

> A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

ARTIFICIAL INTELLIGENCE-BASED EARLY DETECTION OF DENGUE USING CBC DATA

¹Narender Kunta

kuntanarender55@gmail.com

Abstract: An artificial intelligence-based approach for the early detection of dengue using complete blood count (CBC) data is presented. Various feature selection techniques, including Pearson Correlation, Recursive Feature Elimination (RFE) with Random Forest, SelectKBest, Chi-Square (Chi2), and ExtraTree, are applied to identify the most relevant attributes. Multiple machine learning and deep learning algorithms are employed, such as Logistic Regression, Support Vector Machine (SVM), Naive Bayes, Random Forest, AdaBoost, XGBoost, Multi-Layer Perceptron (MLP), LightGBM, and ensemble methods like a Stacking Classifier (XGB + LR + MLP with LightGBM) and Voting Classifier (Boosted Decision Tree + ExtraTree). Deep learning architectures, including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), Bidirectional Long Short-Term Memory (Bi-LSTM), Feedforward Neural Networks (FNN), Transformer, and hybrid models such as CNN + LSTM, are utilized. Combining predictions from individual models through ensemble methods enhances robustness and accuracy, with the Voting Classifier achieving 98% accuracy and F1 Score. The use of hybrid models, particularly CNN + LSTM, further contributes to the system's performance. The approach is designed for user interaction and validation through a Flask-based interface with authentication, ensuring accessibility and security while maintaining high predictive accuracy.

Index Terms - Complete blood count, dengue prediction, explainable AI, feature selection, machine learning, ensemble learning, transformer model.

1. INTRODUCTION

The human body, inherently sensitive, is equipped with its own defense mechanisms to combat external microbial threats. However, humans are often susceptible to viral and bacterial infections that can lead to severe diseases, some of which can be significantly lethal. Dengue fever, a viral disease primarily transmitted by the Aedes mosquito, is one such infection that poses a serious global health threat. Each year, millions of individuals worldwide suffer from dengue, with



> A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

thousands succumbing to its complications [1]. According to the World Health Organization (WHO) and the European Union, over six million people across 92 countries were affected by dengue fever in 2023 alone, with Bangladesh reporting more than 0.31 million cases and over 1,600 deaths from this hemorrhagic fever [2]. Dengue is predominantly prevalent in urban and peri-urban areas of tropical and subtropical regions, a consequence of inadequate sanitation, rapid unplanned urbanization, and haphazard development [3]. The WHO identifies regions in Africa, Southeast Asia, and the Western Pacific as having the highest incidence of dengue, with Bangladesh recording the highest number of cases in the Southeast Asian region during the peak period of June to October in 2023. This year marked the highest number of dengue-related cases and fatalities in Bangladesh in recent decades, emphasizing the urgent need for early, efficient, and rapid detection and response measures [4].

Although dengue is transmitted through the bite of an infected Aedes mosquito, the virus often remains dormant in the body before symptoms manifest. While not inherently fatal, dengue presents a wide range of debilitating symptoms, including intense fever, severe body pain, nausea, loss of appetite, and skin rashes. In many cases, these symptoms overlap with those of other illnesses, making early diagnosis challenging. Furthermore, patients often experience a sudden health decline as the disease progresses, typically within two weeks of infection [5]. Pathological findings frequently reveal a critical drop in platelet levels, signifying severe disease progression. Dengue virus has four serotypes, and while infection with one serotype provides long-term immunity against it, subsequent infection with a different serotype can lead to severe outcomes such as dengue shock syndrome, internal bleeding, or multiple organ failure [6]. Given the escalating number of cases and fatalities, especially in regions with poor preventive measures, it is imperative to implement effective diagnostic systems and preventive strategies to mitigate the impact of dengue and avert surpassing historical records of affected individuals and deaths.

2. RELATED WORK

Dengue detection has seen significant advancements in recent years, transitioning from traditional diagnostic tools to more sophisticated approaches incorporating machine learning and artificial intelligence. Kabir et al. [7] discussed the

International Journal For Advanced Research

In Science & Technology



A peer reviewed international journal ISSN: 2457-0362 www.ijarst.in

evolution of dengue diagnostic tools, highlighting the shift from conventional technologies to point-of-care solutions that offer rapid and accurate detection. These advancements are pivotal in addressing the growing burden of dengue in endemic regions. Davi et al. [8] explored the use of human genome data combined with machine learning for severe dengue prognosis, demonstrating how genetic information improve predictive can accuracy for identifying high-risk patients. Similarly, Sarma et al. [9] applied machine learning algorithms to predict dengue outbreaks, emphasizing the utility of predictive models in public health planning and resource allocation. Their work underscores the potential of datadriven approaches in mitigating the impacts of dengue outbreaks.

