Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5450~5458
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13965

Adaptive confidence-aware ensemble machine learning for viral disease detection and severity

William Mary, Albert Antony Raj
Department of Computer Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received Jul 14, 2026
Revised Sep 11, 2026
Accepted Oct 4, 2026

Keywords:

Confidence-aware soft voting
Multi-disease classification
Severity prediction
Viral disease detection
Weighted adaptive ensemble

ABSTRACT

The rapid emergence and overlapping clinical presentations of viral infectious diseases such as COVID-19 and dengue pose significant challenges for timely diagnosis and risk stratification in resource-limited settings. Conventional single-model approaches often suffer from limited generalizability, class imbalance, and suboptimal feature utilization, resulting in delayed interventions and inefficient resource allocation. This study introduces a weighted adaptive confidence-aware ensemble machine learning framework (WACAM) for accurate early detection and subsequent severity prediction of viral infections using routine hematological and demographic data. A merged dataset comprising 1,369 patient records was preprocessed with median imputation, Z-score outlier handling, Winsorization, and standard scaler normalization. Boruta feature selection identified key clinical markers, followed by synthetic minority over-sampling technique (SMOTE) for class balancing. Seven diverse base learners were integrated through a novel confidence-weighted soft-voting scheme. Model weights were adaptively assigned via exponential softmax of out-of-fold (OOF) area under the curve (AUC) scores, with an optimized probability threshold maximizing Matthews correlation coefficient (MCC). The proposed WACAM achieved superior performance with 86.5% accuracy, 0.8795 F1-score, 0.9156 ROC-AUC, and 0.726 MCC on the test set, outperforming individual models.

This is an open access article under the CC BY-SA license.

Corresponding Author:

William Mary
Department of Computer Applications, Faculty of Science and Humanities
SRM Institute of Science and Technology
Kattankulathur, Chennai, India
Email: lwilliamary@gmail.com

1. INTRODUCTION

Some of the challenges in deployment of a diagnostic model in real-time are medical decision making needs to be highly interpretable, class imbalance of medical data sets and feature redundancy. It may influence the accuracy of the models since the dataset is imbalanced (positives cases' quantity is much higher than negatives cases') and the trained models based on imbalanced data may not be able to generalize well [1]. Building on recent advances in explainable and ensemble modeling [2], the proposed method is a priori effort to better diagnose some infectious diseases that can benefit from diagnosis methodologies with high stakes and objectives aiming at improving accuracy and clinical confidence.

Teng and Wang [3] proposed a stacked deep learning (DL) model for COVID-19 detection from blood samples to capture complex patterns. However, it was limited to a single disease, lacked severity prediction, multi-disease capability, and advanced feature selection methods such as Boruta. Farfán-García *et al.* [4] proposed a diversity-guided stacking ensemble with hybrid feature selection for interpretable COVID-19 mortality prediction. However, the model remained disease-specific and focused only on mortality, lacking

early detection and severity assessment capabilities. Yakovyna *et al.* [5] developed a hierarchical ensemble for COVID-19 case detection and mortality prediction, but it was limited to a single disease and did not incorporate multi-disease detection, Boruta feature selection, or adaptive confidence-aware weighting.

Wei and Zhou [6] suggested credit default prediction with ensemble of layered logistic regressions (LR) method. With the use of the stacking method, more stability has been achieved in a non-medical field with no confidence-aware weighting or advanced feature selection method, using just a conventional base learner. An interpretable real time prediction framework of dengue with clinical blood parameters was proposed by Haque *et al.* [7] which were optimized using genetic algorithm (GA) and generative adversarial network (GAN). Their primary interest was in explainability of their model and accuracy of the prediction of dengue cases. However, the study only covered one disease; no experiments were done on the detection of multiple diseases and the integration of the severity prediction to ensemble methods. Using the population-based data from Puerto Rico region, Niazi and Momand [8] developed machine learning algorithms to predict severe dengue outcomes. But sadly still disease-specific without using any advanced ensemble technology (like confidence aware weighting).

