

## Lassa Fever Predictive Model Using Machine Learning Techniques

**BAKARI, Shehu**

Computer Science Department  
Modibbo Adama University, Yola, Adamawa State, Nigeria  
bakarishehu90@gmail.com

**Dr. Yusuf Musa Malgwi**

Computer Science Department  
Modibbo Adama University, Yola, Adamawa State, Nigeria  
yumalgwi@mau.edu.ng

**Katyo Peter Abu**

Computer Science Department, University of Benin  
E-mail: kayo.peter-abu@customs.gov.ng

**Mohammed, Usman**

Computer Science Department, Federal Polytechnic Bali  
E-mail: shaggyrancy@gmail.com  
DOI: 10.56201/ijcsmt.v10.no4.2024.pg145.173

---

### **Abstract**

*Lassa Fever, a severe hemorrhagic fever caused by the Lassa virus, is notably prevalent in West Africa and can be fatal if not managed properly. Effective outbreak control and mitigation rely heavily on early detection. This study introduces a predictive model for forecasting Lassa Fever using machine learning techniques. The research employs an extensive dataset that includes epidemiological and environmental data collected over multiple years to construct a reliable predictive model. Several machine learning algorithms, such as Random Forest, Support Vector Machines (SVM), and Gradient Boosting, are utilized to analyze historical data and predict potential Lassa Fever outbreaks. To improve model accuracy and minimize false positives, feature engineering and data preprocessing methods are applied. The dataset was divided into 80% for training and 20% for testing, with various algorithms—including SVM, K-Nearest Neighbor, Naïve Bayes, and Logistic Regression—tested to build the predictive models. The performance of these models was assessed using metrics like accuracy, precision, recall, and F1 score. The SVM model demonstrated the highest performance, achieving accuracy, precision, and F1 scores of 90%, 91%, and 96%, respectively. Consequently, SVM was selected as the preferred algorithm for this research. The study recommends the use of this model in Nigeria's medical industry to enhance diagnostic speed and reduce medical errors in healthcare settings. Keywords: Lassa Fever, Predictive Model, Machine Learning, Support Vector Machines (SVM), and Healthcare Diagnostics*

---

## Introduction

Lassa fever is an acute hemorrhagic disease that is more prevalent in West African countries like Nigeria, Sierra Leon, Liberia and Guinea. The illness was discovered in 1969 when two missionary nurses died in Lassa - a town in present day Askira/Uba Local Government Area of Borno state, northeastern Nigeria (Lecompte, Fichet-Calvin, Daffis, Koulemou, & Sylla, 2006). Presently, the disease is a big threat in Sub Sahara Africa. The virus is transmitted to humans through contacts with food or other domestic items contaminated with rodents' urine, saliva, excreta and other body fluids. The virus is a single stranded RNA virus belonging to the Arenaviridae. It presents no definite signs or symptoms and clinical analysis is often problematic especially at the onset of the disease. Accurate diagnosis requires highly specialized laboratories which are expensive, and not readily available to the entire populace. Early diagnosis and treatment of Lassa fever is very vital for survival of the victim patient.

The disease is also associated occasional epidemics, during which the case-fatality rate can reach up to 50% (Omilabu S. *et. al.*, 2012). Number of Lassa fever infections per year in West Africa, according to World Health Organization (W.H.O) is estimated at 100,000 to 300,000 with approximately 5,000 death cases. Unfortunately, such estimates are crude, because surveillance for cases of the disease is not uniformly performed (Amarosa *et. al.*, 2010). In some areas of Sierra Leon and Liberia, it is known that 10% to 16% of people admitted to hospitals have Lassa fever, which indicates the serious impact of the disease on the population of this region (Bausch *et. al.*, 2001).

Healthcare domain is one of the most prominent research areas in the world. With the rapid advancement in technology and availability of data, Machine Learning (ML) in particular is an emerging approach that is used in developing models that aid in prediction and diagnosis of diseases, thereby reducing the hectic workload of doctors and other medical and health workers. Machine Learning is a branch of Artificial Intelligence (AI) that engages the usage of human knowledge to solve complex issues that require human expert (Adedeji A *et. al.*, 2016). Machine Learning technique uses algorithms such as Naive Bayes, Decision Tree, Random Forest to predict the occurrence or otherwise of an event based on provided datasets.

In this work, we have developed a diagnostic model for prediction of Lassa Fever Virus (LFV) based on symptoms using machine learning techniques. Datasets for this research was obtained from the Federal Medical Centre (FMC) Jabi–Abuja. Other secondary sources include Kaggle Machine Learning repository and Viral Hemorrhagic Fever (VHF) medical repository.

## Statement of the problem

It is estimated that the number of casualties of Lassa fever every year is between 10% to 15% of the total number of people infected by the disease, and may even go above 50% during epidemic periods (Omilabu *et. al.*, 2012). Due to the fact that symptoms of Lassa fever are so varied and nonspecific, clinical diagnosis is often difficult and requires highly specialized laboratories in order to obtain an accurate and reliable results, and those specialized laboratory facilities are inadequate because of their expensive nature, hence, they are not readily available in most developing countries like Nigeria - this makes the test results take longer time before they are available, this lead to unfortunate deaths of significant number of victim patients. Because of the above-mentioned problems therefore, a computer based diagnostic model for predicting the disease is a thing of necessity in order to fasten the diagnostic process so as to obtain accurate test results at fastest possible time so as to avoid

loss of life. Another factor is the cost effectiveness. Computer based diagnostic models are cheaper compared to setting up a conventional molecular laboratory.

### **Aim and objectives of the study**

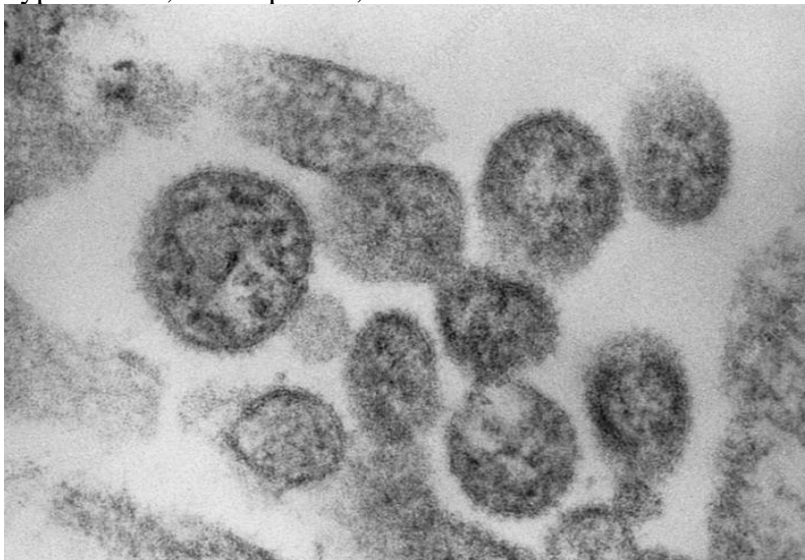
The aim of the study is to develop and implement a model for the prediction of Lassa fever based on certain symptoms using machine learning techniques. The objectives include to:

- i. Build the diagnostic model.
- ii. Train the model by applying part of the dataset on Support Vector Machine (SVM) algorithm.

### **Conceptual Review**

Lassa Fever Virus (LFV) is an acute viral zoonotic illness caused by the Lassa virus, an arenavirus that leads to severe hemorrhagic fever. The virus primarily infects *Mastomys natalensis* (domestic rats), which act as asymptomatic carriers and excrete the virus through urine. Lassa fever is endemic in West Africa, particularly in countries such as Sierra Leone, Guinea, Liberia, and Nigeria, with an estimated 300,000 to 500,000 cases and about 5,000 deaths annually in the region. The virus was first identified in 1969 after two nurses died in Nigeria, with the disease named after the town of Lassa in present-day Borno State, Nigeria. Since then, multiple outbreaks have been reported across Nigeria and other West African countries, with occasional cases involving travelers spreading the virus to other parts of the world.

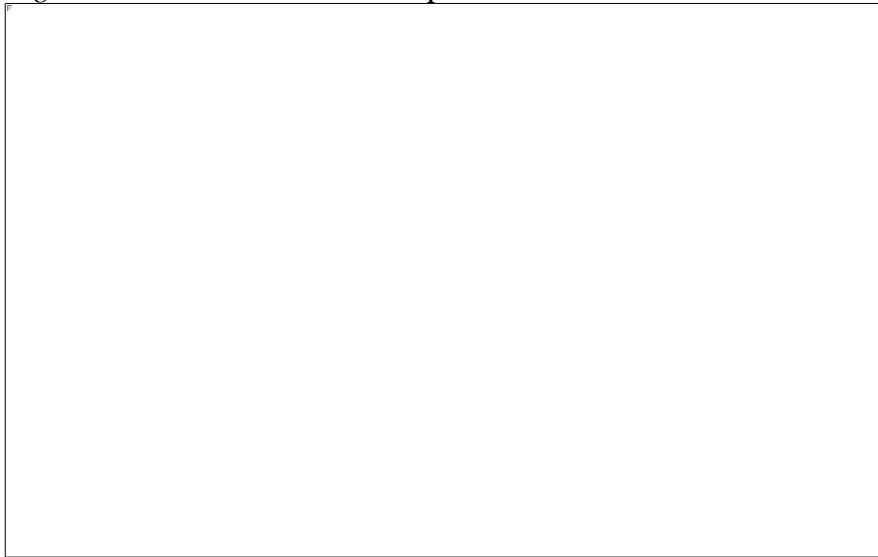
Lassa virus is an enveloped, single-stranded RNA virus from the Arenaviridae family, characterized by a bi-segmented genome. Unlike other arenaviruses, it lacks a conventional negative-strand coding arrangement, and its isolates vary in genetic, serologic, and pathogenic traits. The virus is spherical, with a diameter ranging from 70 to 150 nm, and features a smooth surface envelope with T-shaped glycoprotein spikes. The genome is enclosed within a helical nucleocapsid that measures between 400 and 1300 nm in length. The interior often contains electron-dense granules, identified as host cell ribosomes, giving the virus a "sandy" appearance. Lassa virus can be inactivated by ultraviolet light, gamma irradiation, heating between 56–100°C, and pH levels from 5.5 to 8.5. Chemical agents such as 0.5% sodium hypochlorite, 0.5% phenol, and 10% formalin are also effective in inactivating the virus.