Fernández et al. [10] developed a predictive model to differentiate dengue from other febrile illnesses, leveraging clinical and laboratory data to enhance diagnostic specificity. This approach is particularly valuable in regions where febrile illnesses are prevalent, often leading to misdiagnosis. Mayrose et al. [11] utilized machine learning to detect dengue from blood smear images by analyzing platelet and lymphocyte characteristics. Their research highlights

the potential of image-based diagnostics in complementing traditional laboratory tests, providing a non-invasive and efficient alternative for early dengue detection. Sabrina Prome et al. [12] applied an explainable machine learning approach to predict dengue cases in Bangladesh, integrating epidemiological and environmental data. Their study emphasizes the importance of transparency and interpretability in machine learning models, ensuring their applicability in realworld healthcare settings.

Mello-Román et al. [13] focused on predictive models for dengue diagnosis in Paraguay, demonstrating the applicability of computational methods in diverse geographic and epidemiological contexts. Their findings reinforce the adaptability of machine learning techniques across varying healthcare infrastructures. Dey et al. [14] employed machine learning to predict dengue incidents in Bangladesh, incorporating meteorological, socioeconomic, and patient data. Their comprehensive approach underscores the multifaceted nature of dengue outbreaks and the necessity of integrating diverse data sources for accurate predictions. Collectively, these studies reflect the growing role of machine learning and artificial intelligence in enhancing dengue



A peer reviewed international journal ISSN: 2457-0362 www.ijarst.in

diagnostics and prediction, paving the way for more effective and proactive public health interventions. The integration of clinical, genetic, epidemiological, and environmental data has proven critical in developing robust models that improve early detection and facilitate timely treatment, ultimately reducing the morbidity and mortality associated with dengue fever.

3. MATERIALS AND METHODS

The proposed system leverages artificial intelligence for the early detection of dengue using complete blood count (CBC) data. Feature selection techniques such as Pearson Correlation, Recursive Feature Elimination (RFE) with Random Forest, Chi-Square SelectKBest, (Chi2), and ExtraTree are employed to extract the most relevant attributes. The system integrates various machine learning algorithms [15], including Logistic Regression, Support Vector Machine (SVM), Naive Bayes, Random Forest, AdaBoost, XGBoost, Multi-Layer Perceptron (MLP), and LightGBM, along with ensemble methods [17] like a Stacking Classifier (XGB + LR + MLP with LightGBM) and a Voting Classifier (Boosted Decision Tree + ExtraTree). Advanced deep learning models [19] such as Artificial Neural

Networks (ANN), Convolutional Neural Networks (CNN), Gated Recurrent Units (GRU), Bidirectional Long Short-Term Memory (Bi-LSTM), Feedforward Neural Networks (FNN), Transformer, and hybrid models like CNN + LSTM are utilized for improved performance. The system proposes a robust and accurate predictive model by combining individual model predictions using ensemble techniques. A Flask-based interface with user authentication ensures a user-friendly and secure environment for interaction and testing, aiming to achieve efficient and reliable early dengue detection.



Fig.1 Proposed Architecture

The proposed system architecture for dengue detection begins with the collection of dengue-related data, which undergoes preprocessing and exploratory data analysis (EDA) to understand patterns and distributions. Feature engineering is then applied to enhance the relevant attributes. The data is split into training and testing sets to build and train the



International Journal For Advanced Research

In Science & Technology

A peer reviewed international journal ISSN: 2457-0362 www.ijarst.in

predictive model. Once trained, the model is evaluated using performance metrics to ensure its effectiveness in accurately predicting dengue cases.

i) Dataset Collection:

The dengue dataset consists of 301 entries and 15 features, capturing clinical and demographic information of patients. Key attributes include gender, age, and date of examination, along with complete blood (CBC) parameters like count [1] WBC, neutrophil, haemoglobin, eosinophil, lymphocyte, monocyte, basophil, RBC, and platelets. The dataset also includes ESR values, with some missing entries, and the final diagnostic result indicating dengue presence. The dataset provides comprehensive а foundation for analyzing and predicting dengue through machine learning techniques.

	Serial	Date	Gender	Age	Haemoglobin	ESR	WBC	Neutrophil	Lymphocyte
0	A2308164543	2/8/2023	Female	40.0	11.2	32.0	11.30	65	23.0
1	A2308164502	2/8/2023	Male	13.0	11.6	59.0	7.80	49	42.0
2	A2308164673	3/8/2023	Male	23.0	15.1	NaN	3.85	65	25.0
3	A2308164685	3/8/2023	Male	58.0	8.5	NaN	10.30	85	6.0
4	A2308164626	2/8/2023	Female	35.0	12.0	28.0	6.70	54	20.0

Fig.2 Dataset Collection Table - Dengue

ii) Pre-Processing:

In preprocessing, we discuss data cleaning, handling missing values, performing exploratory data analysis (EDA), extracting relevant features, applying SMOTE [16] sampling to address class imbalance, and selecting important features for model training.