A number of recent research have put forth methods for infectious disease risk assessment and prediction that rely on machine learning. Ma *et al.* [9] and Hsieh *et al.* [10] have used the explainable machine learning (XML) to predict the severity and mortality of COVID-19, respectively, and in [11], [12] and used stacked DL to detect the SARS-CoV-2 from the blood sample. Boundless classification performance for cardiovascular illness and for Alzheimer's illness, respectively, was demonstrated by Manikandan *et al.* [13] and Ramu *et al.* [14] when they used Boruta feature selection method. While forensically most of these publications addressed only one single disease; few merged multiple clinical data and integrated multi-disease identification and severity prediction, no use of adaptive confidence-aware ensemble technique was made.

In recent times, researchers have delved into subjects of explainable models and robust assessment strategies for better prediction of infectious diseases. Therefore, an explainable ML model using nomogram proposed by Zhang *et al.* [15] was suggested to assess the efficacy of long COVID treatment; Itaya *et al.* [16] and Boughorbel *et al.* [17] demonstrated that Matthews correlation coefficient (MCC) which is capable of detecting class imbalance in the problem is superior in binary classes. Researchers working on the dengue side develop boosting techniques called CatBoost, XGBoost and LightGBM for severe dengue predictions [18] and ML models for blood smear based detection of dengue based on platelets and lymphocytes characteristics [19]. While all these studies can be found useful, their disadvantage is that they only take one particular disease into consideration and there's a necessity of multi-disease general structure with adaptive ensemble method which would assist the possible diagnosis of various diseases from a beginning stage.

Class disparity and predictive modelling in the case of infectious diseases have achieved the interest of many studies. However, since the problem of poor classification effect on unbalanced medical samples exists, Li *et al.* [20] proposed an improved synthetic minority over-sampling technique (SMOTE) method, which also generates the samples within a larger space. For instance, the one study, SARS-CoV-2 was efficiently recognized in blood samples by utilizing a stacked DL model as proposed by Wang *et al.* [21]. Another study included prediction of blood transfusion in hospitalized patients with dengue using machine learning as proposed by Ansari *et al.* [22]. Though a significant progress has been achieved so far from the above entries on disease specific predictions and imbalance problems, none of the above work has given a complete framework for multi-diseases early diagnosis and imbalanced severity predictions by adaptive confidence aware ensembles.

While forensically most of these prior works addressed only one single disease; few merged multiple clinical data and integrated multi-disease identification and severity prediction, no use of adaptive confidence-aware ensemble technique was made. While all these studies can be found useful, their disadvantage is that they only take one particular disease into consideration and there's a necessity of multi-disease general structure with adaptive ensemble method which would assist the possible diagnosis of various diseases from a beginning stage.

The proposed weighted adaptive confidence-aware ensemble machine learning framework (WACAM) framework uses a new confidence-weighted soft-voting, SMOTE balance and Boruta feature selection to identify and classify severity of viral infectious diseases. Unlike conventional weighted voting or stacking, WACAM estimates each base learner’s reliability using leakage-free 5-fold out-of-fold (OOF) area under the curve (AUC) scores and converts them into adaptive weights via an exponential softmax function, without training any meta-classifier. The current implementation provides performance measurements for the seven WACAM base learners individually and compares the best individual model [23]–[26], the Hybrid Stacking baseline, and WACAM. The implementation also evaluates WACAM using the OOF-AUC-based adaptive weighting mechanism.

2. METHOD

Proposed WACAM machine learning framework is based on systematic, multi-stage design to facilitate identification and severity prediction of viral infectious illnesses early on. Data is now prepared, which involves a large number of steps, standardization, filling in missing values with the median value, outlier treatment, which is carried out by Winters Winsorization. To find the most important clinical and haematological indicators, the method of Boruta feature selection [13], [14] is used. For class imbalance correction, SMOTE [20] is used and then seven different base learners (random forest (RF), extra trees, gradient boosting, AdaBoost, XGBoost–small configuration, and XGBoost–large configuration) [1] were used as classifiers with parallel training. Then, an exponential softmax function is used to dynamically provide adaptive weights to each base model according to their OOF-AUC performance. These weighted probabilities can be soft voted to maximize the MCC obtained by that, thus getting the ideal threshold for classification. The unified approach allows for both desirable severity risk assessment by means of data from frequent blood tests, and multi-disease classification.