**Figure 2.1:** Lassa virus electron micrograph (Ogbuet *al.*, 2004)

### ***Replication of Lassa virus***

Lassa virus adopts a replication strategy known as “*Ambisense*,” which is very rapid and demonstrates the replication of temporal control (Yuan, Bartlam, Lou, Chen, & Zhou, 2009). The initial stage of transcription of mRNA copies of the negative sense gene, which makes sufficient deposition of virus proteins for the next stage of the replication process. Subsequently, L and NP proteins are translated from the mRNA. Copies of the viral complementary RNA (vcRNA) are made from the positive sense gene. Negative sense progeny is produced by templates of RNA copies while mRNA is synthesized from it. The mRNAs produced from vcRNA are later translated to synthesize the Z and GP proteins. The temporal controls enhance the production of proteins spikes lastly, and therefore, delay the recognition by host immune system. *Fig 2.2*: shows Lassa virus adsorption on cell surface.



**Figure 2.2:** Lassa virus adsorption on cell surface (Yuan *et al.*, 2009)

### ***Symptoms of Lassa fever***

The incubation period of Lassa fever ranges from 7 to 21 days, (Uneke, 2012). The clinical disease begins as a flu-like illness characterized by the following symptoms:

- i. Weakness
- ii. Cough
- iii. Chest pain
- iv. Back pain
- v. Joints pain
- vi. Sore throat
- vii. Dysuria
- viii. Headache
- ix. Abdominal pain
- x. Vomiting
- xi. Diarrhea
- xii. Pharyngitis
- xiii. Conjunctivitis
- xiv. Bleeding
- xv. Abdominal tenderness
- xvi. Rales
- xvii. Facial edema

### ***Lassa virus transmission agents***

The reservoir, or host, of Lassa virus is a rodent called *Mastomys natalensis* also known as the multimammate rat or domestic rat (Tunmibiet. *al*, 2013). Once infected, this rodent is able to excrete virus in urine for an extended time period, maybe for the rest of its life. *Mastomys* rodents breed frequently, produce large numbers of offspring, and are numerous in the savannas and forests of west, central, and east Africa. In addition, *Mastomys* readily colonize human homes and areas where food is stored. All of these factors contribute to the relatively efficient spread of Lassa virus from infected rodents to humans.

Transmission of Lassa virus to humans occurs most commonly through ingestion or inhalation. *Mastomys* rodents shed the virus in urine and droppings and direct contact with these materials, through touching soiled objects, eating contaminated food, or exposure to open cuts or sores, can lead to infection.

Because *Mastomys* rodents often live in and around homes and scavenge on leftover human food items or poorly stored food, direct contact transmission is common. *Mastomys* rodents are sometimes consumed as a food source and infection may occur when rodents are caught and prepared. Contact with the virus may also occur when a person inhales tiny particles in the air contaminated with infected rodent excretions. This aerosol or airborne transmission may occur during cleaning activities, such as sweeping.

Direct contact with infected rodents is not the only way in which people are infected; person-to-person transmission may occur after exposure to virus in the blood, tissue, secretions, or excretions of a Lassa virus-infected individual. Casual contact (including skin-to-skin contact without exchange of body fluids) does not spread Lassa virus. Person-to-person transmission is common in health care settings (called nosocomial transmission) where proper personal protective equipment (PPE) is not available or not used. Lassa virus may be spread in contaminated medical equipment, such as reused needles.



**Figure 2.3:** *Mastomys natalensis*. (Ezeet *al.*, 2010)

### ***Diagnosis of Lassa Fever***

Lassa fever is most often diagnosed by using enzyme-linked immunosorbent serologic assays (ELISA), which detect IgM and IgG antibodies as well as Lassa virus antigen. Reverse transcription-polymerase chain reaction (RT-PCR) can be used in the early stage of the disease. The virus itself may be cultured in 7 to 10 days, but this procedure should only be done in a

high containment laboratory with good laboratory practices. Immunohistochemistry, performed on formalin-fixed tissue specimens, can be used to make a post-mortem diagnosis.

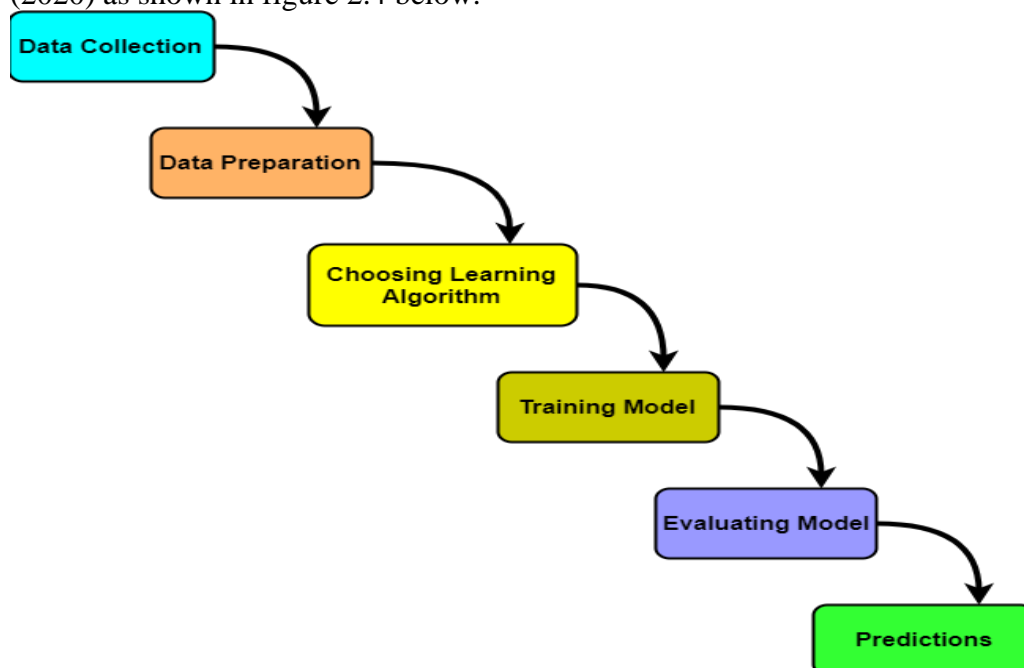
### Machine Learning (ML)

Since their evolution, humans have been using many types of tools to accomplish various tasks in a simpler way. The creativity of the human brain led to the invention of different machines. These machines made the human life easy by enabling people to meet various life needs, including travelling, industries, and computing, and machine learning is one among those numerous human inventions.

Machine learning (ML), according to (Marsland, 2015) is defined as the field of study that gives computers the ability to learn without being explicitly programmed. Machine learning is used to teach machines how to handle data more efficiently. Sometimes after viewing the data, we cannot interpret the extract information from the data. In that case, we apply machine learning. With the abundance of datasets available, the demand for machine learning is on the rise. Many industries apply machine learning to extract relevant data. The purpose of machine learning is to learn from the data. Many studies have been done on how to make machines learn by themselves without being explicitly programmed. Many mathematicians and programmers apply several approaches to find the solution of this problem which are having huge data sets.

Machine Learning relies on different algorithms to solve data problems. Data scientists like to point out that there is no single one-size-fits-all type of algorithm that is best to solve a problem. The kind of algorithm employed depends on the kind of problem you wish to solve, the number of variables, the kind of model that would suit it best and so on.

The task of imparting intelligence to machines seems daunting and impossible. But it is actually really easy. It can be broken down into seven (7) major steps according to Sunil, (2020) as shown in figure 2.4 below.



*Figure 2.4:* Machine Learning Workflow. (Sunil, 2020).

- i. **Collecting Data:** Machines initially learn from the data that you give them. It is of the utmost importance to collect reliable data so that your machine learning model can find the correct patterns.
- ii. **Preparing the Data:** After you have your data, you have to prepare it. You can do this by putting together all the data you have and randomizing it. This helps make sure that data is evenly distributed, and the ordering does not affect the learning process.
- iii. **Choosing a Model:** A machine learning model determines the output you get after running a machine learning algorithm on the collected data. It is important to choose a model which is relevant to the task at hand.
- iv. **Training the Model:** Training is the most important step in machine learning. In training, you pass the prepared data to your machine learning model to find patterns and make predictions. It results in the model learning from the data so that it can accomplish the task set. Over time, with training, the model gets better at predicting.
- v. **Evaluating the Model:** After training your model, you have to check to see how it is performing. This is done by testing the performance of the model on previously unseen data.
- vi. **Parameter Tuning:** Once you have created and evaluated your model, see if its accuracy can be improved in any way. This is done by tuning the parameters present in your model. Parameters are the variables in the model that the programmer generally decides. At a particular value of your parameter, the accuracy will be the maximum. Parameter tuning refers to finding these values.
- vii. **Making Predictions:** In the end, you can use your model on unseen data to make predictions accurately.

### **Machine Learning Algorithms**

Machine learning involves showing a large volume of data to a machine so that it can learn and make predictions, find patterns, or classify data (Mohssen, Muhammad, & Eihab, 2017). Basically, there are three types of machine learning algorithms: Supervised Learning, Unsupervised Learning, and Reinforcement Learning.

#### **Supervised Learning**

Supervised machine learning is a widely used machine learning technique that predicts future outcomes or events. It uses labeled datasets to learn and generate accurate predictions. Supervised learning is classified into two categories, namely, regression and classification.

According to Marsland, (2015), regression is a model that predicts continuous values (numerical), while classification mainly classifies the data. Regression is accomplished by using a linear regression algorithm, and classification is achieved through logistic regression.

In essence, this machine learning type is called “Supervised Learning” because the machine is “supervised” while its learning, which means that you are feeding the algorithm with information to help it learn. The outcome you provide the machine is labeled data, and the rest of the information you give is used as input features (Sarker, 2021). Common algorithms used during supervised learning include Naïve Bias, Neural Networks, Decision Trees, Logistics & Linear Regression, Support Vector Machines (SVM) etc.

Supervised learning is effective for a variety of business processes, including sales forecasting, inventory optimization, and fraud detection. Some examples of use cases include:

- i. Predicting real estate prices
- ii. Classifying whether bank transactions are fraudulent or not

- iii. Finding disease risk factors
- iv. Determining whether loan applicants are low-risk or high-risk
- v. Predicting the failure of industrial equipment's mechanical parts

### **Naive Bayes**

It is a classification technique based on Bayes Theorem with an assumption of independence among predictors (Marsland, 2015). In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature as shown in equation 2.0 below. Naïve Bayes mainly targets the text classification industry. It is mainly used for clustering and classification purpose depends on the conditional probability of happening.

$$P(c|x) = \frac{P(x|c) P(c)}{P(x)} \quad (1)$$

Where:

$P(c|x)$  -Posterior Probability

$P(x|c)$  -Likelihood

$P(c)$  - Class Prior Probability

$P(x)$  - Predictor Prior Probability

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

### **Pseudo Code of Naive Bayes**

**Input:** Training dataset T,

F= (f1, f2, f3 ....fn) // value of the predictor variable in testing dataset.

