**Jurnal Teknik Informatika (JUTIF)** Vol. 6, No. 6, December, 2025, Page. 5944-5955
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.6.5309

# Comparative Analysis Of Machine Learning Algorithms For Dengue Fever Prediction Based On Clinical And Laboratory Features

**Sriyanto*[1], RZ Abdul Aziz[2], Dewi Agushinta Rahayu[3], Zuriati[4], Mohd Faizal Abdollah[5], Irianto[6]**

[1,2,3]Department of Informatics, Institute Informatics and Business Darmajaya, Indonesia
[4]Department of Information Technology, Universitas Gunadama, Indonesia
[5]Department of Internet Engineering Technology, Politeknik Negeri Lampung, Indonesia
[6]Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka, Malaysia

Email: [1]sriyanto@darmajaya.ac.id

Received : Sep 21, 2025; Revised : Sep 28, 2025; Accepted : Sep 29, 2025; Published : Dec 31, 2025

## Abstract

Dengue fever (DF) remains a global health problem requiring accurate early detection to prevent severe complications. This study applies machine learning (ML) algorithms to clinical and laboratory data for improving diagnostic accuracy. Six classifiers were compared: Decision Tree (DT), K-Nearest Neighbor (KNN), Logistic Regression (LR), Naïve Bayes (NB), Neural Network (NN), and Support Vector Machine (SVM). The dataset consists of 1,003 patient records with nine feature columns, of which 989 were used after preprocessing. Class distribution was imbalanced, with 67.6% positive and 32.4% negative cases. Model performance was evaluated using 10-fold cross-validation based on accuracy, precision, recall, F1-score, confusion matrix, and ROC curve analysis. The results indicate that DT achieved the highest performance with 99.4% accuracy, 99.4% precision, 99.7% recall, and 99.6% F1-score, slightly outperforming NN. KNN, LR, and SVM produced comparable results, while NB showed substantially lower accuracy (44.3%) and limited discriminatory power. ROC analysis confirmed these findings, with DT, NN, SVM, and LR achieving AUC values between 0.992 and 0.999, whereas NB performed poorly. These findings highlight the strong potential of ML algorithms, particularly DT, to support medical decision systems, strengthen informatics-based decision support applications, and enhance the accuracy and speed of dengue diagnosis in clinical practice.

***Keywords :*** *Classification Algorithms, Clinical Data, Confusion Matrix, Dengue Fever, Machine Learning, ROC Curve*

## 1. INTRODUCTION

Machine Learning (ML) has become a central approach in medical data analysis due to its ability to uncover hidden patterns in complex datasets and generate accurate predictive models. Recent studies confirm that ML and deep learning (DL) techniques play a vital role in healthcare, supporting disease diagnosis, clinical decision-making, and automation of medical data processing [1],[2]. Applications of ML are evident in infectious diseases, where algorithms have been successfully applied to assist diagnosis, predict disease progression, and classify severity with high accuracy [3].

Dengue fever (DF) remains a major public health challenge in tropical countries, including Bangladesh [4], with cases increasing annually. Early prediction and classification of DF are crucial to prevent severe complications and improve clinical outcomes. Previous studies demonstrated the potential of ML models such as Artificial Neural Networks (ANN), Logistic Regression (LR), Decision Tree (DT), and Multilayer Perceptrons in classifying DF severity with promising results [5],[6], [7].

ML enables the integration of diverse clinical, laboratory, and epidemiological parameters, providing robust prediction models [8]. Beyond clinical data, environmental factors such as weather and population density also enhance predictive accuracy [9]. Despite this progress, selecting the most appropriate algorithm remains a challenge.

No single ML algorithm consistently achieves superior performance across different healthcare datasets. DTs are interpretable but prone to overfitting with small samples, NB is simple and fast but limited by the assumption of feature independence, SVM demonstrates strong performance with high-dimensional data but requires complex parameter tuning, and KNN is intuitive yet computationally expensive for large datasets [10], [11], [12], [13]. This variation underlines the necessity of comparative evaluations to identify the most effective approach for specific datasets and application contexts, particularly in clinical datasets with potential class imbalance, which is a common challenge in medical diagnosis.

Several recent works have compared ML algorithms for DF prediction, reporting diverse findings. Random Forest and SVM are frequently recommended [8], [14], hybrid models such as LSTM–ARIMA or combined learning strategies have shown improved early detection performance [15], and clinical markers such as pulse pressure and AST/platelet ratio are strong predictors of severe dengue [16]. Other studies in Colombia [17] and Latin America [18] also highlighted the advantages of hybrid and integrated approaches, while systematic reviews emphasize that algorithm performance depends heavily on data characteristics and application domains [19].

Recent studies have advanced dengue prediction through hybrid models that integrate meteorological and clinical data [20], systematic reviews of ML and DL methods [21], and regional applications such as case screening with ML in Latin America [7]. Robust multi-model frameworks validated in Brazil and Peru [22], ML-based severity prediction in Puerto Rico [23]. and the use of advanced neural architectures with digital epidemiology tools like Google Trends [24], [25], [26], further demonstrate the diverse approaches enhancing dengue surveillance and forecasting.