2.1. Data preprocessing and balancing

To achieve such quality data input to these models in practice, a pre-planned and well-organized data pretreatment flow was recommended WACAM. As outlined below are the main steps:

a. Handling missing values: due to its resilience in the face of outliers, median imputation was used to fill in the blanks in the clinical and hematological parameters that were lacking data. The suggested system process is given in Figure 1.

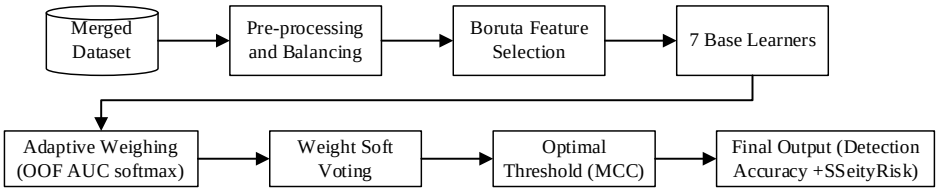


Figure 1. Proposed WACAM framework

The combined dataset consists of 1,369 total patient records formed by merging two distinct, publicly available datasets.(https://github.com/LWilliammary/Explainable-Hybrid-Stacking-Viral-Disease-Prediction) The merged dataset contains 1,369 patient records comprising COVID-19 and dengue cases with demographic and routine hematological parameters, including age, gender, white blood cells (WBC), red blood cells (RBC), hemoglobin (Hb), platelets, lymphocytes, and monocytes. Missing values were handled using median imputation, outliers were treated using Winsorization, features were standardized using standard scaler, Boruta selected informative variables, and SMOTE was applied only to the training partition. The extra trees (large) model was the best with weight 0.1677, followed by the two RF versions. This indicates that the ensemble smartly gave more consideration to the better-performing models, and that all seven learners' input was used. A crucial feature of WACAM's balanced performance is its dynamic weighting approach which distinguishes it from more traditional equal voting or stacking ensembles.

b. Merging procedure: to combine these heterogeneous sources while preserving the integrity of the diagnostic signals, the following pipeline was applied:
c. Feature selection: features were evaluated across both datasets to identify shared complete blood count (CBC) parameters and core demographics.
d. Harmonized variables: demographics: age, sex/gender.
e. Hematological variables: WBC, RBC, Hb, hematocrit (HCT), platelets, neutrophils, lymphocytes, monocytes, eosinophils, and basophils.
f. Preprocessing and missing value handling: records missing crucial diagnostic parameters across all core target features were pruned. Continuous hematological values were standardized/z-score normalized prior to model training to avoid feature scaling imbalances originating from differences in source lab calibration.
g. Target label encoding: a single binary/multiclass target variable (Disease_Label) was constructed, designating class labels explicitly (COVID-19=0, dengue=1).

2.2. Base learners layer

The base learners layer [5], [6], [11] of WACAM machine learning framework consists of seven well-chosen machine learning models that were carefully selected to ensure a diverse set of models with complementary characteristics.

The following base learners were trained in parallel on the SMOTE-balanced training data:
- RF (tuned)–optimized hyperparameters.
- RF (large)–1000 trees for higher capacity.
- Extra trees (large)–1000 trees for additional diversity.
- Gradient boosting machine (GBM).
- AdaBoost (tuned)–optimized version.
- XGBoost (tuned)–hyperparameter-tuned version.
- XGBoost (large)–high-capacity model with different random seed.

The chosen models represent three complementary algorithmic families:
- Bagging-based models (RF tuned, RF large, and extra trees large)–provide high variance reduction and robustness through bootstrap aggregation and random feature selection.
- Boosting-based models (gradient boosting, AdaBoost tuned, XGBoost tuned, and XGBoost large)–focus on sequential residual error correction and capture complex non-linear patterns.

This deliberate combination of bagging and boosting approaches ensures diversity in inductive biases. The adaptive weighting mechanism then assigns higher influence to stronger-performing models while still allowing complementary contributions from all learners, thereby improving overall robustness and generalization. One of WACAM's strongest points is the variety of these base learners; this diversity helps the ensemble to grasp more of the data and lessens the likelihood that any one model's shortcomings would overpower the overall forecast.

