

Thesis Report.....docx

by Taspia Tazri Chaity

Submission date: 19-Sep-2025 07:52PM (UTC+0600)

Submission ID: 2755440006

File name: Thesis_Report.....docx (5.34M)

Word count: 28749

Character count: 159195

ABSTRACT

Dengue fever is a viral infectious disease caused by four distinct serotypes, namely DENV 1, DENV 2, DENV 3, and DENV 4, and it spreads primarily through the bite of an Aedes mosquito that has previously fed on an infected person. It has caused great problems in diagnosis and treatment of the disease at the early stages of it since it has spread to almost all parts of the world thanks to over-urbanization, global warming and poor measures to counter the vectors leading to a mixed picture of Mild Dengue Fever (DF) and severe forms such as Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS). Disease severity classification is by far the most significant tool to minimize complications and maximize care provided to the patient, in addition to efficient use of healthcare resources.

This study is grounded in 1,157 records from Cumilla, Noakhali, and Dhaka in Bangladesh, incorporating clinical and physiological details, demographic information, contextual factors, and cases representing both dengue and non-dengue conditions to ensure balanced representation. Using machine learning and deep learning techniques, the records were classified into four categories: Mild Dengue Fever, Dengue Hemorrhagic Fever, Dengue Shock Syndrome, and No Dengue. To enhance predictive accuracy, a novel hybrid RF-LSTM model was developed by integrating the Random Forest algorithm with the Long Short-Term Memory network. The Random Forest approach was employed for effective feature extraction, while the LSTM network captured and modeled the complex, non-linear temporal patterns in the data. Experimental findings demonstrated that the hybrid RF-LSTM model achieved superior performance, reaching an accuracy of 96.26%, and it consistently outperformed conventional models in terms of accuracy, precision, and recall. Furthermore, the generalizability of the model was validated using unseen datasets, confirming its robustness and reliability for real-world applications. Beyond individual-level classification, the study incorporated geographic analysis to generate risk zone maps, which assist in identifying high-risk areas and formulating targeted intervention strategies. These maps provide valuable insights for public health planning, efficient allocation of resources, and proactive management of dengue outbreaks. Overall, the study establishes a comprehensive, information-driven framework for diagnosis, severity classification, and geographical risk assessment of dengue, while demonstrating how computational modeling combined with epidemiological expertise can enable timely disease detection. Moreover, the approach strengthens population health surveillance and offers a scalable solution adaptable to other vector-borne diseases, thereby contributing to the development of sustainable, data-driven healthcare systems.

Keywords: Dengue, Outbreak, Risk zones, Severity, Public health, Public awareness.

CHAPTER 1

INTRODUCTION

1.1 Background Study

Dengue fever is a highly dangerous infectious disease caused by the dengue virus, and spread by Aedes mosquito bites, which remains a significant health problem worldwide [1]. It is a viral disease that is transmitted by the bite of infected mosquitoes that provide an entry point to the human body to be infected by the dengue virus and is caused by four different serotypes of dengue virus, namely, DENV-1, DENV-2, DENV-3, and DENV-4 [2]. The dengue risk has been increasing globally, with Aedes mosquitos living in conditions conducive to transmission in nearly 86 percent of nations and territories around the world. Each year, tens to hundreds of millions of infections are seen, which indicates the high susceptibility of populations in tropical and subtropical regions [3]. Dengue menace has been on the rise in Bangladesh leading to several deaths, and the number of cases is not showing any improvement and is only becoming threatening to the people as the cases continue to rise [4]. Currently, the most severe and deadliest dengue outbreak in the country has been documented due to an early seasonal outbreak, high-speed geographic expansion, and high mortality rates. Such a serious scenario highlights the high importance of better disease monitoring, developing more effective diagnostic tools, and prioritizing healthcare resources to fight dengue in South Asia [5]. The Ministry of Health and Family Welfare of Bangladesh reported 52,807 dengue cases confirmed by the lab and 230 dengue-related deaths as of November 20, 2022, which reflects the size of the public health crisis [6]. Clinical dynamics of dengue is linked to an unpredictable situation with mild febrile disease on one end and the severe complications such as coagulation disorders, plasma loss, shock and death at the other end [7]. In that respect, it is essential to forecast, anticipate cases of dengue in time, which will assist to make a sensible conclusion and offer a quick medical reaction and preventive measures. Several more sophisticated computational methods have become increasingly popular in recent years, such as machine learning and predictive analytics, which can redefine how epidemics are predicted, diseases are managed, and care is provided [8]. Because of the gravity of this problem, there is an immediate need to develop new methods of precise identification, categorization and predictability of the occurrence of dengue to pursue more effective intervention strategies.

In this research study, the symptoms of dengue fever will be studied closely in order to arrive at a predictive model of dengue fever based on the main information given by the patients. The reason some past research studies were more firmly devoted to one or the other is that neither detected nor generally predicted dengue, and none were systematically determined as clinically significant, nor have they been linked to the analysis of the risk in the region. To be more specific in the diagnoses so as to be able to take specific medical interventions, machine learning, deep learning, and hybrid modeling have been used in this paper to use cases of dengue

to fall into the respective groups. The Cumilla, Dhaka, and Noakhali data are analysed to examine the disease pattern in the region by a spatial risk mapping that determines the areas with high risk of a dengue epidemic. There is also another, invisible, set of data on which they tested the models that also contributed to making the models robust and capable of being generalized; hence, the models could also be used in the context of other data settings. This research bridges the gap in research on absolute classification and regional risk analysis and integrates the merit of good quality computational models with epidemiological analysis to provide a complete, grounded framework to establish, categorize and monitor dengue risks at the regional level. Such a framework should assist in the timely planning of the public health sector, contribute to the effective utilization of resources and enhanced response mechanism of the outbreaks, and, consequently, play a part in the general work of controlling dengue fever spread and management in the affected areas.

¹¹⁶ 1.1.1 Machine Learning

Machine learning is one of the most common fields of artificial intelligence, which gives a computer a chance to learn by experience and make adequate predictions or choices without being explicitly programmed [9]. In both predictive and classification problems in healthcare, machine learning models help greatly since they can identify a relationship between the input features and the outcome based on complex patterns in the data. In this paper, machine learning is used to predict the cases based on the severity using a number of features, such as the clinical symptoms, medical history, and demographic information in the dengue dataset. They enable such models to identify small correlations that human beings would have found difficult to identify, and make the diagnosis more accurately, and contribute towards treatment in areas that are required. Machine learning can be broadly categorized into three main types:

Supervised Learning: In supervised learning, a model is trained using labeled data, with an input and a known output associated with it. The model is able to learn to map inputs against the outputs, hence able to predict for new unknown data [10]. This method works very well in classification and regression problems, including classifying disease severity by using patient characteristics.

Unsupervised Learning: Unsupervised learning is performed on data that is not labeled and the goal of unsupervised learning is to identify unknown structures, patterns or groupings of the data. In the absence of known outcome labels, algorithms such as dimensionality reduction and clustering may select pieces of data or trends that may be useful in unsupervised data analysis [11]. This kind of learning can be applied in exploratory data analysis and abnormality detection of healthcare data.

Reinforcement Learning: The idea of reinforcement learning is based on trial and error, in which the agent is rewarded or punished by the environment, and through this process, the agent learns how to make a decision [12]. The agent then gradually seeks to maximize its actions

regarding the reward accumulation. This learning paradigm has been specifically used in adaptive systems and sequential decision making.

⁹⁴1.1.2 Deep Learning

Deep learning is a more advanced branch of machine learning that builds multi-layer neural networks to learn non-linear and complex correlations between data. This feature of deep learning models is very useful in medical diagnostics and disease identification tasks: unlike more traditional methods, which use manually computed features, feature representations can be automatically learned in a hierarchy. Deep learning has also found significant applications in the healthcare industry as a tool to support the diagnosis of the disease, the forecast of the patient outcome, and estimating the risk [13]. It has an insight into disease progression due to its ability to know more about disease progression owing to the temporal chains and time constraints. Deep learning offers meaningful and generalized solutions, due to the fact that it is not affected by the issue of feature selection being human-biased, and it is able to learn directly from the data. This is an essential element of the hybrid modeling method used in the current study, as it will help to increase the accuracy and validity of dengue severity classification and apply the information to healthcare decisions.

1.2 Problem Statement

¹³⁵Dengue fever is a serious health issue in the rest of the world as it is one of the infectious diseases that are spread by the mosquito, and the rates of this disease have unbelievably grown in the past decade [14]. Millions of people around the world are infected by this disease, particularly in tropical regions and subtropics, and there is an urgency to find the best methods of detection, classification, and treatment [15]. The clinical picture of dengue is characterized by mild dengue fever (DF) observed on one end of the spectrum, and severe and life-threatening dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) on the other end [16]. This presents a complication in that it is both hard to diagnose and difficult to treat, and places a heavy burden on health care services and resources, especially in regions where incidences are often common. Because the number of dengue cases and their gruesome complications has been observed to be on the rise, there is a need to diagnose it early and classify it accordingly in a bid to improve the outcome even further and optimize available healthcare responses to it.

To overcome these limitations, this research takes into consideration the use of developed machine learning and deep learning methods as a hybrid model to promote the early detection of dengue disease and normalization of dengue cases into specific categories. The hybrid approach will be capable of the highest degree of reliability and accuracy in detecting and classifying diseases due to the fact that it is an approach founded on and utilizing the merits of multiple models to obtain an improved impression of disease progression. The study will be based on information gathered in Cumilla, Noakhali, and Dhaka to train and test the models, and also test

their power to generalize on a separate, unknown test set. Besides that, there are also quantitative indicators of the exposure of the areas to future dengue outbreaks, such as spatial risk mapping and rate of positivity per district, that can be used to prioritize and plan the areas that are most likely to be affected by the outbreak. This is broadly framed to enable the continuity of research beyond the formulation of an efficient dengue detection and classification scheme, yet also makes it rational to apply data-driven expertise to support the community in preparedness and resource allocation.

1.3 Research Gap

The past ¹³couple of years have seen advancements in the field of dengue ¹²detection with respect to machine learning as well as deep learning, but the area of research is still rife with gaps. A significant portion of the literary materials available is largely concerned with the distinction between dengue and non-dengue cases, rather than the multi-classification of dengue as Mild Dengue Fever, Dengue Hemorrhagic Fever, and Dengue Shock Syndrome, which would be crucial to the efficient and time-sensitive clinical decision-making process. The vast majority of methods are completely based on ¹³machine learning or deep learning and are not interested in hybrid algorithms that combine feature-based learning with sequential pattern recognition to generate more predictive accuracy. Furthermore, it is not limited by geographic analysis and risk mapping, which would limit the information about risky areas. Often, these models are not trained on obscure or even geographically dissimilar sets of patients. The databases that are available do not usually identify those who are not infected with dengue and thus it gets hard to distinguish between the infected and the non-infected patients. The quality of dengue diagnosis and prediction models can then be improved by using general information about the patients, hybrid, and geo-risk mapping. Such gaps must be reported on, with an attempt to come up with plausible, generalizable, and clinically meaningful dengue classification schemes.

1.4 Motivation of this Research

Dengue fever is a social condition and a severe one, especially in Bangladesh, where it has taken a twist of becoming more common, more severe, and more widespread. Due to the indeterminate character of the course of disease in mild dengue fever (DF) that may progress into ⁵severe and potentially fatal illnesses such as dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS), there is an urgent need to diagnose those diseases as soon as possible and correctly. The normal diagnostic procedures that are usually employed are lengthy, and they can be inaccurate to the point of leading to serious complications and death. This situation is becoming more urgent, as there is a growing sense of urgency in utilizing data-driven and innovative solutions that have the potential to positively impact early diagnosis and the proper distribution and effective healthcare treatment. Recent progress in machine learning, deep learning, and hybrid modeling offers an opportunity to uncover valuable tendencies in patient data and the

surrounding conditions that can present high-quality, compelling, and effective information about the course of the disease.

The central point of this study is to develop this hybrid model that will have the ability to combine the strengths of multiple learning algorithms in a smart way and will therefore produce better classification accuracy, strength, and consistency to classify dengue case. In addition to the detection and classification, the mapping of the spatial risks and the positive rates of the districts of Cumilla, Noakhali, and Dhaka are also part of the study, which makes it possible to note more vulnerable areas and to organize the strategic planning of the public health optimally. To further determine the reliability of the model, performance is evaluated using unknown independent datasets, in a situation closely related to those of the real world, and known to determine whether the predictive models are generalizable. This research will offer an immensely desired insight into understanding the dynamics of the dengue transmission process, improved decision-making, and improved performance in the context of how the dengue outbreaks in the endemic areas can be more significantly monitored and controlled.

1.5 Research Hypothesis

The hypothesis of this study is that the use of the proposed hybrid Random Forest-Long Short-Term memory (RF-LSTM) model will result in a higher quality of the detection and classification of dengue cases, such as Mild Dengue Fever, Dengue Hemorrhagic Fever, Dengue Shock Syndrome, and a person who is not infected with dengue. By integrating feature-based learning and sequential pattern recognition, the learning metrics of accuracy, precision, recall, and F1-score are likely to improve considerably. Moreover, the use of geographic analysis to produce risk maps will allow for the identification of high-risk areas to aid in specific public health interventions. The overall strategy is meant to help in timely medical action, better clinical decision-making, and dengue surveillance and control.

1.6 Research Objectives

This research paper is dedicated to pushing the boundaries of early dengue fever prediction and enhancing the precision of classifying different dengue cases using primary data. To achieve these aims, an array of cutting-edge machine learning, deep learning models, and hybrid methods will be employed. The main objectives of this research are as follows:

1. To gather primary data on the symptoms and medical history of dengue fever in Cumilla, Noakhali, and Dhaka.
2. To create and develop a hybrid model that will effectively recognize the symptoms of dengue and categorize cases with great precision.

3. To categorize the cases of dengue systematically into clearly defined types using data-driven solutions, to maximize the classification strength and accuracy.
4. To analyze and map risk zone areas in order to identify regions most vulnerable to dengue outbreaks.
5. To validate model performance and generalizability on unseen data.

Through the careful implementation and refinement of these objectives, this study aims to contribute significantly to dengue fever research, offering actionable insights for data-driven healthcare planning and improved management of disease outbreaks.

1.7 Scope of this Research

This research paper focuses on the primary data gathered and analyzed on dengue fever in three districts of Bangladesh: Cumilla, Noakhali, and Dhaka. The dataset contains 1,157 records, which are divided into four classes based on the severity of the dengue: No Dengue with 365 samples, Mild Dengue Fever with 305 samples, Dengue Hemorrhagic Fever with 246 samples, and Dengue Shock Syndrome with 241 samples. Each record will contain a detailed combination of features, including symptoms, medical history, and demographic data, which serves as a powerful base to develop and test classification models.

The study uses a range of different machine learning and deep learning models to classify dengue accurately. These machine learning models are Random Forest, XGBoost, and Support Vector Machine, and deep learning models are Multilayer Perceptron, Long Short-Term Memory, and one-dimensional Convolutional Neural Network. Besides that, a hybrid model, combining the strengths of the two methods, the Random Forest and the Long Short-Term Memory, is applied to improve the overall classification accuracy and model robustness. The models are trained and tested on the main dataset and also tested on an independent, unseen dataset to determine their performance on unseen new data.

The study also carries out risk analysis of the areas in detail, along with classification. Spatial risk maps of the three districts were developed, and the proportion positive in each dengue category was computed to provide a quantitative assessment of regional vulnerability and aid in focused planning of public health.

The following specifications were used to undertake the computational analysis:

- **Hardware:** Computer running Windows 10 on an Intel Core i5 processor and 8GB of RAM.
- **Software:** Google Colab .
- **Programming Language:** Python version 3.8.

- **Algorithms:** Random Forest, XGBoost, Support Vector Machine, Multilayer perceptron, Long short-term memory, one-dimensional Convolutional Neural Network, and Hybrid RF-LSTM model.

The study provides evidence-based data to assist in the effective planning of healthcare and enhance the response to future dengue outbreaks through this integrated approach that has confirmed a single framework, precise dengue identification, effective classification, and risk assessment within the region.

1.8 Contribution of this Research

This study holds importance to the diagnosis and control of dengue in several key areas. First, it presents a hybrid RF-LSTM model, wherein the advantages of machine learning and deep learning are combined to achieve superior accuracy and resistance to identifying cases of dengue. Second, the study provides a conceptual grouping of dengue into four clinically significant groups: No Dengue, Mild Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF), and Dengue Shock Syndrome (DSS), and, therefore, enhances health care decision-making. Third, the model is trained on the primary datasets gathered in Cumilla, Noakhali, and Dhaka and tested using them, where an unseen dataset is used to evaluate the model's performance with the new dataset. Furthermore, the analysis of spatial risk mapping is also adopted in the study to demonstrate the vulnerability of the region and provide useful information on the improved preparedness to respond to the outbreaks and health policies of the population, since it combines effective classification methods with the risk analysis into a more complex approach to dengue control. Finally, the results of this research are objective conclusions that have the potential to inform early intervention, adequate allocation of resources, and sound decision-making within health care systems.

1.9 Thesis Structure

This thesis report has the following structure:

Chapter 2 presents a review of available literature on dengue prediction and classification with emphasis on methodologies, datasets, and major findings.

Chapter 3 explains how it will be done, such as data collection, preprocessing, feature selection and how the proposed hybrid Random Forest-Long Short-Term Memory (RF-LSTM) will be developed.

Chapter 4 contains the results and a model performance assessment, and compares it with current methods and has geographic risk zone maps.

Chapter 5 summarizes all the main findings of the research, addresses the implications of the results for the work on population health, and presents limitations and future research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The increasing occurrence and unpredictability of dengue fever remain a significant threat to the health of people worldwide, especially in tropical areas and subtropical regions. Although significant progress has been made in epidemiology, diagnostic testing, and prevention, dengue lacks an easy solution because it is characterized by complicated transmission and lacks a range of highly effective and available treatments or vaccines. Here, machine learning (ML) and deep learning (DL) are some of the data-driven methods that have proven to be effective computational tools capable of improving early detection, severity classification, and more precise forecasting of outbreak patterns. Such approaches have profound potential to supplement conventional clinical and epidemiological practice with the aim of detecting latent trends in patient data, maximizing diagnostic accuracy, and informing prompt intervention plans.

The review will summarize and discuss the most significant research conclusions in the field of dengue transmission, dengue diagnostics, predictive modeling, and dengue control in terms of ways in which the field could be further developed using Machine Learning, Deep Learning, and hybrid methods. The end of the summaries above suggests that a significant step is taken, the blanks that are still blank and, therefore, need to be organized, are filled by the organization of the synthesis. To provide a detailed description, the literature is split into seven portions according to the following principle:

1. The History and Global evolution of dengue fever
2. Dengue Prediction Using Machine Learning
3. Machine Learning-Based Dengue Severity Classification
4. Deep Learning Approaches in Dengue Forecasting
5. Dengue Epidemiology, Risk Factors, and Prevention
6. Novel Approaches and Emerging Trends
7. Dengue Outbreaks and Clinical Studies

These sections of the literature review are arranged in such a way that, on top of synthesising the available information, this paper establishes a framework that connects historical, epidemiological, and computational analyses of dengue. It opens up new and less explored fields, the most prominent ones being the integration of patient records with environmental and geographic data that can produce more situational and practical results. This is specifically relevant in dengue-prone areas like Bangladesh where predictive modelling, severity, and spatial risk analysis can be explicitly transferred and used to make a contribution to healthcare decision-

making. That way, the review will be found to be taken as a summation of prior studies and as a field onto which novel, imaginative and data-intensive techniques can be created that have emerged in the circumstances of the study.

2.2 The History and Global Evolution of Dengue Fever

In this article, Gubler et al. [17] offer a historical and epidemiological overview of dengue fever and its development over time, along with the initial outbreaks and strains of the virus. This chapter recognizes similarities as antigens between Indian and Hawaiian strains, as well as the difference between the New Guinea ones and the difficulty in understanding dengue virus diversity. Among other things, the most significant developments, including the establishment of mosquito-borne transmission in 1903, the reported epidemics as early as the end of the 18th century, and others, are also covered. The paper analyzes all the areas in the world where dengue is spread and identifies the main areas where it is causing outbreaks. It draws attention to the problem of disease control, particularly in the field of vectors and public health control. The work contains useful background on the impact and effects of dengue globally, and guides future research by unifying historical and epidemiological data.

In this paper, Messina et al. [18] use spatial modelling and historical epidemiological data to track the history of worldwide distribution of dengue virus types over a period of 70 years. The spatial distributions and co-circulation of all four dengue virus serotypes are tracked over time throughout the paper. It demonstrates the impact of globalization, urbanization, and weather on the intensity and prevalence of the dengue disease epidemic. Through the investigation, a spatiotemporal profile can be provided to aid the interpretation of epidemiological trends and better monitoring of diseases. The findings in the case above indicate the dengue menace and hyperendemicity in the majority of parts of the world.

As per this research by Stica et al. [19], the world history and dynamics of the dengue viruses can be described in terms of entire genome sequences. The genomic analysis has been applied in the study of the variation of the dengue virus serotype and also in the study of the evolution and distribution of the same virus serotypes in different geographical areas at different times. It has established the mutation and lineage differentiation pattern to provide some insight into the distribution and co-occurrence of various serotypes. It goes more in-depth into how the development of the dengue epidemiology is influenced by the forces of evolution and what consequences the potential of an epidemic outbreak can have. According to the analysis, which is conducted on the basis of a synthesis of the genomic and space and time data, one will receive an idea of tendencies in the world insofar as the issue of virus transmission is concerned. Overall, this has been helpful in creating better vaccines, surveillance plans, and vector control.

According to Guzman and Istituriz et al. [20], a recent survey of the prevalence of dengue in a universal sense has shown that the virus is spreading across geographical boundaries. The authors read between the lines of available epidemiological evidence and comment on some of

the dominant factors that characterize the dengue distribution, such as urbanization, climate, and migration. Their findings show factors that are increasingly susceptible to epidemics and the rising economic health cost of disease. The paper demonstrates some of the issues in the observation, case management, and control of the vectors. This paper recommends that prevention and control strategies can be improved through the development of a superior knowledge base.

Benelli and Mehlhorn et al. [21] describe in this article how malaria is becoming less common and the growing necessity to fight mosquitoes, or, more to the point, vectors, more effectively. Some of the methods of dealing with vectors, such as chemical, biological and genetic methods have been discussed by the authors. They take into consideration extensive management techniques and several interventions to attain a more sustainable control. The analysis has found essential determinants that impact the populations and dynamics of disease transmission in the mosquitoes. This study will aid in offering any practical advice to stop the transmission of new mosquito-borne diseases based on new efforts to control them. On the whole, this article contains a lot of information about how to develop effective programs and strategies to fight dengue, Zika and any other arboviral disease.

2.3 Dengue Prediction Using Machine Learning

Multiple studies have used machine learning algorithms to forecast dengue outbreaks based on patient information, weather conditions, and socio-economic indicators.

Dey et al. [22] tested a significant quantity of data on hospitalized dengue patients' clinical records, meteorological data, and socio-economic data of Bangladesh, and dengue potential occurrence prediction models were developed. Different machine learning algorithms that were used in the evaluation of the predictive performance are as follows: Random Forest, XGBoost, and Support Vector Machine (SVM). It is stated that meteorology, such as temperature, humidity, and rainfall, also contributed significantly to causing the dengue outbreak. XGBoost performed better than the predictive power of Individual models, with the highest score of about 75 percent, compared to random forest and SVM. The findings indicate that when making dengue forecasts, it is important to use as much data as is available. Overall, this study could be useful to inform the design of population health interventions and shape preventive control of outbreaks.

The variables that Kakarala et al. [23] employed in their effort to predict dengue in India included a combination of weather conditions and dengue case history, including temperature, humidity and rainfall, amongst others. In order to determine the effect of a climatic factor on dengue outbreaks, the authors employed different machine learning methods, such as the Random Forest, XGBoost, and Support Vector Machine. They found out that weather information was beneficial to enhancing the prediction quality and that the maximum accuracy of about 82 percent of the models was as a result of XGBoost. The article is very heavy into the part about environmental factors in dengue prediction and creating a machine-based early

warning system. This is generally a good suggestion that can be offered to health authorities within the society to be in a better position to address in future dengue outbreaks.

In this study, Kuo et al. [24] employed epidemiological, environmental, and climatic data to better predict dengue occurrence in Taiwan. To achieve better predictions, they designed a model that makes use of random forests and feature selection to identify which predictors contribute the most. The results revealed feature selection to be an important predictor, and in the Random Forest model, an Area Under the Curve (AUC) of 0.9547. Temperature was identified as the physical factor that exerted the most significant impact and negative effect on the presence of dengue below and above 30 °C, respectively. And there was also the adverse impact of the air quality indices on the prevalence of dengue, which, once again, had not been reported in Taiwan previously.

In this study, Rahman et al. [25] came up with an early warning system of the dengue disease in Bangladesh with the use of climatic data provided by the Bangladesh Meteorological Department, dengue case data provided by the Ministry of Health, and sociodemographic and landscape data provided by the Bangladesh Bureau of Statistics. They used tree based machine learning models, such as Random Forest, XGBoost, and LightGBM with Shapley Additive Explanations (SHAP) to select features. Minimum temperature, agricultural land area, and population density were the major predictors, and LightGBM gave the best predictive results. The paper shows the usefulness of using multi-source official data along with decipherable machine learning to predict the occurrence of dengue outbreaks and aid healthcare policy responses.

In this paper, Prome et al. [26] forecasted the occurrence of dengue in Bangladesh by examining weather statistics of 11 districts. They have used seven machine learning algorithms, among them three ensemble algorithms, and optimized hyperparameters to predict cases of dengue. The Support Vector Regression (SVR) model was the best performing with a mean absolute error of 4.112. The authors used the SHAP (Shapley Additive Explanations) framework to make the results more interpretable by including information about feature contributions. The given approach explains the significance of explainable AI to the field of forecasting in the context of public health.

Saeed et al. [27] trained machine learning models to predict dengue cases and dengue-related deaths in Pakistan based on monthly dengue cases and deaths in 2016-2020. The authors used the methods of Random Forest, Random Tree, Linear Regression, Multi-layer Perceptron (MLP), and Support Vector Machine (SVM) Regression to predict the trends of diseases. Correlation coefficient and error metrics were used to evaluate the models. SVM Regression performed best as it has the highest correlation coefficient of 0.49 in predicting dengue deaths and the highest score in the prediction of cases. The paper illustrates how machine learning can be used in contributing to specific interventions and resources in managing dengue.

Yavari Nejad and Varathan [28] carried out a study that examined key risk indicators of climatic conditions and implemented machine learning algorithms to forecast dengue outbreaks. To

determine the factors that contribute to the spread of dengue, the authors utilized historical data of cases of dengue and climate aspects, including temperature, rainfall, and humidity. Various machine learning algorithms, such as the Random Forest, Support Vector Machine (SVM), and the Gradient Boosting algorithms, were used to predict an outbreak. The evaluation criteria applied to the models were accuracy, precision, recall, and F1-score. Random Forest had the best performance with an accuracy of 92, indicating times when dengue poses a high risk. The present research shows that it is useful when climatic data is integrated with machine learning to ensure the use of the system to monitor dengue proactively and to provide early warning.

Al Mamun et al. [29] described a study that was intended to predict dengue cases in Bangladesh using machine learning methods. The researchers used climatic factors, including temperature, rainfall, and humidity, along with the dengue cases, to develop the prediction model. Several algorithms, such as Decision Tree, Random Forest, and Support Vector Machine, were put to test as they seek to determine their effectiveness. Random Forest performed better than the other, which only reached an accuracy of about 87 percent when predicting the incidence of dengue. The findings stressed that climatic and seasonal factors have a significant role to play in the spread of dengue in Bangladesh. The information in this work can benefit the work of creating early warning systems and recommendations on prevention during future outbreaks.

2.4 Machine Learning-Based Dengue Severity Classification

To improve clinical decision-making, various studies have explored the classification of dengue cases using different ML techniques.