a) Data Processing: The first step in data processing involves removing duplicate entries to ensure that the dataset contains unique records. After that, drop cleaning is performed, where columns with excessive missing values or irrelevant features are removed to reduce noise. Label encoding is then applied to categorical variables, such as 'Gender' and 'Result', transforming them into numerical representations. This step prepares the data for analysis and machine learning modeling, ensuring consistency and accuracy in the dataset.

b) Exploratory Data Analysis (EDA) of **Data:** Exploratory Data Analysis (EDA) involves summarizing the main characteristics of the dataset and distributions. visualizing its Using techniques like histograms, box plots, and correlation matrices, we gain insights into the data's structure and relationships between features. This stage helps identify trends, patterns, and anomalies, as well as understand the distribution of key variables such as age, haemoglobin levels, and platelet count. Visualizations aid in



A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

understanding data patterns and inform preprocessing decisions.

c) Outliers Detection using Z-Score: Outlier detection is an essential step in data preprocessing, as extreme values can distort the results of machine learning models. The Z-score method is used to identify outliers by calculating the number of standard deviations each data point is away from the mean. Data points with a Zscore greater than 3 or less than -3 are considered outliers. These outliers are either removed or adjusted based on the specific characteristics of the dataset, ensuring data quality and model accuracy.

d) Feature Extraction and SMOTE Sampling: Feature extraction begins by selecting the independent variables (X) and the target variable (y) from the dataset. In this case, X consists of the relevant features such as age, haemoglobin, platelets, and others, while y is the dengue diagnosis result (positive/negative). To address class imbalance. **SMOTE** (Synthetic Minority Over-sampling Technique) is applied. SMOTE generates synthetic samples for the minority class by interpolating between existing instances, helping to balance the dataset and improve model performance on underrepresented classes.

e) Feature selection: Feature selection is performed using five techniques to identify the most relevant features for model training. Pearson Correlation is used to remove highly correlated features, ensuring independence. RFE (Recursive Feature Elimination) eliminates less important features iteratively based on model performance. SelectKBest selects the top features based on statistical tests. Chi2 is applied for categorical data to identify features that are most dependent on the target. ExtraTree is used for feature ranking, selecting those with the highest importance. These techniques help improve model accuracy and reduce overfitting.

iii) Training & Testing:

Training and testing the dataset involves dividing the data into two subsets: one for training the model and the other for testing its performance. The training set is used to fit the model and learn the patterns in the data, while the test set serves to evaluate the model's ability to generalize to unseen data. This process ensures that the model is not overfitting and is capable of making accurate predictions on new, previously unseen instances.

iv) Algorithms:



International Journal For Advanced Research

In Science & Technology

A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

Logistic Regression: Used for binary classification to predict the likelihood of dengue based on CBC data, offering interpretability and quick implementation.

SVM (Support Vector Machine): Provides robust dengue prediction by finding the optimal hyperplane that separates data into infected and noninfected categories.

Naive Bayes: A probabilistic model that leverages conditional independence for fast and effective dengue classification, particularly with smaller datasets.

Random Forest: An ensemble method that improves accuracy by building multiple decision trees, ideal for handling complex relationships in the CBC data.

AdaBoost: Boosts the performance of weak learners by iteratively adjusting model weights, improving dengue detection accuracy.

XGBoost: An efficient gradient boosting algorithm that provides highly accurate dengue predictions by minimizing errors in previous models.

MLP (Multi-Layer Perceptron): A deep learning model that learns complex, nonlinear relationships in CBC data for dengue detection. **LightGBM:** A gradient boosting model that provides fast, scalable dengue predictions, especially for large datasets with complex features.

Stacking Classifier: Combines MLP, XGBoost, and Logistic Regression to boost accuracy by leveraging diverse prediction strengths.

Voting Classifier: Aggregates predictions from Boosted Decision Trees and ExtraTree for a more reliable and accurate dengue detection outcome.

ANN (Artificial Neural Network): A deep learning approach that mimics human brain functioning, effectively capturing complex patterns in the CBC data for dengue prediction.

CNN (Convolutional Neural Network): Extracts spatial hierarchies from CBC data for precise dengue prediction, particularly effective in feature extraction.

GRU (Gated Recurrent Unit): A recurrent neural network variant, GRU captures sequential dependencies in CBC data, making it useful for time-series-based dengue predictions.

Bi-LSTM (Bidirectional Long Short-Term Memory): Captures both past and future data sequences to enhance dengue



A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

prediction by understanding sequential patterns in the CBC data.

FNN (Feedforward Neural Network): Processes CBC data in one direction for rapid prediction without relying on past information, suitable for straightforward classification tasks.

Transformer:UtilizesattentionmechanismsforefficientlyhandlingsequentialCBCdata,leadingtoaccuratedenguepredictions.

CNN + LSTM: Combines CNN for feature extraction and LSTM for sequential data processing, improving accuracy in predicting dengue by capturing both spatial and temporal patterns.