Despite these global advances, limited studies have specifically examined clinical and laboratory datasets from South Asia, particularly Bangladesh. This research gap underscores the need for localized comparative analyses. Building on this background, the present study aims to evaluate widely used ML classifiers—DT, KNN, LR, NB, NN, and SVM—for dengue fever prediction using Bangladeshi clinical and laboratory data, in order to determine which algorithm provides the best predictive performance. The evaluation focuses on accuracy, precision, recall, and F1-score, offering both theoretical insights into ML performance for DF and practical guidance for healthcare practitioners and policymakers in early detection and disease control [4].

## 2. METHOD

This study is an applied research in the field of machine learning that aims to compare the performance of six classification algorithms in predicting dengue fever (DF) based on patients' clinical and laboratory data. The research methodology consists of several stages, namely:

### 2.1. Dataset Collection

The dataset used in this study was obtained from Mendeley Data, titled Predictive Clinical Dataset for Dengue Fever Using Vital Signs and Blood Parameters. This publicly accessible dataset was originally collected from patients treated at the Upazila Health Complex in Kalai, Joypurhat, Bangladesh. The raw dataset consists of 1,003 patient records, each comprising nine features representing vital signs and blood parameters. Class distribution in the dataset is moderately imbalanced, with 669 positive cases (67.6%) and 320 negative cases (32.4%). Table 1 provides a detailed description of the attributes, including their data types and example values, which serve as the foundation for the dengue prediction modeling in this study.

Table 1. Dengue Dataset Description

| No | Attribute | Data Type | Example Value |
|---|---|---|---|
| **1** | Age | int64 | 43 |
| **2** | Gender | object | Male, Female |
| **3** | Hemoglobin | float64 | 12.6 |
| **4** | White Blood Cell Count (WBC) | float64 | 2200 |
| **5** | Differential Count | int64 | 1, 0 |
| **6** | Red Blood Cell Panels (RBC Panels) | int64 | 1, 0 |
| **7** | Platelet Count | float64 | 17000 |
| **8** | Platelet Distribution Width (PDW) | float64 | 11.0 |
| **9** | Final Result(Label) | int64 | 1=Positive, 0=Negative |

### 2.2. Data Preprocessing

Data preprocessing consisted of three main steps. First, data cleaning was performed by removing entries without target labels and handling missing values using median imputation, leaving 989 valid samples from the original 1,003 records. Second, categorical variables such as Gender were encoded using Label Encoding. Third, numerical features were standardized with StandardScaler to ensure comparable scales and avoid bias during model training. These steps prepared the dataset to be clean, consistent, and ready for classification.

### 2.3. Classification Algorithm Modeling

In this study, six ML algorithms were employed to build dengue fever (DF) classification models, namely: DT, KNN, LR, NB, NN, and SVM.

DT was selected for its ability to generate interpretable models through decision tree structures (implemented with default parameters without limiting max_depth). KNN was used as a simple instance-based algorithm (with k=5 neighbors and Euclidean distance metric). LR was considered due to its strength in modeling the relationship between predictor variables and binary target variables (L2 regularization, solver=lbfgs, and maximum iteration of 1000). NB was chosen for its lightweight and efficient nature, particularly when dealing with datasets containing a large number of features (GaussianNB variant). NN was applied to explore complex non-linear patterns in patient data, leveraging its capacity to adjust weights across multiple layers (a feedforward neural network with one hidden layer of 100 neurons, ReLU activation, and Adam optimizer, trained for up to 1000 iterations). SVM was included for its strong performance in handling high-dimensional data (with linear kernel and probability estimates enabled).

Each algorithm was implemented using the preprocessed dataset, which had undergone missing value imputation, normalization, and categorical variable encoding. The initial parameters of the algorithms were set to default configurations, with several adjustments applied to improve performance. Although the dataset shows moderate class imbalance (67.6% positive vs. 32.4% negative), no oversampling technique such as SMOTE was applied in order to maintain the original distribution of the data. This limitation is acknowledged as it may affect the generalization of certain classifiers such as NB.

### 2.4. Model Evaluation

One of the commonly used methods to evaluate the performance of machine learning models is the confusion matrix, which presents a comparison between predicted outcomes and actual data [27]. This matrix contains the predicted target values compared against the actual target values. The predicted data are the values obtained from the machine learning model, while the actual data represent the true values. An illustration of the confusion matrix is shown in figure 1. The confusion matrix provides a detailed overview of the number of correct predictions (True Positives/TP, True Negatives/TN) as well as incorrect predictions (False Positives/FP, False Negatives/FN).