In the article by Sarma et al. [30], the issue of dengue fever is addressed, as it is a severe health problem worldwide, where the number of dengue fever victims is 100-400 million per year, and in 2019, a major dengue fever epidemic hit Bangladesh with 101,000 reported dengue cases annually. The purpose of this paper is to make a prediction about dengue development in Bangladesh on the basis of the information regarding the patients observed in health institutions. The machine learning algorithms, including Decision Tree and Random Forest, were used to evaluate clinical and environmental factors and their impact on dengue cases. Among them, the Decision Tree algorithm had proven to be the most precise with a score of 79, which is higher than that of the Random Forest model. These models were also tested within the realms of traditional performance measures and have been found to aid in early dengue detection.

In this study [31], Gambhir et al. outlined a model to present an early diagnosis of dengue depending on decision tree approaches. To develop predictive models that identify dengue, the authors relied on patient clinical data, including patient symptoms and laboratory tests. The data was classified through a decision tree algorithm to identify patients as dengue positive or dengue negative and features were carefully selected to improve the precision of the data model. Some of the measures used in testing the models include accuracy, precision, recall, and F1-score. The

decision tree approach was found to be 82 percent accurate, and it is beneficial in the initial diagnosis of dengue.

Ho et al. [32] directly compared the machine learning programs to the classical case-control method in the identification of confirmed cases of dengue. The authors have used patient statistics, such as clinical manifestations, laboratory findings, and demographics, to develop predictive models. These are Support Vector Machine (SVM) and Gradient Boosting, which are the two types of random forest machine learning algorithms, and their results were compared with the case-control traditional algorithms. The models were evaluated using accuracy and specificity measures and area under the curve (AUC) measures. Random Forest was the most accurate at 87% and performed better than the SVM and case-control method.

In this study, Exebio-Chepe et al. intended to improve the categorization of dengue virus (DENV) by applying machine learning algorithms [33]. These authors evaluated many cases of dengue at a state-owned hospital in Peru, including demographic and clinical conditions. Three machine learning algorithms were used to classify the cases: Support Vector Machine (SVM), random forest (RF), and artificial neural network (ANN). The performance of the models was compared using the measures of accuracy. ANN had the highest score of 86.47, followed by RF and SVM. Machine learning and, in particular, ANN can be used in this study to aid in the appropriate diagnosis of dengue and help the work of the population health segment in a significant manner.

In this paper, the topic of interest was to predict the intensity of dengue during the initial stages using machine learning and feature selection, as stated by Archana et al. [34]. The researchers collected demographic, blood test, and symptoms data, as well as clinical and laboratory data, in a community hospital in India. The feature selection techniques were applied to identify the most significant variables that influence the development of the disease. These identifiable features were trained on the machine learning models, including the Random Forest and Support Vector Machine (SVM), in order to identify the severity of dengue. Random Forest model proved the best with the highest prediction of 88, which is very predictive. The study mentions the opportunities of using a combined approach to feature selection and machine learning to improve the early diagnosis of dengue and support clinical decision-making.

2.5 Deep Learning Approaches in Dengue Forecasting

In capturing temporal dependencies in dengue cases, deep learning methods, specifically Long Short-Term Memory networks, have been the focus of attention.

In this paper, Saleh and Baiwei [35] have tried to make predictions of the dengue cases using deep learning with Long Short-term memory networks. The authors have summarized a compilation of past occurrences of dengue and other environmental parameters of interest, like temperature, rain, and humidity. The LSTM model is made to define the time trend within the information so as to effectively forecast the dengue outbreaks. Accuracy, precision, recall, and F1-score are some of the measures used to ascertain the performance of the model. The LSTM

model could predict the dengue trends accurately (90 percent), meaning that the LSTM model can effectively identify the dengue trends over time.

In this study, Mumtaz et al. [36] created a deep learning model to predict whether dengue virus (DENV) serotyping will occur. The authors sampled a collection of DENV genomic sequences present in open virus sequence repositories. Long Short-Term Memory networks were used to examine the regularities in the sequential prediction on viral genomes. The new serotypes were determined using these sequence features to train the model. Its prediction ability was obtained based on the area under the receiver operating characteristic curve (AUC) to make the prediction result fixed at 0.818.

Chen et al. [37] created a deep learning framework using LSTM to predict dengue cases and outbreaks in Brazil. The model combined previous dengue case records, climate, and human mobility data to capture time and space dynamics. The dataset involved a combination of epidemiological data and mobility-adjusted imported cases to enhance the prediction. Measures of performance were assessed based on MAE, MAPE, CRPS, and measures of outbreak detection, including accuracy, sensitivity, specificity, and F1 score, with LSTM performing better than baseline models. This paper has shown that human mobility and climatic factors combined with deep learning have the potential to dramatically improve dengue early-warning systems and preparedness for human health.

2.6 Dengue Epidemiology, Risk Factors, and Prevention

In the study, Ogieuhi et al. [38] examined the trend in dengue in Bangladesh during the period January-September 2023 based on national surveillance data of 203,406 cases and 989 deaths of dengue. The authors presented the result of the retrospective epidemiological analysis of the demographic, geographic, and climatic factors, urbanization, and health care capacity. More than half of the cases and deaths occurred in Dhaka, where the high population density, rapid urbanization, favorable climate, and lack of awareness among the population were listed as the main factors. The article underscores the burden on the health care sector and the necessity of integrated community-based approaches to dengue disease, such as improved surveillance, control of vectors, and community participation, to reduce dengue.

In this paper, Roy and others [39] attempted to determine the epidemiology, biology, and causation of the dengue virus. Depending on the outcome of the conducted research in the world and in the region, the authors of the paper carefully analyzed the literature regarding the transmission and the molecular biology of the virus and the clinical manifestations. The observations revealed the complex interaction of the virus, Aedes mosquito vectors and the environment. It means that to further control the spread of dengue virus, the pathophysiology of the disease should be better comprehended to improve prevention and control. It also means that further studies are required to close the gaps in the knowledge of the mechanisms of the disease transmission and of its mechanisms.

Hossain et al. [40] aimed to study epidemiology, clinical spectrum, serotypes, and emerging dengue outbreak risks in Bangladesh during the past 22 years. These researchers have studied information about outbreaks in the form of clinical records, serotypes, and epidemiology using the available information in different hospitals and systems contained in national health information. The findings demonstrated an upward trend in dengue epidemics and their clinical severity and alterations in serotype over time. Other demographic, environmental, and urbanization determinants relating to dengue spread could also be measured during the investigation. The authors noted that early warning and effective public health policy could play a very important role in the reduction of future outbreaks.

Saha et al. [41] analyzed information regarding transmission of single-strain dengue and suggested measures that will successfully decrease the transmission of dengue. The authors have constructed a mathematical model that contains significant parameters of epidemiology and approximated the course of the disease and actions against it. Measures to control vectors, population education and early interventions all played significant roles in their results to control the spread of the infection and reduce the rate of infection. In this paper, it has been demonstrated that modeling can be utilized to give an approximation of how the resources should be used to optimally control dengue. It also created the prospect of employing an array of control measures in order to obtain long-term disease profits.

2.7 Novel Approaches and Emerging Trends

Recent research has suggested innovative approaches to predict and contain dengue, including ensemble learning, analytics and epidemiological modeling of social media.

In this study, Zarin et al. [42] introduce non-linear dengue epidemiological model with a convex incidence rate and using the Atangana-Baleanu fractional-order operator. The article considers the potential impact of the fractional-order operator on the dynamics of dengue transmission and infection. The RK-4 and the NSFD schemes were applied to the model to examine the model behavior using qualitative analysis and numerical simulation. The findings show how the impacts of the memory, as the fractional derivatives, influence the development of the diseases. It shows how a more realistic way to address dengue outbreaks can be achieved using fractional-order dynamics. Overall, the study could be of much value in creating effective interventions to combat dengue and curb the spread.

This research paper was written by Bomfim et al. [43] to understand how human mobility influences the transmission of dengue in Fortaleza, Brazil, when predicting dengue using two forecasting systems. The authors integrated bus transport mobility data across 119 neighbourhoods and epidemiological data to model the ebb and flow of dengue. These results indicated that when mobility was included with the forecasts of outbreaks, this was far more accurate than when trends in past cases alone were used. Also identified were high-risk neighborhoods where interventions could be targeted or prioritized. In totality, it highlights the

importance of initializing human mobility details with forecasting examination to entail the population health reaction and release dengue prevention efforts.

In this paper, Pintado Silva et al. [44], provide a review of the present situation with dengue vaccinations and present details on nine vaccines currently being tested. In this paper, some of the complex immunological responses to dengue virus infection, e.g., complications generated by the four serotypes and antibody-dependent stimulation, are explained. Live-attenuated, inactivated and subunit vaccine preparations have been used as a current vaccine candidate and are characterized in terms of effectiveness, safety and limitations similar to those of the approved vaccine Dengvaxia. The authors explore clinical trial outcomes, regulatory problems and the issues of population heterogeneity and pre-existing immunity. Vector-based and multi-epitope vaccines are addressed as innovative approaches on which solutions to current limitations should be premised.

Gupta et al. [45] proposed a prediction and diagnosis model (DDPM) of dengue disease in ¹⁰ this paper based on the fusion of machine learning algorithms and sentiment analysis. The authors collected data regarding patient symptoms and social media content to assess the prudent attitude toward the dengue epidemics. This model applied ⁵⁶ machine learning techniques, which include Random Forest, Support Vector Machine, and a Decision Tree classifier, to predict the dengue incidence and assist in early dengue detection. The results showed the DDPM was quite precise and suggested that clinical information could be combined with sentiment-based knowledge to forecast diseases. The paper reinforces the idea that social media and social sensitivities should be utilised in tracking the trend of dengue as an assistive data set.

In this study, Hasani et al. [46] discuss the current trends in managing dengue fever by applying the use of both herbal and synthetic drug solutions. The authors have conducted a review of literature on antiviral agents, application of plant extracts and synthetic therapeutics to dengue virus infections and pathogenesis. They further indicated potential herbal molecules like flavonoids and alkaloids, and synthetic molecules exhibiting inhibitory viral protein. The study also discusses measures of drug delivery, action mechanisms, and safety profiles of these interventions. These findings provide insights into the potential of combining traditional herbal medicine with synthetic drugs in an effort to enhance the effectiveness of treating dengue disease.

In this article, Bashir et al. [47] pay attention to the current situation of dengue fever in Bangladesh concerning the demographic trend, male/female ratio and pattern. The authors analyzed national and hospital-based dengue data to conclude on differences in affected age groups, the changes in proportions of females, and hepatic aspects. According research, they can say that the number of women is growing and that there is a serious problem with complications that the liver can cause in patients. Similar changes in clinical manifestations are also reported in the study, which vary in severity and symptoms of the disease over time. The latter tendencies fill the role of individual policy with cruel words of health and clinical medical activity.

2.8 Dengue Outbreaks and Clinical Studies

Sharif et al. [48] evaluated the epidemiology and clinical features of the 2023 dengue outbreak in Bangladesh in this study. The authors analyzed hospital data, demographics, and presentation of the patients to identify trends and patterns of severity. This they found in their results, and even worse forms of cases were now more numerous than ever. The paper identifies the relevance of better surveillance, timely health care response, and mobilization of resources. It also captures the demographic and clinical progress of the previous outbreaks.

The article by Yesmin et al. [49] is devoted to an epidemiological issue of the dengue epidemics; i.e. the recent DENV-5 serotype infection. The authors argue that the emergence of the new serotype complicates monitoring of the disease, its diagnosis, and treatment. They underline the danger of the additional dissemination and non-characteristic manifestation, and the issue of the validity of vaccinations because of DENV-5. The research also aims at enhancing the social health infrastructure, the control programmes, as well as the early warning system. It is also concerned with unremitting scrutiny, engagement with society, and study in the endeavor to restrain hazards that occur.

In this article, Hasan et al. [50] assessed the clinical and epidemiological features of the 2024 dengue epidemic in Bangladesh using the data collected in various hospitals. Review also discovered that the serious cases of dengue were also on the increase and the principal forms of the disorder were fever, thrombocytopenia, and liver affection. The alteration in age and gender was also experienced with some degree of susceptibility and severity of the disease. The findings indicate the importance of timely diagnoses, effective case management, and improved surveillance systems. This research reiterates these emerging trends, which can be used to inform local health planning.

This paper by Sansone et al. [51] analyzes dengue outbreaks in Brazil and Latin America in terms of emerging and persistent issues. The authors evaluated epidemiological data to determine incidence trends, geographical distribution, and demographic characteristics of affected populations. Results note the rising occurrence of outbreaks, the presence of several dengue serotypes, and variations in the severity of the disease. Other challenges highlighted in the study include inadequate healthcare infrastructure, underreporting, and environmental issues that promote transmission. It further addresses why integrated surveillance, vector control responses, and interventions by the health sector are necessary to reduce the impact of outbreaks.

2.9 Summary

The existing body of research on dengue prediction and classification demonstrates the potential of machine learning and deep learning approaches, including Random Forest, XGBoost, SVM, and LSTM, in modeling disease patterns. However, the majority of studies rely on limited or singular data sources, such as historical outbreak records, meteorological variables, or isolated

clinical datasets, and often emphasize outbreak forecasting rather than patient-level severity assessment. Comprehensive classification of dengue severity into Mild DF, DHF, and DSS remains inadequately addressed, and few studies explore hybrid learning models that could enhance predictive performance and robustness. Geographic risk mapping and integration of multi-source clinical, demographic, and environmental features are also largely neglected, restricting the practical applicability of existing models for targeted intervention.

To address these limitations, the present study incorporates extensive patient-level and environmental data from Dhaka, Cumilla, and Noakhali, encompassing clinical symptoms, laboratory parameters, demographic details, and environmental factors. A hybrid Random Forest–LSTM framework was developed, achieving superior predictive accuracy while enabling detailed severity classification. Geographic risk areas were also mapped to deliver useful information on undertaking the public health planning. Not only does this integrated approach improve predictive reliability, but it also fills key gaps in the extant literature through the integration of comprehensive features analysis, hybrid modeling, and spatial risk assessment.

96

Table 2.1: Summary of Literature Review

S/No.	Author(s)	Dataset	Methodology Used	Results	Limitations
1	Gubler et al.	Historical dengue outbreak records and viral isolates	Epidemiological & historical review	Provided a detailed overview of dengue evolution and transmission milestones	Only descriptive analysis; no predictive modeling or patient-level severity classification
2	Messina et al.	Seventy years of epidemiological and spatial dengue data	Spatial modeling & epidemiological mapping	Mapped global serotype distribution and co-circulation trends	Limited geographic focus; patient-level severity not considered
2	Dey et al.	Dengue patient records, meteorological data, and socio-economic	Random Forest, XGBoost, SVM	XGBoost achieved the highest prediction accuracy of	Limited to outbreak forecasting; did not differentiate severity types

		information from Bangladesh		75%	
3	Kakarla et al.	Dengue case records with temperature, humidity, and rainfall data from India.	Random Forest, XGBoost, SVM	XGBoost achieved the highest prediction accuracy of 82%	Focused mainly on weather variables
4	Kuo et al.	Epidemiological, climate, and air quality index (AQI) data from Taiwan	Random Forest with Feature Selection	Feature selection improved prediction with AUC = 0.9547	Limited to environmental predictors; severity analysis not included
5	Rahman et al.	Dengue case records, climate information, and sociodemographic data from Bangladesh	Tree-based ML with SHAP	LightGBM achieved the best predictive performance	Focused on early-warning only; no detailed patient-level severity classification.
6	Prome et al.	Weather data from 11 districts in Bangladesh	SVR and six other ML models with SHAP	SVR achieved MAE = 4.112	Only outbreak forecasting; did not classify DF, DHF, or DSS
7	Yavari Nejad & Varathan	Historical dengue data combined with climate information	Random Forest, SVM, Gradient Boosting	Random Forest achieved 92% accuracy	Limited to outbreak detection; severity subtypes not analyzed
8	Al Mamun et al.	Dengue case records and climate information from Bangladesh	Decision Tree, Random Forest, SVM	Random Forest achieved 87% accuracy	Incidence prediction only; no patient-level risk scoring or severity split
9	Sarma et al.	Hospital records from the 2019 dengue outbreak in Bangladesh	Decision Tree, Random Forest	Decision Tree achieved 79% accuracy	Low predictive accuracy; limited dataset, and no hybrid modeling applied
10	Gambhir et al.	Clinical patient data	Decision Tree classifiers	REP Tree achieved	Only binary dengue detection; severity

				82.7% accuracy	classification not explored
11	Ho et al.	Clinical, laboratory, and demographic patient records	Random Forest, SVM, Gradient Boosting	Random Forest achieved 87% accuracy	Focused on confirmed cases only; no hybrid deep learning or risk stratification.
12	Exebio-Chepe et al.	Hospital records from Peru	SVM, Random Forest, ANN	ANN achieved 86.47% accuracy	Classified dengue presence only; did not split cases into multiple severity levels.
13	Archana et al.	Clinical and laboratory data from Indian hospitals	Random Forest, SVM with Feature Selection	Random Forest achieved 88% accuracy	Did not explore hybrid modeling or geographic risk mapping
14	Saleh & Baiwei	Historical dengue cases and climate data	LSTM	LSTM achieved 90% accuracy	Focused on outbreak forecasting; clinical severity classification is absent
15	Chen et al.	Brazil dengue case records, climate information, and population mobility data	LSTM forecasting	LSTM outperformed baseline (higher F1, MAE, sensitivity)	Outbreak-focused; did not classify severity levels for patients
16	Ogieuhi et al.	Bangladesh national surveillance data for 2023 (203,406 cases)	Epidemiological analysis	Identified Dhaka as a hotspot	Descriptive only; no ML/DL or hybrid risk modeling
17	Gupta et al.	Clinical patient records combined with social media sentiment data	Random Forest, SVM, Decision Tree with Sentiment Analysis	High accuracy in predicting dengue incidence	Did not classify patient severity; hybrid models not applied
18	Sharif et al.	Hospital records from the 2023 Bangladesh	Clinical and demographic analysis	Observed a surge in severe dengue	Descriptive only; did not implement predictive modeling

		dengue outbreak			or risk mapping
19	Hasan et al.	Hospital data from the 2024 Bangladesh dengue outbreak	Clinical analysis	Observed severe dengue with liver involvement	Descriptive, hybrid, or predictive severity classification was not applied

95 CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the methodological approach used in this research, which is dedicated to the further development of dengue fever prediction and classification based on the usage of modern machine learning and deep learning algorithms. Dengue has remained a significant public health issue in Bangladesh, which is of more importance in specific districts of Bangladesh, such as Cumilla, Noakhali, and Dhaka, where specific identification and timely classification of cases are of utmost importance. The study begins with a systematic gathering of primary data, which includes both clinical characteristics and the environmental factors such as temperatures and humidity in the chosen areas. After intense preprocessing, machine learning and deep learning methods are used to analyze the datasets to categorize dengue cases into different degrees of severity. In order to boost predictive power, a hybrid structure is also proposed, where feature-based learning is combined in a synergetic manner with non-linear, pattern-based learning, which allows the model to both identify key determinants and more complex interdependencies in the data. As a complement to these predictive analyses, spatial risk mapping is conducted to determine and map high-incidence areas in the three districts, which reveal important information about the geographic heterogeneity of dengue transmission. Together, such an approach to methodology combines clinical, environmental, and spatial aspects to develop a powerful and holistic approach to accurate prediction and categorization of dengue.

3.2 Overview of Methodology

This study adopts a comprehensive methodology to enhance dengue fever detection, accurate classification, and spatial risk assessment. The process begins with the collection of primary data encompassing clinical attributes, environmental variables such as temperature and humidity, and residential and behavioral factors. The datasets are systematically preprocessed to ensure quality and suitability for model development. Analytical modeling employs both machine learning and deep learning approaches: machine learning models include Random Forest (RF), XGBoost, and Support Vector Machine (SVM), while deep learning models comprise Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), and one-dimensional Convolutional Neural Network (1D-CNN). To further enhance predictive performance, a hybrid framework is conceptually incorporated, integrating feature-based learning with non-linear pattern recognition to capture complex relationships within the data. Model performance is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Complementing predictive modeling, spatial risk zone analysis is conducted for Cumilla, Noakhali, and Dhaka, identifying high-incidence areas and providing insights into the geographic distribution of dengue cases.

Collectively, this methodology integrates clinical, environmental, residential, and behavioral dimensions to offer a robust framework for dengue prediction, classification, and risk mapping.

3.3 Methodological Framework: In this section, the complete methodology of this research work has been discussed in detail. The proposed method is shown in Figure 3.1.

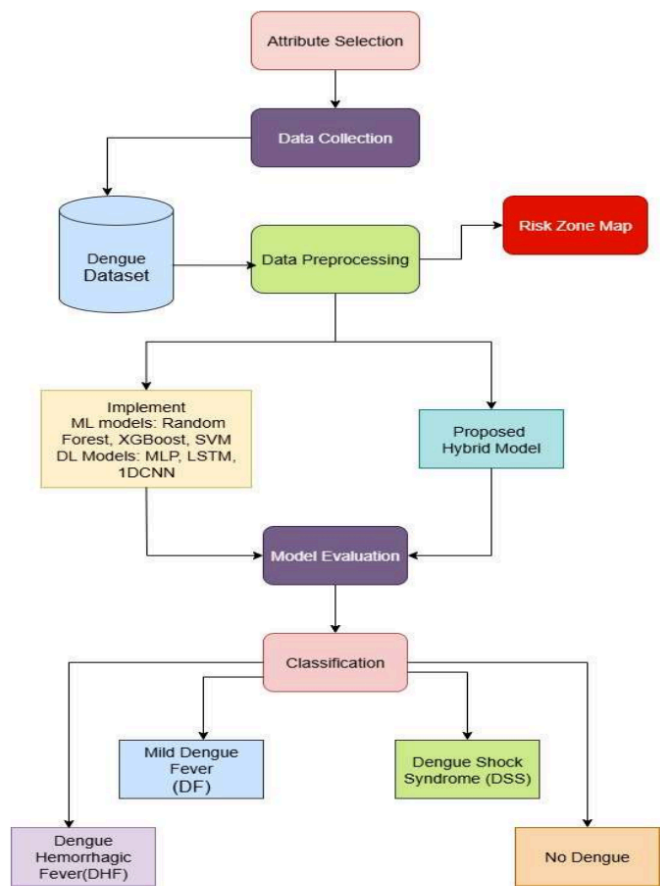


Figure 3.1: Methodological Framework

3.4 Data Acquisition and Preprocessing

3.4.1 Data Acquisition

(i) Parameters

To conduct this study on dengue fever prediction and classification, the following steps were taken:

1. Relevant data was collected from individuals in the districts of Cumilla, Noakhali, and Dhaka, comprising 1157 records.
2. The dataset includes a comprehensive set of 29 parameters, covering demographic information, clinical symptoms, laboratory results, environmental conditions, and residential/behavioral risk factors.
3. Data collection involved direct recording from individuals and relevant sources, ensuring accuracy and completeness.
4. The collected parameters serve as input for machine learning and deep learning models to predict dengue status, classify severity levels, and analyze high-risk zones in the study regions.

Table 3.1: Feature Description

Feature Name	Description	Data Type
Name	Refers to the name of Individuals	Object
Age	Age of Individuals (in years)	Integer
Gender	Gender of Individuals	Object
Residence	Address or Locality of individuals	Object
City	City where the individuals reside	Object
Residence Based On Location	Categorize the individuals' living places as town, semi-town, and rural area.	Object
Environmental Temperature (°C)	Ambient temperature at the location during the time of data collection	Integer
Humidity (%)	Relative humidity at the location during the time of data collection	Integer
Season	Season during data collection	Object
Blood Group	ABO and Rh blood group of the individual	Object

Body Temperature (°F)	Body temperature at the time of data collection	Float
Fever Duration (days)	Number of days the individual experienced fever	Integer
Muscle Pain	Indicates whether the individual experienced muscle pain or not	Object
Joint Pain	Indicates whether the individual has experienced joint pain or not	Object
Skin Rash	Denotes the presence or absence of skin manifestations like rashes	Object
Vomiting	Indicates the occurrence and severity of vomiting	Object
Eye Pain	Denotes the presence or absence of retro-orbital or eye pain	Object
Headache	Indicates the occurrence of headache, including intensity	Object
Heart Rate (bpm)	Quantitative measures of an individual's pulse rate in beats per minute	Integer
⁹⁹ Platelet Count ($\times 10^9/L$)	Laboratory-measured platelet count	Integer
⁹⁹ WBC Count ($\times 10^9/L$)	Laboratory-measured white blood cell count	Float
BP (Condition)	Blood pressure status	Object
Mosquito Bite History	Record of mosquito bites prior to symptom onset	Object
Waterlogging Near Home	Presence of stagnant water near the residence	Object
Recent Travel	Travel history	Object
Pre-existing Condition	Any previously diagnosed medical conditions	Object
Dengue Status (Yes/No)	Indicates whether the individual was diagnosed with dengue or not	Object
Dengue Classification	Severity classification of dengue (Mild Df, DHF, DSS) if positive, otherwise No dengue	Object

⁴⁹
(ii) Data Collection

The data for this research were collected through a structured and systematic process to ensure both diversity and reliability. Clinical information was obtained from hospitals and diagnostic centers in Cumilla, Noakhali, and Dhaka, focusing on confirmed dengue cases with detailed medical profiles. To provide a comparative baseline, additional data were gathered from individuals without dengue through field surveys, covering key demographic, behavioral, and environmental aspects. In total, information was recorded across 29 distinct parameters, encompassing symptoms, laboratory test results, medical history, exposure factors, and living conditions. Complementing the clinical and survey-based data, environmental variables such as temperature and humidity were retrieved from reputable online meteorological sources, aligned with the respective residential areas of the participants. This integration allowed the dataset to reflect not only medical attributes but also contextual environmental factors influencing dengue transmission. The final dataset comprised 1,157 records, distributed across four categories: No Dengue (365), Mild Dengue Fever (305), Dengue Hemorrhagic Fever (246), and Dengue Shock Syndrome (241). This comprehensive and parameter-rich dataset serves as a robust foundation for both classification modeling and regional risk zone analysis.

Table 3.2: Dengue Classification and Data Sample Size

Dengue Classification	Data Sample Size
No Dengue	365
¹⁰⁸ Mild Dengue Fever (Mild DF)	305
Dengue Hemorrhagic Fever (DHF)	246
Dengue Shock Syndrome (DSS)	241

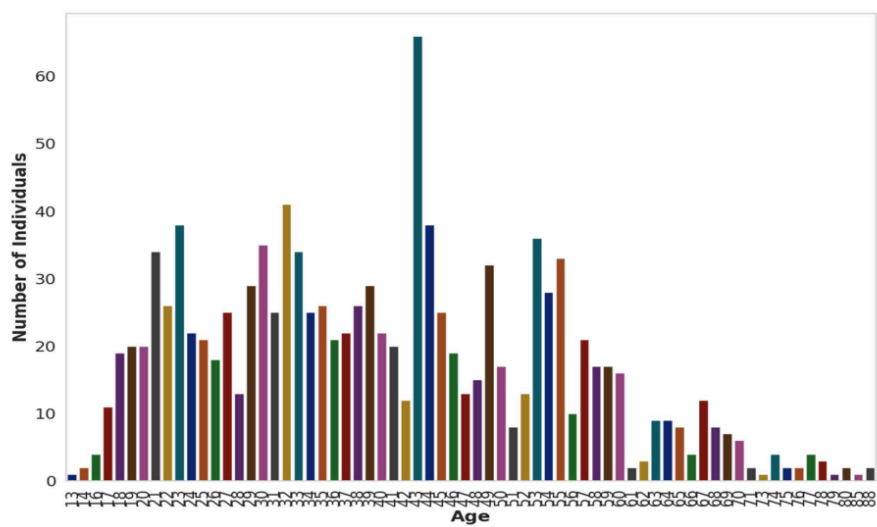
(iii) Dataset Description

In this section, we examine the collected data and provide a comprehensive description of the dataset. A total of ¹⁴29 features were selected to capture both clinical and environmental factors relevant to dengue. The dataset is then prepared and structured for subsequent modeling tasks. Before model training, a detailed exploration of the data is conducted to understand its characteristics and ensure its suitability for analysis.