4. RESULTS & DISCUSSION

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy =
$$\frac{TP + TN}{TP + FP + TN + FN}$$
(1)

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives.

Thus, the formula to calculate the precision is given by:

Precision

 $=\frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}(2)$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}(3)$$

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

F1 Score = $2 * \frac{\text{Recall X Precision}}{\text{Recall} + \text{Precision}} * 100(4)$

AUC-ROC Curve: The AUC-ROC Curve is a performance measurement for classification problems at various



threshold settings. ROC plots the True Positive Rate against the False Positive Rate. AUC quantifies the overall ability of the model to distinguish between classes, where a higher AUC indicates better model performance.

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} (5)$$

In Table 1, 2, 3 & 4 the Voting Classifier (Boosted DT + ExtraTree) achieved the highest accuracy and performance across all sampling techniques— Pearson Correlation, Recursive Feature Elimination (RFE) with Random Forest, SelectKBest, Chi-Square (Chi2), and ExtraTree. It consistently outperformed other algorithms on all metrics, including accuracy, precision, recall, and F1 score.

ML Model	Accuracy	Precision	Recall	F1-Score	AUC ROC
Logistic	0.882	0.893	0.882	0.880	0.963
Regression					
SVM	0.838	0.864	0.838	0.835	0.958
NaiveBayes	0.882	0.893	0.882	0.880	0.947
RandomForest	0.853	0.887	0.853	0.850	1.000
AdaBoost	0.853	0.873	0.853	0.850	0.994
XGBoost	0.868	0.895	0.868	0.865	1.000
MLP	0.824	0.831	0.824	0.821	0.972
LightGBM	0.838	0.864	0.838	0.835	0.976
Proposed	0.794	0.861	0.794	0.792	0.980
Extension	1.000	1.000	1.000	1.000	0.959
FNN	0.868	0.883	0.868	0.865	0.959
ANN	0.368	1.000	0.368	0.538	0.959
CNN	0.794	0.840	0.794	0.791	1.000
GRU	0.632	1.000	0.632	0.775	1.000
Bi-LSTM	0.368	1.000	0.368	0.538	1.000
CNN + LSTM	0.853	0.873	0.853	0.850	1.000
Transformer	0.368	1.000	0.368	0.538	1.000

Table.1 Performance Evaluation Metrics – Chi 2 FS

Graph.1 Comparison Graphs – Chi 2 FS





ML Model	Accuracy	Precision	Recall	F1-Score	AUC
					ROC
Logistic Regression	0.882	0.893	0.882	0.880	0.967
SVM	0.838	0.864	0.838	0.835	0.961
NaiveBayes	0.882	0.893	0.882	0.880	0.931
RandomForest	0.868	0.895	0.868	0.865	1.000
AdaBoost	0.912	0.924	0.912	0.910	0.999
XGBoost	0.882	0.904	0.882	0.880	1.000
MLP	0.912	0.916	0.912	0.911	0.992
LightGBM	0.838	0.864	0.838	0.835	0.974
Proposed	0.853	0.873	0.853	0.850	0.989
Extension	1.000	1.000	1.000	1.000	0.955
FNN	0.897	0.897	0.897	0.897	0.955
ANN	0.368	1.000	0.368	0.538	0.955
CNN	0.632	0.773	0.632	0.637	1.000
GRU	0.632	1.000	0.632	0.775	1.000
Bi-LSTM	0.368	1.000	0.368	0.538	1.000
CNN + LSTM	0.809	0.847	0.809	0.806	1.000
Transformer	0.368	1.000	0.368	0.538	1.000

Graph.2 Comparison Graphs – ET FS





Table.3 Performance Evaluation Metrics – Pearson Correlation

ML Model	Accuracy	Precision	Recall	F1-Score	AUC
					ROC
Logistic	0.853	0.873	0.853	0.850	0.969
Regression					
SVM	0.824	0.855	0.824	0.820	0.959
NaiveBayes	0.882	0.886	0.882	0.881	0.934
RandomForest	0.853	0.887	0.853	0.850	1.000
AdaBoost	0.838	0.852	0.838	0.835	0.997
XGBoost	0.882	0.904	0.882	0.880	1.000
MLP	0.882	0.886	0.882	0.881	0.985
LightGBM	0.838	0.864	0.838	0.835	0.975
Proposed	0.853	0.873	0.853	0.850	0.997
Extension	1.000	1.000	1.000	1.000	0.948
FNN	0.868	0.869	0.868	0.868	0.948
ANN	0.368	1.000	0.368	0.538	0.948
CNN	0.838	0.864	0.838	0.835	1.000
GRU	0.368	1.000	0.368	0.538	1.000
Bi-LSTM	0.368	1.000	0.368	0.538	1.000
CNN + LSTM	0.853	0.855	0.853	0.851	1.000
Transformer	0.368	1.000	0.368	0.538	1.000

