

Machine Learning Based Early Detection of Dengue Using CBC Data

Ambresh Bhadrashetty¹, Manjulabai Bhadrashetty², Pranjali Hunagund³

¹Assistant Professor, Department of Computer Science and Engineering (MCA) Visvesvaraya Technological University, Centre for PG Studies, Kalaburagi, India

²Associate Professor, Department Of Computer Science, Government First Grade College for Women's, Jewargi Colony, Kalaburagi, India

³Student, Department of Computer Science and Engineering (MCA) Visvesvaraya Technological University, Centre for PG Studies, Kalaburagi, India

Abstract - Dengue fever remains a major health concern in tropical and subtropical regions, where delayed diagnosis can lead to severe complications and increased mortality. Conventional diagnostic approaches often rely on laboratory investigations and expert interpretation of Complete Blood Count (CBC) reports, which may not always be available promptly during large-scale outbreaks. This project presents **DengueAI-CBC**, an intelligent dengue risk assessment system that utilizes machine learning techniques to analyze routine CBC parameters and identify potential dengue cases at an early stage.

The proposed system processes hematological indicators such as platelet count, white blood cell count, hemoglobin level, hematocrit, age, and gender to estimate dengue risk. Data preprocessing techniques, including cleaning, normalization, and feature preparation, are applied to improve prediction quality. A machine learning model is trained and evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. The final model is integrated into a user-friendly web application that enables healthcare personnel to obtain rapid risk assessments from CBC values.

The system classifies patients into different risk categories and provides immediate results to support early clinical decision-making. By reducing dependence on manual analysis and enabling fast screening, DengueAI-CBC offers a practical and cost-effective solution for preliminary dengue detection. The application is intended to assist healthcare professionals and should be used as a supportive screening tool rather than a substitute for medical diagnosis.

Key Words: Dengue Fever, Complete Blood Count, Machine Learning, Healthcare Analytics, Risk Prediction, Early Screening, Clinical Decision Support.

1. INTRODUCTION

Dengue fever is a mosquito-borne viral disease that continues to pose a significant public health challenge in many tropical and subtropical countries. The disease is primarily transmitted

through the bite of infected *Aedes aegypti* mosquitoes and affects millions of people every year. Rapid urbanization, population growth, climate change, and inadequate mosquito control measures have contributed to the increasing incidence of dengue worldwide. Early symptoms such as high fever, headache, muscle pain, nausea, and fatigue are often similar to those of other viral infections, making timely diagnosis difficult.

Early identification of dengue infection is essential to prevent severe complications, including Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS). Medical professionals commonly rely on laboratory investigations, particularly Complete Blood Count (CBC) reports, to evaluate important hematological parameters such as platelet count, white blood cell count, hemoglobin, and hematocrit levels. However, manual interpretation of these reports can be time-consuming and may lead to delays during periods of high patient volume.

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have created new opportunities for improving disease prediction and healthcare decision-making. Machine learning algorithms can analyze large volumes of clinical data, identify hidden patterns, and generate accurate predictions within a short time. By utilizing routinely available CBC parameters, AI-based systems can assist healthcare professionals in identifying patients who may be at risk of dengue infection.

The **DengueAI-CBC** system is developed as an intelligent dengue risk assessment tool that uses machine learning techniques to analyze CBC data and provide rapid risk predictions. The system processes patient information and hematological parameters to classify dengue risk levels, helping healthcare workers perform early screening and prioritize medical attention when necessary. A user-friendly web interface further enhances accessibility and usability in clinical environments.

2. PROBLEM STATEMENT

Dengue fever is a rapidly spreading viral disease that affects millions of people worldwide, particularly in tropical and subtropical regions. Early diagnosis is critical for effective treatment and prevention of severe complications. However, the initial symptoms of dengue, such as fever, headache, fatigue, and body pain, are often similar to those of other infectious diseases, making accurate diagnosis challenging during the early stages.

Healthcare professionals commonly depend on laboratory tests and manual analysis of Complete Blood Count (CBC) reports to identify dengue-related abnormalities. During outbreaks, hospitals and diagnostic centers experience a large number of patients, increasing the workload on medical staff and causing delays in diagnosis. Manual interpretation of CBC parameters may also be influenced by human error and varying levels of clinical experience.

Although advanced diagnostic tests are available, they can be expensive, time-consuming, and inaccessible in resource-limited healthcare settings. As a result, there is a growing need for an automated and efficient system capable of analyzing routine CBC data and providing early dengue risk assessment.