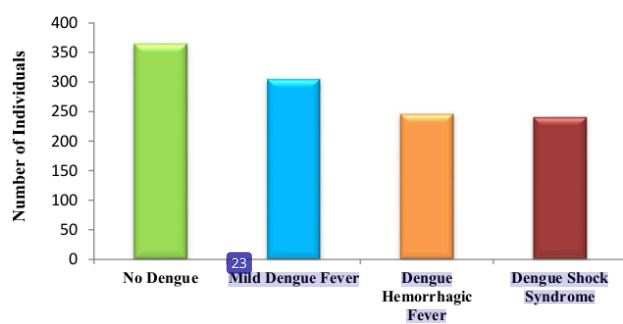
Table 3.3: Data Sample

Patient ID	Name	Age	Gender	Residence	City	Residence Based On Location	Environmental Temperature (°C)	Humidity (%)	Season	Heart Rate (bpm)	Platelet Count (x10 ⁹ /L)	WBC Count (x10 ⁹ /L)	BP (Condition)	Mosquito Bite History	Waterlogging Near Home	Recent Travel	Pre-existing Condition	Dengue Status (Yes/No)	Dengue Classification	
0	1	Kamal Hossain	54	Male	Laksham	Cumila	Town	30	66	Pre-Monsoon	...	90	115	4.3	Normal	No	No	Hypertension	Yes	Mild Dengue Fever
1	2	Fatema Akter	19	Female	Kandirpur	Cumila	Town	32	76	Pre-Monsoon	...	110	124	3.5	Normal	Yes	Yes	Healthy	Yes	Mild Dengue Fever
2	3	Liza Akter	39	Female	Daudkandi	Cumila	Town	27	77	Monsoon	...	110	142	4.1	Normal	No	Yes	Healthy	Yes	Mild Dengue Fever
3	4	Ripa Akter	20	Female	Jambari	Cumila	Rural area	29	87	Post-Monsoon	...	90	99	4.1	Normal	Yes	No	Healthy	Yes	Mild Dengue Fever
4	5	Firoz Ahmed	49	Male	Burichong	Cumila	Semi-Town	33	75	Pre-Monsoon	...	99	150	3.8	Normal	Yes	No	Weak Immunity	Yes	Mild Dengue Fever
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1152	1153	Abdul Mia	59	Male	Jambari	Cumila	Rural area	29	90	Monsoon	...	90	100	4.1	Normal	Yes	Yes	Asthma	Yes	Mild Dengue Fever
1153	1154	Shahjahan Mia	63	Male	Gabtol	Cumila	Rural area	30	84	Monsoon	...	120	155	2.1	slightly low	Yes	Yes	Healthy	Yes	DSS
1154	1155	Miss Nappi	30	Female	Elashpur	Cumila	Rural area	31	85	Post-Monsoon	...	120	140	4.8	slightly low	Yes	No	Healthy	Yes	Mild Dengue Fever
1155	1156	Robin	36	Male	Elashpur	Cumila	Rural area	33	77	Pre-Monsoon	...	90	153	3.2	very low	Yes	Yes	Hypertension	Yes	Mild Dengue Fever
1156	1157	Adnan Khan	22	Male	Shimpur	Cumila	Rural area	31	79	Pre-Monsoon	...	100	159	4.4	very low	Yes	Yes	Weak Immunity	Yes	Mild Dengue Fever

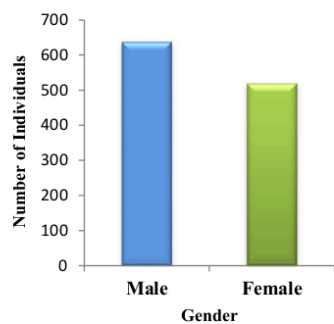
1157 rows x 29 columns



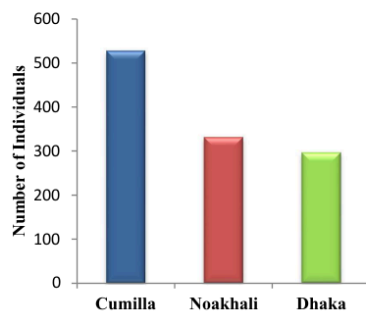
(a) Ages of the collected sample data



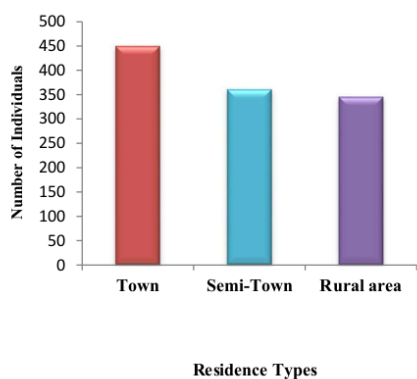
(b) Number of individuals across Dengue Classification



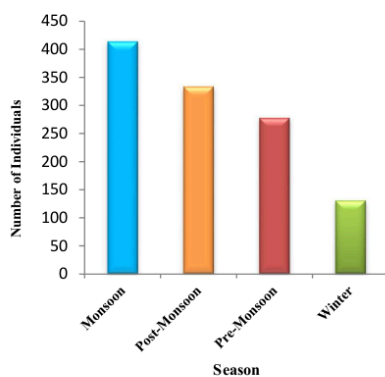
(c) Number of individuals across genders



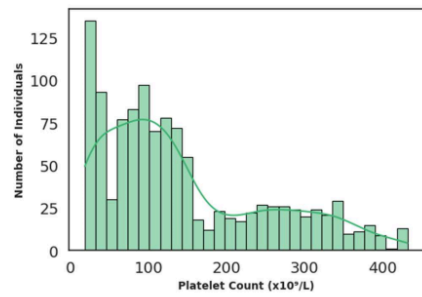
(d) Number of individuals across cities



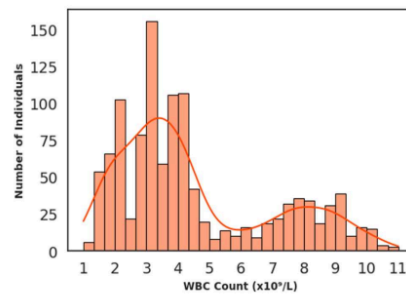
(e) Number of individuals Residence Types



(f) Number of individuals across Season



(g) Platelet count distribution among individuals



(h) WBC count distribution among individuals

Figure 3.2: Frequency distribution of sample data

3.4.2 Data Preprocessing

This is one of the most important stages of the implementation process. As data is collected from multiple sources, it may contain missing values, duplicates, or inconsistencies that can negatively affect analysis. To ensure data quality and reliability, data preprocessing was performed.

The preprocessing steps included handling missing values, removing duplicate records, applying label encoding to categorical variables, and performing feature scaling to bring all values into a comparable range. Additionally, class imbalance was addressed through up-sampling to ensure fair representation of all categories. These steps helped to clean and optimize the dataset, making it suitable for building effective models for dengue classification.

(i) Data Cleaning

Data cleaning is an important step in data preprocessing that involves removing or correcting any inconsistencies, errors, or irrelevant data from the dataset. This step helps to ensure that the data is accurate, complete, and consistent, which is essential for accurate analysis and modeling. One common data cleaning technique is removing duplicate values, which involves identifying and removing any rows or entries that are exact duplicates of one another. This can be done using the `drop_duplicates()` method in pandas. Another technique is removing any rows or entries with missing or NaN values, which can be done using the `dropna()` method in pandas. In addition, the `drop()` method can be used to remove specific rows or columns from the dataset based on certain conditions or labels. The `iloc[]` method can be used to select specific rows or columns from the dataset based on their indices or labels. Lastly, the `replace()` method can be used to replace any values in the dataset with a specified value, such as replacing any 0 values with a unique value to avoid any confusion or bias during analysis. Overall, data cleaning is an essential step in data

preprocessing that involves identifying and correcting any issues in the dataset to ensure accurate analysis and modeling.

(ii) Encoding the Dataset

In machine learning and deep learning applications, datasets often contain categorical values that are represented as labels in text form for human readability. However, algorithms cannot process these textual values directly, as they require numerical input. Therefore, encoding is an essential preprocessing step to convert categorical data into a numeric format. This transformation ensures that the models can effectively interpret and learn from the dataset. Among the common approaches, Scikit-learn provides encoders such as Label Encoder, One-Hot Encoder, and Ordinal Encoder. In this study, we employed Label Encoding to transform categorical attributes into integer values. This step enables both traditional machine learning models and deep learning architectures to process the dataset efficiently. The output of the label encoding process is presented in Table 3.4, which demonstrates how categorical values were converted into numerical representations suitable for model training.

Table 3.4: After label encoding the dataset

115

Patient ID	Name	Age	Gender	Residence	City	Residence Based On Location	Environmental Temperature (°C)	Humidity (%)	Season	Heart Rate (bpm)	Platelet Count (x10 ⁹ /L)	WBC Count (x10 ⁹ /L)	BP (Condition)	Mosquito Bite History	Waterlogging Near Home	Recent Travel	Pre-existing Condition	Dengue Status (Yes/No)	Dengue Classification		
0	1	210	54	1	66	0	2	30	66	2	...	98	115	4.3	1	0	0	0	4	1	2
1	2	104	19	0	61	0	2	32	76	2	...	110	124	3.5	1	1	1	1	2	1	2
2	3	249	39	0	34	0	2	27	77	0	...	110	142	4.1	1	0	1	0	2	1	2
3	4	468	20	0	54	0	0	29	87	1	...	98	99	4.1	1	1	0	1	2	1	2
4	5	108	49	1	22	0	1	33	75	2	...	99	150	3.8	1	1	0	0	5	1	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
1152	1153	3	59	1	54	0	0	29	90	0	...	90	100	4.1	1	1	1	1	0	1	2
1153	1154	534	63	1	46	0	0	30	84	0	...	120	155	2.1	3	1	1	1	2	1	1
1154	1155	309	30	0	43	0	0	31	85	1	...	120	140	4.8	3	1	0	1	2	1	2
1155	1156	471	36	1	43	0	0	33	77	2	...	98	153	3.2	4	1	1	1	4	1	2
1156	1157	12	22	1	95	0	0	31	79	2	...	100	159	4.4	4	1	1	1	5	1	2

1157 rows × 29 columns

(iii) Upsampling the Data

Upsampling is regarded as one of the preprocessing techniques that may be used to address the issue of the imbalance in the classes that may occur due to the underrepresentation of certain groups of the variable of interest. In a non-balanced dataset, the majority category may be somewhat advantaged over the rest of the categories and predict the minority categories in a biased manner. Randomly up-sampling the training data set was one method of overcoming this problem.

Here, training data were up-sampled using `resample()` of the scikit-learn library. Instead, it was to witness the declassification of the classes under the Dengue Classification column, which includes Mild Dengue Fever, Dengue Hemorrhagic Fever, Dengue Shock Syndrome, and No Dengue. The next step was to up-sample all the minority groups, so that the largest minority group has as many cases as the other. The `replace=True` option made it possible to replace sampling by a method that would yield us enough samples when the number of samples parameter is greater than the majority size of the number of classes. The constant random state could also be reproduced. This was repeated until the data was balanced in terms of the number of samples in each category (255 per category) to make it possible to learn the classification models fairly using all of the different categories of samples. Although it has been concluded that random up-sampling will be used due to it not being a very complicated process, as it should not have any effect on the original data, some more advanced methods to generate synthetic samples have been employed as well, including SMOTE (Synthetic Minority Over-sampling Technique).

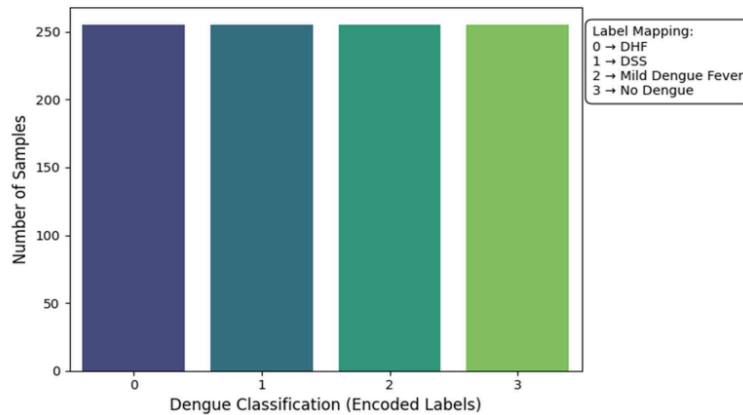


Figure 3.3: Class Distribution after Upsampling

(iv) Feature Scaling

The data was normalized by applying a normal preprocessing approach, Min-Max Scaling. It is a method where the features are converted in a manner that the value of the features becomes within a specific range, usually 0-1. This test and training were required to be performed by the MinMaxScaler type of sklearn.A preprocessing library.

The former is the scaler used on top of the training set with the help of the fit transform algorithm that defined the highest and the lowest possible value of each of the features of the training set and did the transformation. This too was similar to that of the testing set, but again through the transform method that involves scaling of the sets. This will help in the absence of bias in the model because of the difference in the feature size.

The transformation is mathematically defined as:

$$X_scaled = \frac{X - X_min}{X_max - X_min} \quad (3.1)$$

Where X is the original feature value and X_min and X_max are the minimum and maximum values of that feature in the training set, and X_scaled is the normalized value.

This scaling method is particularly important for machine learning and deep learning models, such as MLP, LSTM, and 1D-CNN, as it improves convergence, stabilizes gradients, and ensures that all features contribute proportionally during model training.

(v) Data Analysis and Key Feature Identification

In this work, a complex statistical and exploratory analysis is used to determine the main factors that predetermine the prevalence and severity of dengue. The first was the analysis of data that would help to obtain an idea about its general distribution, tendencies, and important trends in demographic and clinical parameters. A sensitivity analysis was done to estimate the impact of any change in the characteristics of an individual patient on the severity of dengue in the overall sample of data, as well as in subsets (i.e., by age, gender, city (Cumilla, Noakhali, Dhaka), and type of residence (Town, Semi-Town, Rural Area)). One-parameter impact analysis was extremely important in explaining the relationship between the given characteristics (body temperature, duration of fever, platelet count, WBC count, etc) and the progression and severity of the disease. Data balancing and scaling of features created similarity in the variables and ensured equal representation of all classes, which in effect justifies the power of these extra modeling processes.

The statistical significance of the relationship observed during the exploratory and sensitivity analysis was determined through Chi-Square tests of demographic, clinical, and environmental variables such as the severity of dengue. These tests showed the utility of the most important

relationships and offered the possibility to rank the strongest parameters in terms of the choice of features. The fact that all these types of analysis will enable not only to define the most important aspects based on the course of the disease, but also to build an effective, factual base of real, active, and reliable classification of dengue cases. The acquired knowledge in the process can bring a new degree of consciousness of how patient characteristics and the extent of illnesses can interplay in the predictive model and within specific and personalized health promotion initiatives.

⁸3.5 Training and Testing

¹⁰Training and testing are another helpful step that a real dengue prediction and classification model carries out since it provides the algorithm's patterns and its correspondence to the data to work with it. To ensure the generalization, and avoid overfitting, the information is split in ²⁵training and testing set. In this respect, machine learning (ML) and deep learning (DL) models will be used ⁶in this paper, and optimization of the training process will also be modified to the requirements of each model.

The percentage of ³⁸training and testing of the machine learning models will be trained and tested at 70 percent and 30 percent on the dataset, wherein ⁴⁷the machine learning models are being trained on the bigger percentage of the data, and on the data on which the machine learning models are being tested. The deep learning models should include an additional validation set to track the progress in training processes. The three-thirds test is fixed, although the training (70 percent) is fixed as well but further divided into 80 percent training and 20 percent validation. The next step would be a validation step to optimize the hyperparameters and prevent overfitting.

Among the important performance indicators that are observed during the training of the ML and ²³DL to achieve the best and the most stable multi-class dengue severity prediction, there are Accuracy, Precision, Recall, and F1-score. It is a formal model that assists in a good basis where the model will be further explained to consider all the models with completely prepared data to learn positively.

3.6 Model Development for Early Dengue Prediction

In this study, a series of ⁶machine learning and deep learning models are applied to tackle the issues of dengue prediction and classification. They are ⁸Random Forest, Support Vector Machines, XGBoost, ¹⁶Multi-layer Perceptron, Long Short-term memory network, and one-dimensional Convolution Neural Network. Both models are streamlined to tackle the dengue cases at an early age and classify them with a higher diagnostic accuracy. Forecasting capabilities are developed with the intended support and examination to achieve improved disease control strategies.

The fact that dengue cases can be divided into three levels of severity, Mild ⁴²Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF), and Dengue Shock Syndrome (DSS), is arguably one of the

most significant results of the current study. These models make use of machine learning and deep learning to explain patient data and show more complex patterns and relationships, which would otherwise have been hard to detect using more traditional diagnostic tools. The classification process is based on valuable clinical observation, such as platelet count, hematocrit findings, blood pressure, and the presence of certain symptoms.

Mild Dengue Fever (DF): Mild Dengue Fever is defined by the median body temperature at 100.4°F (38 °C), platelet count at 100,000/mmL, and WBC count within the normal range or slightly low. Hematocrit (below 20 percent) is rising, which means a small amount of plasma leakage. The rashes are small and not extensive. This is the most prevalent self-limiting form of dengue, and under supportive care the symptoms can be addressed.

Dengue Hemorrhagic Fever (DHF): DHF is characterized by elevated fever (more than 102°F (38.9 °C), platelet count less than 100,000/ul), and low levels of WBC (leukopenia). The hematocrit is raised by 20 to 30 percent, indicating mild plasma leakage. The patients will have serious rashes such as petechiae and purpura. It involves emergency treatment to deal with bleeding, leaking of plasma, and dehydration.

Dengue Shock Syndrome (DSS): DSS is considered the initial ailment characterized by high fever that may resolve as the illness advances. Platelets are lower than 50,000/uL, and the WBC counts are typically low. Hematocrit exceeds 30 percent, thus indicating extreme plasma leakage. The rashes could have bleeding on them. It is essential to emphasize acute treatment (such as fluid rehydration and acute care) to ensure the absence of complications that can lead to death.

By adding such clinical parameters to the models, it increases the precision of diagnosis, since the healthcare specialists could make informed decisions using the information, leading to a personalised treatment plan. This is not the only such strategy that can be successfully applied to the enhancement and optimization of resource distribution in a healthcare facility, but it also leads to enhanced patient outcomes, fewer complications, and improved survival rates.

3.6.1 Random Forest

Random forest is a supervised learning algorithm. It builds a forest of an ensemble of decision trees, which are usually trained with the bagging method. The fundamental idea behind the bagging approach is that as you combine multiple learning models, the overall output is improved. Random Forest algorithm added additional randomness in splitting a node; the algorithm finds the most optimal feature in a random set of features. This gives rise to a higher tree diversity, which gives up a higher bias in favor of a low variance, typically giving an overall better model. Thus the conditions under which the random forest works well are: There must be some real signal in our features such that models constructed with these features perform better than random guessing. In addition, predictions (and hence errors) by the individual trees must be lowly correlated with each other [52].

The main characteristics of the algorithm are:

1. Shorter training duration than other algorithms.
2. Good predictive power even when using large data sets.
3. Strength in the ability to stay accurate even where there was a substantial amount of missing data.

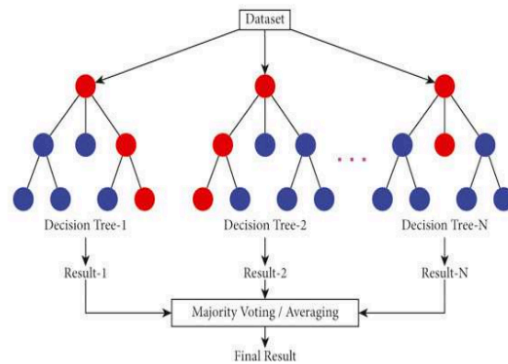


Figure 3.4: Random Forest [53].

The random forest formation has two steps, firstly building the forest by pooling N decision trees and secondly prediction is made using these trees. The process involves:

1. Randomly sampling K data points of the training data.
2. Developing decision trees regarding the chosen data points.
3. Indeed, it is the number N of decision trees required.
4. Iterating Steps 1 and 2.
5. Making predictions using the aggregation of the predictions made by separate trees to give new data point predictions, and giving them the classification of the majority of the votes.

3.6.2 Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning model, which is used in regression and classification problems. The objective of SVM is to identify the optimum decision

boundary or the hyperplane separates the different classes in the feature space. The method SVM has advantages and strength in tackling problems of classification as it can generalise better the unseen data by maximising the margin among the classes.

The concepts of SVM are:

- Hyperplane: A line (or plane) to differentiate the different classes that is represented in a linear classification as $w^T x + b = 0$ where w is the normal and b is the bias.
- Support Vectors: The data points closest to the hyperplane that identifies where it is and the direction it faces.
- margin: The distance between the support vectors and the hyperplane, SVM attempts to make this as large as possible to improve classification.
- Kernel: It is a process to project a collection of data into a more dimensional form in order to process a non-linearly separable data efficiently. Popular kernels include linear, radial basis function (RBF) and polynomial kernels.
- Soft Margin: In order to allow generalization in cases where the data cannot be separated, some misclassifications are allowed.
- Hinge Loss: A hinge loss is a loss function that does not just penalize those points that cross the margin but encourages the model to also learn to correctly classify and maximize the margin.

The simplification of SVM goes as follows: SVM identifies the hyperplane that offers the best margin between the classes. When the data is linearly separable, the algorithm finds the hyperplane that exactly fits the two classes. A soft-margin then allows the margin to be violated in certain instances in order to induce better generalization, and the data are mapped to higher dimensional spaces where one can separate them via a kernel function.

Supposing a binary classification problem where the classes have +1 and -1 labels, and input vectors X and labels Y , the linear hyperplane can be mathematically defined as follows:

$$w^T x_i + b = 0 \quad (3.2)$$

Here, w represents the normal vector that is perpendicular to the hyperplane, and b represents the bias term that represents the distance between the origin and the hyperplane. The proximity of data x_i to the hyperplane may be calculated as:

$$d_i = \frac{|w^T x_i + b|}{\|w\|} \quad (3.3)$$

w is the weight-vector and is the Euclidean norm of the weight-vector w . The hyperplane that maximizes the distance between the nearest support vectors of both of the classes is chosen as the optimum decision boundary.

By way of this approach, SVM is able to classify not only linearly separable data, but also non-linearly separable data, thereby making it a flexible and a powerful tool in this research, in order to classify multi-class dengue severity [54].

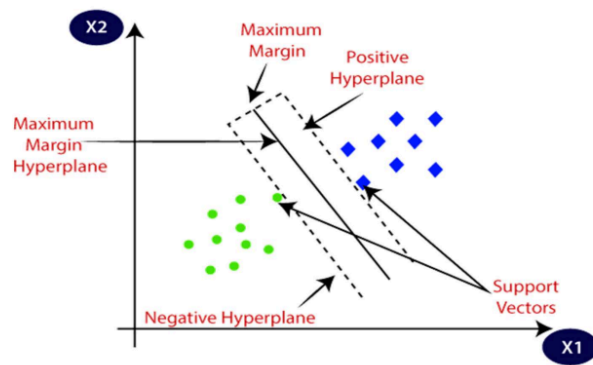


Figure 3.5: Support Vector Machine [55].

3.6.3 XGBoost

XGBoost (eXtreme Gradient Boosting) is a high-speed, scaled and accurate implementation of the Gradient Boosting algorithm, tuned to be best at speed, scale, and accuracy. It is a sub area of ensemble learning whereby a group of models (weak learners) makes a better and stronger model. There is, however, no need to simplify the patterns, as with other traditional machine learning models, such as decision trees and random forests, but instead, XGBoost is capable of working with complex patterns, or large datasets, and it employs parallel computation, regularization, and advanced tree pruning to optimize the learning process without inducing overfitting. XGBoost has been used in data science contests, financial modeling, medical diagnosis, and anomaly detection because it is efficient and capable of more organized data processing (high-dimensional) [56].

XGBoost is a powerful predictive model because it automatically generates decision trees so that the error of any previous tree is minimized by the next tree. The big idea behind XGBoost is the following:

- **Initialize a Base Model:** The algorithm starts by constructing a Base Model: A weak model, typically a naive decision tree. This base learner is able to give an initial guess that predicts the mean of target variable in regression problems.
- **Calculate Residual Errors (Loss Function Computation):** The model will compare the prediction with the actual value after training and the difference between the models will be calculated (residual errors). These residual values indicate how much the model values do not coincide with the real values.
- **Train a New Tree to Correct Errors:** The second decision tree is trained on the mistakes, not on the target variable itself, and attempts to predict the mistakes of the previous tree. This minimizes prediction bias and helps to achieve fine tuning of the model.
- **Update Predictions with a Learning Rate:** An old tree makes predictions that are multiplied by a learning rate (shrinkage factor) and added to the old predictions. This prevents excessive adjustment and offers gradual and gentle improvements.
- **Repeat Until Stopping Criteria is Met:** This process continues, adding new trees iteratively, until one of the stopping criteria is met:
 - A maximum number of trees is reached.
 - The model's improvement becomes negligible after a certain number of iterations.
 - An early stopping condition is triggered (if validation error stops decreasing).
- **Final Prediction:** Final prediction is the accumulation of all the tree output with their contribution. The final result is that since each of the individual trees is designed to decrease the errors of the other existing trees, the final tree will be more accurate and less biased than the individual decision trees.

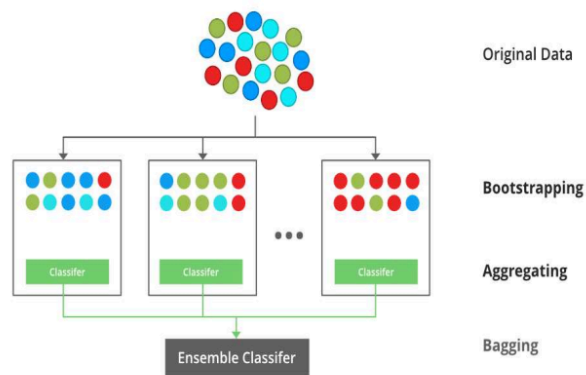


Figure 3.6: XGBoost [57].

3.6.4 Long Short-Term Memory Networks

The limitation of traditional RNNs is that they learn only short-term dependencies, as the time-varying hidden state is only passed through a single hidden state. Long Short-Term Memory (LSTM) addresses this problem by introducing a memory cell that can remember information over a long period of time. This allows LSTM networks to efficiently represent long-term correlations in orderly data, which is why they are useful in applications such as language translation, speech recognition, and time series forecasting. Other neural networks like CNNs which are used to analyse images and videos can also be supported by LSTMs.

There are three gates that control the memory cell: input gate, forget gate, and output gate. These gates control the input, deletion, and deliver of information to the memory cell respectively. The input gate controls the information that enters the memory cell, the forget gate controls information removal, and the output gate controls information that is released by the memory cell. This process enables LSTM networks to be selective when it comes to retaining or discarding information as it passes through the network, which aids the learning of long-term dependencies [58].

LSTM architecture has a chain structure that contains four neural networks and different memory blocks called cells.

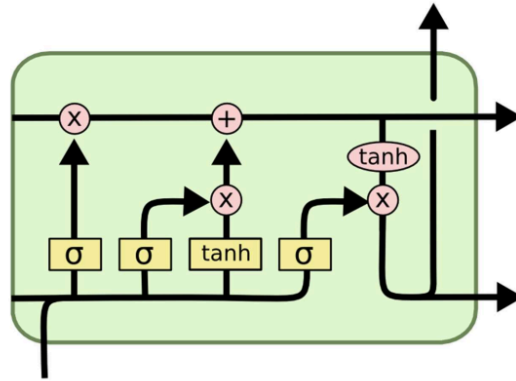


Figure 3.7: Long Short-Term Memory Network [59].

Information is retained by the cells, and the memory manipulations are done by the gates. There are three gates:

Forget Gate: The forget gate is responsible to forget information that otherwise is irrelevant in the cell state. It takes two inputs, x_t (the input during the given time step) and h_{t-1} (the immediate previous cell output) and multiplies them by matrices of weights and sums them out using biases. The resulting values are then passed through an activation function leading to a binary output. It can be said that once a certain kind of cell state has led to the production of 0, then the opposite of this is also true: once the production of 1, and then the same data is remembered to be used later. The forget gate equation

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3.4)$$

Here:

W_f denotes the weight matrix for the forget gate.