**Output:** A class of testing dataset.

**Steps:**

- i. Read the training dataset T;
- ii. Calculate the mean and standard deviation of the predictor variables in each class;
- iii. Repeat Calculate the probability of fi using the gauss density equation in each class;  
Until the probability of all predictor variables (f1, f2, f3,.., fn) has been calculated.
- iv. Calculate the likelihood for each class;
- v. Get the greatest likelihood

### **Support Vector Machine (SVM)**

Another most widely used machine learning technique is Support Vector Machine (SVM). Support-Vector Machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis(Sarker, 2021). In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. It basically, draw margins between the classes. The margins are drawn in such a fashion that the distance between the margin and the classes is maximum and hence, minimizing the classification error.



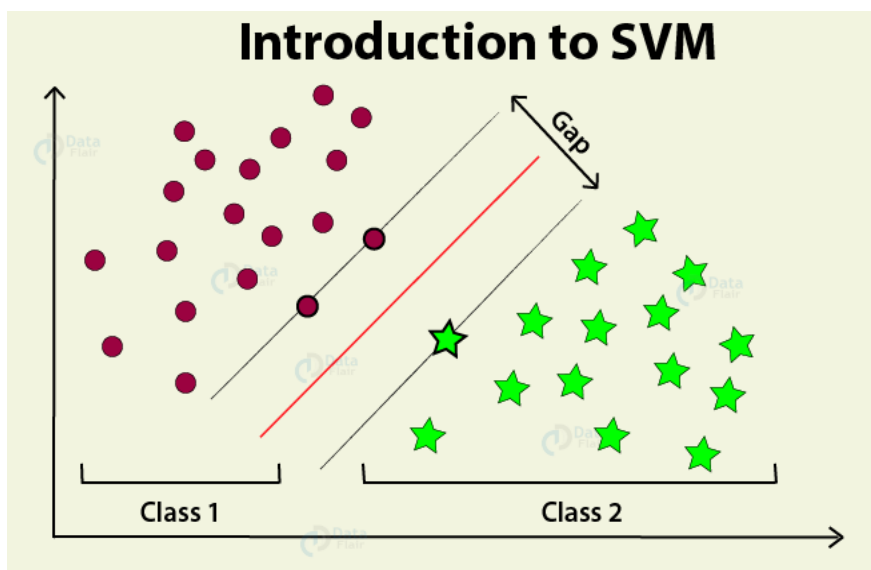


Figure 2.5: Support Vector Machine. (Sunil, 2020).

Support Vector Machines are a type of supervised machine learning algorithm that provides analysis of data for classification and regression analysis. While they can be used for regression, SVM is mostly used for classification. We carry out plotting in the n-dimensional space. Value of each feature is also the value of the specific coordinate. Then, we find the ideal hyperplane that differentiates between the two classes. These support vectors are the coordinate representations of individual observation. It is a frontier method for segregating the two classes.

### Decision Tree

Decision tree is a graph to represent choices and their results in form of a tree (Figure 2.6). The nodes in the graph represent an event or choice and the edges of the graph represent the decision rules or conditions. Each tree consists of nodes and branches. Each node represents attributes in a group that is to be classified and each branch represents a value that the node can take.

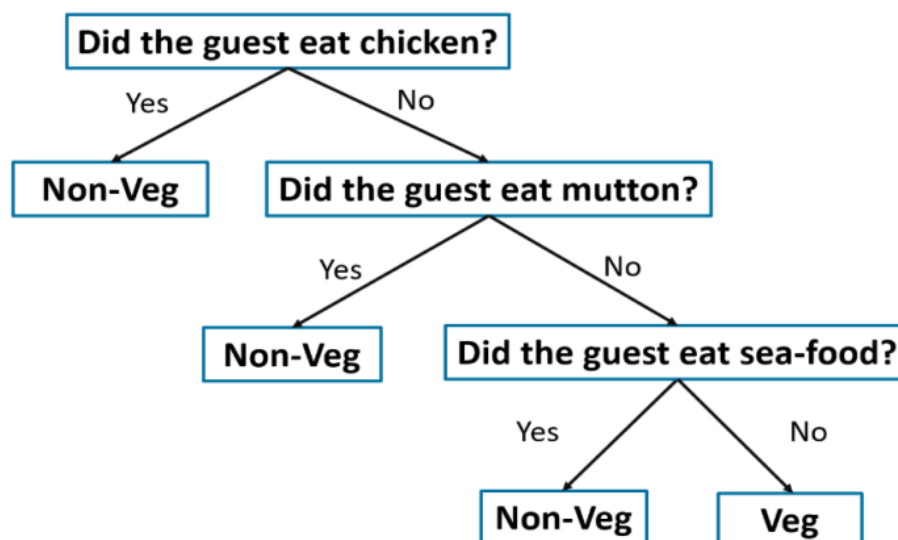
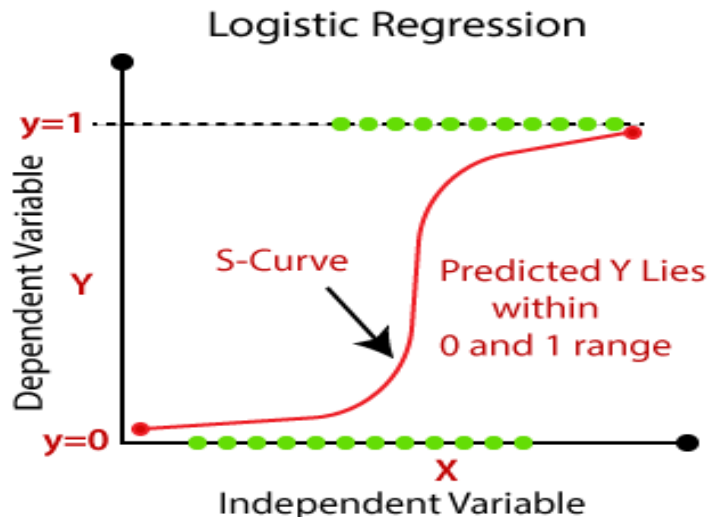


Figure 2.6: Decision Tree(Sunil, 2020).

### ***Logistic Regression***

Logistic regression is a classification algorithm that is used to predict a binary outcome based on a set of independent variables, (Batta, 2018). A binary outcome is one where there are only two possible scenarios - either the event happens (1) or it does not happen (0). Independent variables are those variables or factors which may influence the outcome (or dependent variable).



**Figure 2.7:** Logistics regression graph (Batta, 2018)

### ***Unsupervised Learning***

While supervised learning requires users to help the machine learn, unsupervised learning does not use the same labeled training sets and data (Batta, 2018). Instead, the machine looks for less obvious patterns in the data. This machine learning type is very helpful when you need to identify patterns and use data to make decisions. Common algorithms used in unsupervised learning include Hidden Markov Models, K-Means, Hierarchical Clustering, Gaussian Mixture Models etc.

This type of machine learning is widely used to create predictive models. Common applications also include clustering, which creates a model that groups objects together based on specific properties, and association, which identifies the rules existing between the clusters. A few examples use cases include:

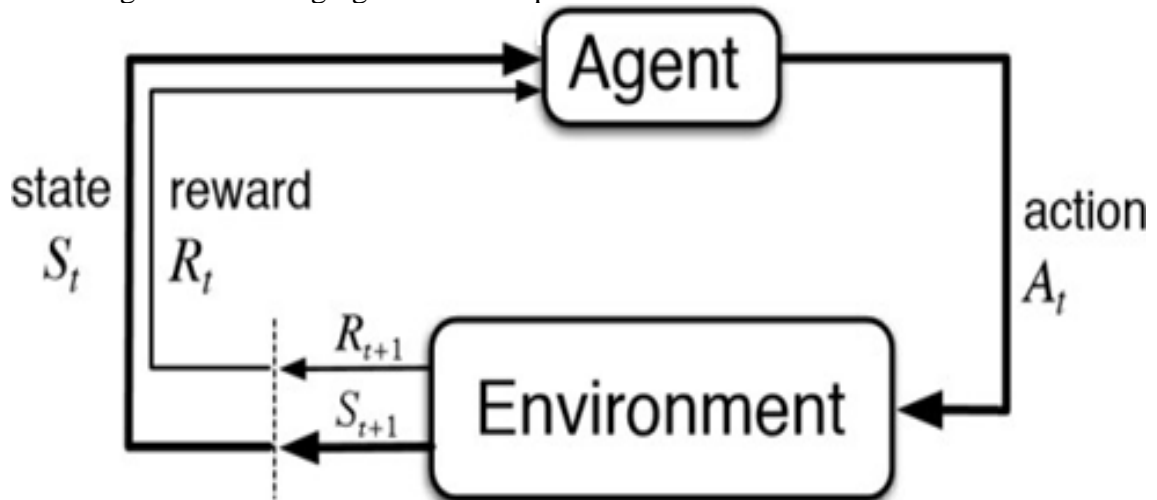
- i. Creating customer groups based on purchase behavior.
- ii. Grouping inventory according to sales and/or manufacturing metrics.
- iii. Pinpointing associations in customer data (for example, customers who buy a specific style of handbag might be interested in a specific style of shoe).

### ***Reinforcement Learning***

Reinforcement learning is the closest machine learning type to how humans learn. The algorithm or agent learns by interacting with its environment and getting a positive or negative reward (Batta, 2018). Common algorithms include Temporal Difference, Deep Adversarial Networks, And Q-Learning.

Olubi, Oniya, & Owolabi, (2021) emphasized that most machine learning platforms do not have reinforcement learning capabilities because it requires higher computing power than most organizations can afford. Reinforcement learning is applicable in areas capable of being fully simulated that are either stationary or have large volumes of relevant data. Because this type of machine learning requires less management than supervised learning, it's viewed as

easier to work with dealing with unlabeled data sets. Practical applications for this type of machine learning are still emerging. Some examples of uses include:



**Figure 2.8:** Reinforcement Learning(Sunil, 2020).

Teaching cars to park themselves and drive autonomously, dynamically controlling traffic lights to reduce traffic jams, and Training robots to learn policies using raw video images as input that they can use to replicate the actions they see.

### Empirical Reviews

In the time past, several studies are being conducted on diagnosing Lassa fever using various Artificial Intelligence techniques. For instance (Tunmibi, Adeniji, Aregbesola, & Dasylya, 2013) presented a rule-based expert system called e-Diagnosis that diagnosed Lassa fever and other types of fever. The system results helped reduce overcrowding at the hospitals, cost-effective, reliable, and had a high-level of detection accuracy of various types of fever.

Osigbeme, Ogunwolu, Omoare, & Inyiama, (2014) proposed an expert system that diagnosed contagious diseases using Fuzzy logic as an approach. The expert system detected contagious diseases such as Lassa fever and Ebola fever with 90% detection accuracy.

Osaseri, and Osaseri, (2016) developed an expert system that employed the Adaptive Neuro-Fuzzy Inference System (ANFIS) for the prediction of Lassa fever. The system predicted Lassa fever with high detection accuracy.