The problem addressed in this project is the lack of a fast, reliable, and accessible screening tool that can assist healthcare professionals in identifying potential dengue cases using commonly available hematological parameters. Therefore, the proposed **DengueAI-CBC** system aims to utilize machine learning techniques to analyze CBC reports and generate accurate dengue risk predictions, supporting timely clinical decision-making and improving patient care.

3. OBJECTIVES

The main objective of the **DengueAI-CBC** project is to develop an intelligent and automated dengue risk assessment system that utilizes machine learning techniques to analyze Complete Blood Count (CBC) parameters and support early disease detection. The proposed system aims to assist healthcare professionals in identifying potential dengue cases quickly and accurately, thereby improving clinical decision-making and reducing diagnostic delays during outbreaks.

To achieve this goal, the project focuses on collecting and preprocessing CBC data, including handling missing values, removing inconsistencies, and preparing features for analysis. The system is designed to learn patterns associated with dengue infection by training machine learning models on hematological parameters such as platelet count, white blood cell count, hemoglobin level, hematocrit, age, and gender. The developed model is evaluated using standard performance measures, including accuracy, precision, recall, and F1-score, to ensure reliable predictions.

Another important objective is to categorize patients into different risk levels, such as Low Risk, Moderate Risk, and High Risk, based on their CBC results. The project also aims to provide a user-friendly web-based interface that enables healthcare personnel to perform rapid dengue risk assessments

without requiring specialized technical knowledge. Furthermore, the system seeks to offer a cost-effective, accessible, and efficient screening solution that can be used in hospitals, clinics, and resource-constrained healthcare environments. While the application is not intended to replace professional medical diagnosis, it serves as a valuable decision-support tool for early dengue screening and timely patient management.

4. METHODOLOGY

The methodology adopted for the **DengueAI-CBC** project follows a structured process that transforms raw hematological data into meaningful dengue risk predictions through machine learning techniques. The proposed framework consists of data collection, preprocessing, model development, evaluation, and deployment stages to ensure accurate and efficient dengue risk assessment.

Initially, a dataset containing Complete Blood Count (CBC) parameters and patient-related information is collected from reliable medical sources. The dataset includes important hematological attributes such as platelet count, white blood cell count, hemoglobin level, hematocrit, age, and gender, which are commonly used in dengue diagnosis. After data acquisition, preprocessing techniques are applied to improve data quality. This stage involves handling missing values, removing inconsistencies, eliminating duplicate records, and normalizing numerical features to ensure uniformity across the dataset.

Following preprocessing, exploratory data analysis is performed to understand the distribution of features and identify relationships between CBC parameters and dengue occurrence. Relevant features are then selected and used to train machine learning models capable of recognizing patterns associated with dengue infection. The trained model learns from historical data and generates predictions based on the input CBC values provided by users.

The performance of the developed model is evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score. These metrics help determine the effectiveness and reliability of the prediction system. Once the optimal model is identified, it is integrated into a web-based application that allows users to enter CBC parameters and receive instant dengue risk predictions.

The final system classifies patients into appropriate risk categories and presents the results through an intuitive user interface. This methodology enables rapid, data-driven dengue screening while maintaining simplicity, accessibility, and practical usability in healthcare environments.

5. LITERATURE SURVEY

Dengue fever remains one of the most prevalent mosquito-borne diseases worldwide, creating a significant burden on healthcare systems, especially in tropical and subtropical regions. Early diagnosis is essential for reducing disease severity and preventing complications such as Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS). Traditional diagnostic methods, including NS1 antigen

tests, antibody detection, and manual interpretation of Complete Blood Count (CBC) reports, often require laboratory infrastructure, trained personnel, and considerable processing time. These limitations have encouraged researchers to explore Artificial Intelligence (AI) and Machine Learning (ML) techniques for rapid and accurate dengue prediction.

Several studies have demonstrated the effectiveness of machine learning models in analyzing clinical and hematological data for dengue diagnosis. Researchers have applied algorithms such as Decision Trees, Support Vector Machines (SVM), Logistic Regression, Random Forest, and Artificial Neural Networks (ANN) to identify patterns associated with dengue infection. These approaches have shown promising results by improving prediction accuracy and supporting early disease detection. However, some traditional models require extensive feature engineering and may struggle to capture complex relationships among clinical parameters.