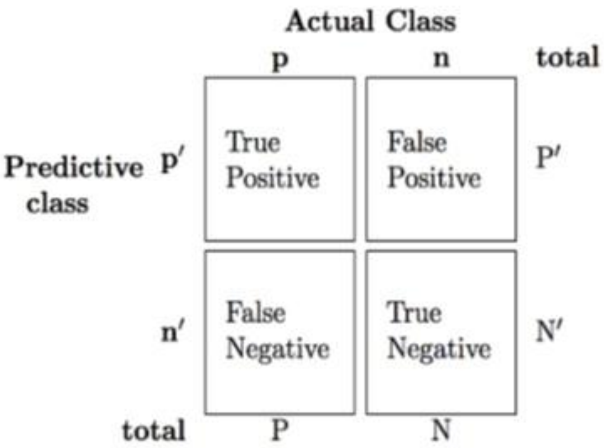

Figure 1. Confusion Matrix [27]

In the context of dengue detection, the positive class represents patients diagnosed with dengue fever, while the negative class indicates patients not diagnosed with dengue fever. Based on this matrix, several evaluation metrics can be calculated to assess the quality of the model, namely: accuracy, precision, recall, and F1-score. The accuracy metric can be computed using Equation (1).

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

Meanwhile, precision is defined by Equation (2).

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Next, recall is calculated using Equation (3).

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

Finally, the F1-score can be computed using Equation (4).

$$F1 - Score = 2\,x\,\frac{Presisi\ x\ Recall}{Presisi\ + Recall} \tag{4}$$

## 3. RESULT

After the training process, model performance was evaluated using the confusion matrix, which illustrates the comparison between predicted outcomes and actual conditions. The following is an explanation of the confusion matrices produced by each ML algorithm.

### 3.1. Evaluation of the Decision Tree Algorithm

The confusion matrix results for the DT algorithm are presented in figure 2, showing that the model successfully classified 667 positive cases as positive (TP) and 316 negative cases as negative (TN). However, there were still 2 positive cases incorrectly classified as negative (FN) and 4 negative cases misclassified as positive (FP). These findings indicate that DT performs fairly well in recognizing both classes, although some errors remain in predicting the negative class.

Using Equations (1) through (4), the accuracy, precision, recall, and F1-score values were calculated. The performance resulted in an accuracy of 99.4%, precision of 99.4%, recall of 99.7%, and an F1-score of 99.6%. This indicates that the DT model is capable of balancing between prediction accuracy for positive cases and the ability to capture all positive instances.

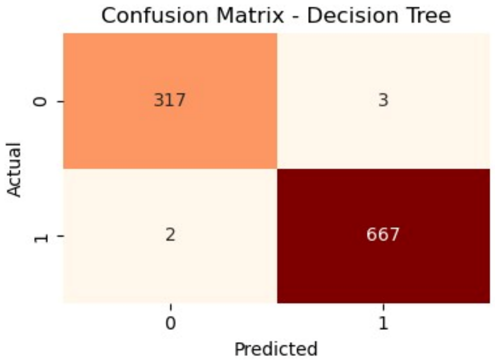

Figure 2. Confusion Matrix of the Decision Tree Algorithm

### 3.2. Evaluation of the K-Nearest Neighbor Algorithm

Figure 3 presents the confusion matrix of the KNN algorithm. The classification results with k = 5 show that 313 samples were correctly classified as negative and 667 samples as positive. However, there were still 7 false positive (FP) cases and 2 false negative (FN) cases.

Using equations (1) to (4), the accuracy, precision, recall, and F1-score can be calculated. The performance yielded an accuracy of 99.1%, precision of 99.0%, recall of 99.7%, and F1-score of 99.3%. Although there were minor misclassifications, particularly in positive cases that should have been detected, KNN still demonstrated very good performance and can be considered reliable in detecting dengue patients.

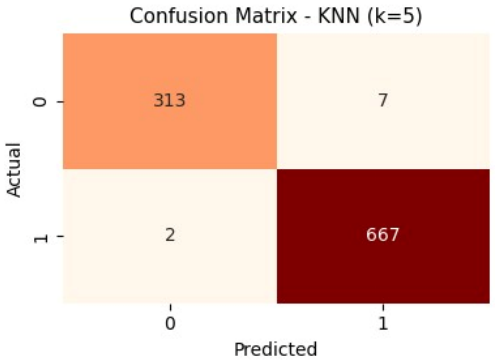

Figure 3. Confusion Matrix of the K-Nearest Neighbor Algorithm

### 3.3. Evaluation of the Logistic Regression Algorithm

Figure 4 presents the confusion matrix of the LR algorithm. This model demonstrates excellent performance, with a low number of type I errors (FP) and no type II errors (FN). The figure shows that LR produced a pattern identical to SVM, with 312 true negative samples, 669 true positive samples, 8 false positive cases, and 0 false negative cases.

Using equations (1) to (4), the accuracy, precision, recall, and F1-score were calculated. The performance achieved an accuracy of 99.2%, precision of 98.7%, recall of 100%, and F1-score of 99.4%.

These results indicate that LR has excellent performance in detecting disease cases, making it a reliable and robust model in this study.