Abiola, Alaba, & Joy, (2017) designed a medical expert system for diagnosing Lassa fever based on rules. The system aimed to replace the manual method of diagnosing Lassa fever still been practiced by medical professionals, with an Expert System (ES) which had the capability of correcting all the limits associated with the manual method of diagnosing Lassa fever. The system detected Lassa fever with high detection accuracy.

Ajenaghughure, Sujatha, & Akazue, (2017) proposed an integrated expert system that diagnosed several types of fevers using Fuzzy Logic as an approach. Nevertheless, the expert system detected the symptom relationship between the various types of fevers, namely: Dengue Hemorrhagic fever, Malaria fever, Typhoid fever, Lassa fever, and yellow fever respectively, with the fuzzy rules applied accurately, determining the type of fever first, and then ascertained the level of infection to be either mild or severe.

Aminu, Abiodun, Isah, Iliyasu, Isah, & Hussaini (2018) presented an expert system diagnosing Lassa fever and related ailments using Fuzzy Logic. The system diagnosed Lassa fever and related ailments with high detection accuracy.

Nnebe, Okoh, Otumu, & Oshoiribhor, (2019) proposed a Neuro-fuzzy system to detect Lassa fever based on observed symptoms. The system detected cases of Lassa fever with high detection accuracy based on the twenty-nine observed symptoms captured.

Steur and Mueller (2019) presented a system that classified Viral Hemorrhagic Fever (VHF) with emphasis on Ebola and Lassa fever using Neural Network (NN). The system detected Ebola and Lassa fever with high detection accuracy.

Egwali and Obi (2015) developed a Neuro-fuzzy system to identify Ebola Hemorrhagic fever (EHF). The system detected EHF with high detection accuracy based on symptoms using 29 clinical signs and indications which were classified into 5 levels.

Govinda and Eswara (2013) (2013) designed a multi-fever expert system to detect Dengue Hemorrhagic fever, Malaria fever, and Typhoid fever using a rule-based technique called Fuzzy Logic. The system was able to determine the level of infection ranging from mild to severe with high detection exactness, and also make recommendations where necessary on food and drugs to be administered to infected patients.

Djam, Wajiga, Kimbi, & Blamah, (2011) developed an expert system for handling malaria termed fuzzy expert system for the management of malaria (FESMM). The system detected malaria with high detection accuracy based on clinical observations, medical diagnosis, and expert knowledge. Besides, the system also served as a decision support system (DSS) for healthcare practitioners in malaria-endemic regions.

Alile (2020) presented a model for detecting Dengue Haemorrhagic fever (DHF) using a machine learning technique called Bayesian Belief Network. The proposed model detected Dengue Haemorrhagic fever using data retrieved from Dengue Haemorrhagic fever medical repository and had 99.84% prediction precision based on test data.

Imanov and Asengi (2021) proposed an expert system for diagnosing Lassa fever using Artificial Intelligence. The system diagnosed Lassa fever with high detection accuracy, with the system's outcomes authenticated and supported by specialized medical personnel based on his analysis. Furthermore, the system could analyze Lassa fever where the number of infected patients is high. On the other hand, the system also assists medical experts with speedy, precise determinations and can be used in situations where is a shortage or no medical expert available.

Alile (2020) presented a model for early diagnosis of cerebral malaria, multi-fever, and mosquito-borne diseases using a supervised machine learning approach called Bayesian Belief Networks. The proposed model detected cerebral malaria, mosquito-borne diseases, and other types of fevers such as Dengue Haemorrhagic fever, Malaria Fever just to name but a couple with an overall prediction exactness of 99.98% based on test data. Notwithstanding, Table 1 below shows a summary of the existing systems discussed above showing the ailment type, the artificial intelligence (AI) Technique employed to diagnose the ailment type, aim, and objectives, as well as the results of those works.

Ijeoma *et. al.*, (2021) Discovered that between April 5, 2018, and March 15, 2020, 534 patients with confirmed Lassa fever were admitted to hospital, of whom 510 (96%) gave consent and were included in the analysis. The cohort included 258 (51%) male patients, 252 (49%) female patients, 426 (84%) adults, and 84 (16%) children (younger than 18 years). The median time between first symptoms and hospital admission was 8 days. At baseline, 176 (38%) of 466 patients had a Lassa fever RT-PCR cycle threshold (Ct) lower than 30. From admission to end of follow-up, 120 (25%) of 484 reached a National Early Warning Score (second version; NEWS2) of 7 or higher, 67 (14%) of 495 reached a Kidney Disease Improving Global Outcome (KDIGO) stage of 2 or higher, and 41 (8%) of 510 underwent dialysis. All patients received ribavirin for a median of 10 days (IQR 9–13). 62 (12%) patients died (57 [13%] adults and five [6%] children). The median time to death was 3 days (1–6). The baseline factors independently

associated with mortality were the following: age 45 years or older (adjusted odds ratio 16.30, 95% CI 5.31–50.30), NEWS2 of 7 or higher (4.79, 1.75–13.10), KDIGO grade 2 or higher (7.52, 2.66–21.20), plasma alanine aminotransferase 3 or more times the upper limit of normal (4.96, 1.69–14.60), and Lassa fever RT-PCR Ct value lower than 30 (4.65, 1.50–14.50)

Olawole & Knonyana (2022) have conducted comprehensive research titled “Forecasting Lassa Fever Outbreak Progression with Machine Learning”. In that work, they concluded that prompt surveillance of Lassa Fever outbreaks and forecasting of future cases are vital for healthcare organizations. The study explored decision tree (DT) and k-nearest neighbor (kNN) algorithms with time-dependent variables for model performance and reliable error estimates of weekly confirmed Lassa Fever cases. Lassa Fever models were benchmarked using the seasonal autoregressive integrated moving average (SARIMA). Results from SARIMA, DT, kNN, grid search of DT, and grid search kNN show that implemented models produced robust performance regardless of the LF confirmed case data complexity. Evaluation from the test set shows that kNN with grid search prediction outperformed other models except in kNN model, where the MAE had a lower error score, and the DT with grid search for MAPE had a lower error when compared to other models.

Samson, Akingbade, & Orija, (2023), established that in Sub-Saharan Africa, particularly in Nigeria, Lassa fever poses a significant infectious disease threat. The research employed count regression and machine learning techniques to model mortality rates associated with confirmed Lassa fever cases. Utilizing weekly data from January 7, 2018, to April 2, 2023, provided by the Nigeria Centre for Disease Control (NCDC), an analytical comparison between these methods was conducted. Overdispersion was indicated ( $p < 0.01$ ), prompting the exclusive use of negative binomial and generalized negative binomial regression models. Machine learning algorithms, specifically Medium Gaussian Support Vector Machine (MGSVM), Ensemble Boosted Trees, Ensemble Bagged Trees, and exponential Gaussian Process Regression (GPR), were applied, with 80% of the data allocated for training and the remaining 20% for testing. The efficacy of these methods was evaluated using the coefficients of determination ( $R^2$ ) and the root mean square error (RMSE). Descriptive statistics revealed a total of 30,461 confirmed cases, 4,745 suspected cases, and 772 confirmed fatalities attributable to Lassa fever during the study period. The negative binomial regression model demonstrated superior performance ( $R^2=0.1864$ ,  $RMSE=4.33$ ) relative to the generalized negative binomial model ( $R^2=0.1915$ ,  $RMSE=18.2425$ ). However, machine learning algorithms surpassed the count regression models in predictive capability, with ensemble boosted trees emerging as the most effective ( $R^2=0.85$ ,  $RMSE=1.5994$ ). Analysis also identified the number of confirmed cases as having a significant positive correlation with mortality rates ( $r=0.885$ ,  $p < 0.01$ ). The findings underscore the importance of promoting community hygiene practices, such as preventing rodent intrusion and securing food storage, to mitigate the transmission and consequent fatalities of Lassa fever.

### Research Gaps of the Study

Previous models developed with various Artificial Intelligence methods have shown low prediction accuracy and sensitivity for the Lassa virus. For instance, Samuel, Ozemoya, Adetokunbo, and Emmanuel (2019) used Fuzzy Logic to create a predictive system for diagnosing Lassa fever. Their model achieved only 70% accuracy and 67% sensitivity, indicating subpar performance. In contrast, this model achieved 90% accuracy and 91% sensitivity to the virus.

Another significant research gap is that many existing systems rely on either Expert Systems or Fuzzy Logic. The limitation of these approaches is that their algorithms do not

adapt or improve over time; they are explicitly programmed to make decisions based on the symptoms included during coding. As a result, if a disease presents an unusual symptom, perhaps due to virus mutation (as with viral diseases like Lassa fever), the system must be manually updated to include the new symptom in the source code. In contrast, this system was developed using Machine Learning techniques, enabling the algorithm to learn from past data and automatically improve its performance.

## **Method**

### **Location of the study**

The research work was conducted at the Nigerian Centre for Disease Control (NCDC), Abuja, and Federal Medical Centre (FMC), Jabi – Abuja.

### **Sources of data**

For the purpose of this research work, medical records of patients diagnosed with Lassa fever virus was collected from the Federal Medical Center (FMC), Jabi - Abuja. Also, datasets to be used in training the algorithm was obtained from Kaggle Machine Learning repository - a data science platform and online community of data scientists and machine learning practitioners under Google LLC. It enables users to find and publish datasets, explore and build models in a web-based data science environment.

### **Method of data collection**

For the purpose of this research work, interviews and direct observations were also conducted as the methods for data collection.

### **Development Tools**

Below are some of the basic development tools used for the development, implementation and testing the accuracy of the system.

#### ***Programming Languages***

- i. ***Python*** – High-Level, general purpose programming language used in building and training the model.

#### ***Integrated Development Environments (IDEs)***

- i. ***Google Collab*** - Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free access to computing resources, including GPUs and TPUs. Colab is especially well suited to machine learning, data science, and education.
- ii. ***Spyder*** - Spyder is a free and open-source scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.
- iii. ***Anaconda Navigator***- Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda Distribution that allows you to launch applications and manage conda packages, environments, and channels without using command line interface (CLI) commands

#### ***Libraries***

- i. ***NumPy***–This stand for Numerical Python. It is an open-source Python library that is used in almost every field of science and engineering. It is the universal standard for working with numerical data in Python, and it is at the core of the scientific Python and PyData ecosystems.

- ii. **Pandas**—This stand for Panel Data. It is a python library used for working with datasets. It has functions for analysing, cleaning, exploring, and manipulating data.
- iii. **Pickle** - Pickle in Python is primarily used in serializing and deserializing a Python object structure. It is the process of converting a Python object into a byte stream to store it in a file/database, maintain program state across sessions, or transport data over the network.
- iv. **Streamlit**—This is an open-source Python library that makes it easy to create and share beautiful, custom web apps for machine learning and data science.