Graph.3 Comparison Graphs - Pearson Correlation





Table.4 Performance Evaluation Metrics – RFE FS

ML Model	Accuracy	Precision	Recall	F1-Score	AUC
	_				ROC
Logistic Regression	0.853	0.862	0.853	0.851	0.968
SVM	0.838	0.864	0.838	0.835	0.957
NaiveBayes	0.853	0.855	0.853	0.851	0.939
RandomForest	0.882	0.904	0.882	0.880	1.000
AdaBoost	0.809	0.812	0.809	0.806	0.994
XGBoost	0.853	0.887	0.853	0.850	1.000
MLP	0.838	0.852	0.838	0.835	0.979
LightGBM	0.824	0.855	0.824	0.820	0.974
Proposed	0.838	0.880	0.838	0.836	0.995
Extension	1.000	1.000	1.000	1.000	0.950
FNN	0.838	0.864	0.838	0.835	0.950
ANN	0.368	1.000	0.368	0.538	0.950
CNN	0.647	0.697	0.647	0.643	1.000
GRU	0.632	1.000	0.632	0.775	1.000
Bi-LSTM	0.368	1.000	0.368	0.538	1.000
CNN + LSTM	0.853	0.862	0.853	0.851	1.000
Transformer	0.368	1.000	0.368	0.538	1.000

Graph.4 Comparison Graphs – RFE FS





Table.5 Performance Evaluation Metrics - SelectkBest

ML Model	Accuracy	Precision	Recall	F1-Score	AUC
					ROC
Logistic Regression	0.868	0.883	0.868	0.865	0.966
SVM	0.824	0.855	0.824	0.820	0.959
NaiveBayes	0.868	0.883	0.868	0.865	0.935
RandomForest	0.882	0.904	0.882	0.880	1.000
AdaBoost	0.882	0.886	0.882	0.881	0.997
XGBoost	0.882	0.904	0.882	0.880	1.000
MLP	0.838	0.880	0.838	0.836	0.987
LightGBM	0.838	0.864	0.838	0.835	0.975
Proposed	0.809	0.847	0.809	0.806	0.991
Extension	1.000	1.000	1.000	1.000	0.960
FNN	0.838	0.864	0.838	0.835	0.960
ANN	0.368	1.000	0.368	0.538	0.960
CNN	0.691	0.716	0.691	0.700	1.000
GRU	0.368	1.000	0.368	0.538	1.000
Bi-LSTM	0.632	1.000	0.632	0.775	1.000
CNN + LSTM	0.838	0.864	0.838	0.835	1.000
Transformer	0.368	1.000	0.368	0.538	1.000

Graph.5 Comparison Graphs – SelectkBest



In Graphs (1,2,3,4, & 5) accuracy is represented in light blue, precision in orange, recall in gray, F1-score in yellow, AUC ROC in blue. The Voting Classifier outperforms the other algorithms in all metrics, with the highest values compared to the remaining models. These details are visually represented in the above graph.



Fig. 3 Dash Board

The image shows a web page titled "Artificial Intelligence-Based Early Detection of Dengue Using CBC Data." It appears to be a dashboard or interface for a project related to using artificial intelligence to detect dengue fever based on complete blood count (CBC) data.

Step 6	
Register	
Username	_
Fullname	_
Email	_
Phone Number	_
Password	_
Forgot Password?	
Register)
Already member! Signin	

Fig. 4 Register page



The Fig. 4 shows a user registration form. It requires a username, full name, email, phone number, and password. It also includes a "Forgot Password?" link and a "Register" button.

	Step 7
	Login
Userno admir	ame
Passw	ord
Forgot	Password?
	Login
	Not a member? Signup

Fig. 5 Login page

The Fig. 5 shows a login page with fields for username and password. The username field is pre-filled with "admin." A "Forgot password?" link is provided below the fields. There is also a button labeled "Login" and a link to "Signup" for new users.



Fig.6 Main page

The Fig.6 shows a web dashboard with the title "Artificial Intelligence-Based Early Detection Of Dengue Using CBC Data." It has a navigation bar with options like "Prediction" and "Graph," and a dropdown menu likely for selecting data analysis methods.

	lest case i
FORM	LYMPHOCYTE:
GENDER:	41.50170764
0	MONOCYTE
	3.561859634
HAEMOGLOBIN:	
12.043473	RBC:
	4.488127135
ESR:	
48.62201163	PLATELETS:
	268
WBC :	
7.726151811	
NEUTROPHIL:	
51	Predict
OUTCOME	
COTOOME	
NEGATIV	E, PATIENT IS NOT

Fig. 7 Test case -1

The Fig. 7 shows a form for predicting dengue fever. It collects patient data like gender, haemoglobin, WBC count, etc. After inputting data, the form predicts the risk of dengue fever. In this case, the prediction is "NEGATIVE, PATIENT IS

