

Original Article

Artificial Intelligence for Symptomatic Detection and Classification of Suspected Cases of Filovirus Disease in Africa: An Explorative Approach

Amadi Gloria Yaa¹, Omotosho Olawale J², Ebiesuwa Oluwaseun³, Olujimi Alao⁴, Oludele Awodele⁵, Eze Monday O⁶, Kuyoro Afolashade⁶, Maitanmi Olusola S⁷, Kalesanwo Olamide⁸, Adigun Taiwo O⁹, Ow'honda Golden¹⁰, Ngofa Reuben¹¹

^{1,2,3,4,5,6,7,8,9,10,11}Department of Computer Science, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

¹Corresponding Author : oyebodeaduragbemi@gmail.com

Received: 12 September 2023

Revised: 19 October 2023

Accepted: 04 November 2023

Published: 18 November 2023

Abstract - Misdiagnosis exists as an inevitable human error factor in healthcare operations; diagnostic errors contribute to numerous complications that result in a high death rate. The problems of misdiagnosis and delays in disease identification, especially highly infectious diseases like Filovirus disease, have drastically increased the death rate, especially during the outbreak, according to International Health Organizations and Statistics on Improving Diagnosis in Healthcare. This overwhelming challenge has contributed majorly to many complications that increase casualties, especially during the Filovirus disease outbreak. However, studies conducted in this area have identified some technological drawbacks associated with computational complexities, such as the curse of dimensionality, overfitting, weight and bias, etc., as major challenges in implementing machine learning techniques. Emerging solutions are recommended to better improve healthcare delivery in this domain. This study aimed to develop a model that can be used to provide precise symptomatic detection and classification of suspected cases of filovirus disease with reference to the Ebola Virus and Marburg virus disease. The study obtained the filovirus dataset from multiple sources involving states, Federal and International Health Organizations over time within Africa. The unstructured data collected was preprocessed on Sklearn with Pandas NumPy and split into 8:2 as train and test datasets. Machine Learning Algorithms like Logistic Regression, Linear Discriminant Analysis, Decision Tree, Support Vector Machine, Random Forest, etc were adopted by the Scikit-learn library for Classification analysis. The Confusion Matrix was implemented as a basic ML Statistical analytical framework, and the Accuracy, Precision, Recall, F1-score and a few other metrics were applied for the model Performance Evaluation. The experimental result of the study obtained the highest accuracy of 96% ratio for Logistic Regression, followed by the Linear Discriminant Analysis and Naïve Bayes that offered relatively 95% and 93% rates of Accuracy performance. This outcome suggests that the best-recommended algorithm suitable for the objective of this study is the Logistic Regression model. However, as discovered, the Logistic Regression can be recommended with Feature Selection methods to provide a better dimensional reduction model for future improvement.

Keywords - Artificial Intelligence, Machine Learning Algorithms, Classification Model, Symptomatic Detection of Suspected Cases, Filovirus diseases, Ebola Virus Disease, Marburg Virus Disease.

1. Introduction

In order to be ready for any unplanned and essentially unavoidable outbreak, there should be proper management intervention measures in place for preparedness for any global health emergency on highly infectious diseases like Filovirus disease, especially in Nigeria and Africa as a continent. [1] [2]. Filovirus Disease embodies the family of Filoviridae - the Single-Stranded Negative-Sense RiboNucleic Acid (SSNS - RNA) Viruses that are categorized as highly infectious and deadly diseases, which poses a serious threat to Global Public Health and often lead to high fatality rate in fewer outbreaks [3] [4]. Ebola Virus Disease (EVD) and Marburg Virus

Disease (MVD) are common viruses classified under the Filovirus disease that share relatively similar clinical characteristics, which are capable of causing the onset of severe haemorrhagic fever in human and non-human primates [5] [6]. Over 50 filovirus disease outbreaks associated with Ebolavirus and Marburgvirus have occurred globally following the emergence of Ebola Virus Disease (EVD) in 1967 and Marburg Virus Disease (MVD) in 1976 respectively [6] [7]. Filovirus diseases should be early and accurately detected, speedily responded to and appropriately treated to reduce the spread and provide improved medical intervention and control measure, which can drastically prevent other



predominant health complications that may lead to death [2] [6]. Due to the life-threatening nature of Public Health Epidemiology, highly infectious diseases like Ebola and Marburg Virus Disease demand a higher level of early diagnosis, proper attention, proactive intervention and prompt response during any outbreak [8] [9] [4].

Disease resurgence and concurrent and/or multiple virus disease outbreaks are inevitable, especially as it concerns viruses that share similar clinical characteristics like the Ebola Virus and Marburg Virus Diseases [10] [11]. Most recently, the World Health Organization (WHO) declared the first newly confirmed cases of Marburg Virus Disease (MVD) outbreak in Ghana as a Global Health Emergency [12] [13]. Similarly, in 2021, the WHO formally declared the first confirmed case of Marburg Virus disease, following the just reported confirmed case of Ebola Virus Disease in Guinea.