$[h_{t-1}, x_t]$ represents the concatenation of the current input and the previous hidden state.

b_f stands for the bias associated with the forget gate.

σ denotes the sigmoid activation function.

Input gate: The input gate aids in the absorption of valuable information within the cell state. The first one will be to use the sigmoid functions processing in the forget gate where h_{t-1} and x_t are the inputs. The function Tanh is then used to create a vector whose values lie between +1 and -1, any potential value of h_{t-1} and x_t can be achieved. Finally, the process amplifies the values of the processed ones to access the relevant information. Input gate equations can be written as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3.5)$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3.6)$$

We then multiply the previous state by f_t , thus excluding the previously chosen ignored information, and include $i_t \cdot \hat{C}_t$. This denotes the updated candidate values, adjusted according to the extent to which we opted to update each state value.

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \quad (3.7)$$

Output gate: The output gate does the task to extract useful information to be presented as output out of the current cell state. The start point is to create a vector by implementing the tanh operation over the cell. Subsequently, the data is controlled based on the sigmoid function and filtered by the data to be recalled based on the inputs h_{t-1} and x_t . Finally, the values of the vector are multiplied with the regulated values and sent as an output and an input to a cell. The output gate equation is:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3.8)$$

3.6.5 Multi-Layer Perceptron

A Multi-Layer Perceptron (MLP), is a device comprised of fully connected dense layers, which convert input data of one dimension to another. It is multi-layer in nature in that it has an input layer, 1 or more hidden layers, and an output layer. An MLP is intended to represent complex inputs versus outputs relationships [60].

Multi-layer Perceptron (MLP) Components:

- **Input Layer:** The input layer has one neuron or node that is associated with an input feature. To give an example, when you have three input features, the input layer will contain three neurons.
- **Hidden Layers:** MLP may have multiple hidden layers with each having any amount of nodes. These layers process the information given on the input layer.
- **Output Layer:** The ultimate forecast or result is manufactured by the output layer. When the number of outputs is more than one, the number of neurons in the output layer will be equal.

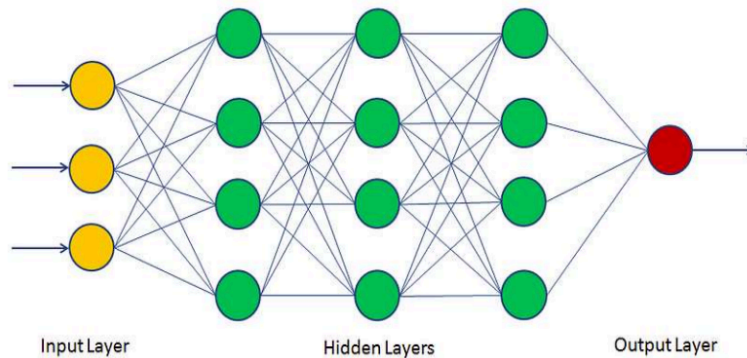


Figure 3.8: Multilayer Perceptron [61].

Every connection in the diagram is a representation of the fully connected nature of an MLP. This means that every node in one layer connects to every node in the next layer. As the data moves through the network, each layer transforms it until the final output is generated in the output layer.

Working of Multi-Layer Perceptron: Let's see the working of the multi-layer perceptron. The key mechanisms such as forward propagation, loss function, backpropagation, and optimization.

1. Forward Propagation

In forward propagation, the data flows from the input layer to the output layer, passing through any hidden layers. Each neuron in the hidden layers processes the input as follows:

(a). **Weighted Sum:** The neuron computes the weighted sum of the inputs:

$$z = \sum_i w_i x_i + b \quad (3.9)$$

Where:

- x_i is the input feature.
- w_i is the corresponding weight.
- b is the bias term.

(b). **Activation Function:** The weighted sum z is passed through an activation function to introduce non-linearity. Common activation functions include:

- Sigmoid: $\sigma(z) = \frac{1}{1+e^{-z}}$ (3.10)

- ReLU (Rectified Linear Unit): $f(z) = \max(0, z)$ (3.11)

- Tanh (Hyperbolic Tangent): $\tanh(z) = \frac{2}{1+e^{-2z}} - 1$ (3.12)

2. Loss Function

Once the network generates an output, the next step is to calculate the loss using a loss function. In supervised learning, this compares the predicted output to the actual label.

For a classification problem, the commonly used binary cross-entropy loss function is:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(y^{\wedge}_i) + (1 - y_i) \log(1 - y^{\wedge}_i)] \quad (3.13)$$

Where:

- y_i is the actual label.
- $y^{\wedge}_i$ is the predicted label.
- N is the number of samples.

For regression problems, the mean squared error (MSE) is often used:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y^{\wedge}_i)^2 \quad (3.14)$$

3. Backpropagation

The goal of training an MLP is to minimize the loss function by adjusting the network's weights and biases. This is achieved through backpropagation:

1. Gradient Calculation: The gradients of the loss function with respect to each weight and bias are calculated using the chain rule of calculus.
2. Error Propagation: The error is propagated back through the network, layer by layer.
3. Gradient Descent: The network updates the weights and biases by moving in the opposite direction of the gradient to reduce the loss:

$$w = w - \eta \cdot \frac{\partial w}{\partial L} \quad (3.15)$$

Where:

- w is the weight.
- η is the learning rate.
- $\frac{\partial w}{\partial L}$ is the gradient of the loss function with respect to the weight.

4. Optimization

MLPs rely on optimization algorithms to iteratively refine the weights and biases during training. Popular optimization methods include:

- Stochastic Gradient Descent (SGD): Updates the weights based on a single sample or a small batch of data: $w = w - \eta \cdot \frac{\partial w}{\partial L}$ (3.16)
- Adam Optimizer: An extension of SGD that incorporates momentum and adaptive learning rates for more efficient training:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \cdot g_t \quad (3.17)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (3.18)$$

Here, g_t represents the gradient at time t , and β_1, β_2 are decay rates.

3.6.6 1D Convolutional Neural Network

One-Dimensional Convolutional Neural Network is a type of neural network, which can be applied to analyze one-dimensional ordered or structured data. Though two-dimensional CNNs are most frequently applied to images due to the grid-like structure, 1D-CNNs are intended to process ordered data, where the sequence of objects matters considerably. They contain text data, audio waveforms, biomedical measurements and time series signals.

The intrinsic nature of 1D-CNNs is that it can learn local dependencies and short-range relationships among data automatically. An illustration of this is that immediate physical reading of physiological measurements may tell about some important patterns in medical data, but when reading text the closeness of words may be used to explain a semantic relationship. 1D-CNNs have convolutional filters along one axis to learn these localized features and require only a few

manual features to learn these features; this is why 1D-CNNs are computationally efficient and effective[62].

The functioning of a One-Dimensional Convolutional Neural Network (1D-CNN) can be described through a series of sequential stages:

- **Input Layer:** This layer accepts one-dimensional sequential data, such as sensor readings, biomedical test results, or encoded signal sequences, serving as the raw input to the network.
- **Convolutional Layer:** A set of filters (kernels) slides along the input sequence, performing element-wise multiplication with localized regions of the data. The results are summed to produce feature maps, which capture localized dependencies and highlight significant patterns such as spikes, fluctuations, or contextual groupings. Mathematically, the convolution operation at position i is expressed as:

$$y_i = \sum_{j=0}^{k-1} x_{i+j} w_j + b \quad (3.19)$$

where y_i is the output feature at position i , x_{i+j} is the input values, w_j are the filter weights, k is the kernel size, and b is the bias term.

$$y_f[i] = \sum_{c=1}^C \sum_{j=0}^{k-1} x_c[i+j] w_{f,c,j} + b_f \quad (3.20)$$

where C is the number of input channels, K is the kernel size, $w_{f,c,j}$ is the weight of filter f at channel c and index j , and b_f is the bias term for filter f .

- **Activation Function:** Non-linear activation functions, most commonly the Rectified SLinear Unit (ReLU), defined as

$$f(x) = \max(0, x)$$

are applied to the feature maps. This step introduces non-linearity into the model, enabling it to capture complex relationships in the data.

- **Pooling Layer:** The pooling operation reduces the dimensionality of feature maps by selecting the most significant information (e.g., through max-pooling or average pooling). This not only lowers computational requirements but also enhances robustness by discarding irrelevant variations.
- **Flattening Layer:** The reduced feature maps are converted into a one-dimensional vector, preparing the data for processing in the fully connected layers.
- **Fully Connected (Dense) Layers:** These layers integrate the extracted features from previous stages, learning high-level abstractions and combining information across the entire input sequence.
- **Output Layer:** The final layer generates predictions. A Softmax activation is typically employed for classification tasks, producing class probabilities, whereas a linear activation is used for regression problems to provide continuous outputs.

The 1D-CNNs are fairly resistant to local sequence-wise and feature-wise correlations. They automatically produce hierarchical feature of input and consequently removes the labor of hand engineering features and enhances the model performance. That enables the application of 1D-CNNs to any research domain of interest, including time-series prediction, speech and audio signals, natural language processing and biomedical signal analysis. Sequential data analysis is a key aspect of the present deep learning as it is simple to derive intricate patterns with ease at computation scale.

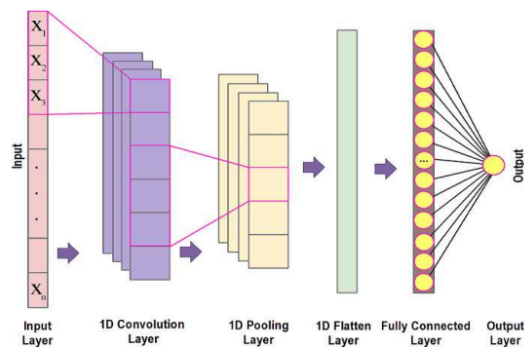


Figure 3.9: 1D Convolutional Neural Network [63].

3.7 Proposed Hybrid RF-LSTM Model

This paper introduces a hybrid RF-LSTM network which will be utilized to identify severity of dengue under four classes, that is, No Dengue, Mild Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS). The hybrid model combines the advantages of the Random Forest and Long Short-Term Memory networks, as it will use RF to effectively select features and predict probabilities and LSTM to learn complicated non-linear patterns. The design will permit the model to focus on the most pertinent characteristics, the complex links among those, and the severity predictions to be highly dependable and precise. The methodological framework of the proposed hybrid RF-LSTM model is illustrated in Figure 3.10.

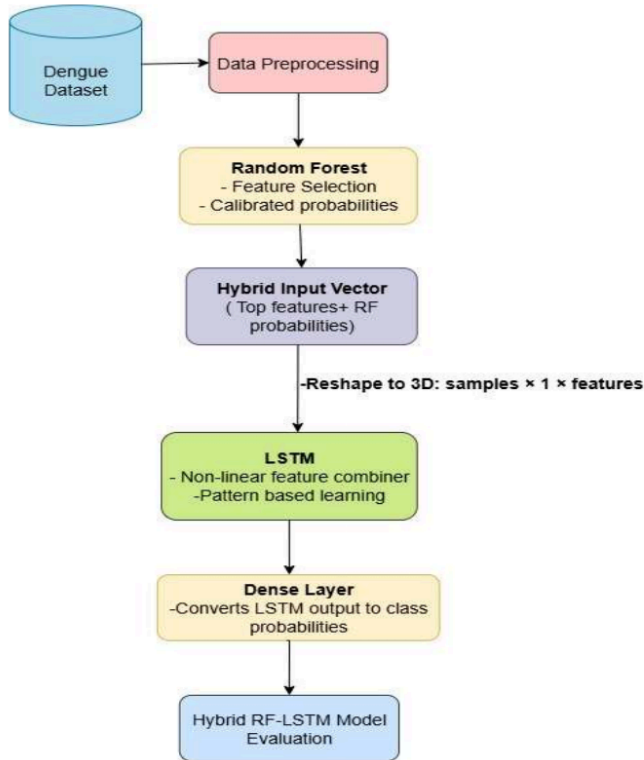


Figure 3.10: Methodological Framework of Proposed Hybrid Random Forest-LSTM Model

This work is based on the pretrained base models; however, this research selects a hybrid strategy between the Random Forest and LSTM. Random Forest determines the most informative features to the severity of dengue and generates probability estimates that are calibrated to each of the classes. These probabilities give a probabilistic explanation on the probability of every event, which has guaranteed realistic forecasts. With a combination of the best features with the calibrated probabilities and RF generates a small and information-rich feature of each patient instance, cutting the noise and highlighting the most pertinent data to be further processed.

The Hybrid Input Vector is formed by the combination of the best features that are chosen by RF with the RF-calibrated probabilities. The raw feature values are combined with the probabilistic knowledge of RF in this vector to create a small and richer representation of each patient case. It takes key patient-level and environmental data, which improves prediction efficiency of the next LSTM network.

In order to process the hybrid vector using LSTM, the vector is reshaped in to a 3D tensor with format: samples timesteps features, timesteps = 1 and each sample is regarded as a single pseudo-timestep. This change is needed since non-sequential tabular features demand LSTM networks to accept 3D inputs so that it can process the data accordingly. The LSTM network is fed with the transformed hybrid input vector, consisting of the best RF-selected features in addition to the RF-calibrated probabilities.

The LSTM is a deep non-linear feature combiner and it learns complicated interactions between features that could not be learned by standard models. Its gated nature has an input, forget and output gate to ensure that only the most informative patterns are stored with redundant or irrelevant patterns being discarded. Besides the combination features, the LSTM automatically finds latent dependencies and interactions in the composite vector. RF probability outputs are learned in an indicative manner and the final forecast relies on the extracted patterns. The LSTM generates a high-dimensional feature representation, which is fed into the dense layer a classifier that changes these representations into probabilities of each severity of dengue class, making final multi-class classification possible. This increases RF outputs and also improves predictive performance of the hybrid model on the whole.

Regular classification measures, including accuracy, precision, recall and F1-score are used to measure the final performance of the hybrid RF-LSTM model. Confusion matrices give a visual representation of correctly and wrongly classified data, and training and validation curves show whether convergence is appropriate and that possible overfitting is identified. The hybrid model can efficiently reduce noise, dimensionality, and render complex feature-to-feature relationships, which with RF-based feature selection and probability calibration, makes the hybrid model a robust, reliable, and highly accurate framework of predicting the severity of dengue.

3.8 Performance Analysis

Performance analysis involves thorough observations in order to improve overall efficiency, help in making data-driven decisions, and provide useful feedback to refine the system. It is a detailed evaluation of the extent to which some situation fits the desired aim. The models developed will be evaluated on the basis of relevant evaluation measures including:

- a) Confusion Matrix
- b) Accuracy
- c) Precision
- d) Recall
- e) F1 score

a) Confusion Matrix: A confusion matrix is viewed as one of the methods to evaluate the work of a classification model with the help of a screen of $N \times N$ where N is assumed to be the classification of classes. This matrix may be compared with the values of the target and estimated machine learning models. The outcomes of a confusion matrix contain one of the graphs which explain the general performance of a classification model or algorithm used in the machine learning processes. It has been drawn out in the following figure:

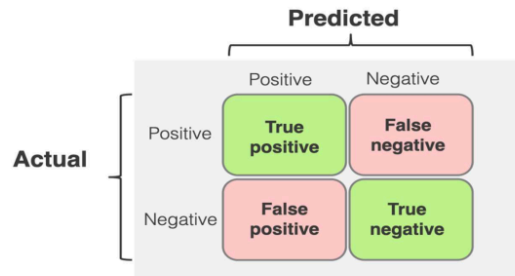


Figure 3.11: Confusion Matrix [37].

Target variable of binary classification is classified into two values Positive or Negative. The actual values of target variable will be in columns and the estimated values of target variable will be in rows.

⁴⁴
True Positive (TP): The predicted outcome aligns with the actual positive value. The model accurately identified and predicted a positive result, where both the predicted and actual values are positive.

True Negative (TN): The predicted value aligns with the actual value, indicating a negative outcome. This implies that the actual value was negative, and the model accurately predicted a negative result.

False Positive (FP) – Type 1 error: The model produced an incorrect prediction, leading to a Type 1 error. More precisely, the actual value was negative, but the model mistakenly predicted a positive outcome.

False Negative (FN) – Type 2 error: The model made a false prediction, leading to a Type 2 error. To clarify, the actual value was positive, yet the model mistakenly predicted a negative outcome.

- b) **Accuracy:** Accuracy, an essential metric in machine learning, gauges the model's effectiveness in predicting outcomes for binary and multiclass classification challenges. It calculates the accuracy rate by comparing the total correct predictions with the overall count of input samples. The mathematical formula for accuracy is:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+F} \quad (3.21)$$

- c) **Precision:** Precision, a vital metric in assessing classification model performance, signifies the proportion of true positive predictions to the total number of positive predictions issued by the model. The mathematical expression for precision is:

$$Precision = \frac{TP}{(TP+F)} \quad (3.22)$$

- d) **Recall:** In the realm of machine learning, recall serves as a performance metric, gauging a model's capability to accurately recognize all pertinent instances of a specific class or category within a dataset. It signifies the model's proficiency in avoiding false negatives, holding particular significance in applications like medical diagnosis or fraud detection. Mathematically, recall is computed as the ratio of true positives (TP) to the sum of true positives and false negatives (FN). Mathematically, recall can be expressed as:

$$Recall = \frac{TP}{(TP+F)} \quad (3.23)$$

- e) **F1 Score:** The F1 score, a consolidated metric merging precision and recall, provides a singular measure for evaluating test accuracy in machine learning. The F1 score, as the harmonic mean of precision and recall, provides a balanced representation, with numerical values limited to the interval of 0 to 1.

$$F1 \text{ Score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (3.24)$$

3.9 Performance Comparison of Distinct Learning Approaches for Dengue Classification

This study includes a systematic comparison of machine learning, deep learning, and a hybrid approach to inform positively about the classification of dengue cases. Consistent with respect to evaluation conditions, we use a suite of benchmark models (Random Forest, XGBoost, Support Vector Machine, Multi-Layer Perceptron, Long Short-Term Memory Network, and 1D Convolutional Neural Network) to establish reliable baseline comparisons. Then, we evaluate the Hybrid RF-LSTM model in the same evaluation context as the other models to provide context for its relative performance. Accuracy, precision, recall and F1-score were used as the main performance metrics to facilitate a fair and rigorous comparison between classical machine learning methods, classical deep learning methods, and a hybrid approach.

3.10 Comparative analysis with existing works

This study not only conducts an internal benchmarking exercise of models but also compares with existing literature on dengue prediction and classification. A systematic review of previous studies that utilized climatic, epidemiological, and hospital-based datasets aims to provide context to the proposed approach's performance. These past studies used methodologies like Random Forest, Support Vector Machine, Extreme Gradient Boosting, Artificial Neural Network, and Long Short-Term Memory, which allows this work to reference their methodology and contextualize its own scientific contribution. The focus of many of these existing methodologies fell primarily on outbreak prediction or early detection of a dengue epidemic, while the current study's focus is to classify patients with dengue into either Mild Dengue Fever, Dengue Hemorrhagic Fever, or Dengue Shock Syndrome. By situating the proposed Hybrid Random Forest-Long Short-Term Memory methodology within the studies reviewed, this study is able to validate methodological rigor, while also contributing to the body of research by improving reliable dengue case classification. The classification data and spatial risk mapping provide even more applied value of the study to public health capacity by indicating places such

as Cumilla, Noakhali, and Dhaka were high-risk areas, leading to improved targeted healthcare planning and intervention.

3.11 Risk Zone Map and Dengue Prevention and Awareness

In this paper, a visualization of the spatial distribution of dengue in the areas of choice in Bangladesh was done using the Risk Zone Mapping. Risk Zone Maps is a type of geospatial information that attempts to investigate the target of diseases in more and less areas of the nation and a much more casual concept of the occurrence of disease and potential hot spots. These maps will allow visualizing the numerical data on the case data and will provide an opportunity to introduce some health interventions to the population because numerical data can easily be transformed into the type that will provide an opportunity to make interpolations of the numerical data more effective.

In this research, the dengue cases in the three study areas were mapped out both in terms of the regional map, which revealed the general distribution, and the local map, which indicated the finer spatial variations with a color-coded scheme showing areas with high and low case concentration. To develop into a useful visual display of dengue incidence, the Risk Zone Maps depict a blend of geographic distribution, clinical, and demographic information. It is also a beneficial method as it will enable the identification of the potentially vulnerable regions, which will be used in the resource prioritization and enable the planning of developing the evidence.

¹⁹ Dengue is one of the biggest public health problems that the world is facing, and prevention should be at the center of the disease management approach. The effectiveness of the prevention will entail the adoption of an integrated approach which combines the vector control, personal protection, and community awareness. Vector control is the removal of mosquito breeding areas such as the removal of water that has been kept in an area or a certain type of container, putting on insecticides or larvicides, and sometimes using some fish that are always near to the larvae like the larva-eating fish. If an individual is using the methods like mosquito nets, repellents and keeping houses clean, they will expose themselves to the bites of mosquitoes to a small extent, while community-focused awareness programs may also be a source of information for early detection and postgraduate behaviors. The programs that targeting communities are delivered through schools, media, and health campaigns which help in early detection and proactive behaviors. Within this research context, the risk zone mapping of Cumilla, Noakhali, and Dhaka provides a best way to pinpoint the areas where the populations are most at risk for dengue outbreaks. The use of geospatial analysis connected with prevention strategies enables the health authorities to make the best use of mosquito control programs, awareness campaigns, and resources to the areas where the need is greatest. This connection guarantees that the preventive measures are not only widespread but also targeted thus being more organized, acting before the situation gets out of hand, and being sustainable in the reduction of the outbreak that may spread far and wide.

3.12 Required Tools

This study is done using different tools. They are all open and free. These tools are briefly described below:

1. **Python:** Python is a general-purpose and high-level programming language, most frequently used in machine learning, deep learning, and data science applications. It forms the basis of the application of models and the management of data in this research.
2. **Virtual Environment (Anaconda & Jupyter Notebook / Google Colab):** Anaconda framework offers Jupyter Notebook and other programs, which can help perform the experiments. A virtual environment will be established within Anaconda in which all the necessary libraries will be installed. Moreover, Google Colab will be applied as a cloud solution to execute models effectively and with the use of a GPU.
3. **Scikit-learn:** Scikit-learn is a popular machine learning library in Python offering easy and effective data preparation, model creation, and assessment. Its application in this study is in carrying out machine learning models and feature selection.
4. **NumPy:** NumPy is a Python numerical computing package. It supports arrays and matrices as well as a set of mathematical functions, which allow it to do scientific computation with large data volumes efficiently.
5. **Matplotlib:** Matplotlib is a Python plotting library that is typically used alongside NumPy and offers both static and interactive plotting. It applies to the plotting of graphs, performance measurements, and visualization of the study outcomes.
6. **Pandas:** Pandas is an open-source Python library that offers high-performance manipulation and analysis of data. It is applied in cleaning, organising, and managing structured patient and environmental data effectively.
7. **Seaborn:** Seaborn is a Python data visualization library created on Matplotlib. It offers a more advanced user interface to create appealing and informative statistical graphics that aid in the evaluation of features and correlation visualization.
8. **SciPy:** SciPy is another scientific computing library that is based on NumPy. It offers basic mathematical functions, optimization, and statistical techniques that can be applied in any higher-level data analysis and pre-processing.
9. **OS:** Python has an OS module that offers operating system interaction functionality. It provides a mobile means of providing operating system-dependent functionality, including file processing, directory processing, and interaction with the environment.

10. **Plotly:** Plotly is a library of interactive and dynamic visualization written in open-source. Here, it is applied to create detailed charts and dashboards in order to obtain an interactive vision of dengue case trends, severity distributions, and time-related patterns.
11. **GeoPandas and Folium:** GeoPandas is an open-source extension of Pandas enabling operations on spatial data, and Folium is an open-source interactive leaflet library. This study uses them jointly to map the distribution of dengue cases in Cumilla, Noakhali, and Dhaka to facilitate the development of interactive risk zone maps.
12. **GeoJSON:** GeoJSON is a format of an open standard based on the encoding of geographic data structures defined in JSON. The district and sub-district boundaries of Bangladesh are represented in this study with the use of GeoJSON files, that enable the spatial mapping of the dengue case information gathered. Interactive maps have been created using GeoJSON and Folium to identify areas on the map that have a high, medium, or low risk of cases.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter describes the results of the research focused on improving dengue fever prediction and classification using modern machine learning and deep learning algorithms. Dengue remains a major public health issue in Bangladesh, especially in the districts of Cumilla, Noakhali, and Dhaka. The proper identification and classification of cases are crucial for timely and effective action. The findings start with a sensitivity analysis of key factors such as age, gender, residence type, seasonal variation, and city of origin. Then, the individual contribution of these parameters is evaluated. To capture the spatial differences of the disease, risk zone mapping is done for the three districts. This identifies areas at higher risk and regions with increased incidence rates. Additionally, the classification performance of several algorithms, including Random Forest, Support Vector Machine, XGBoost, Multi-Layer Perceptron, Long Short-Term Memory, and 1D CNN, is assessed. The chapter also introduces the Hybrid RF-LSTM model, which combines feature-based and pattern-based learning to improve predictive ability. The strength of these models is tested using unseen data. Finally, the chapter concludes with a comparison of performance metrics and existing research, placing the findings within a larger research context.

4.2 Sensitivity Analysis of Distinct Features Influencing Dengue Severity

In this study, sensitivity analysis was conducted to evaluate how variations in clinical, demographic, and environmental features influence dengue severity. This systematic approach identifies the factors that most strongly impact disease progression, enabling researchers and healthcare professionals to prioritize key parameters for assessment and monitoring. Clinical features such as platelet count, blood pressure, and body temperature not only serve as diagnostic indicators but also play a crucial role in guiding clinical decisions.

As shown in Figure 4.1, platelet count emerged as the most sensitive and influential feature, consistently reflecting changes in patient condition. Other laboratory parameters, including white

blood cell (WBC) count and body temperature, also demonstrated significant sensitivity, highlighting their role in systemic and immune responses. Cardiovascular indicators, such as heart rate and blood pressure, showed moderate sensitivity, while symptom-related features headache, vomiting, eye pain, and fever duration contributed moderately but with less consistency across severity levels. Environmental and demographic features such as mosquito bite history, waterlogging near residence, humidity, and gender exhibited lower sensitivity. While important for epidemiological context, these factors were less directly associated with clinical severity in this dataset.

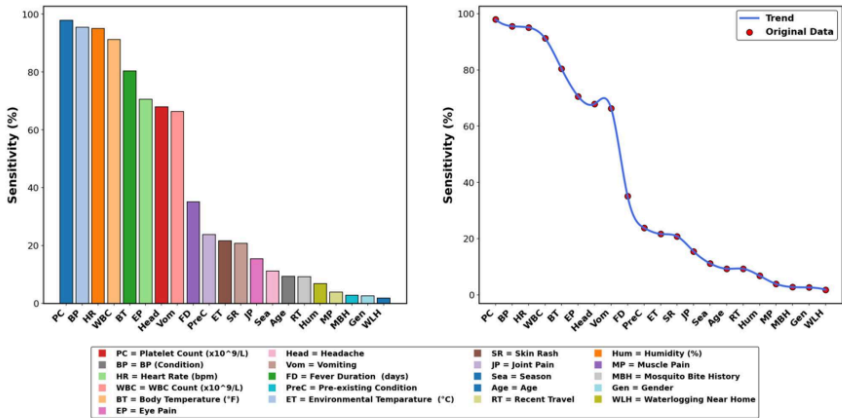


Figure 4.1: Sensitivity analysis of all features influencing Dengue Severity

4.2.1 Sensitivity Analysis for Age Groups

The sensitivity analysis across different age groups highlights how the influence of clinical, physiological, and contextual features varies with age in relation to dengue severity. In Figure 4.2(a), for the age group 15–30 years, Heart Rate, Blood Pressure, and Platelet Count emerge as the most significant features, while Joint Pain and Vomiting contribute the least. In Figure 4.2(b), for the age group 31–45 years, Heart Rate, Platelet Count, and White Blood Cell Count play a dominant role, whereas Fever Duration and Environmental Temperature show lower influence. Figure 4.2(c), for the age group 46–60 years, again highlights Platelet Count and White Blood Cell Count as key features, followed by Blood Pressure and Heart Rate, with increasing importance of Pre-existing Conditions and Vomiting. In Figure 4.2(d), for the age group 61 years and above, Platelet Count, White Blood Cell Count, Season, and Pre-existing Conditions are the most influential, reflecting the growing impact of chronic and environmental factors, while Heart

Rate and Environmental Temperature are relatively less significant. Overall, Platelet Count and White Blood Cell Count remain the most consistent and critical key features across all age groups.

