordering system. Quantitative research on the other hand will be used to check on the numbers of the new customers who would like to use the new system or even thought that it was a perfect idea (Anon., 2012).

### System Architecture

The system will be developed using support vector machines that will be implemented over an online chatbot. The support vector Machines will be trained to classify the water borne diseases based on symptoms that a patient presents.

The patient symptoms will be recorded and requested by the chatbot, the chatbot will ask for the symptoms in a structured manner only asking for the symptoms that are most necessary to assist in the classification of the patient diseases, this will help so as to ensure that the patients only give the most important information required for the model to be able to classify the water borne disease that the patient is suffering from.

### System Analysis

System analysis which consists of three approaches (process-oriented, object-oriented and data-oriented approaches) will be able to study the procedures in order to identify the purposes and create the system that will be achieved in an efficient manner.

Data-Oriented approach will be mainly used in this system analysis as we try to understand the characteristics and symptoms of the different water borne diseases.

### System Design

Approach to be used based on methodology chosen and their purposes

Data preparation techniques such as Feature selection will be used to identify the symptoms of a waterborne disease that are most significant in identifying whether a patient has the waterborne disease or not.

We will also perform other data preparation features such as data cleaning to ensure that we work with complete data.

### System Implementation

The system will be implemented over hospital websites or as a website on its own so that people living in areas with increased levels of flooding and places with high levels of waterborne disease infection have the chance to check their symptoms and Identify if they are suffering from a waterborne disease.

### System Testing

Usability as a test will be done to test the system functionality where it will entail; interactive element validation on the screen (buttons, texts, input), validation navigation, responsiveness and user-friendliness and ease of navigation.

## Target Population

The target population will be the residents of Ahero district which is found in Nyanza province in Kenya. The area is prone to flooding and the stagnant waters lead to outbreaks of the waterborne diseases that we are trying to predict.

## Data Collection Procedure

The data collection methods which will be used include questionnaire, and interview where interview will be prepared to ask doctors and medical practitioners the most significant symptoms that the patients will be exhibiting when it comes to these three main waterborne diseases. Data from the online repository Kaggle will be used as the dataset for training the model on the different waterborne diseases, for unavailable data, synthetic data will be created with the use of the python library faker. These will be generated from information presented by the interviews that were carried out.

## Data Analysis Procedure

Data preparation methods will be used on the data that will be obtained from Kaggle, this will be done to fill in for missing values, after that the categorical data will also be converted to numerical data to aid in the mathematical procedures involved in Support vector Machine Training. Once the data has been prepared, the models will use feature selection together with the information gathered during the interviews with the medical practitioners. This will be done in order to find the most important symptoms that are necessary for classification of the waterborne disease by our SVM model.

## Deliverables

Upon completion of the systems we expect to have the following

### The proposal

This will be a brief documentation of the proposed system that will highlight the need for the system and its validity.

### The Application

This will be the system itself, we expect to have a support vector machine model cable of diagnosing a waterborne disease based on the symptoms that a patient is exhibiting. The SVM model will be connected to a chatbot that will be used to gather patient symptoms

### The System Documentation

This will be a document that will show the design and development procedure for the system to be used by future developers incase improvements want to be made on the system.

# SYSTEM ANALYSIS AND DESIGN

## Introduction

System analysis is the process of collecting and interpreting facts, identifying the problem at hand and dividing the system into components. It is a problem solving technique that not only improves the specific system but also checks the functionality of the system components. System design is the process of planning or replacing an existing system by identifying the components to satisfy the requirements where the major focus is on how the system needs to accomplish its objective analysis. System architecture is the conceptual model that defines the view of the system, its behavior and structure and often organized in a manner that supports the reasoning about the structure and the behavior of a system.

## Requirement gathering

The data was collected by use of a questionnaire; we were able to get interviews with medical practitioners to get the most significant symptoms for the waterborne diseases that the models were trained for. Patient symptom data for hepatitis was gotten from Kaggle database, for typhoid and diarrhea synthetic data was created using faker and this combined data will be used for data training.

## System Requirement

This is basically the components of the computer system that is the hardware and software part which ensures that the system required to install and use the software/hardware are efficient and in this case our focus was on the functional and non-functional requirements.