Statistically, the incidence rate recorded a high Case Fatality Ratio (CFR) of about 28,616 Ebola Virus cases, of which about 11,319 are reported death cases, and while the Marburg Virus is more than 252 cases, out of which 227 are reported deaths in Africa [14]. The occurrence of these two Filovirus Diseases that share similar symptoms relatively in the same country, period and year can be clinically confusing and challenging to distinguish promptly and provide adequately certain suitable medical care and rapid intervention management during the concurrent outbreaks [15], [9], [16].

Artificial Intelligence(AI) techniques are making remarkable impacts and penetrating vastly into the next-generation unknown, and their usefulness in controlling and managing many diseases in healthcare service delivery continues to surge across developing societies, [12, 17-19]. Innovative technological advancement continues to traverse enormously in many developed countries of the world, and far-reaching research is deepened across many developing societies globally [20], [21]. Consequently, there is a lack of preemptive measures, early disease detection, prompt reaction, appropriate diagnosis instruments, and adequate healthcare service delivery [2], [11], [13].

The early detection of virus diseases, classification and analysis of disease symptoms, and prediction of highly infectious diseases like Filovirus diseases with the use of hybrid technologies are emerging and fascinating in this era [22], [24]. The use of AI-based Machine Learning techniques capable of data classification analysis, prediction and performance evaluation are enormously the basis of this study [25], [26]. This study intends to design and develop a symptomatic detection and classification system that allows for predicting EVD and MVD of Filovirus disease. The research also involves using machine learning algorithms and performance evaluation methods to evaluate the general performance of the proposed model.

2. Literature Review

2.1. Artificial Intelligence (AI) in Healthcare

The widespread application of artificial intelligence approaches in healthcare is extremely helpful in preventing and treating a wide range of illnesses. AI offers great potential for enhancing conventional service delivery, particularly regarding healthcare [3], [7]. Danai & Henri (2021) claim that AI-based algorithms that use Natural Language Processing (NLP) can discover information about potential future outbreaks through social media platforms, such as Google Trends, with the potential outbreak zone and fatality cases. As a result, the NLP can be programmed to read, comprehend, and efficiently extract meaningful human languages and conversations. [27].

The ability of a system to correctly understand external input, to learn from that data, and to use those learnings to accomplish particular goals and activities through flexible adaptation is known as artificial intelligence (AI) [26]. Carroll & Whende (2021) claim that Artificial Intelligence (AI) has a lengthy history in many industries and has revolutionized how people, the technology's intended users, communicate, conduct financial and social transactions and access information. Containment to prevent the spread of pandemics is one of the areas where artificial intelligence has been deployed and has a big impact on society. [28].

Infectious disease pandemics, which are widespread outbreaks that can increase morbidity and mortality over a large geographic area and cause significant economic, social, and political disruption, do indeed present special challenges because of their rapid spread and complications, which calls for a quick response on many fronts, from diagnostic methods to drug development to resource allocation and planning in the medical field. In its capacity to slow down and restrict the virulent spread of illness during a pandemic, AI has demonstrated great potential. Artificial Intelligence provides new confidence to not only preempt, prevent and control the spread of infectious diseases but also helps scientists in faster clinical trials and identify effective drug discovery for tackling the diseases and predicting future outbreak occurrences [29].

The AI-based tools of Machine Learning Algorithms have increasingly demonstrated their ability to predict the spread of virus disease as well as in advance use and help in the policy-making to plan, prioritize, and implement diverse, innovative interventions, exemplified during the recent global COVID-19 Lockdown, based on the severity of the virus spread. In addition, AI also helps to check whether the public adheres to the preventive actions taken by policy-makers [30].

The preparedness of healthcare system capabilities to handle any re-emerging disease outbreaks remains the need of the hour. Thus, the Practical application of Artificial Intelligence to contain disease outbreaks can be seen in the

case of the COVID-19 pandemic, where AI-powered technologies and methods were used and are still being adopted to contain the spread of the pandemic. AI has tremendously been very useful in accelerating the development of subunit treatment and nucleic acid vaccines [7], [31]. Furthermore, AI must go in hand with big data analytics and cloud computing to develop viable, thoroughly validated tools that can provide real support to local healthcare providers across many health sectors [32], [33]. Finally, Artificial Intelligence in the form of Machine Learning Techniques are employed in this study to classify the data from suspected cases and predict future outcomes of the virus during any outbreak of Filovirus Disease [12], [34].

2.2. Machine Learning

Machine Learning seeks to build computer software applications that satisfy performance improvement. A focal point of ML is the ability to learn from experience without human intervention. It is a subset of Artificial Intelligence that aims to design systems/models that learn to improve performance given the data [35]. The ability of a computer device to automatically make decisions based on a given data and experience without explicitly being programmed is Machine Learning. The Learning Algorithm involves a process whereby a computer is designed to continually train models and classify test data for improved performance over time [30], [36]. The following discusses the two main types of Machine Learning:

2.3. Supervised Learning

Supervised learning is a type of machine learning that uses labeled data to train models. The output is already known in labelled data, and the model just needs to map the inputs to the respective outputs. This is by looking at a label as an instance of a successful prediction, which the algorithm learns from. This means that target variables are required to use labeled data to train machine learning models. Some of the most popularly used supervised learning algorithms in healthcare and medical administrations are Logistic Regression, Support Vector Machine, and K-Nearest Neighbor. Decision Tree, Random Forest, Naive Bayes and a few others [35], [37].