Recent advancements in deep learning have further enhanced predictive performance by automatically learning intricate patterns from medical datasets. Neural network-based approaches have achieved higher accuracy levels compared to conventional machine learning techniques, making them suitable for healthcare applications. Nevertheless, deep learning models often require larger datasets and higher computational resources, which may limit their deployment in resource-constrained environments.

Researchers have also explored image-based dengue detection using microscopic blood smear images and computer vision techniques. While these methods provide high accuracy, they depend on specialized imaging equipment and may not be practical for widespread screening. In contrast, CBC-based prediction systems offer a more accessible and cost-effective alternative because CBC tests are routinely performed in most healthcare facilities.

The reviewed literature indicates that machine learning can significantly improve dengue screening by utilizing readily available clinical and hematological parameters. However, there remains a need for an efficient, user-friendly, and accessible solution that combines accurate prediction with practical deployment. Motivated by these findings, the proposed **DengueAI-CBC** system employs machine learning techniques to analyze CBC parameters and provide rapid dengue risk assessment through a web-based platform. This approach aims to support healthcare professionals in early screening and timely decision-making while maintaining simplicity and affordability.

6. SYSTEM DESIGN

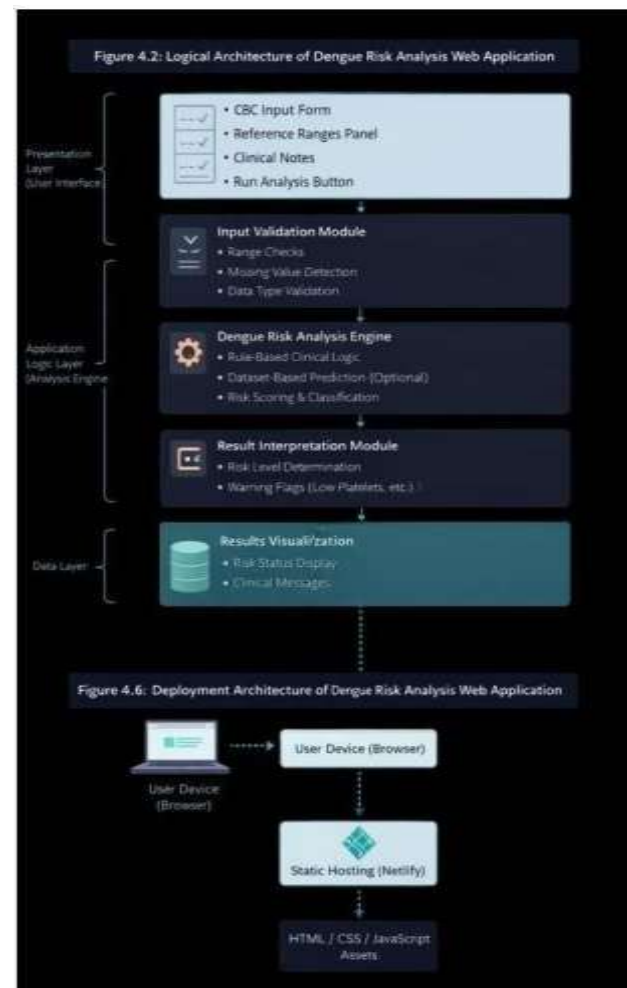


Figure 1: Deployment Architecture of the DengueAI-CBC Web Application

The DengueAI-CBC system follows a layered architecture for dengue risk prediction using CBC data. It includes a user interface for data entry, a processing layer for validation and analysis, and a result layer for displaying predictions. The machine learning model evaluates the input parameters and determines the risk level of dengue infection. The application is deployed as a web-based platform, allowing users to access the system through a browser. This design ensures efficient processing, easy accessibility, and reliable risk assessment.