### Functional Requirements

Functional requirement for the project will enable the system to be able to allow users to access the chatbot using telegram, the chatbot will enable the users to input their symptoms and from there the symptoms will be classified based on the SVM models that have been trained login upon registration where the customers upon registration will place an order of what beverages they may be interested in. The system will have a system administration.

|  |  |
| --- | --- |
| ID | DESCRIPTION |
| FRQ1 | The system should be able to provide an automated chatbot that explains the functioning of the system. |
| FRQ2 | The system should be able to provide the patients with the ability to intput their symptoms . |
| FRQ3 | The system should be able to classify the patient symptoms and Identify whether they have a waterborne disease or not. |

**Table 4.1 Functional Requirements**

### Non-functional Requirements

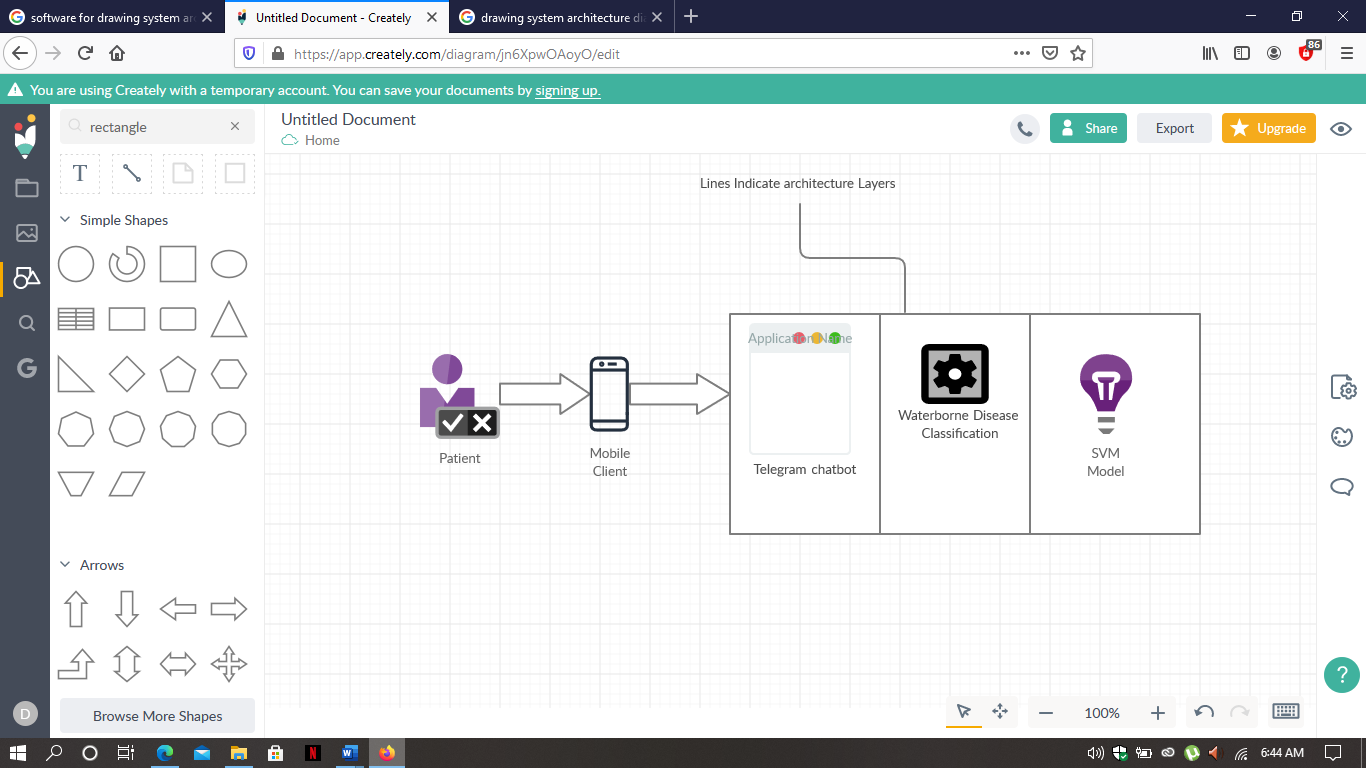
The system will allow for easy access to all users who have access to a telegram account the requirements will include; access to devices that can access the internet like tablets, laptops, smartphones that will enable easy and convenient access to the system. This will enable a fast and easy transfer of information from to the model.

|  |  |  |
| --- | --- | --- |
| ID | CATEGORY | DESCRIPTIONS |
| NFRQ1 | Usability | The system will only be used by any person that has a telegram account. |
| NFRQ2 | Reliability | The system should be able to allow any person to access it as long as they have a valid telegram account that can be used to access the system. |
| NFRQ3 | Performance | The system should be able to accommodate more than one user concurrently. |
| NFRQ5 | Supportability | The system should be able to backup data on the database. In case of a system upgrade the system should be compatible with advancing environment. |

**Table 4.2 Non-functional Requirements**

## System Architecture

The diagram show the interaction between the system and the module where the independent variable was manipulated for purposes of obtaining the dependent variable.