2.4. Unsupervised Learning

Unsupervised learning is a type of machine learning that uses unlabeled data. The unlabeled data has no fixed target output variable; the model learns from the data, discovers patterns and features, and returns the output. [23], [35]. This type of model is very useful in classification tasks, clustering algorithms, anomaly detection and generative tasks. The model is capable of finding hidden associations among objects, resolving every segmentation problem and many others [30], [38]. The training process used in the unsupervised learning techniques requires no supervision to build models. They learn on their own and predict the output. [35].

2.5. Machine Learning Methods: Supervised Learning Techniques

2.5.1. Classification Method

The classification method in the supervised learning technique is a method of finding a function that helps divide the dataset into categories and/or classes based on different parameters or factors. Since a computer program is trained on the training dataset and based on that training, it categorizes the data into different classes. The task of the classification algorithm is to find the mapping function to map the input(x) to the discrete output(y). [3], [39], [23], [22]. Classification is the task of predicting a discrete class label.

2.5.2. Regression

Regression algorithms describe numerical targeted values in place of classes. This is done by estimating numerical variables, and these algorithms are powerful at predicting product demand, sales figures, marketing returns, etc. It involves the practice of finding the connections between dependent and independent variables [37], [40]. Regression fosters continuous variables like the prediction of business or market trends, customers' needs, weather forecasts, etc. Usually, this kind of model is trained with historical data. It focuses on finding the symbolic function in order to map the input variable(x) to the constant output variable (y), respectively, [13], [41]. Regression is the task of predicting a continuous quantity. [42]–[44].

3. Machine Learning Algorithms

3.1. Logistic Regression

Logistic Regression is one of the categories of the Supervised Learning Techniques based on statistical analysis commonly used to predict binary data representation involving Yes (1) or No (0) in the nature of categorical dependent variables, given the set of independent variables. It is very useful in analyzing the relationship between one or more existing data-independent variables by predicting the dependent variable [45]. The Logistic Regression is simply used to predict the probability of a target variable. That, in a nutshell, means it learns the linear relationship between a given dataset and, from the learning, suggests a Sigmoid function as non-linearity; it uses a given set of labeled data to learn the relationship and categorizes the data into discrete classes. It is often used in solving classification problems like Disease prediction and Spam mail detection [30], [46], [47].

3.2. Support Vector Machine

The SVM is one of the Supervised Learning types of Machine Learning Algorithms used for data classification and regression analysis. It is usually applied in high-dimensional cases and effectively solves linear and non-linear practical problems. SVM utilizes a relatively good amount of specified custom kernels and is capable of creating hyperplanes that help in separating data into different classes [25], [29], [48]. In simple terms, the SVM is used to map data points to high dimensional cases and further finds the optimal hyperplane

that helps in dividing the data into classes for better productivity.

3.3. K-Nearest Neighbor

The K-Nearest Neighbors is also part of the Supervised Learning kind of Machine Learning Algorithms often used to solve problems of Regression and Classification of data analysis. The KNN assumes that there can exist close proximities in data classification and regression processes; hence, it finds the most basic similarity in proximity within a given dataset [45]. This, in essence, suggests that KNN focuses on using labeled input datasets to forecast the output of the data points. The KNN decision is based on the number of nearest neighbors.

3.4. Decision Tree

The Decision Tree is another type of supervised learning algorithm that can be used for both classification and regression in data analysis. It is a flowchart represented in a tree-like structure, denoted with an internal node, which signifies attribute and a leaf node, which signifies terminal node. The decision tree represents the probability of the outcome of a decision in a tree-like structure [23], [37]. A decision tree is a diagram in the shape of a tree used to choose a course of action. Each branch of the tree represents a potential choice, event, or response. Based on the supplied training samples, a standard Decision Tree is constructed [38],[31], [40].

3.5. Random Forest

The Random Forest is one of the plethora of classification and regression algorithms of Supervised Learning in Machine Learning used to solve complex problems. It is an algorithm that involves ensemble learning and utilizes the process of decision trees to provide solutions to difficult challenges. It acts like a tree predictor that depends mainly on the random vector values of every individual block of trees to predict the outcome. Random Forest is designed with decision nodes, leaf nodes and root nodes. It is an advanced Decision Tree Algorithm that uses building blocks to provide an improved decision support system in the form of a tree-like structure [5], [27], [45].

3.6. Naive Bayes

The Naïve Bayes is a probabilistic classifier most suitable for big data analytics and founded on the Bayesian Theorem. It is a flexible classification method that beats many other sophisticated classifiers. The Naïve Bayes works by assigning the membership of an object to a given class and further articulating the posterior probability. It is often used in text classification and real-time medical data classification and prediction [23], [37].