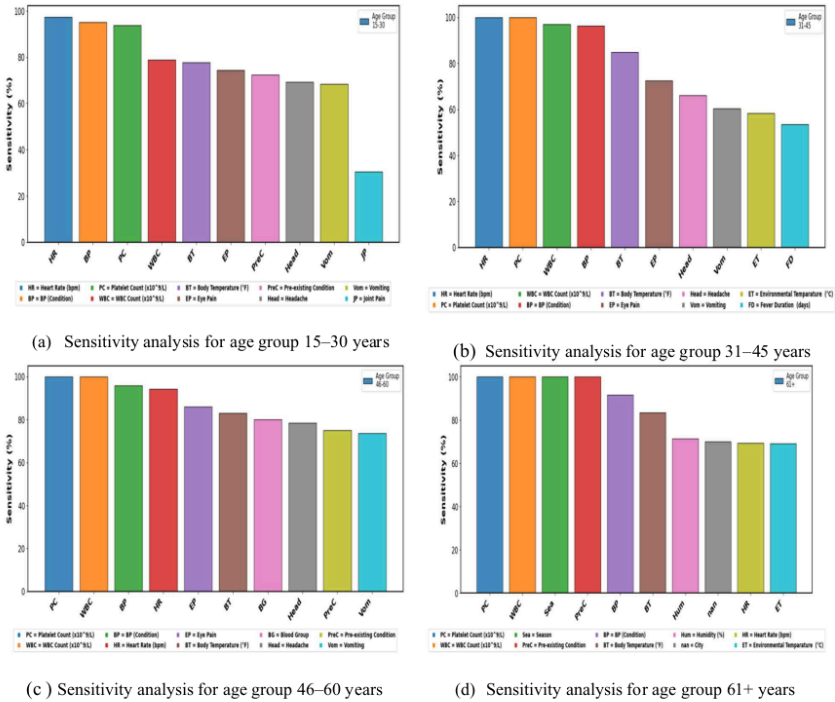


Figure 4.2: Sensitivity Analysis of Age Groups Influencing Dengue Severity

4.2.2 Sensitivity Analysis for Genders

The sensitivity analysis based on gender highlights how the influence of various clinical and physiological features differs between male and female individuals in relation to dengue severity. In Figure 4.3(a), among males, White Blood Cell Count, Platelet Count, Heart Rate, and Blood Pressure emerge as the most significant features, while Pre-existing Conditions and Blood Group are among the least influential. In Figure 4.3(b), for females, Platelet Count, Heart Rate, and

Blood Pressure dominate as key features, followed by White Blood Cell Count and Body Temperature, whereas Blood Group and Fever Duration show minimal contribution. Overall, Platelet Count and Heart Rate consistently appear as dominant features in both genders, but males show slightly more sensitivity to White Blood Cell Count, while females demonstrate relatively higher relevance for Body Temperature and Vomiting.

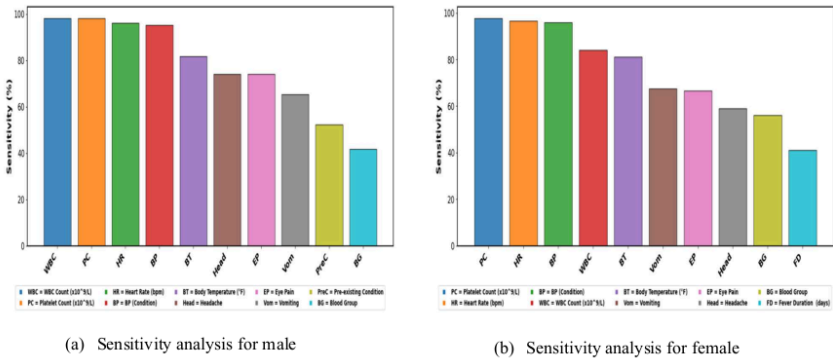


Figure 4.3: Sensitivity Analysis of Genders Influencing Dengue Severity

4.2.3 Sensitivity Analysis for Cities

The sensitivity analysis across different cities highlights distinct variations in how clinical and physiological features influence dengue severity. In Figure 4.4(a), representing Cumilla, platelet count, blood pressure, heart rate, and white blood cell count emerge as the most influential features. Body temperature and eye pain follow in importance, while muscle pain shows the least impact. In Figure 4.4(b), corresponding to Noakhali, heart rate stands out as the top contributing factor, closely followed by platelet count, blood pressure, and white blood cell count. Features such as body temperature, eye pain, and vomiting have moderate significance, whereas fever duration has the lowest influence. In Figure 4.4(c), illustrating Dhaka, platelet count, white blood cell count, and blood pressure are the leading indicators, while heart rate and body temperature maintain moderate relevance. Features like vomiting, eye pain, season, and fever duration contribute comparatively less.

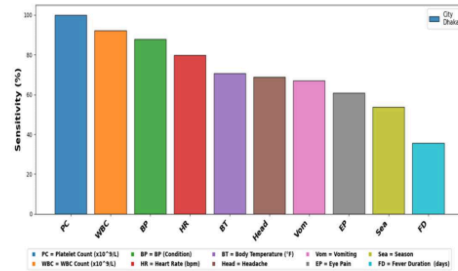
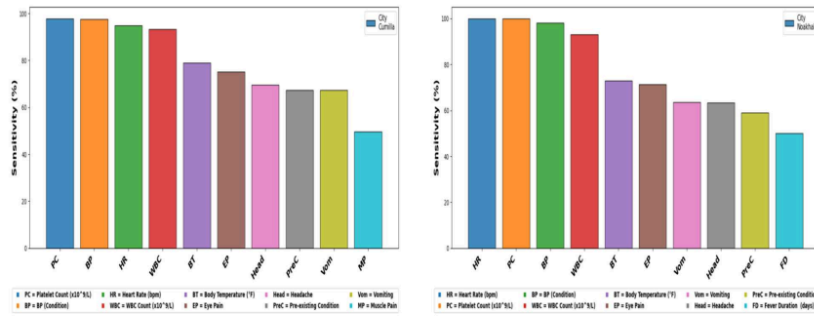
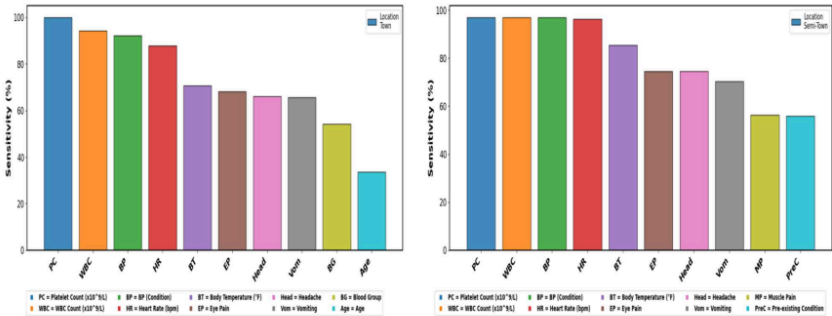


Figure 4.4: Sensitivity Analysis of Cities Influencing Dengue Severity

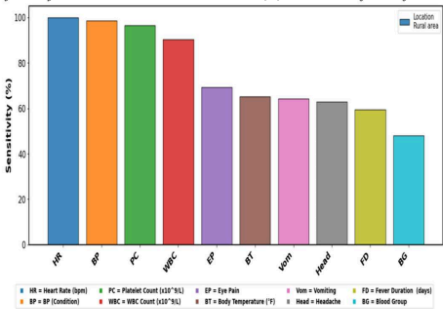
4.2.4 Sensitivity Analysis for types of Residence

The sensitivity analysis of residence types influencing dengue severity reveals notable differences in how clinical and physiological features affect the condition. In Figure 4.5(a), representing Town areas, platelet count emerges as the most influential feature, followed by white blood cell count, blood pressure, and heart rate. Body temperature and eye pain have moderate importance, while headache, vomiting, blood group, and age exhibit lower impact. In Figure 4.5(b), corresponding to Semi-Town regions, platelet count, white blood cell count, blood pressure, and heart rate remain the primary contributors, with body temperature, eye pain, headache, and vomiting showing moderate sensitivity. Muscle pain and pre-existing conditions have the least influence. In Figure 4.5(c), illustrating Rural areas, heart rate is the top contributing factor, followed by blood pressure, platelet count, and white blood cell count. Eye

pain, body temperature, vomiting, and headache demonstrate moderate relevance, whereas fever duration and blood group contribute the least to dengue severity in rural settings.



(a) Sensitivity analysis for Town (b) Sensitivity analysis for Semi-Town



(c) Sensitivity analysis for Rural Area

Figure 4.5: Sensitivity Analysis of Residence Types Influencing Dengue Severity

4.2.5 Sensitivity Analysis for Seasons

The sensitivity analysis across different seasons highlights distinct variations in how clinical and physiological features influence dengue severity. In Figure 4.6(a), representing the Pre-Monsoon season, platelet count emerges as the most sensitive indicator, followed by heart rate, blood group, and WBC count, while environmental temperature and vomiting exhibit the least influence. In Figure 4.6(b), corresponding to the Monsoon season, platelet count, blood pressure, blood pressure, WBC count, and heart rate show the highest sensitivity, with blood group and fever duration

contributing minimally. Figure 4.6(c), illustrating the Post-Monsoon season, maintains platelet count as the top contributing factor, along with blood pressure, heart rate, and WBC count, whereas blood group and fever duration again show low sensitivity. In Figure 4.6(d), representing the Winter season, body temperature, heart rate, platelet count, and WBC count all demonstrate maximum sensitivity, emphasizing their strong diagnostic significance, while vomiting and pre-existing conditions have reduced influence. These findings underscore the importance of adapting clinical focus based on seasonal dynamics to improve the management of dengue severity.

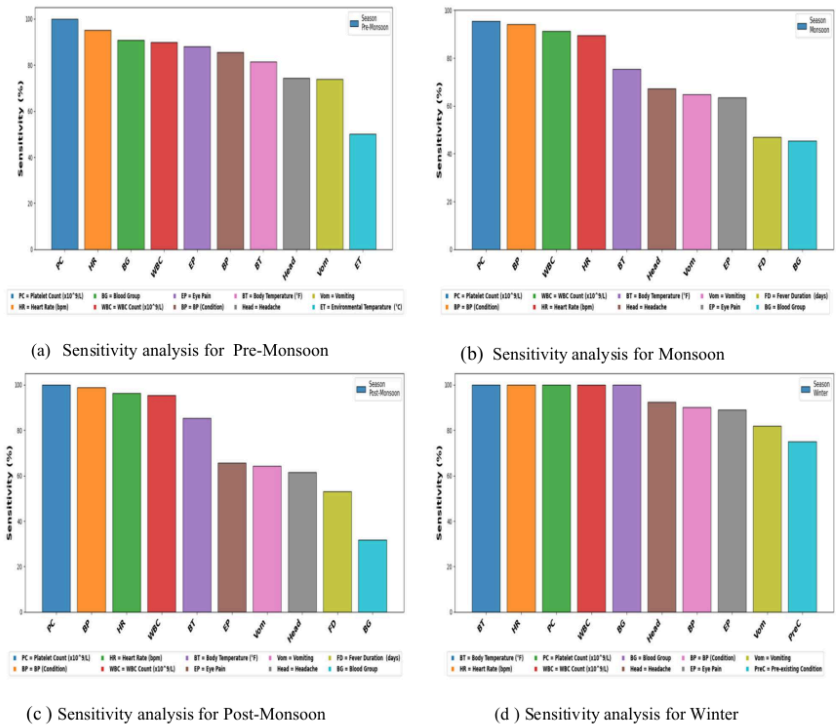


Figure 4.6: Sensitivity Analysis of Seasons Influencing Dengue Severity

4.3 Impact of Individual Parameters on Dengue Severity

²⁷ In this study, we conducted a comprehensive feature-based analysis to examine how different patient and environmental characteristics vary across the four severity levels of dengue: No Dengue, Mild Dengue Fever, Dengue Hemorrhagic Fever, and Dengue Shock Syndrome. This analysis is important because it not only identifies features that show consistent changes as disease severity increases but also highlights those that overlap between severity levels, which can indicate areas of diagnostic ambiguity. By examining these patterns, we can better understand the influence of each feature on dengue severity classification and improve the accuracy and reliability of predictive models. Figures 4.7 (a–l) illustrate these variations across all features.

(i) Platelet Count: In Figure 4.7(a), healthy individuals show stable platelet levels, providing a healthy range for comparison. Platelet count declines progressively with increasing dengue severity. Patients with Mild Dengue Fever generally maintain platelet levels within a relatively normal range, while those with Dengue Hemorrhagic Fever experience a marked reduction, reflecting the disease's impact on coagulation. Patients with Dengue Shock Syndrome present critically low platelet counts, indicative of severe disease progression. This downward trend highlights platelet count as an important marker for differentiating severity levels.

(ii) Body Temperature: In Figure 4.7(b), non-dengue individuals exhibit normal body temperature ranges, serving as a baseline. Body temperature rises steadily with increasing severity. Mild Dengue Fever patients show mild to moderate fever, Dengue Hemorrhagic Fever patients reach higher temperatures, and those with Dengue Shock Syndrome experience the highest temperatures. This trend emphasizes fever as a progressive indicator of disease severity, although some partial overlap exists between Dengue Hemorrhagic Fever and Dengue Shock Syndrome.

(iii) Fever Duration: In Figure 4.7(c), individuals without dengue generally do not experience fever, providing a healthy reference. Fever duration increases as severity rises. Dengue Shock Syndrome patients have the longest febrile periods, followed by Dengue Hemorrhagic Fever, while Mild Dengue Fever patients experience shorter fever durations. This trend demonstrates that prolonged illness is associated with higher severity and provides insight into disease progression over time.

(iv) White Blood Cell Count: In Figure 4.7(d), healthy individuals show WBC counts within the normal range. White blood cell count gradually decreases with increasing dengue severity. Mild Dengue Fever patients typically remain within normal WBC ranges, Dengue Hemorrhagic Fever patients exhibit moderate leukopenia, and Dengue Shock Syndrome patients present the lowest WBC counts, reflecting immune suppression in severe stages. This pattern highlights WBC count as a diagnostic feature for differentiating severe cases.

(v) Blood Pressure: In Figure 4.7(e), individuals without dengue have normal blood pressure ranges, establishing a healthy range. Blood pressure declines progressively with increasing severity. Mild Dengue Fever patients generally exhibit normal blood pressure, Dengue Hemorrhagic Fever patients show moderate hypotension, and Dengue Shock Syndrome patients

experience critically ¹⁸⁴ blood pressure consistent with shock. The continuous decrease underscores its utility as a marker of disease severity.

(vi) Heart Rate: In Figure 4.7(f), healthy individuals display normal heart rate levels, providing a baseline. Heart rate generally rises from Mild Dengue Fever ¹⁰⁶ to Dengue Hemorrhagic Fever and Dengue Shock Syndrome. A slight overlap occurs near Dengue Hemorrhagic Fever ¹⁵⁷ because some Dengue Shock Syndrome patients have heart rate values similar to those in Dengue Hemorrhagic Fever. Once patients with higher heart rates in Dengue Shock Syndrome are included, the trend continues upward. This variability indicates that heart rate alone may not fully distinguish the most severe categories.

(vii) Age: In Figure 4.7(g), age distribution in healthy individuals reflects the general population and serves as a comparison. Younger patients are ¹² more frequently diagnosed with Mild Dengue Fever, while older patients appear more often in Dengue Hemorrhagic Fever and Dengue Shock Syndrome. The overlapping age ranges between Dengue Hemorrhagic Fever and Dengue Shock Syndrome explain the temporary reconnection of the curve before rising again for the oldest groups.

(viii) Environmental Temperature: In Figure 4.7(h), non-dengue areas reflect normal environmental temperature exposure. Environmental temperature increases with severity. Mild Dengue Fever occurs in moderate temperature ranges, Dengue Hemorrhagic Fever is concentrated in higher temperatures, and Dengue Shock Syndrome cases are observed across the upper range, producing temporary overlaps. This suggests temperature influences risk but does not fully separate severe cases.

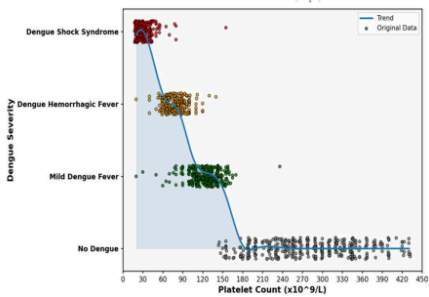
(ix) Vomiting: In Figure 4.7(i), healthy individuals rarely experience vomiting, providing a baseline for comparison. The occurrence of vomiting rises with severity. Mild Dengue Fever cases show occasional vomiting, Dengue Hemorrhagic Fever patients report it more consistently, and Dengue Shock Syndrome patients show slightly higher prevalence. The curve temporarily aligns with Dengue Hemorrhagic Fever due to overlapping values.

(x) Humidity: In Figure 4.7(j), non-dengue areas provide a baseline for environmental humidity. Humidity levels are moderate in Mild Dengue Fever and elevated in more severe cases. Dengue Hemorrhagic Fever shows a higher concentration of cases at elevated humidity, causing partial overlap with Dengue Shock Syndrome. This highlights humidity as a contributing factor to severity.

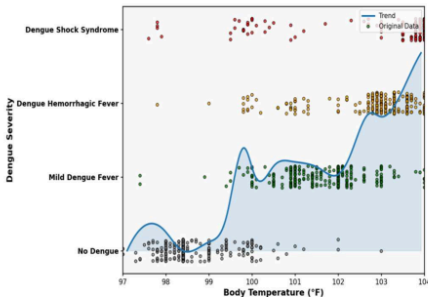
(xi) Skin Rash: In Figure 4.7(k), healthy individuals do not show skin rash, providing a baseline. Rash prevalence is higher in severe cases, particularly Dengue Hemorrhagic Fever. Dengue Shock Syndrome patients also show rash, but the highest concentration remains among Dengue Hemorrhagic Fever cases, explaining the temporary curve alignment.

(xii) Seasonality: In Figure 4.7(l), non-dengue individuals show a relatively uniform seasonal distribution, serving as a baseline. Severe cases peak during the monsoon and post-monsoon periods. Dengue Hemorrhagic Fever cases are densely concentrated in these seasons, resulting in temporary alignment before Dengue Shock Syndrome patients are included. Mild Dengue Fever

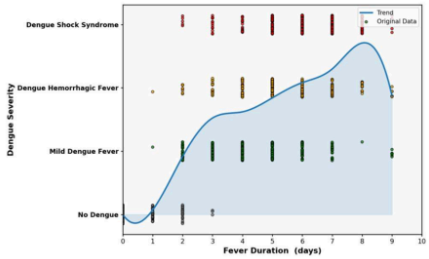
cases are more evenly distributed throughout the year. This seasonal pattern emphasizes the temporal aspect of dengue risk and severity



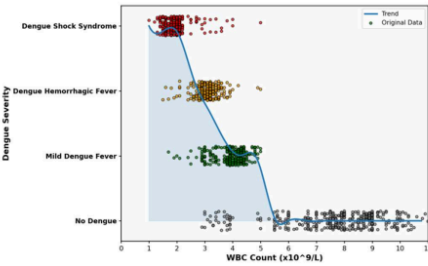
(a) Impact of Platelet Count on Dengue Severity



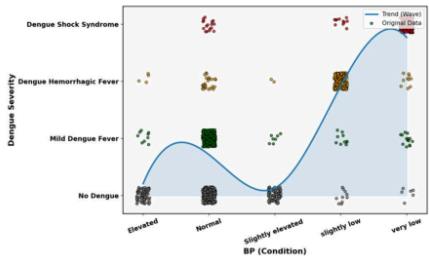
(b) Impact of Body Temperature on Dengue Severity



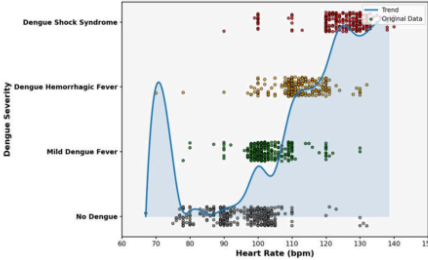
(c) Impact of Fever Duration (days) on Dengue Severity



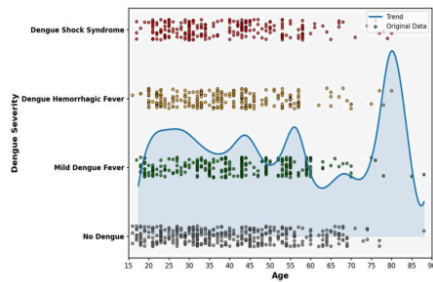
(d) Impact of WBC Count on Dengue Severity



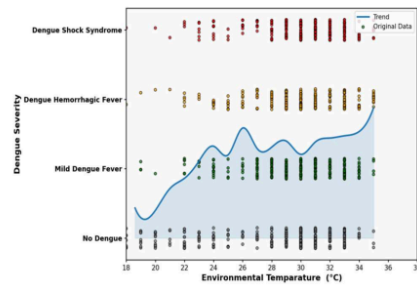
(e) Impact of Blood Pressure on Dengue Severity



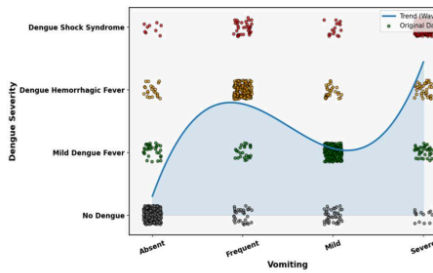
(f) Impact of Heart Rate (bpm) on Dengue Severity



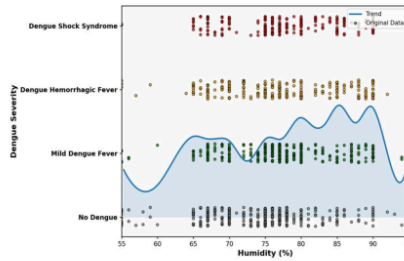
(g) Impact of Age (years) on Dengue Severity



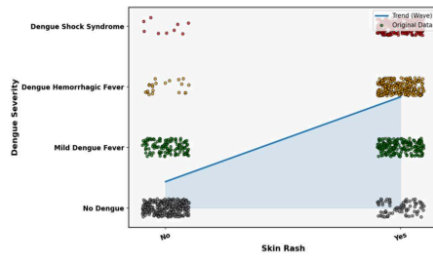
(h) Impact of Environmental Temperature on Dengue Severity



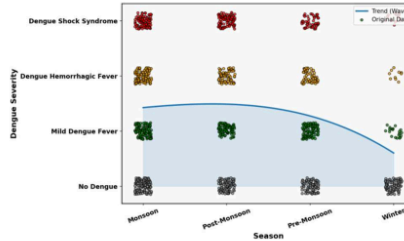
(i) Impact of Vomiting on Dengue Severity



(j) Impact of Humidity (%) on Dengue Severity



(k) Impact of Skin Rash on Dengue Severity



(l) Impact of Season on Dengue Severity

Figure 4.7: Impact of Individual Parameters on Dengue Severity

4.4 Risk Zone Mapping of Dengue Cases

A risk map is a visualization tool that depicts the geographic distribution and intensity of disease occurrence across specific regions. In the context of dengue, risk zone mapping enables the identification of localities with varying case counts and potential transmission risks. Unlike raw statistics or classification outputs, risk maps integrate epidemiological data with geographic information, offering a clear and intuitive representation of how dengue cases are spatially distributed.

Dengue incidence of the three study areas was visualized in two mapping techniques. There are regional maps which show the overall positivity rates and the local maps show the fine variations that exist at the regions. Color-coded scheme (green: 1 to 5 cases, orange: 6 to 10 cases, and red: more than 10 cases) indicates the lower, moderate and higher burden of cases. This practice will aid in identifying the high-risk areas, justification of targeting of the population health interventions, utilization of resources, control of vectors, and public awareness, and the correlation between the disease rates and urbanization, population density, and local environmental conditions.

²⁷ Figure 4.8 shows the positional risk of dengue transmission in the three study locations Dhaka, Comilla and Noakhali as evidenced by the positivity rate. The number of individuals who tested positive out of the number of people who were tested in an area was calculated to get the rate of positivity. The relative risk was categorized by a ranking strategy with highest positivity rate constituting the highest risk in red, the second highest the moderate risk in orange and the lowest the lower risk in green. According to this method, Dhaka is at the highest risk with positivity of 76.77 percent, Comilla at moderate risk with positivity of 74.10 percent and Noakhali at least at risk with positivity of 60.24 percent. This kind of visualization helps to know how strong the transmission is and helps to plan the targeted public health.

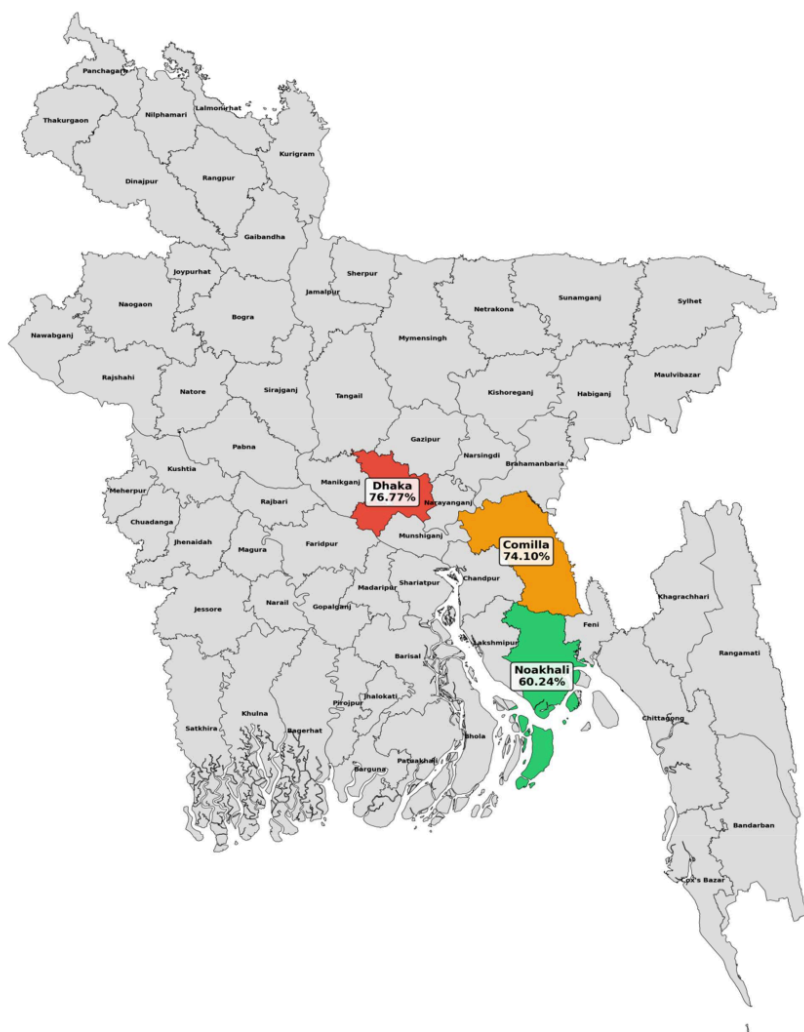


Figure 4.8: Dengue Risk Map of Dhaka, Comilla, and Noakhali

4.4.1 Risk Zone Map of Dengue in Cumilla District

In order to comprehend the localized effects of dengue in Comilla, a more elaborate mapping of the administration districts in the district was conducted, which indicated the distribution of the cases in the various localities. A colour-based system to denote the different case concentration was used whereby red cases denoted areas with over 10 cases, orange denoted 6 to 10 cases, and green denoted 1 to 5 cases.