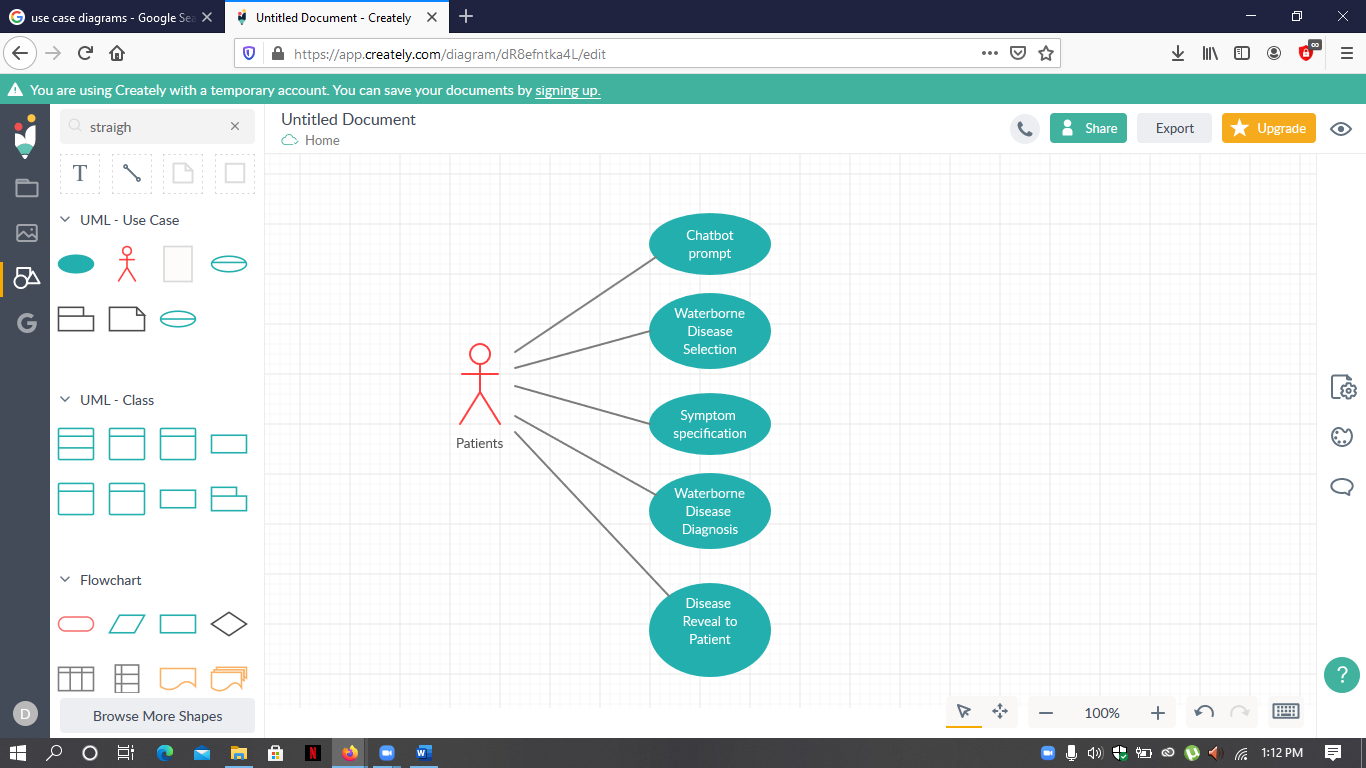


**Figure 4.1 System Architecture**

## System Design

It entails diagrams that represents the design from an interaction point of view of the system component in the online beverage ordering system which is dependant with the methodology adopted which in this case is the Object-Oriented Analysis and Design (OOAD) which deals with the objects that interact with one another to achieve a common goal. The Iterative Methodology which was used delivers the diagnosis of the patients and in a shorter time and hence promote progress that is sustainable for development.