4. Related Works and Technology

Ebrahim et al. [49] developed a machine-learning classification model for Chronic Kidney Disease diagnosis

using Recursive Feature Elimination Techniques. The study obtained 400 patient data with 24 features and applied four classification algorithms to achieve better performance evaluation of the model. The experimental result provided a significant level of accuracy, precision, recall and F1-score. Umang et al. [48] used the ML approach to analyze and improve the accuracy of the prediction model for Ebola disease with many input symptoms on the patient's early condition. The hybrid model applied Decision Tree, KNN, SVM, LR, Boosting and Bagging classifiers, etc, and the highest accuracy of 96% was derived from KNN and Random Forest Classifiers with Neural Network as a hybrid model. Ahsan et al. [50] developed a COVID-19 differential diagnosis model that applied Multilayer Perception and Convolutional Neural Network (ML-CNN) with mixed data involving numerical, categorical and image data to offer efficient and rapid detection of the virus in patients for early diagnosis that leads to timely treatment intervention. Structured and unstructured datasets were considered with various optimization algorithms. The result showed that the trained model could distinguish between COVID-19 and non-COVID-19 patients with a greater overall accuracy of 94% and 96%, respectively. Gomathy and Naidu [37] designed a disease prediction system based on disease symptoms using Machine Learning Algorithms. The system was successfully implemented as a web application with a grails framework, which obtained an average accuracy of 98% in prediction [35].

5. Methodology

5.1. Data Acquisition and Collection

The relevant existing symptomatic data was obtained ethically from multiple sources that involved State, Federal and International Health Organizations and Corporate Partners within Africa. The preliminary data collected involved structured and unstructured datasets that included numerical and categorical forms on EVD and MVD of the Filovirus disease. The data elements are the defined EVD and MVD symptoms like high temperature, hemorrhagic fever, headache, vomiting, diarrhea, nausea, abdominal pain, etc. The data size was considerably large in the range of collections that involved suspected, probable and confirmed cases.

5.2. Data Extraction

The datasets gathered for this investigation were tabular data that needed to be converted because they contained numerous repeating and unstructured columns. It includes 21 days' worth of temperature data, which must be synchronized to provide a specific column for the temperature reading, just like the virus's other symptoms. To get the average percentage of the overall temperature reading, this dataset section was represented as the total sum of the 21-day temperature readings divided by 21, which is the number of days. The value of the average temperature reading was further modified to show values less than or equal to 36.9 as negative (0) and

values greater than or equal to 37.0 as positive (1), respectively, for binary classification.

5.3. Data Preprocessing

The data from many sources used in this study included information that was missing or unavailable, which was deemed an unstructured dataset that needed a substitute value to prevent bias and the loss of important data. The Data

Imputation method replaced the missing values with the Most Frequent-Value in the column for the categorical and numerical data variables. In an effort to balance the dataset between the two disease kinds, all the records that included NIL were subsequently removed. The Scikit-Learn Library's Pandas and Numpy tools were used to clean and preprocess the data. The data was divided into train and test datasets in 8:2 ratios.