The map identifies some of the red hotspots, such as Daudkandi Paurashava, Laksam, Kandirpar, Homna Union, Burichang, and Muradnagar, which reported more than 20 cases each. These red zones are as such with high concentration of outbreaks and therefore there is an urgent need to have health intervention intervention. Regions like Shubhapur, Debidwar, Panchthubi, Comilla Dakshin Paurashava, and Gobindapur seem orange which indicates moderate clusters of cases and which should be closely monitored. In the meantime, the relatively lower cases in green-colored regions such as Durgapur and Shibirbazar need to be kept under constant surveillance to avoid further infections.

Altogether, this color-coded map offers a clear and intuitive depiction of the dengue distribution across Comilla, allowing health authorities to recognize the hotspots, distribute resources effectively, and develop specific vector control and community awareness interventions.

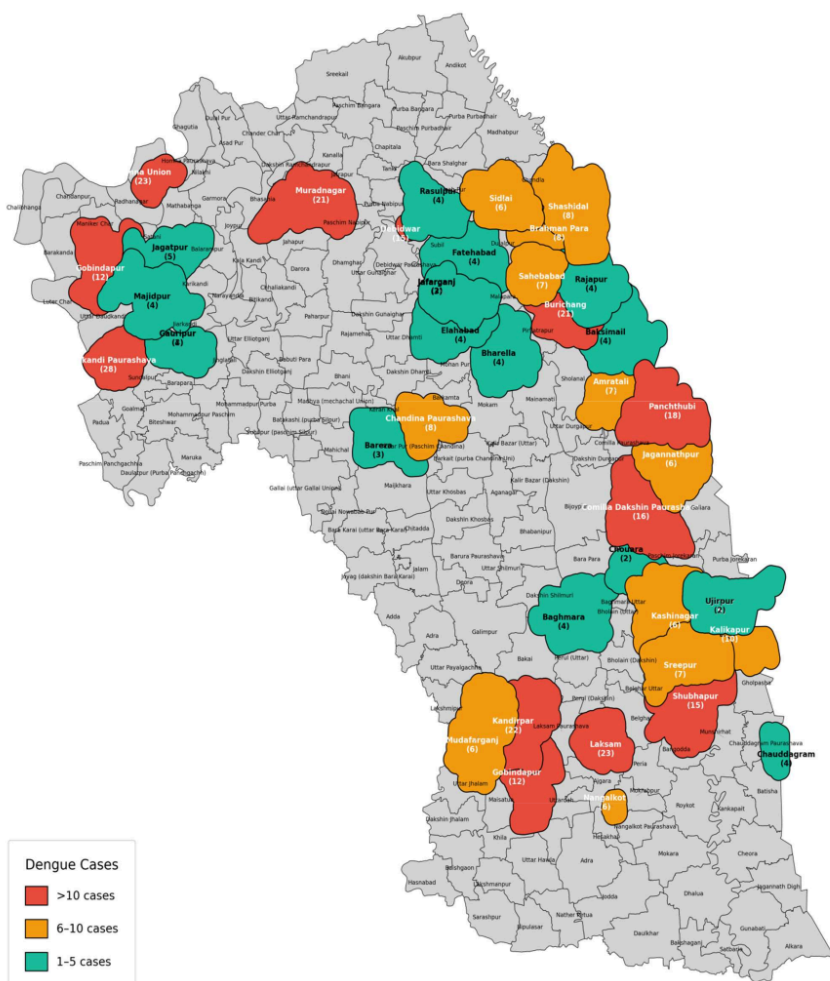


Figure 4.9: Risk Zone Map of Dengue Cases in Comilla

4.4.2 Risk Zone Map of Dengue in Noakhali District

The case mapping of dengue in Noakhali district was done to represent the spatial distribution of infections by the administrative areas of the district. The color-coded system was used to depict different densities of the cases: red represented high density (more than 10 cases), orange represented the middle density (between 6 and 10 cases), and green represented the low density (between 1 and 5 cases).

The map shows multiple hotspots marked in red, such as Alexander, Chaumohani Paurashava, Maijdee, and Noakhali Paurashava, where most cases in the district are reported. These are areas of special interest and need urgent and stringent measures to control the vectors, provide healthcare and community education. Areas that are marked in orange such as Amanullapur, Sonapur, Sonaimuri Paurashiva, Basurhat Paurashiva, and Kabirhat Paurashiva are moderate clusters that should be closely monitored. In the meantime, the green colored regions, including Ramnarayanpur, Companiganj, Dotterhat, and Musapur demonstrate a comparatively less number of cases but still need to be under surveillance to avoid additional infection. This color-coded risk zone map offers an overview of dengue incidence in Noakhali in a clear and intuitive format, allowing health authorities to quickly identify hotspots that are critical, plan interventions effectively, and boost prevention efforts throughout the district.

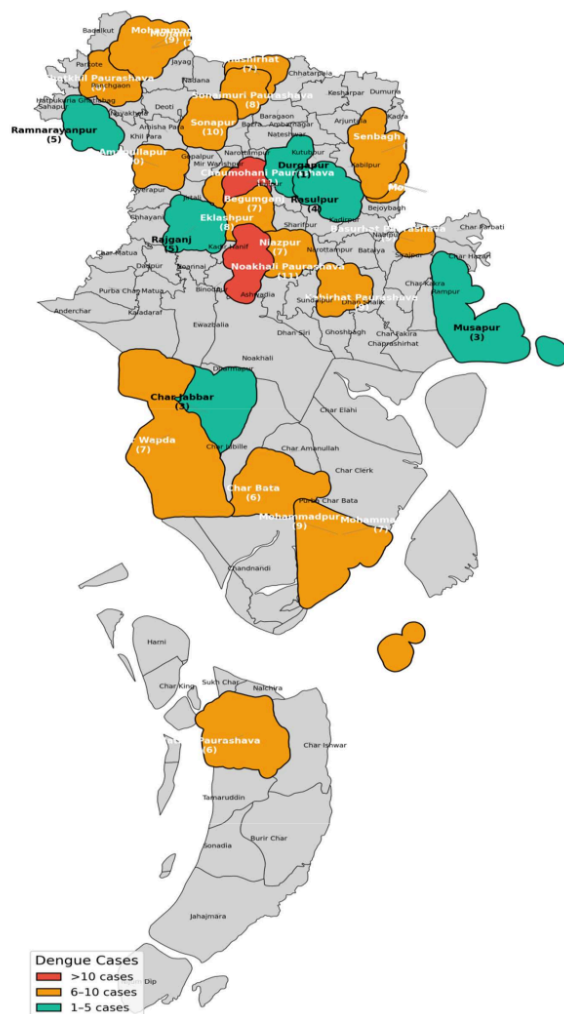


Figure 4.10: Risk Zone Map of Dengue Cases in Noakhali

4.4.3 Risk Zone Map of Dengue in Dhaka District

The geographical distribution of dengue cases in the Dhaka district has been represented in a color-coded map indicating the numbers of confirmed cases. The classification is similar to those of the other districts: red is areas with over 10 cases, orange is 6 to 10 cases, and green sources 1 to 5 cases.

The most noticeable spots on the map are several red-marked urban centers, such as Mirpur, Rampura, Motijheel, Kotwali, Gulshan and Paltan, all of which reported more than 10 cases each, and suggest high levels of concentration of the transmission. These areas are especially alarming because of their urban construction, which can help breed mosquitos and transmit the virus. Regions highlighted in orange, including Kalabagan, Dhanmondi, Mohammadpur, Shyampur, Ashulia, and Badda, have middle-level case density and thus need constant observation and prevention. In the meantime, other areas such as Demra, Keraniganj, Tejgaon, and Uttara are colored green indicating low rates of reported cases though they are vulnerable to an outbreak unless well controlled.

It is a well-constructed color-coded map that shows the spatial distribution of dengue disease in Dhaka in a critical manner that can help the city health authorities to locate the areas with the highest burden, to move resources appropriately and to initiate the specific measures tackling vectors, as well as, raising awareness among the population.

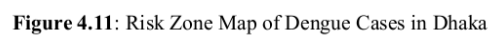


Figure 4.11: Risk Zone Map of Dengue Cases in Dhaka

4.5 Performance Analysis of Dengue Prediction using distinct Learning Algorithms

This paper has used multiple machine learning (ML) and deep learning (DL) algorithms in dengue classification. The ML models used are: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), whereas the DL models are Multi-layer Perceptron (MLP), One-dimensional convolutional neural network (1D-CNN), and the Long short-term memory (LSTM). Besides this, a hybrid Random Forest-LSTM model was built to combine feature-based learning with sequence modeling. All of the models were trained and tested on the processed dataset, obtained in Dhaka, Cumilla, and Noakhali and included patient health records, clinical symptoms, and environmental factors.

The outcomes of these models are presented through a combination of tables and figures, summarizing overall performance metrics, training and validation metrics for deep learning models, and confusion matrices. Table 4.1, Performance Metrics of Several Machine Learning and Deep Learning Models, provides a comparative overview of accuracy, precision, recall, and F1-score for all models. Table 4.2, Training and Validation Metrics of Deep Learning Models, presents the training and validation accuracy and loss of deep learning models. Figure 4.12 visualizes the performance metrics for each dengue classification, showing class-wise accuracy, precision, recall, and F1-score. Figure 4.13 shows the training and validation accuracy curves of deep learning models, while Figure 4.14 illustrates the corresponding loss curves over epochs. To further evaluate model predictions, Table 4.3, Confusion Matrices of All Models, shows the correctly and incorrectly classified instances along with accuracy and error percentages, and Figure 4.15 provides a graphical representation of these confusion matrices for comparison across all dengue classes.

4.5.1 Random Forest

The Random Forest model demonstrated strong predictive performance, achieving an accuracy of 93.98%, along with balanced precision, recall, and F1-score across all classes as shown in Table 4.1. From the confusion matrix in Table 4.3 and Figure 4.15, the model correctly classified 64 out of 74 DHF cases, with 10 misclassified mainly as Mild DF, resulting in 86.49% accuracy for this class. For DSS, 69 out of 72 cases were correctly identified with only minor errors, while Mild DF showed 84 correct predictions out of 92, yielding 91.3% accuracy. All 110 No Dengue cases were predicted correctly with 100% accuracy. The overall error rate of the model was 6.03%, with most errors occurring between DHF and Mild DF due to their overlapping clinical features. Overall, these results highlight the robustness of the Random Forest model in effectively classifying dengue severity levels and distinguishing dengue from non-dengue cases.

4.5.2 Support Vector Machine

⁵² The Support Vector Machine model demonstrated strong predictive performance, achieving an accuracy of 91.1%, along with balanced precision, recall, and F1-score across all classes as shown in Table 4.1. From the confusion matrix in Table 4.3 and Figure 4.15, the model correctly classified 60 out of 74 DHF cases, with 14 misclassified mainly as Mild DF and a few as DSS, resulting in 81.08% accuracy for this class. For DSS, 63 out of 72 cases were correctly identified, with minor misclassifications into DHF and Mild DF, yielding 87.5% class accuracy. Mild Dengue Fever showed 84 correct predictions out of 92, with an accuracy of 91.3%, while all 110 No Dengue cases were predicted correctly, achieving 100% accuracy. The overall error rate of the model was 8.9%, with most misclassifications occurring between DHF and Mild DF due to overlapping clinical features. Overall, ¹⁷ these results highlight the robustness of the SVM model in effectively classifying dengue severity levels and accurately distinguishing dengue-negative cases.

4.5.3 XGBoost

The XGBoost model showed strong predictive ability, achieving an accuracy of 92.82% with balanced precision, recall, and F1-score across all classes, as shown in Table 4.1. From the confusion matrix in Table 4.3 and Figure 4.15, 66 out of 74 Dengue Hemorrhagic Fever cases were correctly classified, with the remaining 8 misclassified as Mild Dengue Fever or Dengue Shock Syndrome, resulting in 89.18% accuracy for this class. Dengue Shock Syndrome was correctly predicted in 63 out of 72 instances, resulting in 87.5% accuracy, while Mild Dengue Fever had 84 correct predictions out of 92, resulting in 91.3% accuracy. As for the No Dengue class, the model showed complete reliability in distinguishing these instances from dengue-positive cases, demonstrating a clear boundary in prediction. The overall error rate was 7.18%, with most misclassifications occurring between Dengue Hemorrhagic Fever and Mild Dengue Fever. These results indicate that XGBoost provides reliable dengue severity classification while maintaining strong discrimination for non-dengue cases.

¹⁵⁵ 4.5.4 Multi-Layer Perceptron

The Multi-Layer Perceptron model demonstrated strong predictive performance, achieving a training accuracy of 94%, validation accuracy of 95%, and testing accuracy of 91.1%, along with balanced precision, recall, and F1-score across all classes as shown in Table 4.1. The training and validation losses decreased steadily, reaching 0.15 and 0.11, respectively, indicating effective learning and convergence. The accuracy curves showed a steady upward trend, while the loss curves declined smoothly, reflecting reduced prediction errors over time. The close alignment between training and validation performance confirms that no overfitting occurred, highlighting the MLP model's effectiveness in learning relevant patterns for accurate dengue

case classification. From the confusion matrix in Table 4.3 and Figure 4.15, 57 out of 74 Dengue Hemorrhagic Fever cases were correctly classified, with 5 misclassified as Dengue Shock Syndrome and 12 as Mild Dengue Fever, resulting in 77% accuracy for this class. Dengue Shock Syndrome was correctly identified in 64 out of 72 instances, while Mild Dengue Fever had 86 correct predictions out of 92, resulting in 88.9% and 93.5% accuracy, respectively. As for the No Dengue class, the model showed complete reliability in distinguishing these instances from dengue-positive cases, demonstrating a clear boundary in prediction. The overall error rate was 8.9%, with most misclassifications occurring between Dengue Hemorrhagic Fever and Mild Dengue Fever. Overall, these results highlight the robustness of the Multi-Layer Perceptron model in accurately classifying dengue severity levels and effectively separating dengue-positive from dengue-negative cases.

4.5.5 Long-Short Term Memory

The Long Short-Term Memory model demonstrated strong predictive performance, achieving a training accuracy of 96%, validation accuracy of 95%, and testing accuracy of 92.24%, along with balanced precision, recall, and F1-score across all classes as shown in Table 4.1. Both training and validation losses steadily decreased over the epochs, reaching 0.12 and 0.13, respectively, indicating effective learning and proper convergence. The accuracy curves showed a consistent upward trend, reflecting progressive improvement in prediction correctness, while the loss curves declined over epochs, showing that prediction errors reduced as training progressed. The close alignment between training and validation performance confirms that no overfitting occurred, highlighting the LSTM model's ability to capture temporal dependencies in patient data and reliably classify dengue cases across different severity levels. From the confusion matrix in Table 4.3 and Figure 4.15, 61 out of 74 Dengue Hemorrhagic Fever cases were correctly classified, with 13 misclassified as Dengue Shock Syndrome or Mild Dengue Fever, resulting in 82.43% accuracy for this class. Dengue Shock Syndrome was correctly predicted in 66 out of 72 instances, while Mild Dengue Fever had 85 correct predictions out of 92, resulting in 91.67% and 92.4% accuracy, respectively. As for the No Dengue class, the model was predicted with complete accuracy, demonstrating the model's strong ability to distinguish between dengue-positive and dengue-negative samples with absolute confidence. The overall error rate was 7.48%, with most misclassifications occurring between Dengue Hemorrhagic Fever and Mild Dengue Fever. Overall, these results highlight the robustness of the LSTM model in accurately classifying dengue severity levels and effectively separating dengue-positive from dengue-negative cases.

4.5.6 ¹⁷ 1D Convolutional Neural Network (1D CNN)

The 1D Convolutional Neural Network model demonstrated effective predictive capability for dengue ⁸³ case classification using a preprocessed dataset from Dhaka, Cumilla, and Noakhali. It achieved a training accuracy of 95%, validation accuracy of 94%, and testing accuracy of 90.8%, indicating strong generalization and reliable performance on unseen data. Both training and validation losses decreased steadily throughout the 20 epochs, reaching 0.16 for each, which demonstrates effective learning and convergence. The accuracy curves showed a consistent upward trend, ⁵⁹ reflecting improved prediction capability, while the loss curves declined over time, indicating a reduction in prediction errors. The close match between training and validation metrics confirms that no overfitting occurred, highlighting the 1D CNN's ⁷⁸ effectiveness in capturing relevant spatial patterns in the input data for accurate dengue classification. From the confusion matrix in Table 4.3 and Figure 4.15, 59 out of 74 Dengue Hemorrhagic Fever cases were correctly classified, with 15 misclassified as Dengue Shock Syndrome or Mild Dengue Fever, resulting in 79.8% accuracy for this class. Dengue Shock Syndrome was correctly predicted in 65 out of 72 instances, while Mild Dengue Fever had 82 correct predictions out of 92, resulting in 90.28% and 89.13% accuracy, respectively. As for the No Dengue class, 108 out of 110 instances were correctly classified, demonstrating reliable distinction between dengue-positive and dengue-negative samples. The overall error rate was 9.78%, with most ¹² misclassifications occurring between Dengue Hemorrhagic Fever and Mild Dengue Fever. Overall, these results highlight the robustness of the 1D CNN model in accurately classifying dengue severity levels and effectively separating dengue-positive from dengue-negative cases.

Table 4.1: Performance Metrics of Several Machine Learning, Deep Learning Models

Model	Class	Precision	Recall	F1 Score
Random Forest	Dengue Hemorrhagic Fever	0.89	0.86	0.88
	Dengue Shock Syndrome	0.97	0.96	0.97
	Mild Dengue Fever	0.91	0.93	0.92
	No Dengue	1.0	1.0	1.0
XGBOOST	Dengue Hemorrhagic Fever	0.80	0.89	0.85
	Dengue Shock Syndrome	0.95	0.88	0.91
	Mild Dengue Fever	0.93	0.91	0.92
	No Dengue	1.00	1.00	1.00
Support Vector Machine	Dengue Hemorrhagic Fever	0.82	0.81	0.82
	Dengue Shock Syndrome	0.93	0.88	0.90
	Mild Dengue Fever	0.88	0.92	0.90
	No Dengue	1.00	1.00	1.00
Multi-Layer Perceptron	Dengue Hemorrhagic Fever	0.79	0.82	0.81
	Dengue Shock Syndrome	0.91	0.88	0.89
	Mild Dengue Fever	0.90	0.90	0.90
	No Dengue	1.00	1.00	1.00
Long Short-Term Memory	Dengue Hemorrhagic Fever	0.85	0.81	0.83
	Dengue Shock Syndrome	0.92	0.93	0.92
	Mild Dengue Fever	0.89	0.91	0.91
	No Dengue	1.00	1.00	1.00
1D CNN	Dengue Hemorrhagic Fever	0.83	0.77	0.80
	Dengue Shock Syndrome	0.87	0.92	0.89
	Mild Dengue Fever	0.89	0.89	0.89
	No Dengue	0.97	0.98	0.98

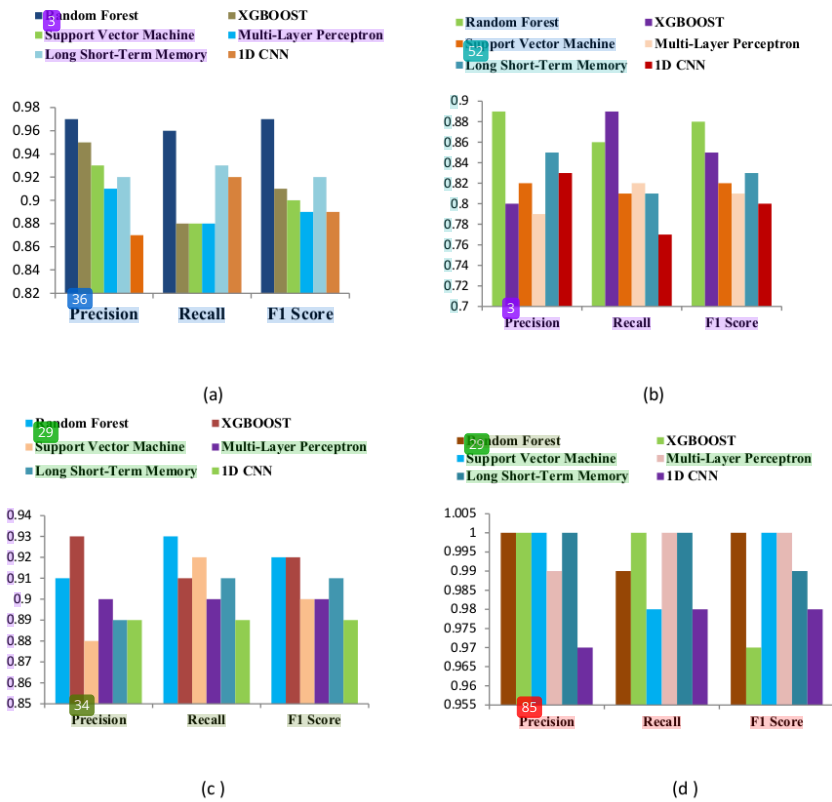
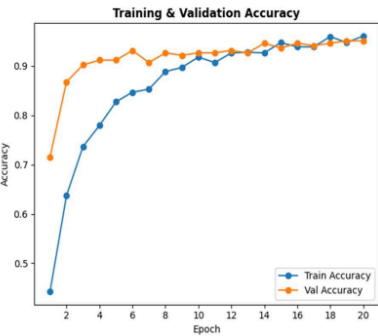


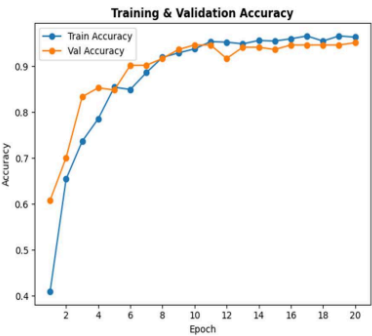
Figure 4.12: Performance metrics of distinct learning algorithms for (a) Dengue Shock Syndrome (DSS), (b) Dengue Hemorrhagic Fever (DHF), (c) Mild Dengue Fever (Mild DF), and (d) No Dengue.

13
Table 4.2: Training and Validation Metrics of Deep Learning Models

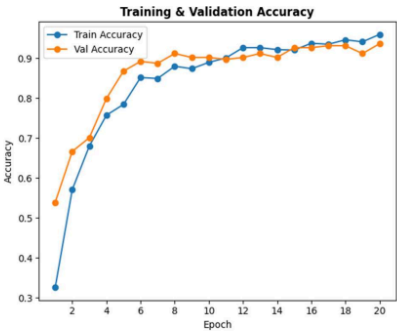
Model		Accuracy	Loss
Multi-Layer Perceptron	Training	94	0.15
	Validation	95	0.11
Long Short-Term Memory	Training	96	0.12
	Validation	95	0.13
1D CNN	Training	95	0.16
	Validation	94	0.16



(a) MLP Training & Validation Accuracy

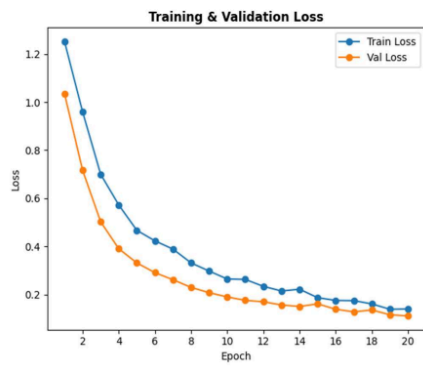


(b) LSTM Training & Validation Accuracy

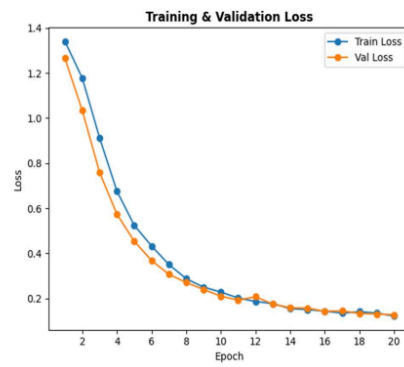


(c) 1D CNN Training & Validation Accuracy

10
Figure 4.13: Training & Validation Accuracy



(i) MLP Training & Validation Loss

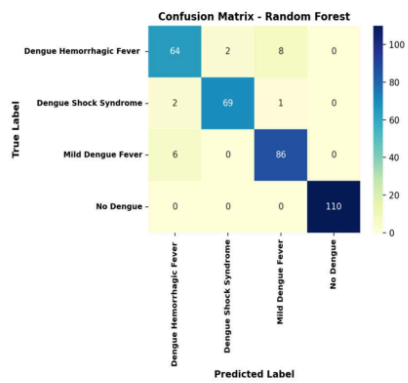


(ii) LSTM Training & Validation Loss

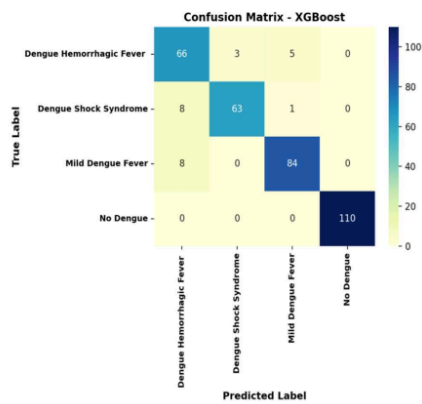


(iii) 1D CNN Training & Validation Loss

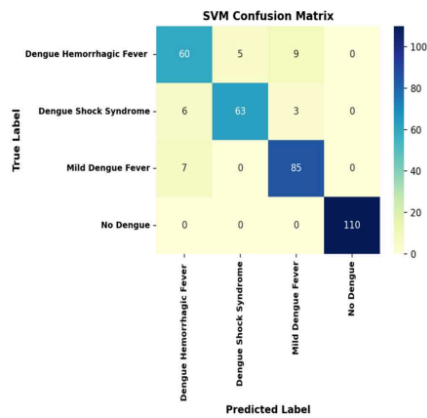
Figure 4.14: Training & Validation Loss



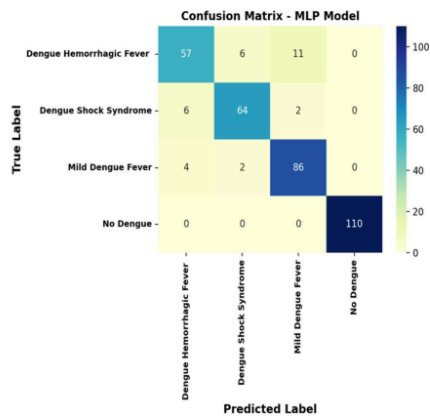
(a) Confusion Matrix of Random Forest



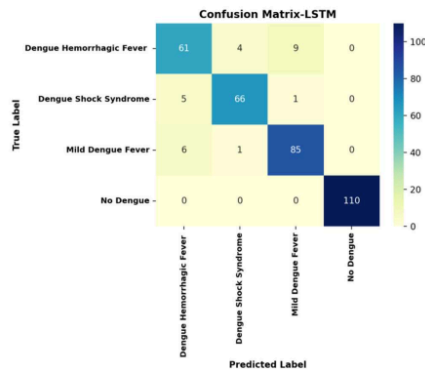
(b) Confusion Matrix of XGBOOST



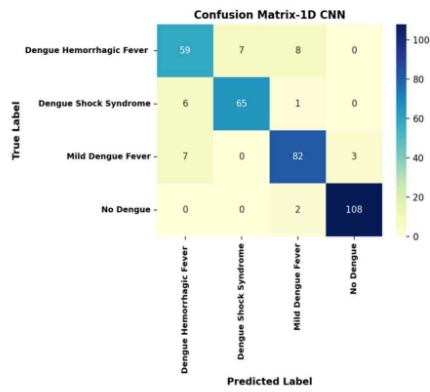
(c) Confusion Matrix of Support Vector Machine



(d) Confusion Matrix of Multi-Layer Perceptron



(e) Confusion Matrix of LSTM



(f) Confusion Matrix of 1D CNN

Figure 4.15: Confusion Metrics

Table 4.3: Accuracy Results of Several Machine Learning, Deep Learning Models

Model	Class	No. Of Instance	Correctly Predicted Instances	Incorrectly Predicted Instances	Error (%)	Accuracy (%)
Random Forest	Dengue Hemorrhagic Fever	74	64	10	13.51	86.49
	Dengue Shock Syndrome	72	69	3	4.17	95.83
	Mild Dengue Fever	92	84	8	8.7	91.3
	No Dengue	110	110	0	0	100
	Total	348	327	21	6.03	93.97
XGBOOST	Dengue Hemorrhagic Fever	74	66	8	10.82	89.18
	Dengue Shock Syndrome	72	63	9	12.5	87.5
	Mild Dengue Fever	92	84	8	8.7	91.3
	No Dengue	110	110	0	0	100
	Total	348	323	25	7.18	92.82
Support Vector Machine	Dengue Hemorrhagic Fever	74	60	14	18.92	81.08
	Dengue Shock Syndrome	72	63	9	12.5	87.5
	Mild Dengue Fever	92	84	8	8.7	91.3
	No Dengue	110	110	0	0	100

	Total	348	317	31	8.9	91.1
Multi-Layer Perceptron	Dengue Hemorrhagic Fever	74	57	17	23	77
	Dengue Shock Syndrome	72	64	8	11.1	88.9
	Mild Dengue Fever	92	86	6	6.5	93.5
	No Dengue	110	110	0	0	100
	Total	348	317	31	8.9	91.1
Long Short-Term Memory	Dengue Hemorrhagic Fever	74	61	13	17.57	82.43
	Dengue Shock Syndrome	72	66	6	8.33	91.67
	Mild Dengue Fever	92	85	7	7.6	92.4
	No Dengue	110	110	0	0	100
	Total	348	322	26	7.48	92.52
1D CNN	Dengue Hemorrhagic Fever	74	59	15	20.2	79.8
	Dengue Shock Syndrome	72	65	7	9.72	90.28
	Mild Dengue Fever	92	82	10	10.17	89.13
	No Dengue	110	108	2	1.8	98.2
	Total	348	314	34	9.78	90.22

4.5.7 Proposed Hybrid RF-LSTM Model

The Hybrid Random Forest-LSTM model demonstrated exceptional predictive performance for dengue classification using preprocessed data from Dhaka, Cumilla, and Noakhali. It achieved a training accuracy of 99.1%, validation accuracy of 99.9%, and testing accuracy of 96.26%, along with balanced precision, recall, and F1-score across all classes as shown in Table 4.4. The training and validation losses dropped rapidly, stabilizing at extremely low values of 0.0055 and 0.0022, respectively as shown in table 4.5 indicating highly effective learning and convergence. The accuracy curves quickly reached a plateau, while the loss curves approached near zero, reflecting precise learning of underlying patterns shown in Figure 4.17. The close alignment between training and validation performance confirms that no overfitting occurred, emphasizing the hybrid model's capability to capture both temporal and spatial features for accurate dengue case classification.