7. SCREENSHOTS

1. Home Page



Figure 2 : Home Page

2. Patient CBC Data Entry Interface

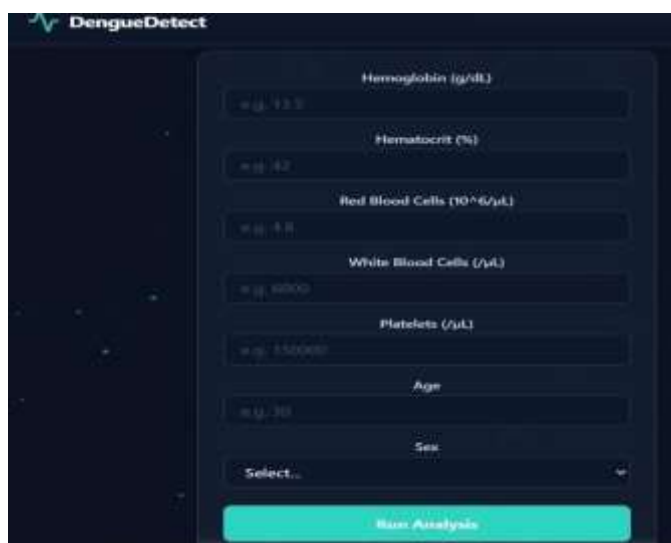


Figure 3: Patient CBC Data Entry Interface

3. CBC Reference Range Display Module

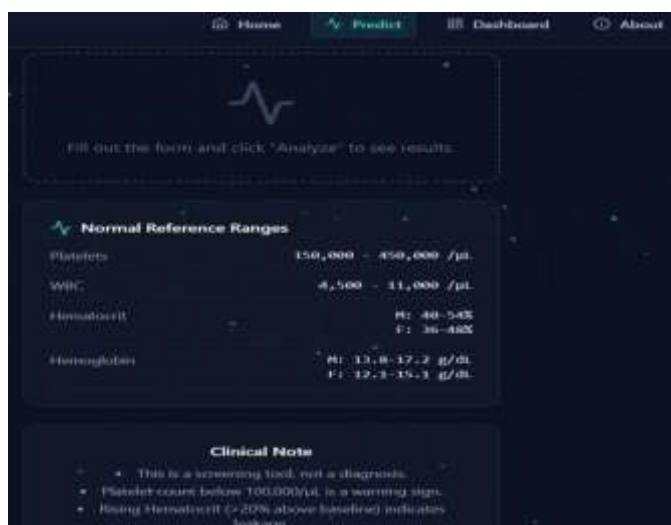


Figure 4: CBC Reference Range Display Module

4. Dengue Risk Classification and Analysis Dashboard



Figure 5: Dengue Risk Classification and Analysis Dashboard

5. Analytics Dashboard of the DengueAI-CBC System



Figure 6: Analytics Dashboard of the DengueAI-CBC System

8. CONCLUSION & FUTURE SCOPE

This research presents a machine learning-based framework for dengue risk assessment using Complete Blood Count (CBC) parameters. The developed system examines routine blood test values to identify patterns associated with dengue infection and generate timely risk predictions. A web-based platform enables users to access the system easily and obtain rapid screening results. The study demonstrates the effectiveness of data-driven techniques in supporting early dengue detection and assisting healthcare professionals during preliminary evaluation. By providing quick and reliable risk assessment, the proposed solution can contribute to improved screening efficiency and informed clinical decision-making.

Future enhancements may focus on incorporating additional clinical indicators and laboratory measurements to strengthen prediction performance. The framework can also be extended to support the identification of other infectious diseases through a unified prediction model. Integration with hospital information systems and electronic health records would enable automated data handling and real-time monitoring. Furthermore, mobile application deployment can improve accessibility in remote and resource-limited areas. The inclusion of explainable artificial intelligence techniques may provide greater transparency in predictions and increase confidence in the system's recommendations.

9. REFERENCES

1. World Health Organization (WHO), *Dengue and Severe Dengue*, WHO Fact Sheet, 2023.
2. Centers for Disease Control and Prevention (CDC), *Dengue: Epidemiology and Prevention*, CDC Publications, 2023.
3. Davi, C., et al., "Machine Learning Approaches for Dengue Prediction Using Clinical and Genomic Data," *Journal of Medical Systems*, vol. 45, no. 3, pp. 1–12, 2021.
4. Sarma, J., and Bhattacharyya, D. K., "A Decision Tree-Based Dengue Prediction System," *International Journal of Computer Applications*, vol. 176, no. 22, pp. 25–31, 2020.
5. Fernández, J., et al., "Logistic Regression Model for Dengue Diagnosis Using Hematological Parameters," *BMC Infectious Diseases*, vol. 19, no. 1, pp. 1–10, 2019.
6. Mayrose, J., et al., "Automated Dengue Detection Using Blood Smear Images and Support Vector Machines," *IEEE Access*, vol. 8, pp. 115021–115030, 2020.
7. Mello, C., et al., "Deep Neural Networks for Dengue Diagnosis Using Clinical Data," *Expert Systems with Applications*, vol. 158, 2020.
8. Pedregosa, F., et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.