6. Proposed Design

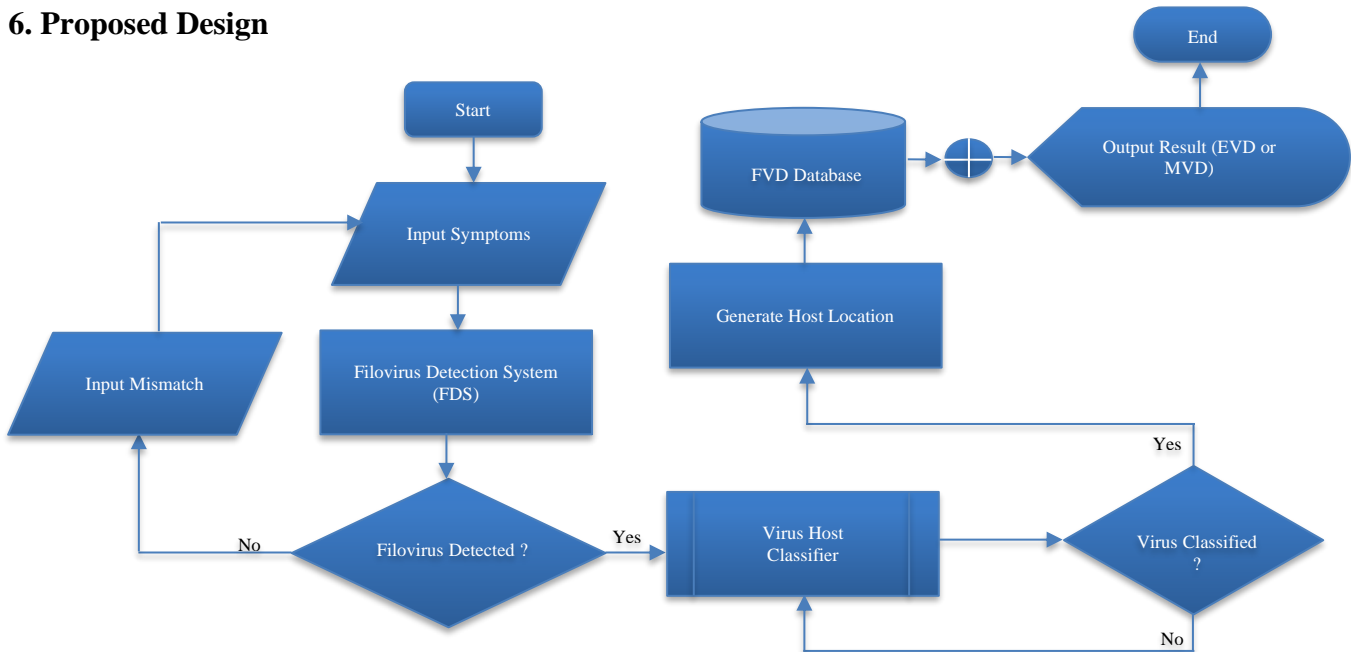


Fig. 1 Diagram of the proposed model

6.1. Analysis of the Proposed Model

The proposed model is logically designed such that the virus can be predicted in an individual only if it returns a given high-level proportion of the input data, which can trigger either a positive or negative response as a detection signal of the filovirus disease upon receiving the data, the model first checks to ascertain the proportion of data required based on a

predefined threshold in order to determine the existence of the virus and further categorize the observed symptoms to classify the type of filovirus disease exist in such individual, in this case, it can either be Ebola virus or Marburg virus disease; generate location and communicate result to Stakeholders for prompt response.

6.2. Experimental Results: Compared Model

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)	
lr	Logistic Regression	0.9580	0.9698	0.6759	0.9323	0.9426	0.9196	0.9215	2.1200
ridge	Ridge Classifier	0.9577	0.0000	0.6644	0.9203	0.9386	0.9185	0.9211	0.0260
lda	Linear Discriminant Analysis	0.9553	0.9724	0.6698	0.9291	0.9390	0.9143	0.9164	0.0780
nb	Naive Bayes	0.9375	0.9779	0.6767	0.9298	0.9241	0.8802	0.8822	0.0290
lightgbm	Light Gradient Boosting Machine	0.9046	0.9739	0.9337	0.9049	0.9046	0.8222	0.8225	0.6480
qda	Quadratic Discriminant Analysis	0.8953	0.9594	0.6322	0.8724	0.8816	0.7996	0.8013	0.0490
gbc	Gradient Boosting Classifier	0.8839	0.9718	0.6508	0.8845	0.8720	0.7773	0.7795	3.4270
ada	Ada Boost Classifier	0.8694	0.6937	0.9094	0.8702	0.8694	0.7566	0.7573	0.4010
et	Extra Trees Classifier	0.8310	0.9126	0.5762	0.7994	0.8143	0.6743	0.6772	1.0730
rf	Random Forest Classifier	0.8084	0.8934	0.5602	0.7782	0.7920	0.6304	0.6339	0.5170
svm	SVM - Linear Kernel	0.7393	0.0000	0.5383	0.8181	0.7278	0.5384	0.5918	0.1350
dt	Decision Tree Classifier	0.5456	0.5866	0.3950	0.5554	0.5502	0.1652	0.1655	0.0560
dummy	Dummy Classifier	0.4922	0.5000	0.3333	0.2423	0.3247	0.0000	0.0000	0.0300
knn	K Neighbors Classifier	0.4820	0.5174	0.3557	0.4743	0.4743	0.0082	0.0082	0.2880

Fig. 2 The Compared model

6.3. Experimental Results: Cross-Validation of the Model

Model Scores								Model Fine-Tuned							
	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC		Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Fold								Fold							
0	0.9536	0.9665	0.6758	0.9308	0.9406	0.9114	0.9128	0	0.9625	0.9649	0.6818	0.9639	0.9457	0.9279	0.9303
1	0.9607	0.9688	0.6667	0.9230	0.9415	0.9243	0.9270	1	0.9554	0.9706	0.6630	0.9228	0.9388	0.9144	0.9164
2	0.9607	0.9761	0.7084	0.9474	0.9498	0.9250	0.9264	2	0.9554	0.9732	0.6629	0.9228	0.9388	0.9144	0.9164
3	0.9607	0.9758	0.6667	0.9230	0.9415	0.9243	0.9270	3	0.9625	0.9768	0.7097	0.9514	0.9510	0.9283	0.9299
4	0.9464	0.9731	0.6569	0.9224	0.9343	0.8981	0.8991	4	0.9589	0.9676	0.6655	0.9229	0.9406	0.9210	0.9234
5	0.9607	0.9650	0.6806	0.9442	0.9447	0.9246	0.9268	5	0.9589	0.9651	0.6793	0.9376	0.9437	0.9213	0.9232
6	0.9589	0.9620	0.6655	0.9230	0.9406	0.9211	0.9235	6	0.9607	0.9607	0.6806	0.9443	0.9447	0.9246	0.9268
7	0.9589	0.9709	0.6794	0.9376	0.9437	0.9213	0.9232	7	0.9554	0.9700	0.6769	0.9322	0.9417	0.9147	0.9162
8	0.9607	0.9703	0.6945	0.9458	0.9474	0.9248	0.9266	8	0.9589	0.9699	0.6794	0.9376	0.9437	0.9213	0.9232
9	0.9589	0.9699	0.6642	0.9262	0.9422	0.9209	0.9229	9	0.9571	0.9690	0.6630	0.9261	0.9413	0.9176	0.9194
Mean	0.9580	0.9698	0.6759	0.9323	0.9426	0.9196	0.9215	Mean	0.9586	0.9688	0.6762	0.9362	0.9430	0.9206	0.9225
Std	0.0044	0.0043	0.0149	0.0099	0.0040	0.0081	0.0085	Std	0.0026	0.0043	0.0135	0.0131	0.0035	0.0050	0.0051