Table 4.4: Performance of Metrics Hybrid RF-LSTM Model

Model	Class	Precision	Recall	F1 Score
Hybrid RF-LSTM Model	Dengue Hemorrhagic Fever	0.91	0.93	0.92
	Dengue Shock Syndrome	0.97	0.96	0.97
	Mild Dengue Fever	0.96	0.95	0.95
	No Dengue	1.00	1.00	1.00

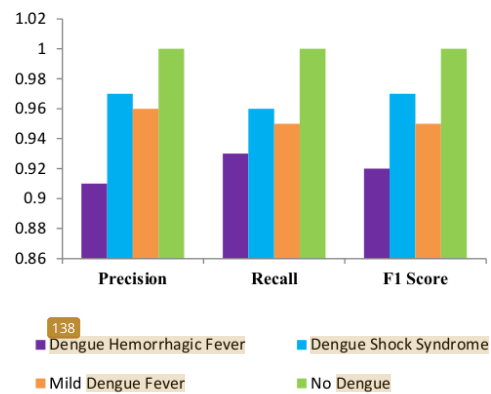


Figure 4.16: Performance Metrics of Hybrid RF-LSTM Model

Table 4.5: Training and Validation Metrics of Hybrid RF-LSTM Model

Model		Accuracy	Loss
Hybrid RF-LSTM Model	Training	99.1	0.0055
	Validation	99.9	0.0022

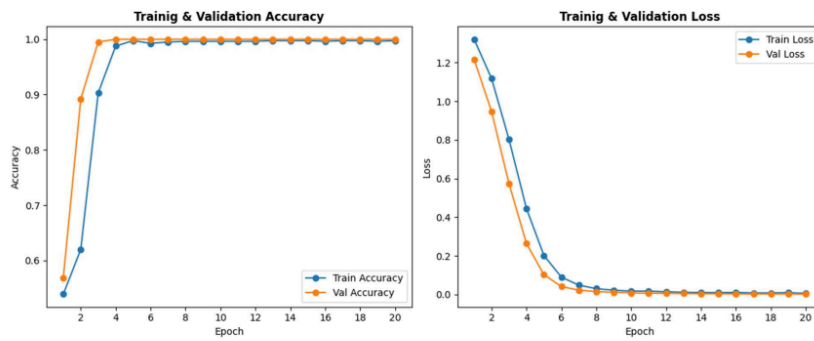


Figure 4.17: Training & Validation Accuracy and Loss of Hybrid RF-LSTM Model

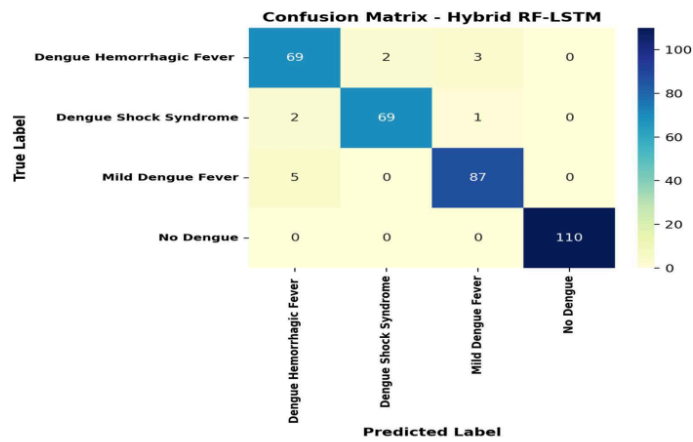


Figure 4.18: Confusion Matrix of Hybrid RF-LSTM Model

Table 4.6: Accuracy Results of Hybrid RF-LSTM Model

Model	Class	No. Of Instance	Correctly Predicted Instances	Incorrectly Predicted Instances	Error (%)	Accuracy (%)
Hybrid RF-LSTM Model	Dengue Hemorrhagic Fever	74	69	5	6.75	93.25
	Dengue Shock Syndrome	72	69	3	4.17	95.83
	Mild Dengue Fever	92	87	5	5.43	94.57
	No Dengue	110	110	0	0	100
	Total	348	335	13	3.74	96.26

From the confusion matrix in Table 4.3 and Figure 4.18, 69 out of 74 Dengue Hemorrhagic Fever cases were correctly classified, with 5 misclassified as Mild Dengue Fever, resulting in 93.25% accuracy for this class. Dengue Shock Syndrome was correctly identified in 69 out of 72 instances, while Mild Dengue Fever had 87 correct predictions out of 92, resulting in 95.83% and 94.57% accuracy, respectively. As for the No Dengue class, all 110 instances were correctly classified, demonstrating the model's reliability in distinguishing dengue-negative samples. The overall error rate was 3.74%, with most misclassifications occurring between Dengue Hemorrhagic Fever and Mild Dengue Fever. Overall, these results demonstrate that the proposed Hybrid Random Forest–LSTM model not only effectively classifies dengue severity levels and distinguishes dengue-negative cases but also outperforms the individual base models, highlighting its superior predictive capability.

4.6 Testing On Unseen Dataset

In order to test the generalization of the models, an unseen dengue dataset was tested. In this dataset there were 54 records, and none of them had been previously used. The models used were: Random Forest (RF), XGBoost, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), Multi-layer Perceptron (MLP), 1D Convolutional Neural Network (1D CNN) and Hybrid RF-LSTM. Their performance on this dataset is summarised in the following tables.

Table 4.7: Prediction performance of different models on the unseen dengue dataset

S/No	Method Used	Actual Class	Predicted Class	Error (%)	Accuracy (%)	Average Error (%)	Average Accuracy (%)
1	Random Forest	Dengue Shock Syndrome	Dengue Shock Syndrome	0	100	12.96	87.04
2		Mild Dengue Fever	Mild Dengue Fever	0	100		
3		No Dengue	No Dengue	0	100		
4		No Dengue	No Dengue	0	100		
5		Dengue	No Dengue	0	100		
6		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
8		Mild Dengue Fever	No Dengue	100	0		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
10		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
11		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
12		Dengue Fever	Mild Dengue Fever	0	100		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		No Dengue	No Dengue	0	100		
16		Mild Dengue Fever	No Dengue	100	0		
17		Mild Dengue Fever	Mild Dengue Fever	0	100		
18		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
19		No Dengue	No Dengue	0	100		
20		No Dengue	No Dengue	0	100		
21		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
23		Mild Dengue Fever	Mild Dengue Fever	0	100		
24		Mild Dengue Fever	Mild Dengue Fever	0	100		
25		Dengue Shock Syndrome	No Dengue	100	0		
26		Mild Dengue Fever	Mild Dengue Fever	0	100		
27		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	Mild Dengue Fever	0	100		
29		No Dengue	No Dengue	0	100		
30		Dengue	No Dengue	0	100		
31		Mild Dengue Fever	Mild Dengue Fever	0	100		
32		Mild Dengue Fever	Mild Dengue Fever	0	100		
33		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
34		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
35		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
37		Mild Dengue Fever	Mild Dengue Fever	0	100		
38		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	No Dengue	0	100		
41		No Dengue	No Dengue	0	100		
42		Mild Dengue Fever	Mild Dengue Fever	0	100		
43		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
44		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		Mild Dengue Fever	Mild Dengue Fever	0	100		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		Mild Dengue Fever	Mild Dengue Fever	0	100		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		

54		Dengue Hemorrhagic Fever ⁷	Dengue Hemorrhagic Fever	0	100		
1		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
2		Mild Dengue Fever	Mild Dengue Fever	0	100		
3		No Dengue	Dengue Shock Syndrome	100	0		
4		No Dengue	Dengue Hemorrhagic Fever	100	0		
5		No Dengue	Mild Dengue Fever	100	0		
6		Dengue Hemorrhagic Fever ⁷	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
8		Mild Dengue Fever ⁵	Mild Dengue Fever	0	100		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
10		Dengue Hemorrhagic Fever ⁶	Mild Dengue Fever	100	0		
11		Dengue Hemorrhagic Fever ⁴	Dengue Hemorrhagic Fever	0	100		
12		Mild Dengue Fever	Mild Dengue Fever	0	100		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		No Dengue ²	No Dengue	0	100		
16		Mild Dengue Fever ³	Mild Dengue Fever	0	100		
17		Mild Dengue Fever	Mild Dengue Fever	0	100		
18		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
19		No Dengue	No Dengue	0	100		
20		No Dengue	Mild Dengue Fever	100	0		
21		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		Mild Dengue Fever ⁴	Mild Dengue Fever	0	100		
23		Mild Dengue Fever ⁷	Mild Dengue Fever	0	100		
24		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
25		Dengue Shock Syndrome ²	Dengue Shock Syndrome	0	100		
26		Mild Dengue Fever ¹²	Mild Dengue Fever	0	100		
27		Dengue Shock Syndrome ⁴	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	Mild Dengue Fever	0	100		
29	XGBOOS T	No Dengue	No Dengue	0	100	22.22	77.78
30		No Dengue ²	No Dengue	0	100		
31		Mild Dengue Fever ³	Mild Dengue Fever	0	100		
32		Mild Dengue Fever ⁷	Mild Dengue Fever	0	100		
33		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
34		Dengue Shock Syndrome ⁷	Dengue Shock Syndrome	0	100		
35		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome ²⁰	Dengue Hemorrhagic Fever	100	0		
37		Mild Dengue Fever	Mild Dengue Fever	0	100		
38		Dengue Hemorrhagic Fever ¹²	Dengue Hemorrhagic Fever	0	100		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	Mild Dengue Fever	100	0		
41		No Dengue	Mild Dengue Fever	100	0		
42		Mild Dengue Fever ⁷	Dengue Hemorrhagic Fever	100	0		
43		Dengue Shock Syndrome ¹²	Dengue Shock Syndrome	0	100		
44		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		Mild Dengue Fever ¹²	Mild Dengue Fever	0	100		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		Mild Dengue Fever ⁵	Mild Dengue Fever	0	100		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
54		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
1		Dengue Shock Syndrome	No Dengue	100	0		
2		Mild Dengue Fever	No Dengue	100	0		
3		No Dengue	No Dengue	0	100		
4		No Dengue	No Dengue	0	100		

5	SVM	No Dengue	No Dengue	0	100	68.52	31.48
6		Dengue Hemorrhagic Fever	No Dengue	100	0		
7		Dengue Shock Syndrome	No Dengue	100	0		
8		Mild Dengue Fever	No Dengue	100	0		
9		Dengue Shock Syndrome	No Dengue	100	0		
10		Dengue Hemorrhagic Fever	No Dengue	100	0		
11		Dengue Hemorrhagic Fever	No Dengue	100	0		
12		Mild Dengue Fever	No Dengue	100	0		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		No Dengue	No Dengue	0	100		
16		Mild Dengue Fever	No Dengue	100	0		
17		Mild Dengue Fever	No Dengue	100	0		
18		Dengue Shock Syndrome	No Dengue	100	0		
19		No Dengue	No Dengue	0	100		
20		No Dengue	No Dengue	0	100		
21		Dengue Shock Syndrome	Dengue Shock Syndrome	100	0		
22		Mild Dengue Fever	No Dengue	100	0		
23		Mild Dengue Fever	No Dengue	100	0		
24		Mild Dengue Fever	No Dengue	100	0		
25		Dengue Shock Syndrome	No Dengue	100	0		
26		Mild Dengue Fever	No Dengue	100	0		
27		Dengue Shock Syndrome	No Dengue	100	0		
28		Mild Dengue Fever	No Dengue	100	0		
29		No Dengue	No Dengue	0	100		
30		No Dengue	No Dengue	0	100		
31		Mild Dengue Fever	No Dengue	100	0		
32		Mild Dengue Fever	No Dengue	100	0		
33		Mild Dengue Fever	No Dengue	100	0		
34		Dengue Shock Syndrome	No Dengue	100	0		
35		Dengue Hemorrhagic Fever	No Dengue	100	0		
36		Dengue Shock Syndrome	No Dengue	100	0		
37		Mild Dengue Fever	No Dengue	100	0		
38		Dengue Hemorrhagic Fever	No Dengue	100	0		
39		Dengue Shock Syndrome	No Dengue	100	0		
40		No Dengue	No Dengue	0	100		
41		No Dengue	No Dengue	0	100		
42		Mild Dengue Fever	No Dengue	100	0		
43		Dengue Shock Syndrome	No Dengue	100	0		
44		Dengue Shock Syndrome	No Dengue	100	0		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		Mild Dengue Fever	No Dengue	100	0		
51		Dengue Shock Syndrome	No Dengue	100	0		
52		46 d Dengue Fever	No Dengue	100	0		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	100	0		
54		Dengue Hemorrhagic Fever	No Dengue	100	0		
1		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
2		Mild Dengue Fever	Mild Dengue Fever	0	100		
3		No Dengue	Dengue Hemorrhagic Fever	100	0		
4		No Dengue	Dengue Hemorrhagic Fever	100	0		
5		No Dengue	Mild Dengue Fever	100	0		
6		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
8		5 Mild Dengue Fever	Mild Dengue Fever	0	100		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		

10	LSTM	72 Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100	20.37	79.63
11		4 Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
12		4 Mild Dengue Fever	Mild Dengue Fever	0	100		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		4 No Dengue	No Dengue	0	100		
16		4 Mild Dengue Fever	Mild Dengue Fever	0	100		
17		4 Mild Dengue Fever	Mild Dengue Fever	0	100		
18		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
19		No Dengue	No Dengue	0	100		
20		No Dengue	Mild Dengue Fever	100	0		
21		4 Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		4 Mild Dengue Fever	Mild Dengue Fever	0	100		
23		23 Mild Dengue Fever	Mild Dengue Fever	0	100		
24		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
25		4 Dengue Shock Syndrome	Mild Dengue Fever	100	0		
26		Mild Dengue Fever	Mild Dengue Fever	0	100		
27		4 Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	Mild Dengue Fever	0	100		
29		No Dengue	No Dengue	0	100		
30		4 No Dengue	Mild Dengue Fever	100	0		
31		4 Mild Dengue Fever	Mild Dengue Fever	0	100		
32		Mild Dengue Fever	Mild Dengue Fever	0	100		
33		5 Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
34		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
35		12 Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
37		Mild Dengue Fever	Mild Dengue Fever	0	100		
38		12 Dengue Hemorrhagic Fever	Dengue Shock Syndrome	100	0		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	Mild Dengue Fever	100	0		
41		No Dengue	Mild Dengue Fever	100	0		
42		5 Mild Dengue Fever	Mild Dengue Fever	0	100		
43		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
44		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		12 Mild Dengue Fever	Mild Dengue Fever	0	100		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		5 Mild Dengue Fever	Mild Dengue Fever	0	100		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
54		Dengue Hemorrhagic Fever	41 Dengue Hemorrhagic Fever	0	100		
1		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
2		Mild Dengue Fever	No Dengue	100	0		
3		No Dengue	No Dengue	0	100		
4		No Dengue	No Dengue	0	100		
5		No Dengue	No Dengue	0	100		
6		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
8		46 Mild Dengue Fever	No Dengue	100	0		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
10		Dengue Hemorrhagic Fever	No Dengue	100	0		
11		Dengue Hemorrhagic Fever	No Dengue	100	0		
12		Mild Dengue Fever	No Dengue	100	0		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		

15	MLP	No Dengue	No Dengue	0	100	44.44	55.56
16		Mild Dengue Fever	No Dengue	100	0		
17		Mild Dengue Fever	No Dengue	100	0		
18		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
19		No Dengue	No Dengue	0	100		
20		No Dengue	No Dengue	0	100		
21		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		Mild Dengue Fever	No Dengue	100	0		
23		Mild Dengue Fever	No Dengue	100	0		
24		Mild Dengue Fever	No Dengue	100	0		
25		Dengue Shock Syndrome	No Dengue	100	0		
26		Mild Dengue Fever	No Dengue	100	0		
27		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	No Dengue	100	0		
29		No Dengue	No Dengue	0	100		
30		No Dengue	No Dengue	0	100		
31		Mild Dengue Fever	No Dengue	100	0		
32		Mild Dengue Fever	No Dengue	100	0		
33		46 d Dengue Fever	No Dengue	100	0		
34		Dengue Shock Syndrome 7 e	Dengue Shock Syndrome	0	100		
35		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
37		Mild Dengue Fever 7	No Dengue	100	0		
38		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	No Dengue	0	100		
41		No Dengue	No Dengue	0	100		
42		5 Mild Dengue Fever	No Dengue	100	0		
43		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
44		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		12 d Dengue Fever	No Dengue	100	0		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		5 Mild Dengue Fever	No Dengue	100	0		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
54		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
1		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
2		Mild Dengue Fever	Mild Dengue Fever	0	100		
3		No Dengue	Dengue Hemorrhagic Fever	100	0		
4		No Dengue	Dengue Hemorrhagic Fever	100	0		
5		No Dengue	Mild Dengue Fever	100	0		
6		Dengue Hemorrhagic Fever 7	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
8		5 Mild Dengue Fever	Mild Dengue Fever	0	100		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
10		72 Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
11		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
12		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		20 Dengue	No Dengue	0	100		
16		3 Mild Dengue Fever	Mild Dengue Fever	0	100		
17		Mild Dengue Fever	Mild Dengue Fever	0	100		
18		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
19		No Dengue	No Dengue	0	100		

20	1D CNN	No Dengue	Mild Dengue Fever	100	0	29.63	70.37
21		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		Mild Dengue Fever	Mild Dengue Fever	0	100		
23		Mild Dengue Fever	Mild Dengue Fever	0	100		
24		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
25		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
26		Mild Dengue Fever	Mild Dengue Fever	0	100		
27		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
29		No Dengue	No Dengue	0	100		
30		No Dengue	No Dengue	0	100		
31		Mild Dengue Fever	No Dengue	100	0		
32		Mild Dengue Fever	Mild Dengue Fever	0	100		
33		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
34		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
35		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
37		Mild Dengue Fever	Mild Dengue Fever	0	100		
38		Dengue Hemorrhagic Fever	Dengue Shock Syndrome	100	0		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	Mild Dengue Fever	100	0		
41		No Dengue	Mild Dengue Fever	100	0		
42		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
43		Dengue Shock Syndrome	Dengue Hemorrhagic Fever	100	0		
44		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
54		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
1		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
2		Mild Dengue Fever	Mild Dengue Fever	0	100		
3		No Dengue	No Dengue	0	100		
4		No Dengue	No Dengue	0	100		
5		No Dengue	No Dengue	0	100		
6		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
7		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
8		Mild Dengue Fever	No Dengue	100	0		
9		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
10		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
11		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
12		Mild Dengue Fever	Mild Dengue Fever	0	100		
13		No Dengue	No Dengue	0	100		
14		No Dengue	No Dengue	0	100		
15		No Dengue	No Dengue	0	100		
16		Mild Dengue Fever	No Dengue	100	0		
17		Mild Dengue Fever	Mild Dengue Fever	0	100		
18		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
19		No Dengue	No Dengue	0	100		
20		No Dengue	No Dengue	0	100		
21		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
22		Mild Dengue Fever	Mild Dengue Fever	0	100		
23		Mild Dengue Fever	Mild Dengue Fever	0	100		
24		Mild Dengue Fever	Mild Dengue Fever	0	100		

25	Hybrid RF-LSTM	Dengue Shock Syndrome	No Dengue	100	0	9.26	90.74
26		Mild Dengue Fever	Mild Dengue Fever	0	100		
27		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
28		Mild Dengue Fever	Mild Dengue Fever	0	100		
29		No Dengue	No Dengue	0	100		
30		No Dengue	No Dengue	0	100		
31		Mild Dengue Fever	Mild Dengue Fever	0	100		
32		Mild Dengue Fever	Mild Dengue Fever	0	100		
33		Mild Dengue Fever	Dengue Hemorrhagic Fever	100	0		
34		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
35		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		
36		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
37		Mild Dengue Fever	Mild Dengue Fever	0	100		
38		Dengue Hemorrhagic Fever	Dengue Shock Syndrome	100	0		
39		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
40		No Dengue	No Dengue	0	100		
41		No Dengue	No Dengue	0	100		
42		Mild Dengue Fever	Mild Dengue Fever	0	100		
43		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
44		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
45		No Dengue	No Dengue	0	100		
46		No Dengue	No Dengue	0	100		
47		No Dengue	No Dengue	0	100		
48		No Dengue	No Dengue	0	100		
49		No Dengue	No Dengue	0	100		
50		Mild Dengue Fever	Mild Dengue Fever	0	100		
51		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
52		Mild Dengue Fever	Mild Dengue Fever	0	100		
53		Dengue Shock Syndrome	Dengue Shock Syndrome	0	100		
54		Dengue Hemorrhagic Fever	Dengue Hemorrhagic Fever	0	100		

4.7 Performance Comparison of Distinct Learning Approaches for Dengue Classification

The Hybrid RF-LSTM⁸ model achieved the highest accuracy of 96.26%, followed by Random Forest at 93.97%, outperforming the other models. XGBoost, Support Vector Machine, Multi-Layer Perceptron, LSTM, and 1D CNN recorded accuracies between 90.22% and 92.82%, showing the differences in performance among the approaches.

Table 4.8: Performance Metrics Among Different Dengue Classification Approaches

Model	⁹² Accuracy (%)	Precision	Recall	F1 Score
Random Forest	93.97	0.94	0.94	0.94
XGBoost	92.82	0.92	0.92	0.92
⁸ Support Vector Machine	91.10	0.91	0.91	0.91
Multi-Layer Perceptron	91.10	0.90	0.90	0.90
¹⁵² Long Short-Term Memory (LSTM)	92.52	0.91	0.91	0.91
1D CNN	90.22	0.89	0.89	0.89
Hybrid RF-LSTM	96.26	0.96	0.96	0.96

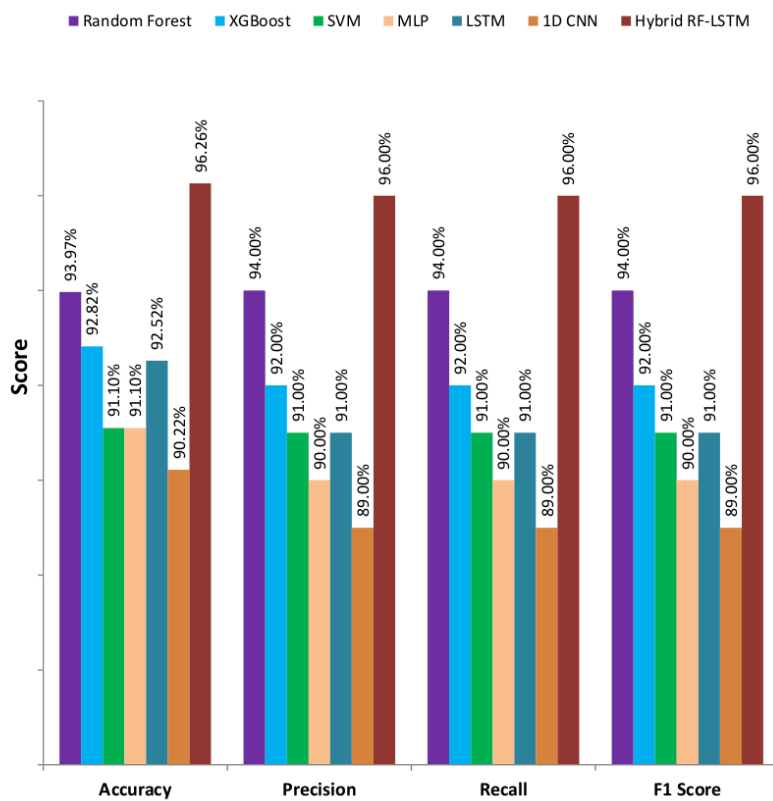


Figure 4.19: Performance Comparison of Distinct Learning Approaches for Dengue Classification

4.8 Discussion

This section evaluates the performance of our proposed model against existing methodologies in the field.

4.8.1Comparative analysis with existing works

The proposed Hybrid RF–LSTM model was evaluated against existing studies on dengue prediction, as summarized in Table X. Prior approaches employing Random Forest, SVM, XGBoost, and LSTM reported accuracies ranging from 75% to 92%, largely based on climatic, hospital, or epidemiological data. While these studies demonstrated the potential of machine learning for outbreak forecasting and early detection, their predictive performance remained limited. In contrast, the Hybrid RF–LSTM model leverages patient-level clinical, physiological, demographic, and contextual features through a combined feature-based and temporal sequence learning framework, achieving an accuracy of 96.26% with balanced precision, recall, and F1-score of 0.96. This underscores its superior capability in reliable and accurate dengue case classification relative to existing methods.