Volume 13, Issue 02, Feb 2023

International Journ In Science IJARST	nal For Advanced Research ice & Technology www.ijarst.in SSN: 2457-0362			
NOT SUFFERING FROM DENGUE DISEASE!"	Step 10 Prediction ~ Graph ~ Notebook ~ Logout			
Step 9 Test case 2	Pearson Coff Chi2 RFE FS SelectKBest ExtraTree FS Dece-Based			
ESR: 64 WBC : PLATELETS:	Fig. 9 Home page			
26.9 15 NEUTROPHIL: 88 LYMPHOCYTE: 8 Bradict	The Fig. 9 shows the home page of a web application for dengue fever detection using AI. It has a navigation bar with			
OUTCOME POSITIVE, PATIENT IS DETECTED OF DENGUE DISEASE!	options like "Prediction" and "Graph," and a dropdown menu for selecting data analysis methods. The background image			

Fig.8 Test case - 2

The Fig.8 shows a form for predicting dengue fever. It collects patient data like gender, haemoglobin, WBC count, etc. After inputting data, the form predicts the risk of dengue fever. In this case, the prediction is "POSITIVE, PATIENT IS DETECTED OF DENGUE DISEASE!"

ESR: 48.62201163 NEUTROPHIL: 51 I YMPHOCYTE: 41.50170764 MONOCYTE : OUTCOME 3.561859634 NEGATIVE, PATIENT IS NOT SUFFERING FROM DENGUE DISEASE! PLATELETS: 268

Fig. 10 Test case - 1

includes medical tools



The Fig. 10 shows a test case scenario for a dengue prediction model. It displays input values for ESR, Neutrophil, Lymphocyte, Monocyte, and Platelets. The model predicts a "NEGATIVE" outcome, indicating the patient is not suffering from dengue disease.

	Step II Test case 2
ESR:	
64	
NEUTROPHIL:	
88	
LYMPHOCYTE:	
8	
MONOCYTE :	OUTCOME
3	
PLATELETS:	POSITIVE, PATIENT IS DETECTE
15	OF DENGUE DISEASE!

Fig. 11 Test case – 2

The Fig. 11 shows a test case scenario for a dengue prediction model. It displays input values for ESR, Neutrophil, Lymphocyte, Monocyte, and Platelets. The model predicts a "POSITIVE" outcome, indicating the patient is detected with dengue disease.



Fig. 12 Home page

The Fig. 12 shows the home page of a web application for dengue fever detection using AI. The user is selecting the "RFE FS" option from a dropdown menu under the "Prediction" tab.

	Step 13 Test case 1
FORM	
GENDER:	
0	EOSINOPHIL:
	3.812713455
HAEMOGLOBIN:	
12.24948771	RBC:
	4.576151811
WBC:	
8.063267591	PLATELETS:
	275
NEUTROPHIL :	
52	
MONOCYTE:	
3.187286545	Predict
OUTCOME	
NEGATIVE, I SUFFERING DISEASE!	PATIENT IS NOT FROM DENGUE

Fig. 13 Test case - 1



The Fig. 13 shows a test case scenario for a dengue prediction model. It displays input values for gender, haemoglobin, WBC count, neutrophil, and other blood parameters. The model predicts a "NEGATIVE" outcome, indicating the patient is not suffering from dengue disease.

GENDER:	
0	EOSINOPHIL:
	1
HAEMOGLOBIN:	
8.6	RBC:
	2.78
WBC:	
26.9	PLATELETS:
	15
NEUTROPHIL :	
88	
MONOCYTE:	
3	
	Predict
011700145	
OUTCOME	

Fig.14 Test case - 2

The Fig.14 shows a test case scenario for a dengue prediction model. It displays input values for gender, haemoglobin, WBC count, neutrophil, and other blood model predicts parameters. The а "POSITIVE" outcome, indicating the patient is detected with dengue disease.



Fig. 15 Home page

The Fig. 15 shows the home page of a web application for dengue fever detection using AI. The user is selecting the "SelectKBest" option from a dropdown menu under the "Prediction" tab.

	Step 15
FORM	Test case I
HAEMOGLOBIN:	
8.6	MONOCYTE
	3
ESR:	
64	RBC:
	2.79
WBC:	
26.9	PLATELETS:
	15
NEUTROPHIL :	
88	
LYMPHOCYTE:	
8	Predict
OUTCOME	
OUTCOME	
POSITIVE I	PATIENT IS DETECTED
OF DENGUE	DISFASEI

Fig.16 Test case – 1



The Fig.16 shows a test case scenario for a dengue prediction model. It displays input values for haemoglobin, ESR, WBC, neutrophil, lymphocyte, monocyte, and platelets. The model predicts a "POSITIVE" outcome, indicating the patient is detected with dengue disease.

	Step 15
FORM	lest case 2
HAEMOGLOBIN:	MONOCYTE:
14.2	4
ESR:	RBC:
6	4.95
WBC:	PLATELETS:
9.9	280
NEUTROPHIL :	
66	
LYMPHOCYTE:	
27	Predict
OUTCOME	
NEGATIVE	, PATIENT IS NOT G FROM DENGUE
DISEASE!	