#### **Database Software**

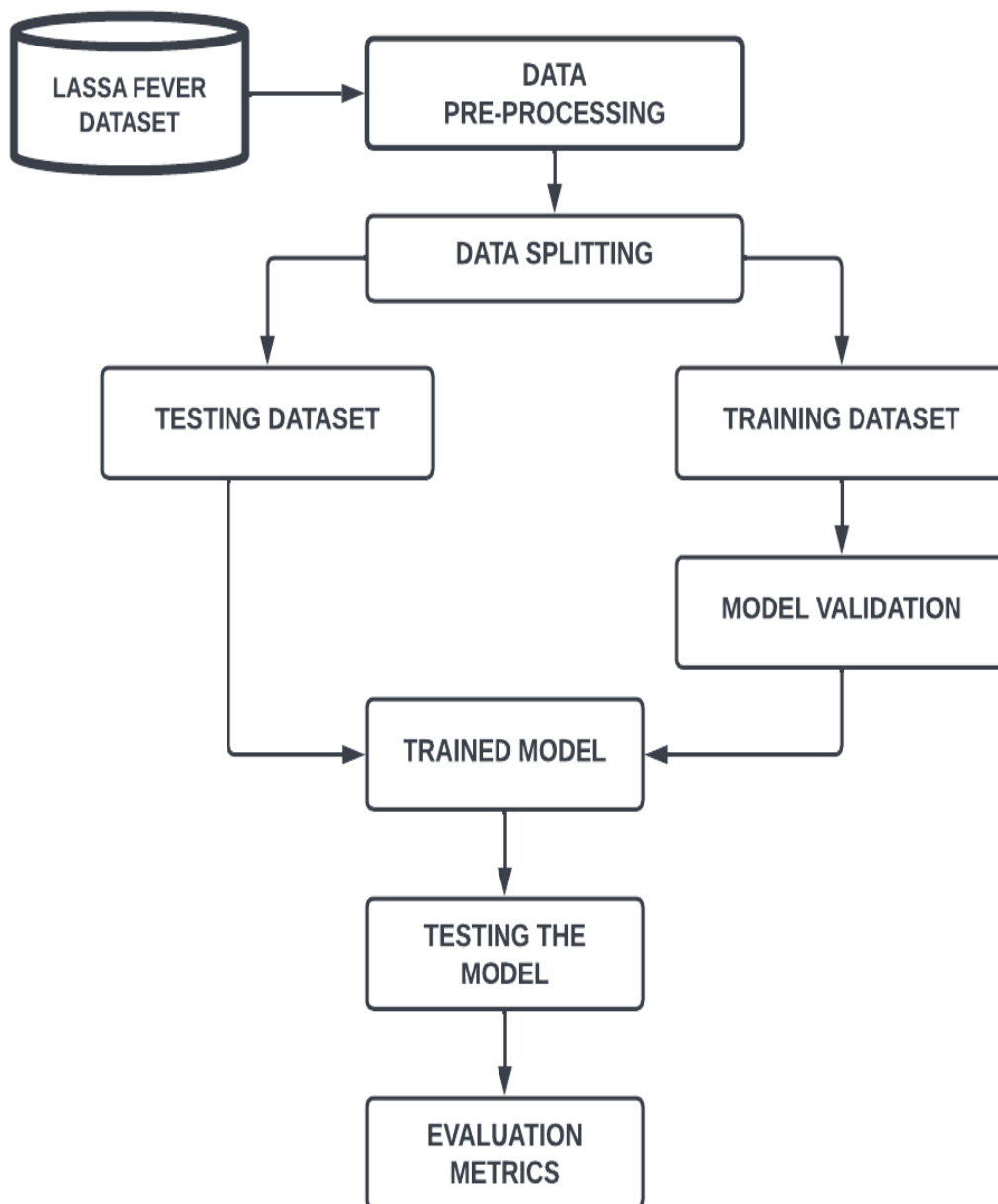
- i. **SQLite** - SQLite is a C-language library that implements a small, fast, self-contained, high-reliability, full-featured, SQL database engine. *SQLite* is the most used. It was used in storing the database containing the login details of the users as well as the classification report.

#### **Testing Tools**

- i. **Microsoft Edge**- Microsoft Edge is a proprietary, cross-platform web browser created by Microsoft. It was used in running the simulation of the model. The application also runs on any other internet browser.
- ii. **Notepad**- Windows Notepad is a simple text editor for Windows used for creating and editing of plain text documents. It was used for manipulating the datasets.

#### **The Model Architecture**

Figure 3.1 shows the model architecture. It is the diagrammatic representation of the algorithm along with the underlying structure or design of the machine learning model. Model architecture consists of layers of interconnected nodes or neurons, where each layer of the model performs a specific function, such as data preprocessing, feature extraction or prediction.



**Figure 3.1:**Architecture of the model

### ***Lassa Fever Dataset***

This is the collection of data related to Lassa fever that will be used in training and testing the model. For the purpose of this work, our dataset will be the collection of clinical records of patients diagnosed with Lassa fever and other related fevers obtained from the hospital.

### ***Data pre-processing***

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step in developing a machine learning



model. It involves refining the data into a more meaningful format before feeding it to the algorithm, elimination of redundant data as well as null data elements.

#### ***Training Dataset***

This is the part of the dataset that will be used in training the model. In the course of this research, 80% of the dataset will be used as training dataset.

#### ***Testing Dataset***

This is the subset of the dataset that will be used in testing the performance of the model. For the purpose of this research, 20% of the dataset would be used in testing the model.

#### ***Model Validation***

Model validation is the set of processes and activities intended to verify that a model is performing as expected. It determines whether the trained model is trustworthy or not.

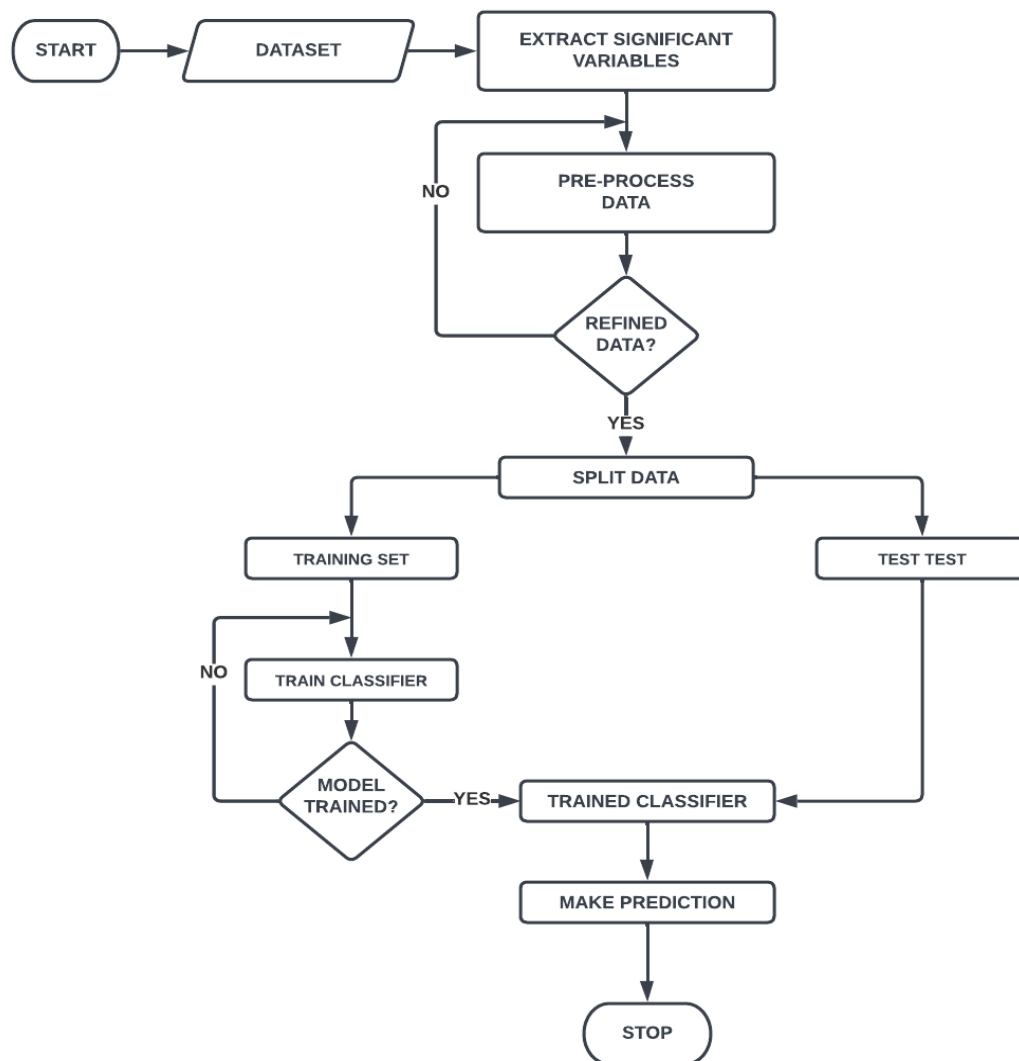
There are many techniques of Model validation such as Train/Test Split, k-Fold Cross-Validation, Leave-One-Out Cross-Validation, Leave-One-Group-Out Cross-Validation, Nested Cross-Validation, Time-Series Cross-Validation etc. For the purpose of this work, k-fold cross-validation will be adopted.

#### ***Evaluation Metrics***

This is where the quality of the model is being extensively measured. There are many types of evaluation metrics available to test a model such as classification accuracy, logarithmic loss, confusion matrix. In this work we intend to use confusion matrix.