2.3. Weighted adaptive confidence-aware ensemble phase in WACAM

The core of the proposed framework is the weighted adaptive confidence-aware ensemble phase, which intelligently combines the predictions of the seven base learners. The softmax scaling factor $\alpha$was fixed at 20 as a design parameter and was not optimized using the independent test set. This parameter controls the concentration of the OOF-AUC-derived weights: a larger value produces a stronger preference for higher-AUC learners, whereas a smaller value produces weights closer to equal voting. The value of 20 provides meaningful differentiation among learners while retaining contributions from complementary models. The independent test set was not used to determine or tune this parameter.

Weights are not assigned equally or manually. Instead, an adaptive weighting mechanism is used as follows:
a. Each base learner generates OOF probability predictions using 5-fold cross-validation on the SMOTE-balanced training data.
b. The AUC score is calculated for each model from its OOF predictions.
c. These AUC scores are converted into final ensemble weights using an exponential softmax function:

$$w_i = \frac{OOF\ AUC_i}{\sum_{i=1}^{7} OOF\ AUC_i} \quad (1)$$

where:
- AUC_i: OOF ROC-AUC of the *i*-th base learner,
- w_i: adaptive confidence weight assigned to learner *I*,
- N: total number of base learners (N=7),
- α: softmax scaling factor,
- $P_i(x)$: predicted probability produced by learner *I*,
- $P_{ensemble}(x)$: weighted ensemble probability, and
- τ: optimized decision threshold.

This method gives significantly higher weight to stronger performing models while still allowing weaker but diverse models to contribute

2.4. Threshold optimization

After computing the weighted soft voting probability, the optimal threshold is selected by evaluating over 4000 values between 0 and 1 and choosing the one that maximizes the MCC on the OOF validation data. This threshold (0.612 in our experiments) serves as the probability cutoff for classifying a patient as Infected. A higher threshold makes the model more conservative (fewer false positives), while a lower one increases sensitivity. Unlike conventional methods that use a fixed threshold of 0.5, WACAM's MCC-based optimization ensures a balanced trade-off between precision and recall, which is especially important for imbalanced medical datasets. This adaptive weighting combined with optimized threshold selection enhances the reliability and clinical relevance of the WACAM ensemble [25].

Algorithm 1. WACAM-weighted adaptive confidence-aware model

```
Input : X_train, X_test, y_train, y_test  (from Algorithm 1, step 1)
Output: ŷ_test, evaluation metrics
X_sm, y_sm ← SMOTE(X_train, y_train, sampling_strategy=1.0)   ▷ full resample
W ← {RF_tuned(best_rf), RF_large(1000 trees), ET_large(1000 trees),
     GradBoost, Ada_tuned(best_ada), XGB_tuned(best_xgb), XGB_large(seed≠)}
for each learner w_i ∈ W:
     p_oof[i] ← CV-Predict(w_i, X_sm, y_sm, k=5, proba)        ▷ leakage-free
     AUC[i]   ← ROC_AUC(y_sm, p_oof[i])
     w_i.fit(X_sm, y_sm)                                       ▷ full-data fit
weight[i] ← exp(20 · AUC[i]) / Σ_j exp(20 · AUC[j])            ▷ softmax, scale=20
p_oof_ens ← Σ_i weight[i] · p_oof[i]
τ* ← argmax_τ  MCC(y_sm, 1[p_oof_ens ≥ τ])                    ▷ sweep, 4001 steps
p_test ← Σ_i weight[i] · w_i.predict_proba(X_test)
ŷ_test ← 1[p_test ≥ τ*]
return ŷ_test, Metrics(y_test, ŷ_test, p_test)
```

WACAM algorithm performs the following steps:

a. Collect prediction probabilities from all seven base models.
b. Evaluate each model's confidence and reliability.
c. Assign adaptive weights to each model.
d. Compute a weighted probability by combining all model outputs.
e. Apply threshold optimization.
f. Generate the final prediction (positive/negative).

3. RESULTS AND DISCUSSION

3.1. Experimental setup

In this work the proposed WACAM weighted soft-voting ensemble framework is implemented using Python 3.x in Jupyter Notebook environment. Open source machine learning and data analysis libraries were used in the implementation for data preparation, feature selection, model training and assessment, and visualization. The confusion matrix in Figure 2 shows that WACAM correctly classified 135 true positives (TP) and 102 true negatives (TN), with only 18 false positives (FP) and 19 false negatives (FN). This balanced error distribution (high sensitivity of 87.66% and specificity of 85.0%) demonstrates the model's reliability in potential clinical screening application. The relatively low number of FP is particularly valuable in clinical settings to avoid unnecessary burden on healthcare resources.