Fig. 17 Test case – 2

The image shows a test case scenario for a dengue prediction model. It displays input values for haemoglobin, ESR, WBC, neutrophil, lymphocyte, monocyte, and platelets. The model predicts a "NEGATIVE" outcome, indicating the patient is not suffering from dengue disease.



Fig. 18 Home page

The Fig. 18 shows the home page of a web application for dengue fever detection using AI. The user is selecting the "ExtratreeFs" option from a dropdown menu under the "Prediction" tab.

ORM	LYMPHOCYTE:
EOSINOPHIL:	
1	ESR:
	64
RBC:	
2.78	MONOCYTE:
	3
AGE:	
22	NEUTROPHIL:
	88
WBC :	
26.9	PLATELETS:
	15
HAEMOGLOBIN:	
8.6	



A peer reviewed international journal ISSN: 2457-0362

Fig.19 Test case – 1

The Fig.19 shows a test case scenario for a dengue prediction model. It displays input values for age, haemoglobin, WBC count, neutrophil, lymphocyte, monocyte, and platelets. The model predicts a "POSITIVE" outcome, indicating the patient is detected with dengue disease.

Step 17 Test case 2	
FORM	LYMPHOCYTE:
	41.50170764
EOSINOPHIL:	
3.625426911	ESR:
	48.62201163
RBC:	
4.488127135	MONOCYTE:
	3.561859634
AGE:	
45	NEUTROPHIL:
	51
WBC :	
7.726151811	PLATELETS:
	268
HAEMOGLOBIN:	
12.04347251	



Fig. 20 Test case - 2

The Fig. 20 shows a test case scenario for a dengue prediction model. It displays input values for age, haemoglobin, WBC count, neutrophil, lymphocyte, monocyte, and platelets. The model predicts a "NEGATIVE" outcome, indicating the patient is not suffering from dengue disease.

www.ijarst.in

5. CONCLUSION

conclusion, the proposed In system demonstrates a highly effective approach for the early detection of dengue [18] using CBC data. By applying a range of feature selection techniques and leveraging various machine learning and deep learning models, the system successfully identifies critical patterns for accurate prediction. Among all the algorithms Voting Classifier [18], tested, the combining a Boosted Decision Tree and achieved ExtraTree, outstanding 100% performance with accuracy, outperforming other models in both prediction accuracy and reliability. This result highlights the potential of ensemble methods to enhance the robustness and precision of predictive systems. The integration of a user-friendly interface developed using Flask ensures accessibility for real-time interaction and testing, with built-in authentication for secure usage. The system's exceptional accuracy and reliability make it a promising tool for early dengue detection, offering a valuable resource for healthcare



International Journal For Advanced Research

In Science & Technology A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

professionals and researchers. By effectively combining state-of-the-art algorithms and techniques, this approach contributes to the growing body of AIdriven solutions in the medical field, providing timely insights to improve patient outcomes and streamline healthcare decision-making.

The future scope of this system includes expanding its applicability to detect other infectious diseases using various clinical datasets. Further enhancement can be incorporating additional achieved by feature engineering techniques and exploring advanced ensemble methods to improve prediction accuracy. Integration with real-time healthcare monitoring systems and mobile applications can facilitate immediate diagnosis, enabling quicker intervention and better patient management. Additionally, incorporating explainable AI techniques can enhance trust and transparency in medical predictions.

REFERENCES

[1] M. B. Khan, Z.-S. Yang, C.-Y. Lin, M.-C. Hsu, A. N. Urbina, W. Assavalapsakul, W.-H. Wang, Y.-H. Chen, and S.-F. Wang, "Dengue overview:Anupdatedsystemicreview,"J.Inf ectionPublicHealth,vol.16, no. 10, pp. 1625–1642, Oct. 2023.

[2] N. Ali, "The recent burden of dengue infection in bangladesh: A serious public health issue," J. Infection Public Health, vol. 17, no. 2, pp. 226–228, Feb. 2024.

[3] S. Roy, A. Biswas, M. T. A. Shawon, S. Akter, and M. M. Rahman, "Land use and meteorological influences on dengue transmission dynamics in Dhaka city, Bangladesh," Bull. Nat. Res. Centre, vol. 48, no. 1, pp. 1–24, Mar. 2024, doi: 10.1186/S42269-024-01188-0.

[4] N. Sharif, N. Sharif, A. Khan, and S.
K. Dey, "The epidemiologic and clinical characteristics of the 2023 dengue outbreak in Bangladesh," Open Forum Infectious Diseases, vol. 11, no. 2, pp. 1–29, Feb. 2024, doi: 10.1093/OFID/OFAE066.

[5] M. E. H. Kayesh, I. Khalil, M. Kohara, and K. Tsukiyama-Kohara, "Increasing dengue burden and severe dengue risk in Bangladesh: An overview," Tropical Med. Infectious Disease, vol. 8, no. 1, p. 32, Jan. 2023.