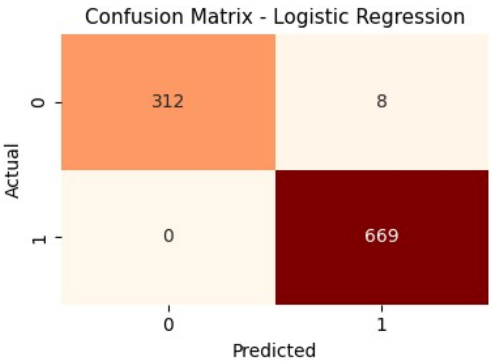

Figure 4. Confusion Matrix of the Logistic Regression Algorithm

### 3.4. Evaluation of the Naive Bayes Algorithm

Figure 5 presents the confusion matrix of the NB algorithm. This model performed very well in avoiding type I errors (FP = 0), meaning no healthy patients were misdiagnosed as dengue positive. However, its main weakness lies in the extremely high number of type II errors (FN), where 551 positive patients failed to be detected and were instead classified as negative. Out of a total of 989 test samples, NB correctly classified only 320 negative samples and 118 positive samples.

Using equations (1) to (4), the accuracy, precision, recall, and F1-score were calculated. The results yielded an overall accuracy of 44.3%, precision of 100%, recall of 17.6%, and an F1-score of only 30.3%. These findings confirm that although the model avoids misdiagnosis of healthy patients, its performance is very poor in recognizing patients truly infected with dengue, making it unsuitable for application to this dataset.

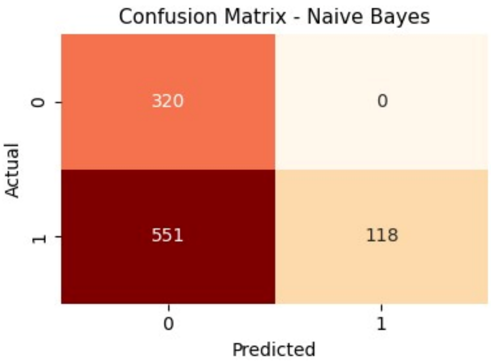

Figure 5. Confusion Matrix of the Naive Bayes Algorithm

### 3.5. Evaluation of the Neural Network Algorithm

Figure 6 presents the confusion matrix of the NN algorithm. This model demonstrates excellent performance with a balance between type I and type II errors. Out of the total test data, 314 samples were correctly classified as negative and 668 samples as positive. Misclassifications were relatively small, with only 6 false positives and 1 false negative.

Using equations (1) to (4), the accuracy, precision, recall, and F1-score were calculated. The performance results yielded an accuracy of 99.4%, precision of 99.1%, recall of 99.9%, and an F1-score of 99.5%. With these results, NN stands out as the best algorithm in this study, as it was able to detect nearly all cases with a very low error rate, both for healthy patients and those infected with dengue.

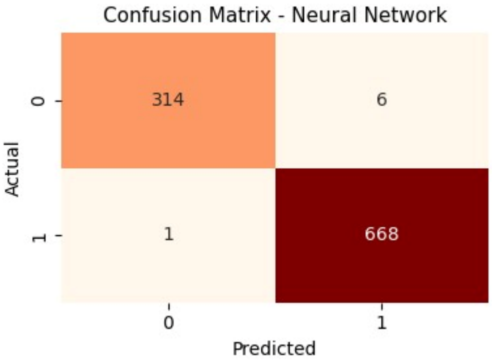

Figure 6. Confusion Matrix of the Neural Network Algorithm

### 3.6. Evaluation of the Support Vector Machine Algorithm

Figure 7 shows the confusion matrix of the SVM algorithm. Out of the total test data, 312 samples were correctly classified as negative and 669 samples as positive. There were only 8 false positive cases and no false negatives.

Using equations (1) to (4), the accuracy, precision, recall, and F1-score were calculated. The performance results yielded an accuracy of 99.2%, precision of 98.7%, recall of 100%, and an F1-score of 99.4%. These results confirm that SVM can detect all positive cases without any being missed, with a very minimal type I error. Its performance is highly competitive and nearly on par with KNN, making SVM one of the strongest models in this study.

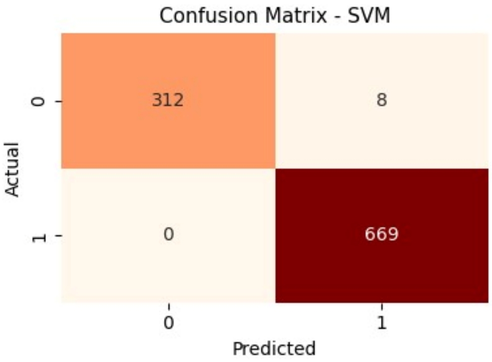

Figure 7. Confusion Matrix of the Support Vector Machine Algorithm

These confusion matrix results are further validated through ROC curve analysis, which provides insight into each algorithm's discriminative ability.