Figure 3 shows the receiver operating characteristic (ROC) curve, which shows that the suggested WACAM model has strong discriminative capabilities. The model has been also presented a great competence in differentiating between the infected and non-infected patients in different classification criteria with the AUC of the model was 0.9156. The red dot represents the ideal working point of threshold value 0.612, which gives high sensitivity (true positive rate (TPR)) of 87.7% and low false positive rate (FPR) of 15.0% based on the independent test-set confusion matrix. The use of adaptive weighting technique and the design of ensemble have proved successful for probability-based predictions for early detection of viral illness because this result is far better than random classification (AUC=0.5).

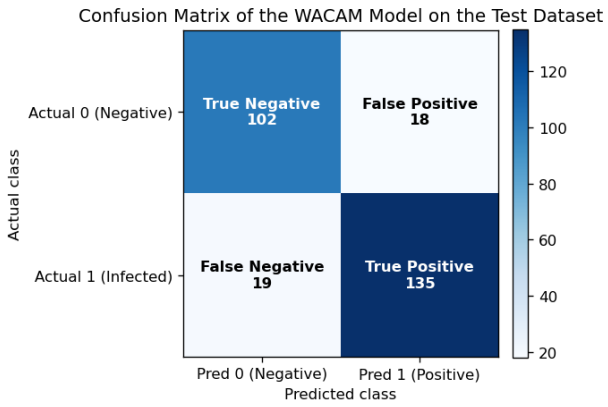


Figure 2. Confusion matrix of the WACAM model on the independent test set (N=274)

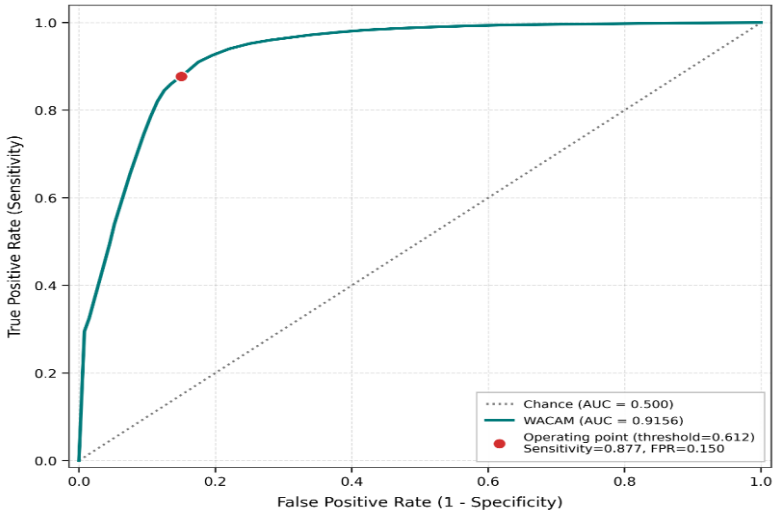


Figure 3. ROC curve of the proposed WACAM model on the independent test set

3.2. Evaluation metrics

An evaluation formula is given for each metric. The Table 1 provides formulas for the various metrics.

Table 1 Evaluation metrics formulas

Metric	Formula
Accuracy	Correct/total
Precision	TP/(TP+FP)
Sensitivity	TP/(TP+FN)
Specificity	TN/(TN+FP)
F1-score	2×(Prec×Sens)/(Prec+Sens)
MCC	$MCC=\frac{(TP+TN)-(FP+FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN}}$
Predicted probability	$P(x)=\sum_{i=1}^{7} w_i p_i(x)$

where TP is true positive; TN is true negative; FP is false positive; FN is false negative; W_i is adaptive weight of the i-th base learner; and $P_i(x)$ is probability output (predict_proba) of the i-th base learner for sample x.

Table 2 shows that the suggested WACAM framework outperforms the competition and provides many benefits over the current models.