#### **The Model Flow Chart**

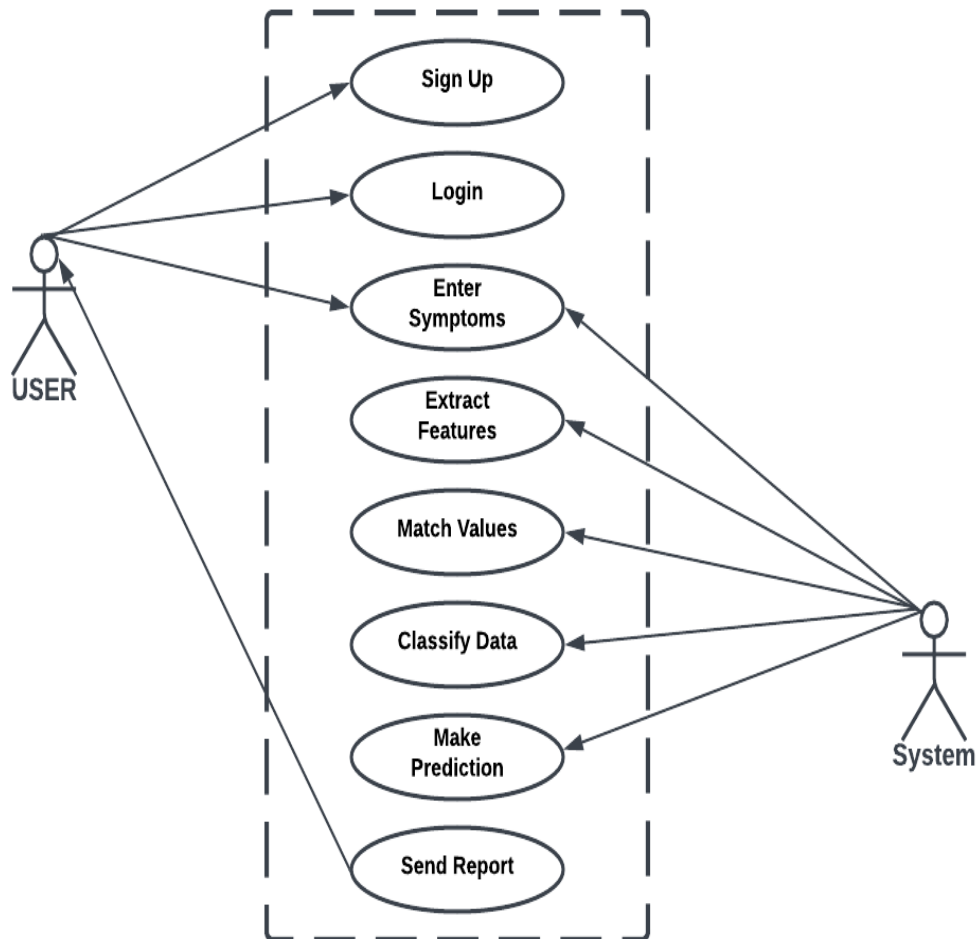
The flowchart of the model is shown in figure 3.2. It describes the number of steps, processes and procedures the algorithm follows in order to solve the problem. It is the diagrammatic representation of the algorithm.



**Figure 3.2:** Flow chart of the model

### Use Case Diagram of the System

The use case diagram, as shown in figure 3.3 models the behavior of the system and helps to capture the requirements of the system. The diagram describes the high-level functions and scope functions and scope of the system.



**Figure 3.3:**Use case diagram of the system

**Description of the System Function (Use Case)**

- i. **Sign up:** This where a new user of the system is being given access to the systems by introducing his details such as username and password to the system.
- ii. **Login:** This authenticates the access of the user of the system every time he wants to use the system.
- iii. **Enter Symptoms:**Stage where the system is being fed into the system in order to perform the prediction.
- iv. **Extract Features:**Here, the system transforms raw data into numerical features that can be processed while preserving the information into original dataset.
- v. **Match Values:**The system matches the available data with the dataset pattern in order to make prediction.
- vi. **Classify Data:**The system evaluates the classifiers and returns its performance to the admin such as the accuracy of the model.
- vii. **Make Prediction:**The system predicts the outcome based on the specified pattern of the dataset available.
- viii. **Send Report:**The system returns the prediction outcome based on the parameters issued.

### Sequence Diagram of the system

These are another form of UML diagrams that show how objects in a system or classes within code interact with each other. Particularly these diagrams show interactions in the order they take place. In another word, they show the sequence of events within the system. Sequence diagrams are primarily used by developers and business professionals to document processes or understand the requirements of a new program. Figure 3.4 show the sequence diagram of the model.

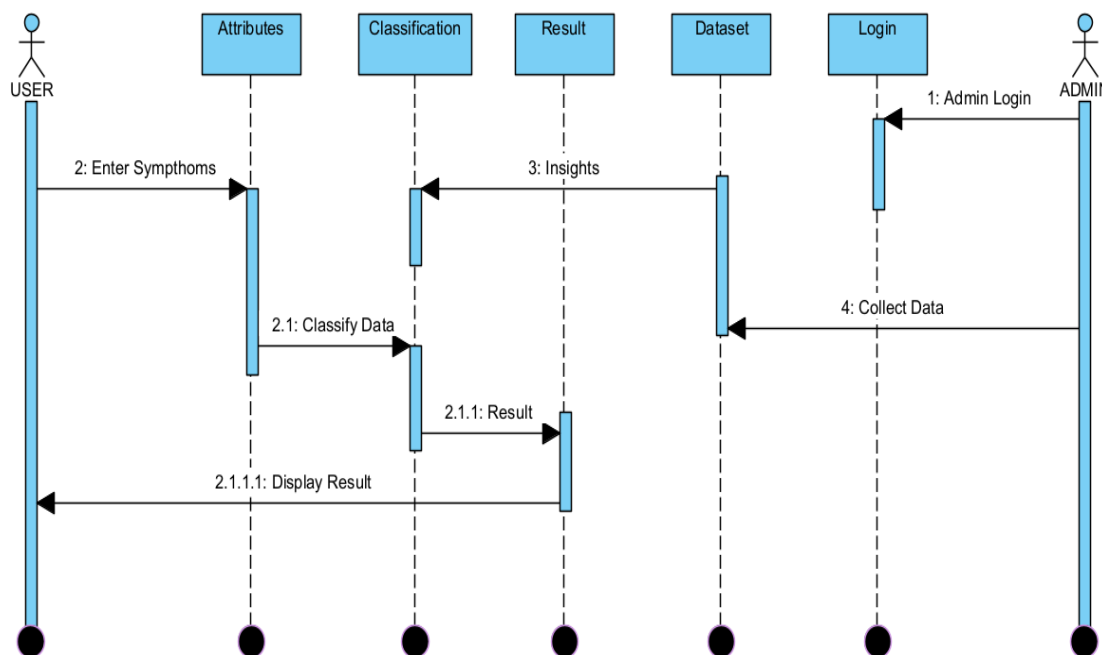


Figure 3.4: Sequence diagram of the system

#### Description of the System Function (Sequence)

- i. **Login:** This authenticates the access for the user into the system.
- ii. **Enter Parameters:** The user inputs the behavioral entities to the learning model to predict the presence of heart disease.
- iii. **Insights:** The model gets insight or pattern from the database in order to make prediction.
- iv. **Classify Data:** The system classifies data according to pattern obtained from the dataset.
- v. **Result:** The system returns the prediction outcome based on the parameters issued and the pattern on the dataset.

#### Materials and Method

The dataset as shown in Figure 4.5, which was obtained from clinical records of Federal Medical Centre (FMC) Jabi, Abuja, and Kaggle Machine Learning repository for Hemorrhagic fevers, was utilized to train our chosen machine learning algorithm called Support Vector Machine (SVM) to predict Lassa fever in patients. The machine learning model was trained, tested and simulated using Python programming language and Streamlit framework for machine learning and data science respectively. Below are the step-by-step details of how the model was designed and implemented.

- i. Data set was collected, refined and stored as Comma Separated Values (csv).
- ii. The dataset was loaded on Pandas library for Python using Google Collaboratory platform. Google Collaboratory or simply Google Colab is an online cross-platform Integrated Development Environment (IDE) that allows developing and running Python codes any installations. It also provides features like Graphical Processing Unit (GPU) and Tensor Processing Unit (TPU) support which are very useful when running machine learning and deep learning models.
- iii. The dataset was split into training and testing data using “train-test-split” function of Python. The dataset was split into 80% training data and 20% testing data.
- iv. Our Support Vector Machine (SVM) model was trained using the 80% training data, and tested using the 20% testing data.
- v. Various predictions were made using the testing data against the training data, and evaluation metrics such as accuracy score, precision, recall and F1\_score were run in order to ascertain the performance of the model.
- vi. The trained model was saved using pickle library as write binary (wb) file.
- vii. Now, the application interface was developed using Spyder IDE and was run on streamlit platform using command prompt.

## **Results and Discussion**

### **Introduction**

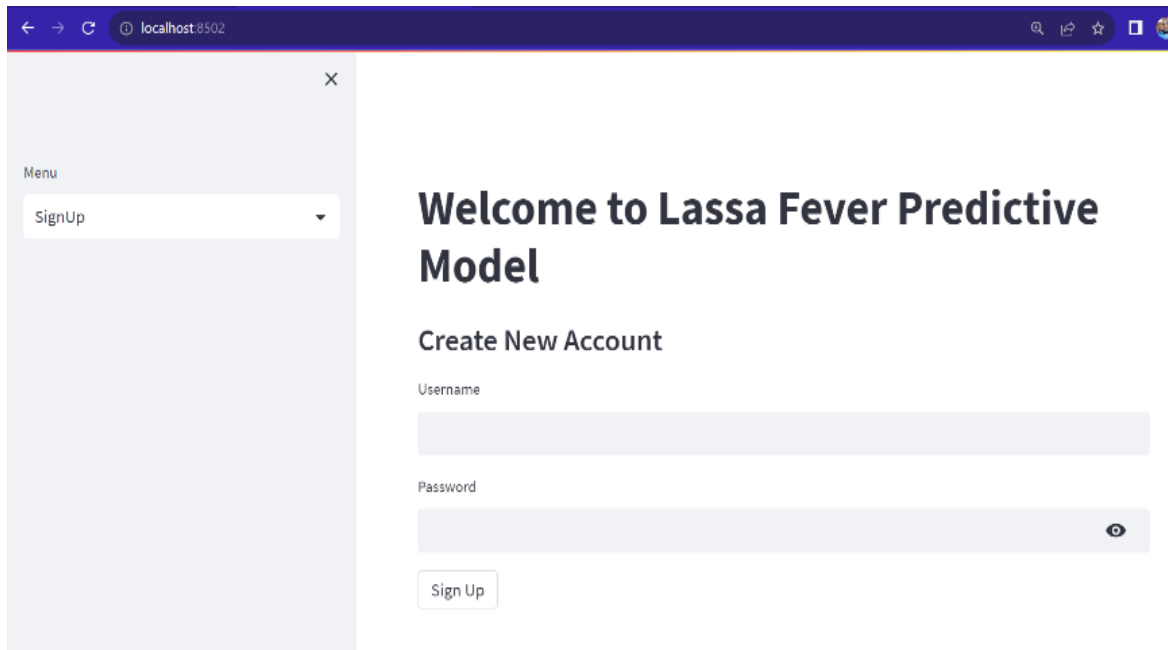
The implementation and outcome of this research work is presented in this chapter. The interfaces used in the implementation are presented here. The results obtained in this study were presented with performance evaluation of the model. The researcher also discussed the study’s outcome, findings and the comprehensive research result for the extended classifiers used in this research.

### **Implementation of the diagnostic system**

During the implementation of this work, the system interfaces along with their respective source codes is a result of the deployment phases at the realization.