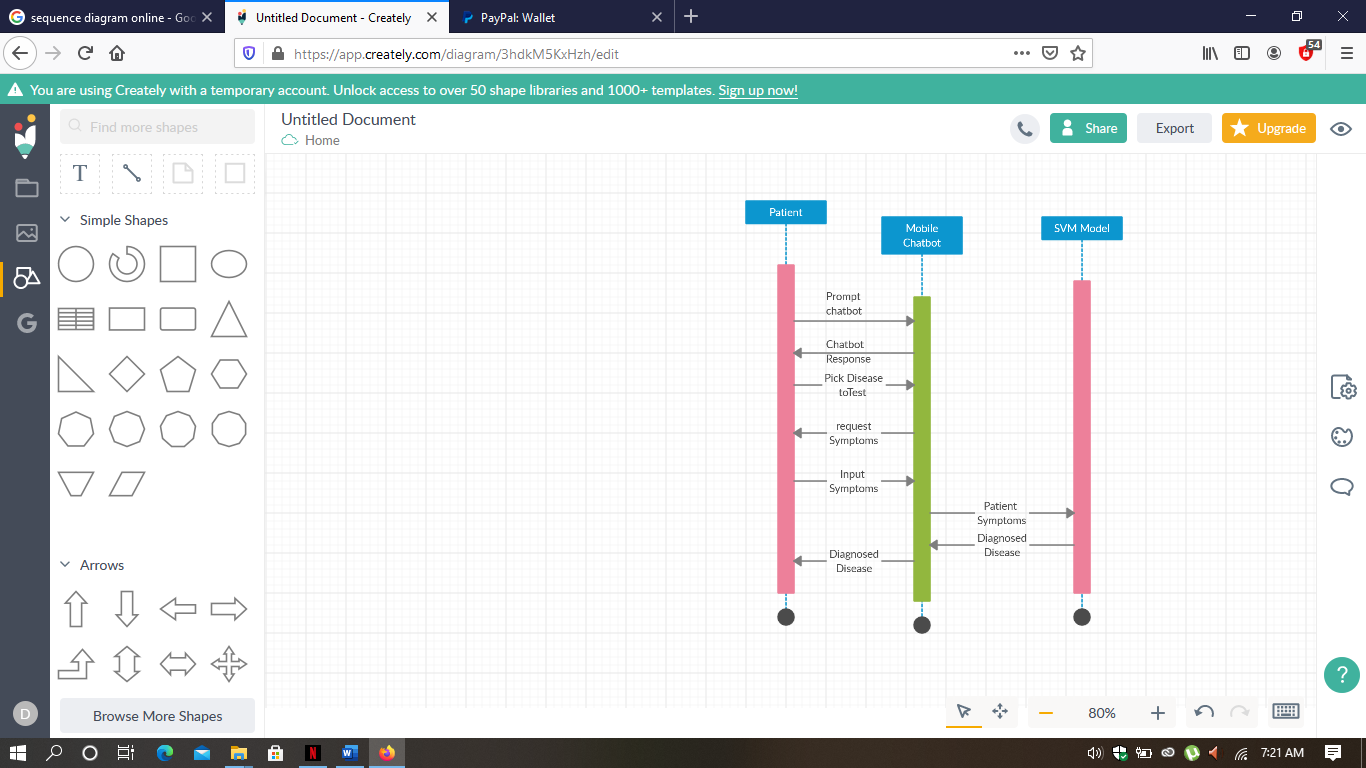
### Use Case Diagram



**Figure 4.2 Use Case Diagram**

In the diagram above, there is a representation of user interaction with the system that then shows relationship between users and different use-case in which the user is involved. We have the patient as the users where the patient is able to prompt the chatbot, select the water borne disease that they want to test for, specify the symptoms that they are exhibiting, send the information to the model for diagnosis and finally receive diagnosis from the SVM models.