Table 2. Performance comparison with state-of-the-art models

Study	Model type	Disease scope	Accuracy	Precision	Sensitivity (recall)	Specificity	F1-score	ROC-AUC	MCC
Teng and Wang [3]	Stacked DL	COVID-19 only	0.852	0.861	0.845	0.838	0.853	0.892	Not reported
Farfán-García *et al*. [4]	Hybrid stacking ensemble	COVID-19 only	0.871	0.865	0.882	0.845	0.873	0.918	0.712
Haque *et al*. [7]	GAN+genetic optimization	Dengue only	0.859	0.874	0.851	0.862	0.862	0.903	Not reported
Madewell *et al.* [18]	Boosting models (XGBoost)	Dengue only	0.847	0.855	0.863	0.831	0.859	0.887	Not reported
Proposed WACAM (ours)	Weighted adaptive confidence ensemble	COVID-19+Dengue	0.8650	0.8824	0.8766	0.8500	0.8795	0.9156	0.7260

As shown in Table 2, the proposed WACAM framework achieves highly competitive performance while offering significant advantages over existing models [3], [4], [7], [8]. It attains the highest ROC-AUC (0.9156) and MCC (0.7260) among the compared studies. Notably, WACAM outperforms most single-disease models in precision and F1-score, despite handling the more challenging multi-disease classification task (COVID-19 and dengue combined). While Farfán-García *et al.* [4] reported a slightly higher accuracy, the proposed model demonstrates better balance across Precision, Specificity, and overall discrimination ability. These results highlight the effectiveness of the adaptive confidence-aware weighting mechanism and Boruta feature selection in building a more robust and generalizable solution clinical decision-support system for preliminary screening and risk stratification rather than as an autonomous diagnostic system. Routine demographic and hematological measurements can be supplied to the trained model, after which the seven base learners independently generate infection probabilities. These probabilities are combined using the OOF-AUC-based adaptive weights, and the optimized decision threshold produces the screening classification. The resulting prediction can support clinical triage and prioritization, while definitive diagnosis and treatment decisions remain under the responsibility of qualified healthcare professionals. Prospective validation using multi-centre clinical data and integration with electronic health-record systems are required before routine clinical deployment. Beyond immediate clinical utility, WACAM contributes to precision-medicine approaches by enabling individualized risk stratification from routine blood parameters and supports pandemic-preparedness efforts by offering a low-cost, scalable screening tool suitable for resource-limited settings, thereby promoting more sustainable diagnostic strategies during future viral outbreaks.

4. CONCLUSION

The key contribution of this work is a (WACAM) that enables reliable multi-disease early detection and severity risk stratification of viral infections (COVID-19 and dengue) from routine hematological data. By converting leakage-free OOF-AUC scores into adaptive soft-voting weights via exponential softmax, WACAM outperforms conventional stacking and equal-weight ensembles without requiring a meta-classifier.The proposed WACAM successfully integrated seven heterogeneous machine learning classifiers using an adaptive weighting strategy based on 5-fold OOF) ROC-AUC scores. The adaptive ensemble assigned higher weights to models demonstrating superior validation performance, thereby improving the reliability of the final prediction.Using an optimal decision threshold of 0.612, WACAM achieved an accuracy of 86.50%, precision of 88.24%, sensitivity of 87.66%, specificity of 85.00%, F1-score of 0.8795, ROC-AUC of 0.9156, and an MCC of 0.7260 on the independent test dataset. The confusion matrix showed 135 TP, 102 TN, 18 FP, and 19 FN, indicating balanced predictive performance across both infected and non-infected patients. Compared with the other stacking ensemble model, WACAM achieved higher precision, specificity, and ROC-AUC, demonstrating improved discrimination capability and fewer false-positive predictions. The comparative results indicate that the integrated feature-selection and adaptive-weighting pipeline contributes to the overall predictive performance of WACAM. However, because a complete factorial ablation of all individual pipeline components was not performed, the independent contribution of each component should not be interpreted as causally established. Future studies involving more diverse clinical datasets are therefore recommended to further confirm the generalizability of the proposed weighting mechanism.

FUNDING INFORMATION

No funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
William Mary	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
Albert Antony Raj		✓				✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to animal use has been complied with all the relevant national regulations and institutional policies for the care and use of animals.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author WM upon reasonable request.