### 3.7. ROC Curve Analysis

In addition to using the confusion matrix and evaluation metrics (accuracy, precision, recall, and F1-score), this study also utilizes the Receiver Operating Characteristic (ROC) curve to analyze the ability of each algorithm to distinguish between the positive class (having dengue fever) and the negative class (not having dengue fever). The ROC curve maps the relationship between True Positive Rate (TPR/Recall) and False Positive Rate (FPR), while the Area Under the Curve (AUC) value is used as a quantitative indicator of the model's discriminative performance. The closer the value is to 1, the better the model's ability to correctly classify both classes. Figure 8 presents the ROC curves of the 6 algorithms used.

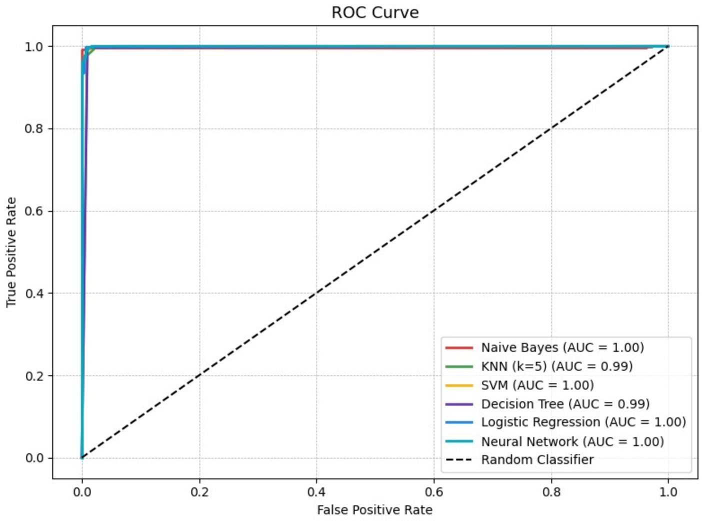

Figure 8. ROC Curve

In the DT algorithm, the ROC curve shows fairly good performance, but there are still fluctuations in several parts of its curve. This indicates that although DT can produce high predictions under certain conditions, there is a potential for classification errors when faced with complex data or uneven feature distributions. The curve not fully approaching the ideal line (top left corner) suggests that this model is still quite vulnerable to overfitting and data variation.

In the KNN algorithm, the ROC curve shows reasonably good performance, although not as high as SVM or DT. Its distance-based characteristics make performance highly influenced by the data distribution and the chosen value of k. While the model can classify with adequate accuracy, recall (sensitivity) may decrease when test data points are inconsistently spaced. Nevertheless, overall, KNN still demonstrates an acceptable discriminative ability.

The LR algorithm displays an ROC curve that is quite close to the ideal area, although slightly below SVM. This indicates that LR, as a linear model, can provide effective class separation in binary classification cases. The advantage of LR lies in its stability in generating interpretable probability predictions, so even though its curve does not always dominate compared to more complex algorithms like NN or SVM, LR remains a highly reliable baseline model.Overall, the ROC curve analysis emphasizes that each algorithm has its own strengths and limitations. SVM and NN occupy the top positions in terms of discriminative ability, while LR provides a balance between stability and accuracy. DT and KNN still yield good results despite certain limitations, while NB offers a simple yet effective solution. Thus, the ROC curve serves as a comprehensive visual proof that machine learning-based models can be significantly utilized in supporting the early detection of dengue fever.

On the other hand, the NB algorithm shows a moderate performance in its ROC curve. Its pattern demonstrates good discriminative ability, but not as strong as algorithms like SVM or NN. This is understandable given that the independence assumption among features used by NB is often not fully met in complex medical data. Nevertheless, NB still offers competitive results with its main advantages being model simplicity and high computational speed.

Further analysis of the NN algorithm reveals that the generated ROC curve approaches optimal performance. The NN curve is generally smooth and tends to align closely with the ideal line, indicating that this model can effectively learn complex patterns from the data. The relatively high AUC for NN demonstrates its ability to identify dengue fever patients with excellent accuracy. This advantage is supported by the network structure that captures non-linear relationships among features, allowing NN to provide highly competitive performance compared to other algorithms.

In contrast to the DT algorithm, SVM presents an excellent ROC curve with nearly perfect proximity to the ideal point. This reflects SVM's strength in constructing the optimal hyperplane to separate the two classes of data with maximum margin. The high ROC curve for SVM serves as evidence that this algorithm is very effective in detecting both truly positive and negative patients, thereby minimizing classification errors. Thus, SVM stands out as one of the most superior algorithms in this research context.

Overall, the ROC curve analysis emphasizes that each algorithm has its own strengths and limitations, and these observations provide a foundation for further evaluation using learning curves.

### 3.8. Learning Curve Analysis

Figure 9 shows the combined learning curves of all six machine learning algorithms. Decision Tree (DT) and Neural Network (NN) consistently achieve high training and validation scores, indicating strong predictive performance. Naive Bayes (NB) shows a notable gap between training and validation, suggesting underfitting. K-Nearest Neighbors (KNN), Logistic Regression (LR), and Support Vector Machine (SVM) display stable curves with minor differences, reflecting balanced learning without overfitting. These curves complement the ROC analysis by illustrating how models learn over increasing training sizes, highlighting their stability and generalization potential.