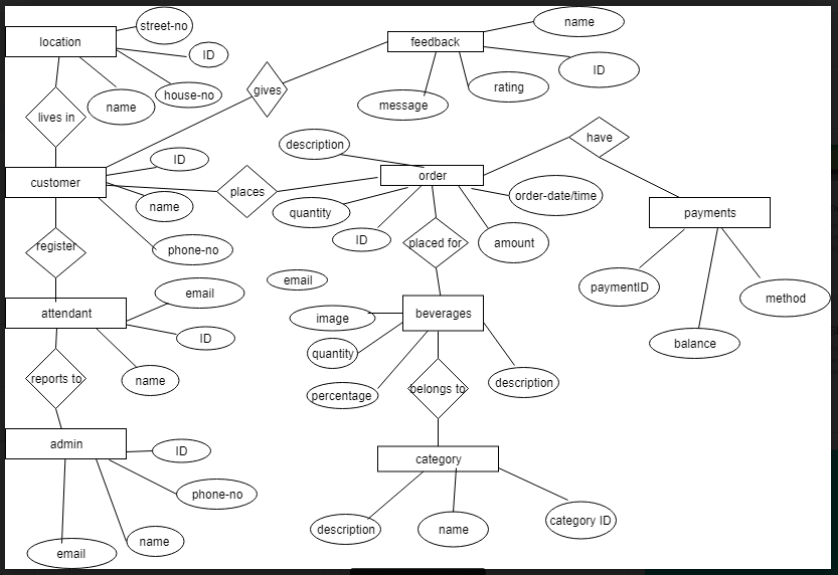
### Sequence Diagram



**Figure 4.3 Sequence Diagram**

In the diagram above, there is an interaction between objects in a sequential manner. The patient prompts the chatbot and begins the session. The chatbot responds and requests the patient of the waterborne disease that they want to be diagnosed for, the patient then gives a suitable response among the choices provided by the chatbot. After the disease is selected the chatbot will request the patient to input the symptoms they are exhibiting. The symptoms will then be sent to the SVM model for classification, once the symptoms have been classified a response will be sent back to the client from the SVM model and finally the mobile client will represent the SVM diagnosis to the patient.

### ER-Diagram



Figure

The above diagram is a high-level conceptual data-model that is based on the notion of real-world entity and relationship in them. It basically displays relationship of entities that are stored in a database. With different entities from the class diagram, the ER-diagram samples the attributes of the entities that is a conceptual data-model stored in a database that enhances relationship to enhance the functionality of the beverage system.

# References

Adetunmbi A.O, Oguntimilehin A, Falaki S.O, 2012. Web-Based Medical Assistant System for Malaria Diagnosis and Therapy [Accessed 10 June 2020].

Doxy. 2020. What Are the Advantages Of Telemedicine? | Doxy.Me. [online] Available at: <https://doxy.me/advantages-of-telemedicine> [Accessed 19 June 2020].

Edition, E., news, 2020. Almost Half of Kenyans At Risk of Water-Related Diseases. [online] www.efe.com. Available at: <https://www.efe.com/efe/english/life/almost-half-of-kenyans-at-risk-water-related-diseases/50000263-3561306> [Accessed 20 June 2020].

Ecatalog.calstatela.edu. 2020. California State University, Los Angeles - Acalog ACMS™. [online] Available at: <http://ecatalog.calstatela.edu> [Accessed 31 August 2020].

Gandhi, R., 2020. Support Vector Machine — Introduction to Machine Learning Algorithms. [online] toward data science. Available at: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47> [Accessed 22 June 2020].

Indexmundi.com. 2020. Kenya Major Infectious Diseases - Demographics. [online] Available at: https://www.indexmundi.com/kenya/major\_infectious\_diseases.html [Accessed 7 June 2020].

Mayo Clinic -Patient care and health Info. [online] Available at: < <https://www.mayoclinic.org//>> [Accessed 4 June 2020].

National Institute of Environmental Health Sciences. 2020. Waterborne Diseases. [online] Available at: <https://www.niehs.nih.gov/research/programs/geh/climatechange/health_impacts/waterborne_diseases/index.cfm> [Accessed 19 June 2020].

Namale, D., 2020. KENYA: Water And Sanitation Ailments Kill 10 000 Infants Annually. [online] Mtaani Insight. Available at: <https://mtaaniinsight.wordpress.com/2014/02/21/kenya-water-and-sanitation-ailments-kill-10-000-infants-annually/> [Accessed 14 June 2020].

Nyamai, F. and omulo, c., 2020. Over 200 Kenyan Schools Shut As Water-Borne Diseases Rise. [online] Daily Nation. Available at: <https://www.nation.co.ke/news/Over-200-schools-shut-as-water-borne-diseases-rise/1056-4571930-r6ipq6/index.html> [Accessed 14 June 2020].

Oguntimilehin, A., Adetunmbi, A. and Abiola, O., 2013. A Machine Learning Approach to Clinical Diagnosis Of Typhoid Fever [Accessed 10 June 2020].

ReliefWeb. 2020. Kenya: Floods Flash Update No. 1 (7 May 2020) - Kenya. [online] Available at: <https://reliefweb.int/report/kenya/kenya-floods-flash-update-no-1-7-may-2020> [Accessed 10 June 2020].

The Water Project. 2020. Water in Crisis - Spotlight Kenya. [online] Available at: <https://thewaterproject.org/water-crisis/water-in-crisis-kenya> [Accessed 1 June 2020].

WebMD - Better information. Better health. [online] Available at: < <https://www.webmd.com/>> [Accessed 4 June 2020].