Table 4.9: Comparative analysis with existing works

S/No.	Author(s), Year	Dataset	Methodology	Performance Metrics
1	Yavari Nejad & Varathan (2021)	Historical dengue cases with climatic variables	Random Forest, SVM, Gradient Boosting	Random Forest achieved 92% accuracy
2	Al Mamun et al. (2023)	Bangladesh dengue incidence with climate data	Decision Tree, Random Forest, SVM	Random Forest achieved 87% accuracy
3	Exebio- Chepe et al. (2024)	Peruvian hospital data with clinical and demographic features	Random Forest, SVM, ANN	Artificial Neural Network achieved 86.47% accuracy
4	Dey et al. (2022)	Bangladesh hospital data combined with meteorological and socio- economic factors	Random Forest, XGBoost, SVM	XGBoost achieved 75% accuracy

5	Archana (2023)	Indian hospital clinical and laboratory data	Random Forest, SVM with feature selection	Random Forest achieved 88% accuracy
6	Saleh & Baiwei (2021)	Historical dengue incidence with environmental parameters	LSTM	Achieved 90% accuracy
7	Proposed Hybrid RF–LSTM (2025)	Bangladesh patient dataset from Cumilla, Noakhali, and Dhaka including clinical, physiological, demographic, and contextual attributes	Hybrid Random Forest–LSTM	Achieved 96.26% accuracy, 0.96 precision, 0.96 recall, and 0.96 F1-score

4.9 Dengue Prevention and Awareness

Dengue is a major public health issue that poses significant risks to human health and the socioeconomic well-being of affected communities. Its prevention is therefore essential to reduce disease burden, prevent severe cases, and minimize potential outbreaks. Dengue prevention is a multi-pronged, multi-faceted strategy with mosquito control, program education, and surveillance. The control of the vectors is the most significant defense mechanism and the primary objective in controlling the population of *Aedes aegypti* mosquitoes that are the key vectors of dengue. They involve clearing of stagnant water where the mosquitoes breed in garbage pits, tires, and open tanks, and use of larvicides and insecticides in some of the high-risk areas. Another option for controlling these mosquitoes in a biological way that includes fish species that feed on the larvae can be introduced. Meanwhile, individual measures to prevent mosquito bites, which include mosquito nets, repellent sprays, window screens, and appropriate attire, can play a significant role in mitigating the risks of being bitten by mosquitoes.

The other factor is that it will be necessary to involve the community and sensitize the people to the population because the long-term success of the prevention programs will be due to community involvement. Among the means of acquiring the responsibility towards the surroundings and consulting the medical experts at the opportune time, one may mention training people on how to keep the surroundings clean, and how to practice preventive and early detection of dengue. This kind of awareness is best spread at school, through media campaigns and community health programs. The other useful one is the location of places that may be at high risk according to the environmental factors, weather, and previous case data. These regions may be mapped, and health authorities will be given a priority in resource allocation so that they may plan even more active surveillance and preventive measures in the areas where they are

most needed. This proposed intervention strategy would ensure the optimal utilization of limited healthcare resources to minimize the likelihood of uncontrolled outbreaks.

Integration of strategies, i.e., mosquito control, community education, and risk-based surveillance, can assist in ensuring dengue prevention becomes more systematic, proactive, and sustainable. Not only does this combined structure contribute to a reduction in the occurrence of dengue, but it also creates policy and community resilience, which ultimately safeguards human health and the socioeconomic burden of the disease.

The study has yielded new insights that can improve the targeted dengue prevention efforts. Through a sensitivity analysis, clinical indicators such as platelet count, ⁴⁵white blood cell count, heart rate, and blood pressure were identified, as they show a strong correlation with disease severity. Timely monitoring of these parameters can thus help guide rapid clinical interventions. Besides, several environmental and demographic factors were found to have an impact on dengue severity; these are mosquito bite history, waterlogging near residences, urbanization patterns, and population density.

Therefore, when these findings are combined with the pictures of the hazard areas of Dhaka, Comilla, and Noakhali, they indicate the places that are the most dangerous and where the prevention activities should be given the first priority. The urban areas of Dhaka and the hotspot areas of Comilla and Noakhali are the places where such steps as mosquito control, public awareness, and the allocation of resources in a proactive way should be practised to a greater extent so as to decrease the risk of an outbreak. Consequently, health officials can implement scientifically rigorous interventions that are not only for the clinic but also for the community by focusing on these areas.

4.10 Summary

This chapter analyzed the severity, distribution, and prediction modeling of dengue comprehensively with clinical, demographic, and environmental data. Sensitivity analyses led to the identification of main features that not ⁴⁶only influence the severity of the disease but also such environmental factors as platelet count, white blood cell count, heart rate, blood pressure, mosquito bite history, waterlogging, urbanization patterns, and population density. Besides that, age, gender, city, residence type, and seasonal variations have been found to influence the extent of these parameters, thus, emphasizing the need for local context-specific evaluations.

The impact analysis of individual parameters visualized the gradual changes of clinical indicators across severity levels, their diagnostic relevance being thus confirmed. Risk zone mapping of Comilla, Noakhali, and Dhaka allowed for a spatial understanding of dengue prevalence, thus discovering high-risk hotspots where targeted interventions are highly necessary.

In terms of predictive performance, different traditional ⁶machine learning methods like Random Forest, Support Vector Machine, and Extreme Gradient Boosting were able to achieve

satisfactory accuracies, while ¹⁶deep learning models such as Multi-Layer Perceptron, Long Short-Term Memory, and One-Dimensional Convolutional Neural Network exhibited comparable results. It is noteworthy that ⁵⁵the hybrid Random Forest–Long Short-Term Memory model that was proposed tremendously surpassed ³⁶all the single models, thereby attaining not only the highest accuracy but also balanced precision, recall, and F1-score, thus indicating its stability in dengue ⁶⁵classification. Moreover, performance on a new dataset without any prior training further attests to the model's ability to generalize.

Indeed, the confluence of sensitivity results, risk mapping, and predictive modeling can greatly facilitate the deployment of targeted measures to prevent the occurrence of dengue, together with public awareness initiatives. As a result of the identification of the areas that are at high risk and of the recognition of the most important clinical and environmental parameters, it will be possible for the health authorities to put into practice proactive, evidence-based interventions aimed at not only alleviating the outbreak of diseases but also at reinforcing the capacity of the community for enduring such situations.

CHAPTER 5²¹

CONCLUSION AND FUTURE WORK

Conclusion

¹⁹Dengue fever remains a significant public health issue in Bangladesh. Its unpredictable nature, widespread occurrence, and potential for serious complications, such as severe hemorrhaging and shock, contribute to this concern. Timely and correct classification of dengue cases is vital for managing patients effectively. This classification helps in focusing interventions and reducing illness and death rates. Traditional diagnostic methods often encounter challenges due to overlapping symptoms and delays in lab results. This underscores the need for better computational methods that can analyze complex patient data. In this study, ¹⁴⁹multiple machine learning and deep learning models were evaluated to classify dengue cases based on detailed records information, including clinical features, physiological data, demographic details, and environmental context. The models included Random Forest, Support Vector Machine, Extreme Gradient Boosting, Multi-Layer Perceptron, Long Short-Term Memory, and one-dimensional Convolutional Neural Networks. In this study, a Hybrid Random Forest and Long Short-Term Memory model is introduced, combining feature-based and sequence-based learning to capture key determinants and temporal patterns in patient data, and it demonstrates ¹⁰superior performance compared to individual models. The Random Forest part pinpoints the most important factors affecting disease severity. The Long Short-Term Memory part tracks temporal trends and complex nonlinear relationships, allowing for accurate classification of cases into categories such as No Dengue, Mild Dengue Fever, Dengue Hemorrhagic Fever, and Dengue Shock Syndrome. To confirm reliability and generalizability, we evaluated the hybrid model on a separate dataset of new patient cases. The model performed consistently across different situations. Additionally, we conducted geographic and contextual analyses to create risk maps for Cumilla, Noakhali, and Dhaka. These maps identify high-risk areas and regions with higher incidence rates, aiding targeted public health actions. By combining hybrid modeling with data-driven spatial risk assessment, this research offers a strong framework that improves diagnostic accuracy, guides evidence-based public health choices, and supports effective prevention and management of dengue in Bangladesh.

Future Work

To improve dengue prediction and management, future research could include environmental, climatic, and mobility data. This would help identify high-risk areas better. It's also important to broaden the dataset across different geographic locations to make the model more general.

Thesis Report.....docx

ORIGINALITY REPORT

18%

SIMILARITY INDEX

14%

INTERNET SOURCES

14%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	1 %
2	www.geeksforgeeks.org Internet Source	1 %
3	www.mdpi.com Internet Source	1 %
4	www.ncbi.nlm.nih.gov Internet Source	1 %
5	www.humenhealth.com Internet Source	1 %
6	doctorpenguin.com Internet Source	<1 %
7	www.ezesauna.co.kr Internet Source	<1 %
8	S.P. Jani, M. Adam Khan. "Applications of AI in Smart Technologies and Manufacturing", CRC Press, 2025 Publication	<1 %
9	worldwidescience.org Internet Source	<1 %
10	H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Computer Science Engineering", CRC Press, 2024 Publication	<1 %
11	www.medchemexpress.cn Internet Source	<1 %

12	Devareddy, Dharmanand Reddy S.. "Clinical Profile and Outcome of Dengue Fever in Children", Rajiv Gandhi University of Health Sciences (India), 2023 Publication	<1 %
13	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dharendra Kumar Shukla. "Recent Trends in Intelligent Computing and Communication", CRC Press, 2025 Publication	<1 %
14	eitca.org Internet Source	<1 %
15	Shajid Hossain Hemal, Md. Ashikur Rahman Khan, Ishtiaq Ahammad, Masudur Rahman, Md. Ahnaf Sa'd Khan, Sabbir Ejaz. "Predicting the impact of internet usage on students' academic performance using machine learning techniques in Bangladesh perspective", Social Network Analysis and Mining, 2024 Publication	<1 %
16	mdpi-res.com Internet Source	<1 %
17	www2.mdpi.com Internet Source	<1 %
18	dokumen.pub Internet Source	<1 %
19	www.science.gov Internet Source	<1 %
20	offlineclinic.com Internet Source	<1 %
21	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in	<1 %

Information, Communication and Computing Technology", CRC Press, 2024

Publication

22	acris.aalto.fi Internet Source	<1 %
23	www.narayanahealth.org Internet Source	<1 %
24	assets-eu.researchsquare.com Internet Source	<1 %
25	link.springer.com Internet Source	<1 %
26	Ashok Kumar, Geeta Sharma, Anil Sharma, Pooja Chopra, Punam Rattan. "Advances in Networks, Intelligence and Computing - International Conference on Networks, Intelligence and Computing (ICONIC-2023)", CRC Press, 2024 Publication	<1 %
27	www.frontiersin.org Internet Source	<1 %
28	dk.um.si Internet Source	<1 %
29	www.ebm-journal.org Internet Source	<1 %
30	www.fastercapital.com Internet Source	<1 %
31	"Data Science, Classification, and Artificial Intelligence for Modeling Decision Making", Springer Science and Business Media LLC, 2025 Publication	<1 %
32	research-portal.uu.nl Internet Source	<1 %

33	www.medwonders.com Internet Source	<1 %
34	Shankar Babu, Mahesh Babu Kota. "Synergies in Smart and Virtual Systems using computational intelligence", CRC Press, 2025 Publication	<1 %
35	origin.geeksforgeeks.org Internet Source	<1 %
36	oulurepo.oulu.fi Internet Source	<1 %
37	energyinformatics.springeropen.com Internet Source	<1 %
38	ftp.math.utah.edu Internet Source	<1 %
39	nceca.in Internet Source	<1 %
40	"Artificial Intelligence for System Oriented Design", Springer Science and Business Media LLC, 2025 Publication	<1 %
41	Md Jahangir Alam, Reaz Mobarak, Md Kamruzzaman, Abu Tayab, Probir Kumar Sarkar, ABM Mahfuj Hassan Al Mamun. "Clinical Evaluation of Dengue Cases in Bangladesh Shishu Hospital & Institute", Dhaka Shishu (Children) Hospital Journal, 2024 Publication	<1 %
42	pmc.ncbi.nlm.nih.gov Internet Source	<1 %
43	www.acc.org Internet Source	<1 %
44	academia.wascal.org	

<1 %

45

academic-accelerator.com

Internet Source

<1 %

46

data.cdc.gov.tw

Internet Source

<1 %

47

ebin.pub

Internet Source

<1 %

48

www.diva-portal.org

Internet Source

<1 %

49

www.isteonline.in

Internet Source

<1 %

50

Nhan, Nguyen Trong. "Time Series and Spatial Analysis of Dengue Fever Before and During the Coronavirus Disease (COVID-19) Pandemic in Phu Yen Province, Vietnam.", National Yang Ming Chiao Tung University, 2024

Publication

<1 %

51

repository.ju.edu.et

Internet Source

<1 %

52

www.medrxiv.org

Internet Source

<1 %

53

Danish, Syed Hasan. "Optimizing Delivery Rider Allocation Using Genetic Algorithm: A Demand Driven Approach for Efficient Order Fulfillment and Fair Earnings", Hamad Bin Khalifa University (Qatar), 2025

Publication

<1 %

54

dspace.bracu.ac.bd:8080

Internet Source

<1 %

55

Atheer Hammad. "Random Forest and LSTM Hybrid Model for Detecting DDoS Attacks in

<1 %

Healthcare IoT Networks", CyberSystem Journal, 2024

Publication

56	arxiv.org Internet Source	<1 %
57	scpe.org Internet Source	<1 %
58	web.realinfo.tv Internet Source	<1 %
59	www.sciencegate.app Internet Source	<1 %
60	Mamta Tikaria, Vineeta Saxena. "Dual-Input Multi-Layered Attention Model for Enhanced Path Loss Prediction", Springer Science and Business Media LLC, 2025 Publication	<1 %
61	Rajganesh Nagarajan, Senthilkumar Narayanasamy, Ramkumar Thirunavukarasu, Pethuru Raj. "Intelligent Systems and Sustainable Computational Models - Concepts, Architecture, and Practical Applications", CRC Press, 2024 Publication	<1 %
62	e-archivo.uc3m.es Internet Source	<1 %
63	ijsrset.com Internet Source	<1 %
64	opentransportationjournal.com Internet Source	<1 %
65	robots.net Internet Source	<1 %
66	Yagyanath Rimal, Sunil Parihar, A. Samson Arun Raj, Purvee Bhardwaj, Anjali Rawat,	<1 %

Abhishek Sharma. "chapter 13 Screen Time and Safety", IGI Global, 2025

Publication

67	eprints.nottingham.ac.uk Internet Source	<1 %
68	iieta.org Internet Source	<1 %
69	ijain.org Internet Source	<1 %
70	jbiomedsci.biomedcentral.com Internet Source	<1 %
71	www.hindawi.com Internet Source	<1 %
72	wwwnc.cdc.gov Internet Source	<1 %
73	Bokaba, Tebogo Emma. "A Data-Driven Framework for Forecasting and Analyses of Road Traffic Incidents", University of Johannesburg (South Africa), 2023 Publication	<1 %
74	Mehdi Ghayoumi. "Generative Adversarial Networks in Practice", CRC Press, 2023 Publication	<1 %
75	aaai.org Internet Source	<1 %
76	hdl.handle.net Internet Source	<1 %
77	stars.library.ucf.edu Internet Source	<1 %
78	www.nature.com Internet Source	<1 %
79	Nealis, Lakely Bea. "Machine Learning Methods for Accuracy Evaluation of Aviation	<1 %

Weather Reports", California State University, Long Beach

Publication

80 Vivek Kumar Singh, Jitendra Kumar, Deepika Saxena, Abhishek Verma, T. Akilan, Ashutosh Kumar Singh. "Machine Vision Analysis in Industry 5.0 - Fundamentals, Applications, and Challenges", CRC Press, 2025

Publication

81 didescholar.acemap.info <1 %

Internet Source

82 do Carmo, Felipe Dantas. "Overcoming Data Scarcity in Load Forecasting: A Transfer Learning Approach for Commercial Buildings", Universidade do Porto (Portugal), 2024

Publication

83 idss.iocspublisher.org <1 %

Internet Source

84 naos-be.zcu.cz <1 %

Internet Source

85 spectrum.library.concordia.ca <1 %

Internet Source

86 www.upgrad.com <1 %

Internet Source

87 Abdi-Sargezeh, Bahman. "Identification of Brain Epileptiform Discharges from Electroencephalograms", Nottingham Trent University (United Kingdom), 2023

Publication

88 Khumukcham Robindro Singh, Nazrul Hoque, Arnab Kumar Maji, Sabyasachi Mondal et al. "Emerging Trends and Future Directions in Artificial Intelligence, Machine Learning, and

Internet of Things Innovations - A proceeding of NEIAIS — 2025", CRC Press, 2025

Publication

-
- | | | |
|----|---|------|
| 89 | Momoh, Oludare Boon. "Evaluating the Efficacy of AI Algorithms in Cyber Threat Detection Across Diverse Datasets", National University
Publication | <1 % |
| 90 | etd.uum.edu.my
Internet Source | <1 % |
| 91 | hajj.uqu.edu.sa
Internet Source | <1 % |
| 92 | Swati V. Shinde, Darshan V. Medhane, Oscar Castillo. "Applied Computer Vision and Soft Computing with Interpretable AI", CRC Press, 2023
Publication | <1 % |
| 93 | erevistas.uacj.mx
Internet Source | <1 % |
| 94 | mindthegraph.com
Internet Source | <1 % |
| 95 | publications.aston.ac.uk
Internet Source | <1 % |
| 96 | tudr.thapar.edu:8080
Internet Source | <1 % |
| 97 | www.aeaweb.org
Internet Source | <1 % |
| 98 | Markapurapu John Dana Ebinezer, Bondalapu Chaitanya Krishna. "Life Insurance Fraud Detection: A Data-Driven Approach Utilizing Ensemble Learning, CVAE, and Bi-LSTM", Applied Sciences, 2025
Publication | <1 % |
-

99 Mehta, Ravi. "The Neurological Manifestations of Zika and Chikungunya Viruses.", The University of Liverpool (United Kingdom), 2020 <1 %
Publication

100 Toai Kim Tran, Roman Senkerik, Hanh Thi Xuan Vo, Huan Minh Vo, Adam Ulrich, Marek Musil, Ivan Zelinka. "Exploring Hybrid Models For Short-Term Local Weather Forecasting in IoT Environment", MENDEL, 2023 <1 %
Publication

101 bmccresnotes.biomedcentral.com <1 %
Internet Source

102 curis.ku.dk <1 %
Internet Source

103 openrepository.aut.ac.nz <1 %
Internet Source

104 pub.epsilon.slu.se <1 %
Internet Source

105 www.researchgate.net <1 %
Internet Source

106 bmcinfectdis.biomedcentral.com <1 %
Internet Source

107 library.oapen.org <1 %
Internet Source

108 medcraveonline.com <1 %
Internet Source

109 ojs.bonviewpress.com <1 %
Internet Source

110 test-jicce.inforang.com <1 %
Internet Source

111 www.earthobservatory.sg <1 %
Internet Source

112	www.expressclinics.in Internet Source	<1 %
113	Anakweze, Charles C.. "Enhancing Network Resilience in the Telecommunication Sector: A Data-Driven Approach to Malware Detection in Android Phones", The George Washington University Publication	<1 %
114	Jin-Xin Meng, Qiu-Ming Hu, Li-Ming Zhang, Nan Li, Yu-Wen He, Zhen-Xing Yang, Yi Sun, Jing-Lin Wang. "Isolation and Genetic Evolution of Dengue Virus from the 2019 Outbreak in Xishuangbanna, Yunnan Province, China", Vector-Borne and Zoonotic Diseases, 2023 Publication	<1 %
115	Pingkan Mayosi Fitriana, Jumadil Saputra, Zairihan Abdul Halim. "The Impact of the COVID-19 Pandemic on Stock Market Performance in G20 Countries: Evidence from Long Short-Term Memory with a Recurrent Neural Network Approach", Big Data, 2023 Publication	<1 %
116	Rashmi Gupta, Shashi Shekhar, Jeetendra Kumar, Gaurav Dhiman. "Ambient Assisted Living - Technological Advances", CRC Press, 2025 Publication	<1 %
117	en.alegsaonline.com Internet Source	<1 %
118	his.diva-portal.org Internet Source	<1 %
119	lirias.kuleuven.be Internet Source	<1 %

120	repository.kaust.edu.sa Internet Source	<1 %
121	www.degruyterbrill.com Internet Source	<1 %
122	www.diva-portal.se Internet Source	<1 %
123	www.magiran.com Internet Source	<1 %
124	"13th European Congress of Clinical Microbiology and Infectious Diseases", Clinical Microbiology and Infection, 2003 Publication	<1 %
125	Hossain, Azfar. "ML vs QML Models: Anomaly Detection on Scalable Network Traffic", Capitol Technology University Publication	<1 %
126	P. M. Kikin, A. A. Kolesnikov, A. M. Portnov. "USE OF MACHINE LEARNING TECHNIQUES FOR RAPID DETECTION, ASSESSMENT AND MAPPING THE IMPACT OF DISASTERS ON TRANSPORT INFRASTRUCTURE", ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019 Publication	<1 %
127	agencia.fapesp.br Internet Source	<1 %
128	bastina.anubih.ba Internet Source	<1 %
129	digitalcommons.wayne.edu Internet Source	<1 %
130	docplayer.net Internet Source	<1 %

131	inmyglobe.blogspot.com Internet Source	<1 %
132	ir.cwi.nl Internet Source	<1 %
133	www.teses.usp.br Internet Source	<1 %
134	www.tnsroindia.org.in Internet Source	<1 %
135	www.weeklyblitz.net Internet Source	<1 %
136	Azumah, Sylvia Worlali. "Cyberbullying on Social Network Sites: Examining Ghanaian Cultural Perspective, Psychological Impact and Detection Technologies", University of Cincinnati Publication	<1 %
137	Iacopo Carnacina, Mawada Abdellatif, Manolia Andredaki, James Cooper, Darren Lumbroso, Virginia Ruiz-Villanueva. "River Flow 2024", CRC Press, 2025 Publication	<1 %
138	Muhammad Asif, Jari S. Algethami, Mohsen A.M. Alhamami, Peng Lie, Shunag Shaomin, Ayesha Aziz. "Redefining nanobiosensors in rapid detection of viral infections: Where are we now?", Trends in Environmental Analytical Chemistry, 2025 Publication	<1 %
139	Richards, Heather. "Is the Application of a Vulnerability Framework Effective in Determining Patterns of the Incidence of Dengue Disease on the Island of Dominica?:The Water Associated Disease	<1 %

Index (WADI) Model", Lancaster University
(United Kingdom), 2020

Publication

140	blockmd.com Internet Source	<1 %
141	bmcsurg.biomedcentral.com Internet Source	<1 %
142	espace.etsmtl.ca Internet Source	<1 %
143	huggingface.co Internet Source	<1 %
144	ir.lib.uth.gr Internet Source	<1 %
145	omeka.ibu.edu.ba Internet Source	<1 %
146	publikationen.uni-tuebingen.de Internet Source	<1 %
147	scholarworks.umass.edu Internet Source	<1 %
148	umpir.ump.edu.my Internet Source	<1 %
149	www.computersciencejournals.com Internet Source	<1 %
150	www.iaeng.org Internet Source	<1 %
151	www.scholarsresearchlibrary.com Internet Source	<1 %
152	www.scitepress.org Internet Source	<1 %
153	www.teikyomedicaljournal.com Internet Source	<1 %

154 Abhishek S. Rao, B. H. Karthik Pai, K. Adithi, Lakshmi Belur Keshav, Karan Malhotra, Sneha Nayak, H. K. Sachidananda. "Scoring model for exploring factors influencing mortality in dengue patients at a tertiary care hospital: a retrospective study", Discover Applied Sciences, 2024

Publication

<1 %

155 Alfardus, Asma. "Evaluating Machine Learning for Intrusion Detection in CAN Bus for In-Vehicle Security", Howard University

Publication

<1 %

156 Brij Kishore Tyagi. "Mosquitoes of India - Mosquito-Borne Diseases, Diagnosis and Control", CRC Press, 2025

Publication

<1 %

157 Charles H. Calisher. "SCIENTIST IN A STRANGE LAND", The Nonproliferation Review, 2009

Publication

<1 %

158 Felipe de Lima Peressim. "Machine learning for the prediction of 28-day cement compressive strength", Universidade de São Paulo. Agência de Bibliotecas e Coleções Digitais, 2025

Publication

<1 %

159 Najafi, Zhubin. "Analysis of Incident Duration and Real-Time Incident Delay Estimation Using Big Data Analytics and Machine Learning Techniques", University of Nevada, Las Vegas

Publication

<1 %

160 Salihi, Ali Abdulrahman Hussein. "Wind Energy Forecasting Methods: A Case Study of the Long Short Term Memory Model (LSTM).", Kirsehir Ahi Evran University (Turkey), 2024

Publication

<1 %

161	Thomas P. Monath. "The Arboviruses: Epidemiology and Ecology", CRC Press, 2019 Publication	<1 %
162	Zaydi Mounia, Maleh Yassine, Gabriel Chênevert, Hayat Zaydi, Amina El Yaagoubi. "Intelligent Cybersecurity and Resilience for Critical Industries: Challenges and Applications", River Publishers, 2025 Publication	<1 %
163	admin.uni-sofia.bg Internet Source	<1 %
164	dspace.unza.zm Internet Source	<1 %
165	eureka.patsnap.com Internet Source	<1 %
166	icandecide.org Internet Source	<1 %
167	ijirt.org Internet Source	<1 %
168	ijnrd.org Internet Source	<1 %
169	ijrpr.com Internet Source	<1 %
170	impactfactor.org Internet Source	<1 %
171	iris.uniss.it Internet Source	<1 %
172	learn.g2.com Internet Source	<1 %
173	sear.unisq.edu.au Internet Source	<1 %
174	sergioescalera.com	

<1 %

175

staging-mental.jmir.org

Internet Source

<1 %

176

weeditspotspray.com

Internet Source

<1 %

177

www.gnedenko.net

Internet Source

<1 %

178

www.zbw.eu

Internet Source

<1 %

179

Handbook of Disease Burdens and Quality of Life Measures, 2010.

Publication

<1 %

180

J. Meiyazhagan, S. Sudharsan, A. Venkatesan, M. Senthilvelan. "Prediction of occurrence of extreme events using machine learning", The European Physical Journal Plus, 2021

Publication

<1 %

181

NADHIRAT SANGKAWIBHA, SUNTHAREE ROJANASUPHOT, SOMPOP AHANDRIK, SUKHO VIRIYAPONGSE et al. "RISK FACTORS IN DENGUE SHOCK SYNDROME: A PROSPECTIVE EPIDEMIOLOGIC STUDY IN RAYONG, THAILAND", American Journal of Epidemiology, 1984

Publication

<1 %

182

Nguyen Thi Cam Huong, Nguyen Thi Ngan, Abdullah Reda, Vinh Dong et al. "Association of self-reported allergic rhinitis with dengue severity: A case-control study", Acta Tropica, 2022

Publication

<1 %

183

P.V. Mohanan. "Artificial Intelligence and Biological Sciences", CRC Press, 2025

Publication

<1 %

184 Singhal, Mayank. "Role of Steroids in Modulating Progression of Dengue Fever in Patients With Warning Signs to Dengue Haemorrhagic Fever/ Dengue Shock Syndrome", Postgraduate Institute of Medical Education and Research, Chandigarh (India), 2024
Publication

185 "Data Science", Springer Science and Business Media LLC, 2020
Publication

186 Anurag Tiwari, Manuj Darbari. "Emerging Trends in Computer Science and Its Application - Proceedings of the International Conference on Advances in Emerging Trends in Computer Applications (ICAETC-2023) December 21–22, 2023, Lucknow, India", CRC Press, 2025
Publication

187 D. Sivabalaselvamani, G. Revathy, Ranjit Singh Sarban Singh. "Advanced AI and Data Science Applications", CRC Press, 2025
Publication

188 G. S. Karthick. "Sensor Data Analytics for Intelligent Healthcare Delivery", CRC Press, 2025
Publication

189 Satish Mahadevan Srinivasan, Phillip A. Laplante. "What Every Engineer Should Know About Data-Driven Analytics", CRC Press, 2023
Publication

190 Sepp Hochreiter, Jürgen Schmidhuber. "Long Short-Term Memory", Neural Computation, 1997
Publication

Yeza Febriani, Muhammad Lutfi Mahasinul Akhlak, Mohammad Apriniyadi, Arif Rahman Saleh. "Evaluation of Deep Learning Model Performance for Water Quality Classification: Comparison of CNN, LSTM, and Hybrid CNN-LSTM", Springer Science and Business Media LLC, 2025

Publication

<1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off