These results confirm the effectiveness of DT and NN in this classification task, as illustrated in Figure 9, and provide further evidence of model robustness across different training sizes.

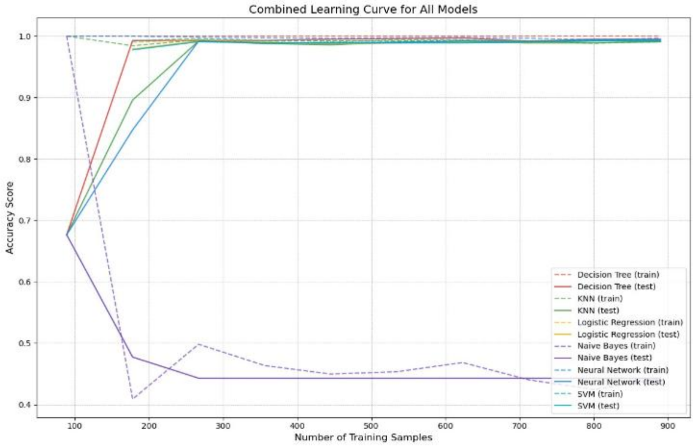

Figure 9: Combined Learning Curve of All Models

### 3.9. Performance Comparison Evaluation of Algorithms

Table 2 presents a comparison of the performance of six machine learning algorithms in detecting dengue fever cases based on four evaluation metrics: accuracy, precision, recall, and F1-score. The results indicate the varying capabilities of each model in classifying the dengue dataset, ranging from DT, KNN, LR, NB, NN, to SVM.

Table 2. Performance Comparison of ML Algorithms

| Model | Accuracy (%) | Precision (%) | Recal l (%) | F1-score (%) |
|---|---|---|---|---|
| DT | 99.4 | 99.4 | 99.7 | 99.6 |
| KNN | 99.1 | 99.0 | 99.7 | 99.3 |
| LR | 99.2 | 98.7 | 100.0 | 99.4 |
| NB | 44.3 | 100.0 | 17.6 | 30.3 |
| NN | 99.4 | 99.1 | 99.9 | 99.5 |
| SVM | 99.2 | 98.7 | 100.0 | 99.4 |

The performance comparison in Table 2 aligns with both ROC and learning curve analyses, confirming DT and NN as the top-performing models and NB as the least effective.

NB obtained the lowest results with only 44.3% accuracy, 17.6% recall, and 30.3% F1-score. Although its precision reached 100%, the model failed to detect most positive cases, making it unsuitable for this dataset.

In contrast, DT, KNN, LR, NN, and SVM all demonstrated strong and competitive performance. SVM and LR achieved perfect recall (100%), ensuring no positive cases were missed. DT recorded the highest accuracy (99.4%) with a strong balance between precision and recall, while KNN and NN also showed excellent performance close to the best values.

## 4. DISCUSSIONS

The findings of this study show that Decision Tree (DT), Neural Network (NN), Logistic Regression (LR), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) achieved excellent results in dengue fever detection, each with accuracy above 99%. DT and NN obtained the highest accuracy (99.4%) with balanced precision and recall, highlighting the interpretability of DT and the capability of NN to capture complex patterns.

LR and SVM also performed competitively, both achieving perfect recall (100%) and successfully detecting all positive cases. Although their precision was slightly lower than DT and NN, this characteristic is particularly valuable in clinical applications where minimizing false negatives is critical for early diagnosis. KNN showed stable performance, though slightly below DT and NN.

By contrast, Naïve Bayes (NB) performed poorly, with accuracy of 44.3% and recall of only 17.6%. Despite its perfect precision, NB failed to recognize most positive cases, which can be attributed to its strong independence assumption that does not fit clinical data complexity.

The ROC curve analysis and learning curve results further support these findings, showing that SVM, NN, and LR not only have strong discriminative ability but also maintain stable learning across increasing training sizes, whereas NB exhibits underfitting and limited generalization. These results suggest that model selection in medical diagnosis must prioritize recall, since failing to detect positive patients carries serious risks. Moreover, the combination of ROC and learning curve analyses provides a comprehensive understanding of each algorithm's performance, highlighting DT and NN as robust choices for accurate and reliable dengue detection.

The results align with previous studies [5], [6], which emphasized the effectiveness of DT, NN, and SVM in infectious disease prediction. Kaur et al. [14] further highlighted the reliability of ensemble and tree-based methods for dengue prognosis. Similarly, Zhao et al. [17] demonstrated the strength of neural networks for dengue forecasting in Colombia, while Bohm et al. [7] and Madewell et al. [23] confirmed that ML-based case screening can outperform traditional diagnostic indicators. These comparisons reinforce that DT, NN, and SVM consistently emerge as high-performing classifiers across different epidemiological contexts.

Threats to validity should also be noted. Internally, although 10-fold cross-validation reduces variance, the relatively small dataset size may still risk overfitting, particularly for DT and NN. Externally, the dataset originates from Bangladesh and may not fully represent clinical populations in other countries, limiting generalizability. Construct validity is also a concern, since NB's poor performance suggests that underlying assumptions of feature independence may not hold in clinical data. Recognizing these limitations is crucial for interpreting the results and guiding future applications.