Fig. 3 The model Cross-validation scores

7. Analysis of Results

The research provided a different relative means of validating the model performance with the supervised machine learning algorithms used for disease classification and

prediction, and this is to ascertain the appropriate algorithm for the research. Thus, based on the experiments, the study concluded that the best-recommended algorithm suitable for accuracy performance is Logistic Regression, which offered the greatest accuracy of 96% ratio.

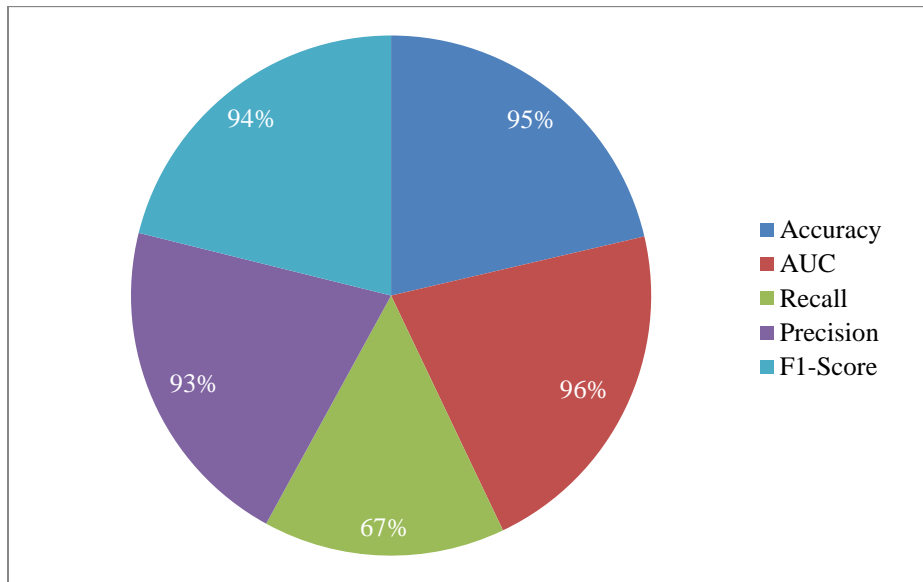


Fig. 4 Logistic regression model performance

8. Discussion of Findings

The input data is a large dataset with many features coming from similar clinical characteristics and many distinctive symptoms of EVD and MVD diseases, which resulted in correlated variables. The situation brought about the need for data extraction and preprocessing to minimize the difficulties for the training set to be visualized, data complexity, and achieve greater computational time. And also

eliminate some redundant factors that can render poor performance of the model. Also, the sigmoid function allowed for the transformation of the input variables into a probability value of 0 and 1. This model is considered when the output depends on the sum of the input parameters. Hence, the outcome is mostly based on the probability that a sample size belongs to a class A(1) or class B(0). For example, a class of symptoms belonging to the target virus A(EVD) or a class of

symptoms belonging to a target virus B(MVD) respectively. Thus, the logistic regression classifier model maps the symptom values as a set of independent variables input into a value between 0 and 1 as EVD and MVD. The Machine Learning statistical approach provides a reliable measure for which different parameters can be well calculated to achieve better results for a given model performance. Binary classification metrics selected are the Accuracy, Precision, Recall, F1-Score, ROC-AUC, Kappa and MCC metrics based on the goal that this study experiment intends to achieve. The Confusion Matrix is a key player in ascertaining more scientific data analysis and presentation was used to create a two-dimensional table that displays actual and expected values, distinguishing between infected and non-infected patients. The four key components of the Confusion matrix, namely True Positive, False Positive, True Negative, and False Negative, of statistical measures were employed to determine the outcome of model performance.