#### ***The Sign-Up Interface***

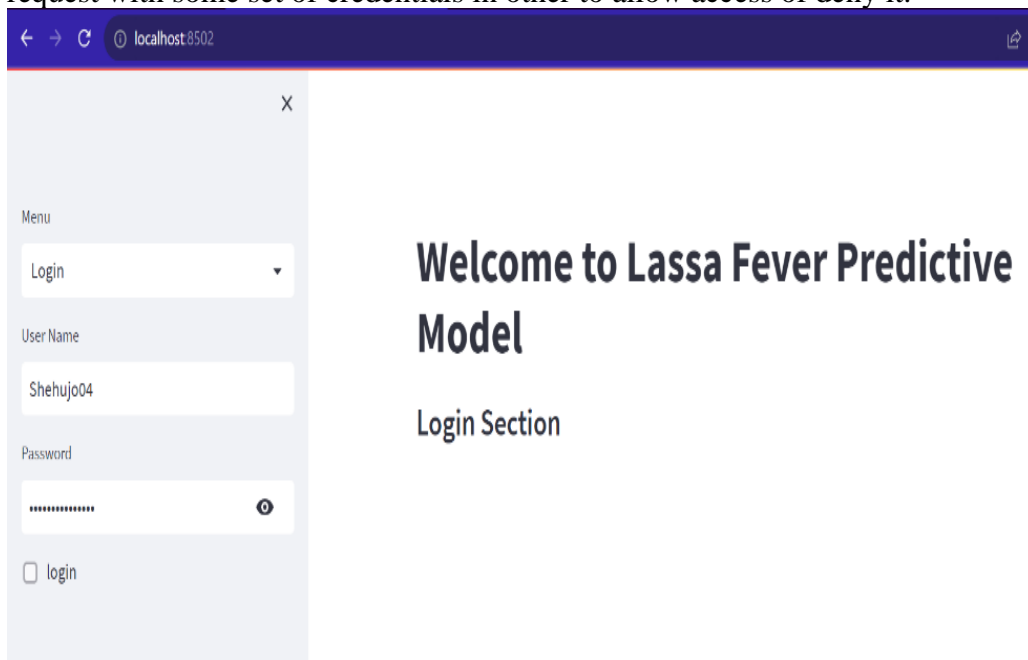
This interface enables system users register their credentials to create accounts that will be used within the system at the authentication phase as shown in figure 4.1.



**Figure 4.1:** The Sign-Up Interface

### ***The Authentication Interface***

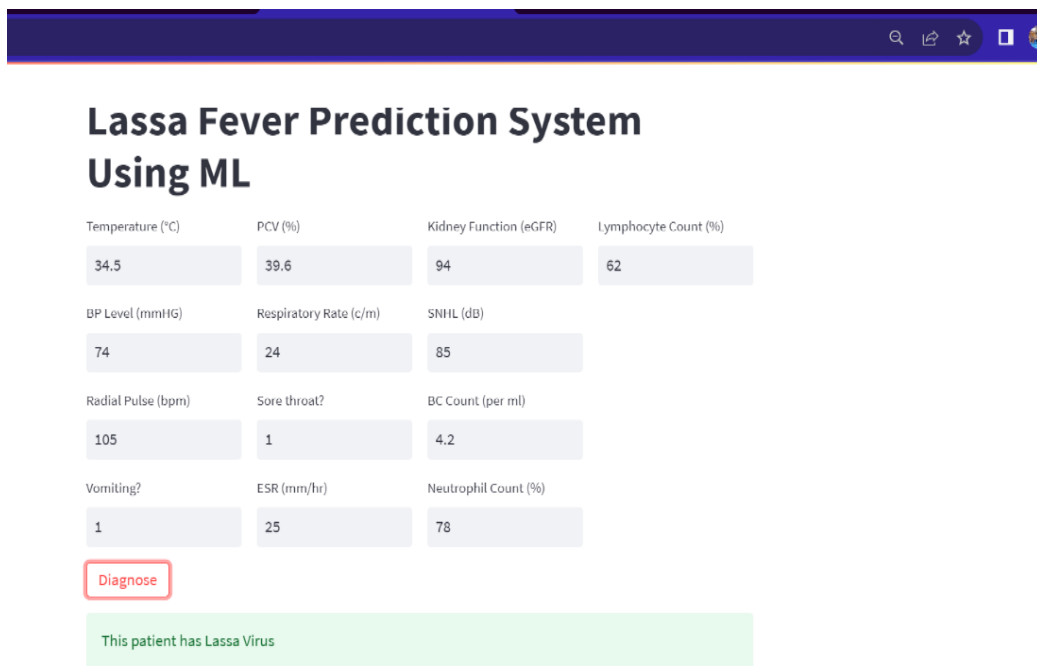
Figure 4.2 shows the authentication page. This interface enables system users be authenticated before having access to the system. The system compares the incoming login request with some set of credentials in order to allow access or deny it.



**Figure 4.2:** The Authentication Interface

### ***Data Capture and Diagnosis Interface***

This interface allows users of the system input disease attribute of patient for prediction of Lassa Fever Virus (LFV). The diagnosis interface is depicted in figure 4.3.



**Figure 4.3:**

Diagnosis Interface

**Description of Dataset Features**

Table 4.1 shows the description of the various elements of the dataset showing the data feature, unit of measurement and normal range.

**Table 4.1:** Description of the dataset features

SN	Feature	Unit of measurement	Normal Range
1.	Temperature	°C	36.1 - 37.2
2.	BP Level	mmHG	120/80
3.	Radial Pulse	bpm	60 - 90
4.	Vomiting	Y/N	1,0
5.	Packed Cell Volume (PCV)	%	Males: 8.3 - 48.6 Females - 35.5 - 44.9 Pregnant Women: 33-38
6.	Respiratory Rate	c/m	12 - 20
7.	Sore Throat	Y/N	1,0

8.	Erythrocyte Sedimentation Rate (ESR)	mm/hr	Males: 0 – 15 Females: 0 - 20
9.	Estimated Glomerular Filtration Rate (eGFR)	ml/m	60 - 90
10.	Sensorineural Hearing Loss SNHL	dB	26 to 40
11.	WBC Count	Per ml	4,500 - 11,000
12.	Neutrophil Count	Per ml	2,500 - 7,000
13.	Lymphocyte Count	Per ml	1,000 and 4,800
14.	Outcome	Y/N	1 – Sick 0 - Healthy

### Sample of the Lassa Fever Dataset

Table 4.2 shows the sample of dataset used in arriving at the outcome of the research work. The dataset consists of 14 data elements and 200 instances as shown in figure 4.2 below.

**Table 4.2:** Sample of the dataset used in training the model.

```
1 #Printing the first five rows of the dataset
2 Lassa_Fever_Dataset.head()
```

	Temperature	BP Level (mmHG)	Radial Pulse(bpm)	Vomiting	Packed Cell Volume (%)	Respiratory Rate (c/m)	Sore Throat	ERS (mm/h)	Kidney Function (eGFR)	SNHL (dB)	WBC Count (ml x 10**9)	Neutrophil Count	Lymphocyte Count	Target
0	39.4	60	110	1	33.5	22	1	40	52	68	2.40	30	70	1
1	36.9	50	112	0	34.2	20	1	38	48	70	7.30	82	13	1
2	37.8	72	93	0	39.2	22	0	11	51	62	8.84	80	38	0
3	39.0	66	90	0	44.4	16	1	12	56	74	11.02	84	67	1
4	36.9	64	93	1	35.8	20	1	24	63	15	10.45	69	76	1

### Summary

Lassa fever is a contagious life-threatening disease caused by the Lassa virus, which belongs to the Arenaviridae virus family whose idiosyncratic attribute is its method of spread to humans which can be direct or indirect through contact with food and household items contaminated with rodents' urine, feces, and secreted body fluids from a Lassa virus-infected person.

Lassa virus has an incubation period of 2 to 21 days in humans after contact before manifestations begin to show in the newly infected persons. Due to the aforesaid distinctive feature, it is quite difficult to detect owing to the overlapping symptoms the disease shares with other viral hemorrhagic fevers.

In the recent past, numerous clinical and machine learning techniques have been employed to diagnose Lassa fever with the intent of curbing the untimely deaths of patients - area medical and IT professionals are making frantic efforts to improve on. Hence, there was



the need to proffer a solution to curb this anomaly. Thus, in this research, we utilized a supervised machine learning approach called Support Vector Machine (SVM) to predict Lassa fever based on its symptoms.

Dataset used in developing the model has 14 data elements with each element representing an exclusive symptom of the disease, and other factors that influence the diagnosis. The model was trained and tested and had an overall accuracy of 90% and sensitivity of 96% in predicting Lassa Virus disease.

### **Conclusion**

According WHO, over 4, 702 people were suspected with Lassa fever disease as at April, 2023, where 877 confirmed cases were detected. Among the confirmed cases, there have been 152 deaths cases, representing 17% of the total infected persons. To mitigate this serious public health threat therefore, a machine leaning predictive model for prediction of Lassa fever diseases based on symptoms has been developed with the medical expert diagnostic dataset which was obtained in from the Federal Medical Centre Jabi, Abuja - Nigeria.

The dataset was partitioned into 80% training set and 20% testing set, respectively, where the models were trained with the 80% and tested with 20% dataset. The dataset was applied on different machine learning algorithms including support Vector Machine (SVM), K-Nearest Neighbor, Naïve Bayes, and Logistic Regression algorithms to build the predictive models.

The model was evaluated based on accuracy, precision, recall, and F1\_score performance evaluation techniques. In terms of accuracy, precision and F1\_Score, SVM model emerged to be the best model with 90% and 91% and 96% respectively. Therefore, SVM was chosen as the best model for implementing this research work.

### **Recommendations**

Based in the findings of this research, we hereby make the following recommendations:

- i. That this research titled Lassa fever predictive model using machine learning techniques is recommended for usage in the medical industry in Nigeria in order to address healthcare challenges in our hospitals.
- ii. That deployment of this model will aid medical practitioners in anticipating the presence of Lassa fever in patients in a very efficient and effective manner, as well as ensuring the medical expert's judgment system is free of challenges.
- iii. That the remarkable comparative results of this study suggest that the SVM should be used on any medical illnesses where prediction conditions are difficult due to its performance in terms of prediction accuracy.
- iv. That researchers should broaden their research beams on aspects not addressed in this study.

### **Contributions to Knowledge**

- i. The study employed actual patient data from the Federal Medical Centre Jabi, Abuja, to forecast positive cases of Lassa fever disease using Support Vector Machine algorithm.
- ii. Several machine learning classification methods were trained using a k-fold cross-validation method, it was discovered that SVM is the most effective method for predicting positive cases of Lassa fever diseases.
- iii. A fully automated machine learning model was developed for predicting Lassa fever disease based on symptoms.