REFERENCES

[1] M. Salmi, D. Atif, D. Oliva, A. Abraham, and S. Ventura, "Handling imbalanced medical datasets: review of a decade of research," *Artificial Intelligence Review*, vol. 57, no. 10, Sep. 2024, doi: 10.1007/s10462-024-10884-2.

[2] R. Dong, H. Yao, T. Chen, W. Yang, Q. Zhou, and J. Xu, "Machine Learning-Based Prediction of In-Hospital Mortality in Severe COVID-19 Patients Using Hematological Markers," *Canadian Journal of Infectious Diseases and Medical Microbiology*, p. 6606842, May 2025 doi: 10.1155/cjid/6606842.

[3] X. Teng and Z. Wang, "Online COVID-19 diagnosis prediction using complete blood count: an innovative tool for public health," *BMC Public Health*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12889-023-17477-8.

[4] A. E. Farfán-García, I. P. Hernández-Peñaranda, O. G. Gómez-Duarte, E. J. Rincón-Barón, G. A. Torres-Rodríguez, and J. M. Flórez-Castillo, "Structural and functional analysis of Escherichia coli membrane disruption by Ib-M peptides," *PLOS One*, vol. 20, no. 10, Oct. 2025, doi: 10.1371/journal.pone.0334029.

[5] V. Yakovyna, N. Shakhovska, and A. Szpakowska, "A novel hybrid supervised and unsupervised hierarchical ensemble for COVID-19 cases and mortality prediction," *Scientific Reports*, vol. 14, no. 1, Apr. 2024, doi: 10.1038/s41598-024-60637-y.

[6] F. Wei and H. Zhou, "Credit default prediction via stacked logistic regression ensemble," *Advances in Economics, Management and Political Sciences*, vol. 185, no. 1, pp. 24–34, Jun. 2025, doi: 10.54254/2754-1169/2025.LH23964.

[7] M. E. Haque, M. N. Absur, F. Al Farid, J. Uddin, and H. A. Karim, "A novel interpretable and real-time dengue prediction framework using clinical blood parameters with genetic and GAN-based optimization," *Frontiers in Artificial Intelligence*, vol. 8, Oct. 2025, doi: 10.3389/frai.2025.1626699.

[8] M. Niazi and S. Momand, "Dengue Severity Level Prediction Using Hematology-Driven Machine Learning Models," *Scientific Journal of Engineering, and Technology*, vol. 3, no. 1, pp. 44–50, Feb. 2026, doi: 10.69739/sjet.v3i1.1542.

[9] Y. Ma *et al.*, "Lung involvement percentage in patients with COVID-19 during the Omicron wave in China: a SHAP-explained machine learning study from a single center," *Frontiers in Public Health*, vol. 13, Jan. 2026, doi: 10.3389/fpubh.2025.1728282.

[10] Y.-C. Hsieh, S. Chen, S.-Y. Tsao, J.-R. Hu, W.-T. Hsu, and C.-C. Lee, "Using machine learning methods to investigate the impact of comorbidities and clinical indicators on the mortality rate of COVID-19," *Frontiers in Medical Technology*, vol. 7, Sep. 2025, doi: 10.3389/fmedt.2025.1621158.

[11] F. Mohtasham, S. S. H. Nazari, M. A. Pourhoseingholi, K. Kavousi, and M. R. Zali, "Hybrid feature-selection and diversity-guided stacking framework for interpretable ensemble learning: Application to COVID-19 mortality prediction," *PLOS One*, vol. 21, no. 4, Apr. 2026, doi: 10.1371/journal.pone.0341198.

[12] H. N. West-Page, K. McGoff, H. Latimer, I. Olufadewa, and S. Chen, "Evaluating biomedical feature fusion on machine learning's predictability and interpretability of COVID-19 severity types: Model development, interpretation, and validation," *JMIR Formative Research*, vol. 10, pp. e76542–e76542, Apr. 2026, doi: 10.2196/76542.

[13] G. Manikandan, B. Pragadeesh, V. Manojkumar, A. L. Karthikeyan, R. Manikandan, and A. H. Gandomi, "Classification models combined with Boruta feature selection for heart disease prediction," *Informatics in Medicine Unlocked*, vol. 44, 2024, doi: 10.1016/j.imu.2023.101442.