References

- [1] V. Oluigbo Ikenna et al., “Revolutionizing the Healthcare Industry in Nigeria: The Role of Internet of Things and Big Data Analytics,” *International Journal of Scientific Research in Computer Science and Engineering*, vol. 5, no. 6, pp. 1–12, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Testimony J Olumade et al., “Infectious Disease Outbreak Preparedness and Response in Nigeria: History, Limitations and Recommendations for Global Health Policy and Practice,” *AIMS Public Health*, vol. 7, no. 4, pp. 736–757, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Romany Fouad Mansour et al., “Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems,” *IEEE Access*, vol. 9, pp. 45137–45146, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Brigitte Beer, and Reinhard Kurth, “Characteristics of Filoviridae: Marburg and Ebola Viruses,” *Springer-Verlag*, vol. 86, pp. 8-17, 1999. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Ayushe Gangal et al., “Prediction Models for Healthcare Using Machine Learning: A Review,” *Handbook of Research on Disease Prediction Through Data Analytics and Machine Learning, IGI Global*, pp. 70–91. 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Sylvester Languon, and Osbourne Quaye, “Filovirus Disease Outbreaks: A Chronological Overview,” *Virology: Research and Treatment*, vol. 10, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Sungho Sim, and Myeongyun Cho, “Convergence Model of AI and IoT for Virus Disease Control System,” *Personal and Ubiquitous Computing*, vol. 27, pp. 1209-1279, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Aaron G. Buseh et al., “The Ebola Epidemic in West Africa: Challenges, Opportunities, and Policy Priority Areas,” *Nursing Outlook*, vol. 63, no. 1, pp. 30–40, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Jens H. Kuhn et al., “New Filovirus Disease Classification and Nomenclature,” *Nature Reviews Microbiology*, vol. 17, no. 5, pp. 261–263, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Güler Kalem, and Çiğdem Turhan, “Mobile Technology Applications in the Healthcare Industry for Disease Management and Wellness,” *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 2014–2018, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Daniel Tom-Aba et al., “Innovative Technological Approach to Ebola Virus Disease Outbreak Response in Nigeria Using the Open Data Kit and Form Hub Technology,” *PLoS ONE*, vol. 10, no. 6, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Devy M. Emperador et al., “Diagnostics for Filovirus Detection: Impact of Recent Outbreaks on the Diagnostic Landscape,” *BMJ Global Health*, vol. 4, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Grit Schubert et al., “The African Network for Improved Diagnostics, Epidemiology and Management of Common Infectious Agents,” *BMC Infectious Diseases*, vol. 2, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Bagbe Atinuke, Badejo Oduyomi Micheal, and Ayodeji Samson Bagbe, “Statistical Analysis of Ebola Virus Disease outbreak in Some West Africa Countries using S-I-R Model,” *Annals of Biostatistics & Biometric Applications*, vol. 2, no. 3, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] E.M. Leroy, J.P. Gonzalez, and S. Baize, “Ebola and Marburg Haemorrhagic Fever Viruses: Major Scientific Advances, But a Relatively Minor Public Health Threat for Africa,” *Clinical Microbiology and Infection*, vol. 17, no. 7, pp. 964–976, 2011. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

9. Conclusion

In conclusion, this study avails other researchers in this area valuable insight towards developing and implementing medical/healthcare innovative support tools for disease spread control, early detection and prediction of highly infectious diseases like Filovirus. The research also provided a partial prototype of distinguishing between clinically related diseases such as Marburg and Ebola virus disease and a geospatial way of improving contact tracing of suspected cases.

9.1. Authors' Contributions.

The main Author conceived the idea for the research area and wrote the initial draft of the manuscript. The proposed system design, problem formulation, suggestive guide and influence on the shape of the research. All authors suggested related works and discussed the structure of the paper and results. All authors read and approved the final manuscript.