## 5. CONCLUSION

Based on the evaluation metrics, five classification algorithms—Decision Tree (DT), K-Nearest Neighbor (KNN), Logistic Regression (LR), Neural Network (NN), and Support Vector Machine

Vol. 6, No. 6, December, 2025, Page. 5944-5955
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.6.5309

(SVM)— achieved excellent performance with accuracy above 99% and F1-scores between 99.3% and 99.6%. These results confirm their effectiveness in detecting dengue cases, with differences mainly in the trade-off between precision and recall.

SVM and LR achieved perfect recall (100%), ensuring no positive cases were missed, although their precision was slightly lower. DT and NN attained the highest accuracy with better balance between precision and recall, while KNN performed consistently but slightly below them. In contrast, Naïve Bayes (NB) showed poor recall and low overall accuracy, making it unsuitable for this dataset. ROC analysis reinforced these findings: SVM, NN, and LR had the best discriminative ability, DT and KNN remained competitive, while NB performed poorly, consistent with its other metrics.

This study contributes to the field of Informatics by demonstrating how ML classifiers can be integrated into computer-based decision support systems for early dengue detection, reducing diagnostic errors in clinical practice. For future work, larger and multi-center datasets, as well as the integration of environmental or temporal features, should be explored to enhance generalizability and robustness across diverse populations.

## REFERENCES

[1] A. Amir, A. Iqbal, F. Jaweed, A. F, H. H, and A. Z, “Exploring the current and prospective role of artificial intelligence in disease diagnosis,” *Annals of Medicine & Surgery*, vol. 86, no. 2, pp. 943–949, 2024, doi: 10.1097/ms9.0000000000001700.

[2] B. C. J. Cheah, C. R. Vicente, and K. R. Chan, “Machine Learning and Artificial Intelligence for Infectious Disease Surveillance, Diagnosis, and Prognosis,” *Viruses*, vol. 17, no. 7, pp. 1–32, 2025, doi: 10.3390/v17070882.

[3] N. Sakinah, T. Badriyah, I. Syarif, and P. Korespondensi, “Analisis Kinerja Algoritma Mesin Pembelajaran Untuk Klasifikasi Penyakit Stroke Menggunakan Citra Ct Scan Performance Analysis Machine Learning Algorithms for Classification of Stroke Using Ct Scan Images,” *Jurrnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 4, pp. 833–844, 2020, doi: 10.25126/jtiik.202073482.

[4] M. A. S. Mim, Md. Minhajul Hayat; Assaduzzaman, Md; Mahmud, Arif; Nirob, “Predictive Clinical Dataset for Dengue Fever Using Vital Signs and Blood Parameters,” 2024. doi: 10.17632/xrsbyjs24t.1.

[5] P. Silitonga, A. Bustamam, and H. Muradi, “Comparison of dengue predictive models developed using artificial neural network and discriminant analysis with small dataset,” *Applied Sciences (Switzerland)*, vol. 11, no. 3, pp. 1–16, 2021, doi: 10.3390/app11030943.

[6] S. Hayat, R. Anwar, and S. Aziz, “An automated diagnosis & classification of dengue using advance artificial neural network,” *Computing and Artificial Intelligence*, vol. 3, no. 1, p. 1489, 2024, doi: 10.59400/cai1489.

[7] B. C. Bohm, F. E. de M. Borges, S. C. M. Silva, A. T. Soares, and D. D. Ferreira, “Utilization of machine learning for dengue case screening,” *BMC Public Health*, vol. 24, no. 1, pp. 1–9, 2024, doi: 10.1186/s12889-024-19083-8.

[8] 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, 2023, doi: 10.3390/diagnostics13020220.

[9] C. Edussuriya, S. Deegalla, and I. Gawarammana, “An accurate mathematical model predicting number of dengue cases in tropics,” *PLoS Neglected Tropical Diseases*, vol. 15, no. 11, pp. 1–15, 2021, doi: 10.1371/journal.pntd.0009756.

[10] D. Setiadi, A. Arif, and A. Oktaria, “Comparison of K-Nearest Neighbor and Naive Bayes Algorithms for Tuberculosis Diagnosis Classification,” *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 1, p. 176, 2025, doi: 10.30811/jaise.v5i1.6456.

[11] J. Rusman, B. Z. Haryati, and A. Michael, “Optimisasi Hiperparameter Tuning pada Metode Support Vector Machine untuk Klasifikasi Tingkat Kematangan Buah Kopi,” *Jurnal Komputer dan Informatika*, vol. 11, no. 2, pp. 195–202, 2023, doi: 10.35508/jicon.v11i2.12571.

[12] Z. Zuriati and N. Qomariyah, “Klasifikasi Penyakit Stroke Menggunakan Algoritma K-Nearest

Neighbor (KNN)," *ROUTERS: Jurnal Sistem dan Teknologi Informasi*, vol. 1, no. 1, pp. 1–8, 2022, doi: 10.25181/rt.v1i1.2665.