[14] S. Ramu, N. Naik, and S. S. Bagalkot, "Boruta feature selection and deep learning for Alzheimer's disease classification," *International Journal of Statistics in Medical Research*, vol. 14, pp. 145–152, Mar. 2025, doi: 10.6000/1929-6029.2025.14.15.

[15] J. Zhang *et al.*, "Explainable machine learning model and nomogram for predicting the efficacy of Traditional Chinese Medicine in treating long COVID: A retrospective study," *Frontiers in Medicine*, vol. 12, Mar. 2025, doi: 10.3389/fmed.2025.1529993.

[16] Y. Itaya, J. Tamura, K. Hayashi, and K. Yamamoto, "Asymptotic properties of matthews correlation coefficient," *Statistics in Medicine*, vol. 44, no. 1–2, Jan. 2025, doi: 10.1002/sim.10303.

[17] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric," *PLOS ONE*, vol. 12, no. 6, Jun. 2017, doi: 10.1371/journal.pone.0177678.

[18] Z. J. Madewell *et al.*, "Machine learning for predicting severe dengue in Puerto Rico: A population-based analysis," *Infectious Diseases of Poverty*, vol. 14, no. 1, Feb. 2025, doi: 10.1186/s40249-025-01273-0.

[19] 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, Jan. 2023, doi: 10.3390/diagnostics13020220.

[20] Y. Li, Y. Yang, P. Song, L. Duan, and R. Ren, "An improved SMOTE algorithm for enhanced imbalanced data classification by expanding sample generation space," *Scientific Reports*, vol. 15, no. 1, pp. 1–21, Jul. 2025, doi: 10.1038/s41598-025-09506-w.

[21] W. Wang, F. Harrou, A. Dairi, and Y. Sun, "Stacked deep learning approach for efficient SARS-CoV-2 detection in blood samples," *Artificial Intelligence in Medicine*, vol. 148, Feb. 2024, doi: 10.1016/j.artmed.2024.102767.

[22] M. S. Ansari, D. Jain, and S. Budhiraja, "Machine-learning prediction models for any blood component transfusion in hospitalized

dengue patients," *Hematology, Transfusion and Cell Therapy*, vol. 46, pp. S13–S23, Nov. 2024, doi: 10.1016/j.htct.2023.09.2365.
[23] Y. Zhang, M. Lin, Z. Wu, Z. Han, L. Cui, and J. Zheng, "Development of a risk prediction model for secondary infection in severe/critical COVID-19 patients," *BMC Infectious Diseases*, vol. 25, no. 1, May 2025, doi: 10.1186/s12879-025-11112-0.
[24] M. Al Mobin, "Forecasting dengue in Bangladesh using meteorological variables with a novel feature selection approach," *Scientific Reports*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-83770-0.
[25] X. Yang, H. Ma, M. Zhu, X. Song, and J. Feng, "Integrated deep learning surveillance of unknown pathogens with pandemic potential using pneumonia of unknown etiology," *Pathogens*, vol. 15, no. 4, Apr. 2026, doi: 10.3390/pathogens15040413.
[26] B. Liu, M. F. Hossain, and S. Hossain, "A comparative evaluation of multiple machine learning approaches for forecasting dengue outbreaks in Bangladesh," *Scientific Reports*, vol. 15, no. 1, Oct. 2025, doi: 10.1038/s41598-025-19752-7.

BIOGRAPHIES OF AUTHORS



William Mary is currently pursuing her Ph.D. in Computer Applications at SRM Institute of Science and Technology, Kattankulathur, Chennai. She has authored and co-authored over fifteen publications in reputed peer-reviewed venues, including IEEE conferences, Scopus-indexed journals, and book chapters. She can be contacted at email: lwilliamary@gmail.com.



Albert Antony Raj is currently serving as professor and deputy dean in the Department of Computer Applications, Faculty of Science and Humanities, at SRM Institute of Science and Technology (SRMIST), Kattankulathur, Chennai, with over 25 years of academic and research experience. He has authored and co-authored more than fifty publications in reputed peer-reviewed venues, including IEEE conferences, Scopus-indexed journals, and book chapters. He can be contacted at email: alberts@srmist.edu.in.