## Future Research

While this research contributes valuable insights in predicting the presence of Lassa fever in patients, several areas warrant further explorations. It is therefore advised that future researchers should consider the following areas:

- i. Explore more diverse data sources in order to enhance the robustness and accuracy of the predictive model. One of the challenges faced when developing this model is lack of robust data. There were only 200 data instances in the dataset which drastically reduced the performance of the model it was established that the more the robustness of the dataset the more accurate the prediction and vice versa.
- ii. Develop a real-time monitoring system that continuously updates the predictive model with new data, enabling timely identification of potential Lassa fever outbreaks thereby allowing swift intervention strategies.
- iii. The system should be developed in such a way that it can be able to prescribe medicine based on the type and extent of the disease.

## REFERENCES

- Adeboye, K.R., & Haruna M. (2015). A mathematical model for the transmission and control of malaria and typhoid co-infection using SIRS approach. *Research Journal of Mathematics*, 2(2), 1-24.
- Adedeji, A., & Smith, S. (2016). The influence of the Nigerian newspaper on Lassa fever reportage. *International Journal of Community Health*, 11, 76–86.
- Abiola, H.M., Alaba, A.A., & Joy, D.L. (2017). Expert System for Lassa Fever Diagnosis Using Rule-Based Approach. *Annals. Computer Science Series*, 15(2), 68-74.
- Abiola, O.O., Idowu, P.A, & Akinyemi, I.O. (2020). Development of a Fuzzy Logic Based Predictive Model for The Risk of Lassa Fever. *Journal of Vector Borne Diseases*, 3(11), 28-32.
- Ajenaghughrure, I.B., Sujatha, P., & Akazue, M.I. (2017). Fuzzy Based Multi-Fever Symptom Classifier Diagnosis Model. *International Journal of Informational Technology & Computer Science*, 9(10), 13-28.
- Aminu, E.F., Abiodun, A.A., Isah, O.R., Iliyasu, A., Isah, A.O., & Hussaini, A.Z. (2018). A Diagnosis System for Lassa Fever and Related Ailments Using Fuzzy Logic. *Journal of Science, Technology, Mathematics and Education (JOSTMED)*, 4(2), 18-30.
- Amarosa, V., MacNeil, A., McConnell, R., Patel, A., Dillon, K.E., Hamilton, K., Erickson, B.R., Campbell, S., Knust, B., Cannon, D., Miller, D., Manning, C., Rollin, P.E., & Nichol, S.T. (2010). Imported Lassa fever, Pennsylvania, USA. *Journal of Emergency Infectious Diseases*, 16, 1598–1600.
- Alile, S.O. (2020). Application of Bayesian Networks in early diagnosis of cerebral malaria and mosquito-borne diseases based on observed symptoms. *International Journal of Scientific Research & Multidisciplinary Studies*, 6(6), 1-14.
- Batta M., (2018). Machine Learning Algorithms - A Review. *International Journal of Science and Research*, 6(2), 11-14.

- Djam, X.Y., Wajiga, G. M., Kimbi, Y.H., & Blamah N.V. (2011). A Fuzzy Expert System for the Management of Malaria. *International Journal of Pure Applied Science and Technology*, 5(2), 84-108.
- Egwali, A., & Obi, J.C. (2015). An Adaptive Neuro-Fuzzy Inference System for Diagnosis of EHF. *Pacific Journal of Science & Technology*, 16(1), 251-261.
- Eze, K.C., Salami, T.A., Eze I.C., Pogoson, A.E., Omordia, N., & Ugochukwu, O.E. (2010). High Lassa Fever Activity in Northern Part of Edo State, Nigeria: Re-Analysis of Confirmatory Test Results. *Journal of Emergency Infectious Diseases*, 17, 1698– 1700.
- Govinda, S.R., Eswara, R.D., (2013). Siva Prasad Diagnosis Rule-Based Expert Systems. *Internal Journal of Engineering Research & Technology*, 2(8), 551-561.
- Hambali, M.A., Akinyemi, A.A., & Luka, J.D. (2017). Expert System for Lassa Fever Diagnosis Using Rule Based Approach. *Anale. Seria Informatică*. 8(2), 254-259.
- Holland, J.J., & Domingo, E. (1997). RNA Virus Mutations and Fitness for Survival. *Annual Review Microbiology*. 51, 151-78
- Igodan, C., Akinyokun, O., & Olatubosun, O. (2013). Online Fuzzy-Logic Knowledge Warehousing and Mining Model for The Diagnosis and Therapy of HIV/AIDS. *International Journal of Computational Science and Information Technology (IJCSITY)*, 1(3), 27-41.
- Ijeoma, E.C., Ajobogah C., Oladele, O.A., Ahmed L.A., & Adesu L.A., (2021). Lassa Fever Outcomes and Prognostic Factors in Nigeria (LASCOPE): A Prospective Cohort Study. *The Lancet Journal of Global Health*. 9(4), 469-478.
- Imanov, E., & Asengi, F.J. (2020). Artificial Intelligence for Lassa Fever Diagnosis System. In Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., & Sadikoglu, F.M. (2020): 14th International Conference on Theory and Application of Fuzzy Systems and Soft Computing. *Springer Nature*, pp: 540-547.
- Lecompte, E., Fichet-Calvin, E., Daffis, S., Koulemou, K., & Sylla, O. (2006). *Mastomys Natalensis* and Lassa fever, West Africa. *Journal of Emergency Infectious Diseases*, 12, 1971-1974.
- Marsland, S. (2015). *Machine Learning: An Algorithmic Perspective*. Florida, USA: CRC press.
- Mohssen, M., Muhammad, B. K., & Eihab B. (2017). *Machine Learning Algorithms and Applications*. London: CRC Press.
- Nnebe, S.E., Okoh, M., Otumu, J.M., & Oshoibhor, E.O. (2019). A Neuro-Fuzzy Case-Based Reasoning Framework for Detecting Lassa Fever Based on Observed Symptoms. *American Journal of Artificial Intelligence (AJAI)*, 3(1), 9-16.
- Olubi, O., Oniya, E., & Owolabi, T. (2021). Development of Predictive Model for Radon Estimation in The Atmosphere Using Stepwise Regression and Grid Search Based-

- Random Forest Regression. *Journal of the Nigerian Society of Physical Sciences*, 2, 132-139.
- Oguntimilehin, A., Abiola, O.B., & Olatunji, K.A. (2015). Computer Aided Diagnostic Systems for Managing Typhoid Fever: A Review of Diagnosis Techniques. *International Journal of Computer Applications*, 2(3), 126-128.
- Oladipo, O., Olayinka, C.T., & Popoola, O.L. (2014). Mobile Compactable Expert System for The Treatment of Typhoid Fever in Developing Countries. *International Journal of Computer Applications*, 5(2), 16-19.
- Ogbu, O., Ajuluchukwu, E., & Uneke, C.J. (2004). Lassa fever in West African sub-region: An Overview. *Journal of Vector Borne Disease*, March 2007, pp. 1–11.
- Olawole, A.S., & Knonyana, T., (2022). Forecasting Lassa Fever Outbreak Progression with Machine Learning. *International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. 3(2), 12-16.
- Olutola, A., & Olufunke, C.O. (2020). Prediction of Mortality Rate of COVID-19 Patients using Machine Learning Techniques in Nigeria. *International Journal of Computer Science and Software Engineering*, 9(5), 30-34.
- Omilabu, S.A., Badaru, S.O., Okokhere, P., Asogun, D., Drosten, C., & Emmerich, P., (2012). Lassa fever, Nigeria. *Journal of Emergency Infectious Diseases*. 11, 1642–1644.
- Osarumwense, A.S., & Eromosele, B.M. (2020). A Machine Learning Based Approach for Detecting Dengue Haemorrhagic Fever. *International Journal of Academic & Applied Research*, 4(5), 13-25.
- Osaseri, R. O., Onibere, E. A., & Usiobiafo, A. R. (2014). Fuzzy Expert Model for Diagnosis of Lassa Fever. *Journal of the Nigerian Association of Mathematical Physics*, 27, 533-540.
- Osaseri, R.O., & Osaseri, E.I. (2016) Soft Computing Approach for Diagnosis of Lassa Fever. *International Journal of Engineering & Computer Science*. 10(11), 67-69.
- Osigbemeh, M.S., Ogunwolu, F.O., Omoare, A.A., & Inyama, H.C. (2014). A linguistic fuzzy expert system for contagious diseases detection and isolation. *Journal of Material Science and Technology*, 2(1), 1-6.
- Putu, M. P., & Iketut, G. D. (2012). Fuzzy Knowledge-Based System with Uncertainty for Tropical Infectious Disease Diagnosis. *International Journal of Computer Science Issues*. 2(9), 157-163.
- Sarker, H. I. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*. 2(3) 1–21.

- Samuel, E.N., Ozemoya, A.O., Adetokunbo, M.G., & Emmanuel, O.O. (2019). A Neuro-Fuzzy Case Based Reasoning Framework for Detecting Lassa fever Based on Observed Symptom. *American Journal of Artificial Intelligence*. 3(1), 9-16.
- Samuel, O.W., & Omisore, M.O. (2013). Genetic Neuro-Fuzzy System for The Diagnosis of Typhoid Fever. *Conference on Medical Innovation and Computing Service*, Pp. 1-13.
- Samson, K.T., Akingbade, T. & Orija, J., (2023). Comparative Analysis of Mortality Predictions from Lassa Fever in Nigeria: A Study Using Count Regression and Machine Learning Methods. *Acadlore Transactions on AI and Machine Learning*. 2(4), 204-211.
- Steur, N.A., & Mueller, C. (2019). Classification of viral hemorrhagic fever focusing Ebola and Lassa fever using neural networks. *International Journal of Machine Learning & Computer*, 9(3), 334-343.
- Sunil, V.A. (2021). Diseases Prediction Model using Machine Learning Technique. *International Journal of Scientific Research in Science and Technology*, 8(2), 121-123.
- Tunmibi, S., Adeniji, O., Aregbesola, A., & Dasyilva, A. (2013). A Rule Based Expert System for Diagnosis of Fever. *International Journal of Advanced Research*. 1(7), 343-348.
- Umoh, U.A., & Ntekop, M.M. (2013). A Proposed Fuzzy Framework for Cholera Diagnosis and Monitoring. *International Journal of Computer Applications*, 82, 1-10.
- Uneke, C.J. (2012). Lassa virus infection of rhesus monkeys: pathogenesis and treatment with ribavirin. *Journal of Infectious Diseases* 141: 580–9.
- Vamsi, B., Debnath, B.A., & Divya, M. (2020). Prediction of Brain Stroke Severity Using Machine Learning. *International Journal of Information and Engineering Technology*, 71(1): 25-32.
- Yuan, P., Bartlam, M., Lou, Z., Chen, S., & Zhou, J. (2009). Crystal Structure of An Avian Influenza Polymerase Pa(N) Reveals an Endonuclease Active Site. *Nature* 458: 909–913.