[6] D. C. Kajeguka, F. M. Mponela, E. Mkumbo, A. N. Kaaya, D. Lasway, R. D.

Volume 13, Issue 02, Feb 2023



A peer reviewed international journal ISSN: 2457-0362

Kaaya, M. Alifrangis, E. Elanga-Ndille, B. T. Mmbaga, and R. Kavishe, "Prevalence and associated factors of dengue virus circulation in the rural community, Handeni district in Tanga, Tanzania," J. Tropical Med., vol. 2023, pp. 1–9, Nov. 2023.

[7] M. A. Kabir, H. Zilouchian, M. A. Younas, and W. Asghar, "Dengue detection: Advances in diagnostic tools from conventional technology to point of care," Biosensors, vol. 11, no. 7, p. 206, Jun. 2021.

[8] C. Davi, A. Pastor, T. Oliveira, F. B. d.
L. Neto, U. Braga-Neto, A. W. Bigham,
M. Bamshad, E. T. A. Marques, and B.
Acioli-Santos, "Severe dengue prognosis using human genome data and machine learning," IEEE Trans. Biomed. Eng., vol. 66, no. 10, pp. 2861–2868, Oct. 2019.

[9] D. Sarma, S. Hossain, T. Mittra, Md.
A. M. Bhuiya, I. Saha, and R. Chakma,
"Dengue prediction using machine learning algorithms," in Proc. IEEE 8th
R10 Humanitarian Technol. Conf., Dec.
2020, pp. 1–6.

[10] E. Fernández, M. Smieja, S. D.Walter, and M. Loeb, "A predictive model to differentiate dengue from other febrile illness," BMC Infectious Diseases, vol. 16, no. 1, pp. 1–7, Dec. 2016.

www.ijarst.in

[11] H. Mayrose, G. M. Bairy, N. Sampathila, S. Belurkar, and K. Saravu, "Machine learning-based detection of dengue from blood smear images utilizing platelet and lymphocyte characteristics," Diagnostics, vol. 13, no. 2, p. 220, Jan. 2023.

[12] S. Sabrina Prome, T. Basak, T. Islam Plabon, and R. Khan, "Prediction of dengue cases in Bangladesh using explainable machine learning approach," in Proc. Int. Conf. Inventive Comput. Technol. (ICICT), Apr. 2024, pp. 1–5.

[13] J. D. Mello-Román, J. C. Mello-Román, S. Gómez-Guerrero, and M. García-Torres, "Predictive models for the medical diagnosis of dengue: A case study in Paraguay," Comput. Math. Methods Med., vol. 2019, pp. 1–7, Jul. 2019.

[14] S. K. Dey, M. M. Rahman, A. Howlader, U. R. Siddiqi, K. M. M. Uddin, R. Borhan, and E. U. Rahman, "Prediction of dengue incidents using hospitalized patients, metrological and socio-economic data in Bangladesh: A machine learning approach," PLoS One, vol. 17, no. 7, Jul. 2022, Art. no. e0270933.



> A peer reviewed international journal ISSN: 2457-0362

www.ijarst.in

[15] B. Abdualgalil, S. Abraham, and W.
M. Ismael, "Early diagnosis for dengue disease prediction using efficient machine learning techniques based onclinical data," J. Robot. Control (JRC), vol. 3, no. 3, pp. 257–268, May 2022.

[16] S. Q. Ong, P. Isawasan, A. M. M. Ngesom, H. Shahar, A. M. Lasim, and G. Nair, "Predicting dengue transmission rates by comparing different machine learning models with vector indices and meteorological data," Sci. Rep., vol. 13, no. 1, pp. 1–10, Nov. 2023.

[17] J. K. Chaw, S. H. Chaw, C. H. Quah, S. Sahrani, M. C. Ang, Y. Zhao, and T. T. Ting, "A predictive analytics model using machine learning algorithms to estimate the risk of shock development among dengue patients," Healthcare Anal., vol. 5, pp. 1–17, Jul. 2024.

[18] M. T. Sarwar and M. A. Mamun, "Prediction of dengue using machine learning algorithms: Case study Dhaka," in Proc. 4th Int. Conf. Electr., Comput. Telecommun. Eng. (ICECTE), Dec. 2022, pp. 1–6.

[19] T. Akter, M. T. Islam, M. F. Hossain, and M. S. Ullah, "A comparative study between time series and machine learning technique to predict dengue fever in Dhaka city," Discrete Dyn. Nature Soc., vol. 2024, pp. 1–12, May 2024.

[20] M. A. Majeed, H. Z. M. Shafri, Z. Zulkafli, and A. Wayayok, "A deep learning approach for dengue fever prediction in Malaysia using LSTM with spatial attention," Int. J. Environ. Res. Public Health, vol. 20, no. 5, p. 4130, Feb. 2023.

Volume 13, Issue 02, Feb 2023