[13] W. Sholihah, A. S. Handayani, and Sarjana, "Revolutionizing Healthcare: Comprehensive Evaluation and Optimization of SVM Kernels for Precise General Health Diagnosis," *Scientific Journal of Informatics*, vol. 10, no. 4, p. 445, 2023, doi: 10.15294/sji.v10i4.46430.

[14] S. Kaur, S. Sharma, A. U. Rehman, E. T. Eldi, and N. A. Ghamry, "Predicting Infection Positivity, Risk Estimation, and Disease Prognosis in Dengue Infected Patients by ML Expert System," *Sustainability (Switzerland)*, vol. 14, no. 20, pp. 1–20, 2022, doi: 10.3390/su142013490.

[15] Y. P. Bria, P. A. Nani, Y. C. H. Siki, N. M. R. Mamulak, E. M. Meolbatak, and R. D. Guntur, "Leveraging a Random Forest Classifier and SVMSMOTE for an Early-stage Dengue Prediction," *Engineering, Technology and Applied Science Research*, vol. 15, no. 3, pp. 23436–23442, 2025, doi: 10.48084/etasr.10762.

[16] Q.-U.-A. Usmani, Muqtadir Jameel;, Arshad Yousra;, and Ayub Hudaubiya, "Clinical Predictors of Severe Dengue," *Pakistan Journal of Medicine and Dentistry*, vol. 14, no. 2, pp. 294–300, 2025, doi: 10.36283/ziun-pjmd14-2/045.

[17] N. Zhao, K. Charland, M. Carabali, M. Giroux, M. Yuan, and C. G. Balaquera, "Machine learning and dengue forecasting: Comparing random forests and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia," *PLoS Neglected Tropical Diseases*, vol. 14, no. 9, pp. 1–16, 2020, doi: 10.1371/journal.pntd.0008056.

[18] M. E. Francisco, T. M. Carvajal, and K. Watanabe, "Hybrid Machine Learning Approach to Zero-Inflated Data Improves Accuracy of Dengue Prediction," *PLoS Neglected Tropical Diseases*, vol. 2024-Octob, pp. 1–22, 2024, doi: 10.1371/journal.pntd.0012599.

[19] E. Sylvestre, C. Joachim, E. C. Joseph, G. Bouzille, and B. C. Gimenes, "Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review," *PLoS Neglected Tropical Diseases*, vol. 16, no. 1, pp. 1–22, 2022, doi: 10.1371/journal.pntd.0010056.

[20] Y. Cheng, R. Cheng, T. Xu, X. Tan, Y. Bai, and J. Yang, "Integrating meteorological data and hybrid intelligent models for dengue fever prediction," *BMC Public Health*, vol. 25, no. 1, 2025, doi: 10.1186/s12889-025-22375-2.

[21] D. C. Andrade Girón, W. J. Marín Rodriguez, F. de M. Lioo-Jordan, and J. L. Ausejo Sánchez, "Machine Learning and Deep Learning Models for Dengue Diagnosis Prediction: A Systematic Review," *Informatics*, vol. 12, no. 1, pp. 1–20, 2025, doi: 10.3390/informatics12010015.

[22] M. Cabrera, J. Leake, J. Naranjo-Torres, N. Valero, J. C. Cabrera, and A. J. Rodríguez-Morales, "Dengue Prediction in Latin America Using Machine Learning and the One Health Perspective: A Literature Review," *Tropical Medicine and Infectious Disease*, vol. 7, no. 10, 2022, doi: 10.3390/tropicalmed7100322.

[23] Z. J. Madewell, D. M. Rodriguez, M. B. Thayer, V. Rivera-Amill, G. Paz-Bailey, and L. E. Adams, "Machine learning for predicting severe dengue in Puerto Rico," *Infectious Diseases of Poverty*, vol. 14, no. 1, pp. 1–17, 2025, doi: 10.1186/s40249-025-01273-0.

[24] A. Sebastianelli, D. Spiller, R. Carmo, and J. Wheeler, "A reproducible ensemble machine learning approach to forecast dengue outbreaks," *Scientific Reports*, vol. 14, no. 1, pp. 1–17, 2024, doi: 10.1038/s41598-024-52796-9.

[25] X. Chen and P. Moraga, "Forecasting dengue across Brazil with LSTM neural networks and SHAP-driven lagged climate and spatial effects," *BMC Public Health*, vol. 25, no. 1, 2025, doi: 10.1186/s12889-025-22106-7.

[26] Y. Xiao, G. Soares, and L. Bastos, "Dengue nowcasting in Brazil by combining official surveillance data and Google Trends information," pp. 1–21, 2025, doi: 10.1371/journal.pntd.0012501.

[27] M. Gong, "A Novel Performance Measure for Machine Learning Classification," *International Journal of Managing Information Technology*, vol. 13, no. 1, pp. 11–19, 2021, doi: 10.5121/ijmit.2021.13101.