- [16] Suwon Kim, and Seongcheol Kim, “User Preference for an IoT Healthcare Application for Lifestyle Disease Management,” *Telecommunications Policy*, vol. 42, no. 4, pp. 304–314, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Setiyo Budiyo et al., “Design and Monitoring Body Temperature and Heart Rate in Humans Based on WSN Using Star Topology,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 1, pp. 304-309, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Seunghoon Lee, and Young Hoon Lee, “Article Improving Emergency Department Efficiency by Patient Scheduling Using Deep Reinforcement Learning,” *Healthcare*, vol. 8, no. 2, pp. 1-17, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Arlene C. Chua et al., “The Case for Improved Diagnostic Tools to Control Ebola Virus Disease in West Africa and How to Get There,” *PLoS Neglected Tropical Diseases*, vol. 9, no. 6, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Fernando Cornelio Jiménez González et al., “Smart Multi-Level Tool for Remote Patient Monitoring Based on a Wireless Sensor Network and Mobile Augmented Reality,” *Sensors*, vol. 14, no. 9, pp. 17212-17234, 2014. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Luka Mangveep Ibrahim et al., “Electronic Reporting of Integrated Disease Surveillance and Response: Lessons Learned from Northeast, Nigeria, 2019,” *BMC Public Health*, vol. 21, 2021 [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Nataliya Boyko, and Iryna Dosiak, “Analysis of Machine Learning Algorithms for Classification and Prediction of Heart Disease,” *IDDM-2021: 4th International Conference on Informatics & Data-Driven Medicine, CEUR Workshop Proceedings (CEUR-WS.org)*, pp. 233-249, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [23] F.Y. Osisanwo et al., “Supervised Machine Learning Algorithms: Classification and Comparison,” *International Journal of Computer Trends and Technology (IJCTT)*, vol. 48, no. 3, pp. 128–138, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Franck Katembo Sikakulya et al., “Safe and Dignified Burial of a Deceased from a Highly Contagious Infectious Disease Ebolavirus: Socio-Cultural and Anthropological Implications in the Eastern DR Congo,” *One Health*, vol. 13, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Gokulnath Chandra Babu, and S.P. Shantharajah, “Survey on Data Analytics Techniques in Healthcare Using IoT Platform,” *International Journal of Reasoning-based Intelligent Systems*, vol. 10, no. 3-4, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Nitin Mittal, “Data Monitoring and Gathering in Sensor Networks for Healthcare Applications,” *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 1, pp. 3521-3539, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Madhav Erraguntla, Josef Zapletal, and Mark Lawley, “Framework for Infectious Disease Analysis: A Comprehensive and Integrative Multi-Modeling Approach to Disease Prediction and Management,” *Health Informatics Journal*, vol. 25, no. 4, pp. 1170–1187, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Afshan Hassan et al., “Gauging the Impact of Artificial Intelligence and Mathematical Modeling in Response to the COVID-19 Pandemic: A Systematic Review,” *Biomed Research International*, vol. 2022, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Md Manjurul Ahsan, Shahana Akter Luna, and ssZahed Siddique, “Machine-Learning-Based Disease Diagnosis: A Comprehensive Review,” *Healthcare*, vol. 10, no. 3, pp. 1-30, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [30] Xiaokang Zhou et al., “Deep-Learning-Enhanced Human Activity Recognition for Internet of Healthcare Things,” *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6429-6438, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [31] Sanjeev P. Bhavnani, Jagat Narula, and Partho P. Sengupta, “Mobile Technology and the Digitization of Healthcare,” *European Heart Journal*, vol. 37, no. 18, pp. 1428–1438, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [32] Joel J.P.C. Rodrigues et al., “Enabling Technologies for the Internet of Health Things,” *IEEE Access*, vol. 6, pp. 13129–13141, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [33] Said Agrebi, and Anis Larbi, “Use of Artificial Intelligence in Infectious Diseases,” *Artificial Intelligence in Precision Health, Elsevier*, pp. 415–438, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [34] Iqbal H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Computer Science*, vol. 2, no. 3, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [35] Samir S. Yadav, and Shivajirao M. Jadhav, “Machine Learning Algorithms for Disease Prediction Using IoT Environment,” *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6, pp. 4303-4307, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [36] C.K. Gomathy, and A. Rohith Naidu, “The Prediction of Disease Using Machine Learning,” *International Journal of Scientific Research in Engineering and Management*, vol. 5, no. 10, pp. 1-7, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [37] Panagiotis Radoglou-Grammatikis et al., “Modeling, Detecting, and Mitigating Threats against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 2041-2052, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [38] Nicola Orsini et al., “A Tool for Deterministic and Probabilistic Sensitivity Analysis of Epidemiologic Studies,” *The Stata Journal: Promoting communications on statistics and Stata*, vol. 8, no. 1, 2008. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [39] Toshita Sharma, and Manan Shah, “A Comprehensive Review of Machine Learning Techniques on Diabetes Detection,” *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, no. 1, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [40] Oluwadare Adepeju Adebisi, John Adedapo Ojo, and Olatunde Michael Oni, “Comparative Analysis of Deep Learning Models for Detection of COVID-19 from chest X-Ray Images,” *International Journal of Scientific Research in Computer Science and Engineering*, vol. 8, no. 5, pp. 28–35, 2020. [[Google Scholar](#)] [[Publisher Link](#)]
- [41] Tom M. Mitchell, *The Discipline of Machine Learning*, CMU School of Computer Science, 2006. [[Google Scholar](#)] [[Publisher Link](#)]
- [42] Sabina Tomkins et al., “IntelligentPooling: Practical Thompson Sampling for Mhealth,” *Machine Learning*, vol. 110, no. 9, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [43] C. Fährnich et al., “Surveillance and Outbreak Response Management System (SORMAS) to Support the Control of the Ebola Virus Disease Outbreak in West Africa,” *Euro Surveillance*, vol. 20, no. 12, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [44] Umang Soni, Nishu Gupta, and Sakshi, “An Artificial Intelligence Approach for Forecasting Ebola Disease,” *Journal of Physics: Conference Series*, vol. 1950, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [45] Zeeshan Ahmed et al., “Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine,” *Database, Oxford University Press*, vol. 2020. 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [46] Gaber H. Alsayed et al., “Mathematical Modeling and Classification of Viruses from Herpesvirus Family,” *International Journal of Computer Applications*, vol. 87, no. 12, pp. 5–13, 2014. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [47] Akshaya Karthikeyan et al., “Machine Learning Based Clinical Decision Support System for Early COVID-19 Mortality Prediction,” *Frontiers Public Health*, vol. 9, 2021, [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [48] Shahadat Uddin et al., “Comparing Different Supervised Machine Learning Algorithms for Disease Prediction,” *BMC Medical Informatics and Decision Making*, vol. 19, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [49] Ebrahime Mohammed Senan et al., “Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques,” *Journal of Healthcare Engineering*, vol. 2021, pp. 1-10, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [50] Md Manjurul Ahsan et al., “Deep MLP-CNN Model Using Mixed-Data to Distinguish between COVID-19 and Non-COVID-19 Patients,” *Symmetry*, vol. 12, no. 9, pp. 1-14, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]