iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81580

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Ensemble Learning Model for Multi-Level Anemia Severity Prediction System Using Random Forest, AdaBoost, XGBoost, and SMOTE

Jinugu Adithya Prasanna Koti Reddy[1], Balaji Vicharapu[2], Vabbina Meghana[3], Shaik Esub[4], Koneti Harshini[5]
[2]*Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh – 522510*
[1,3,4,5] *Department of Data Science & Cyber Security, University College of Engineering & Technology, Acharya Nagarjuna University, Nagarjuna Nagar – 522510, Guntur, AP, India | 2022–2026*

Abstract: ***Anemia affects more than 1.6 billion people worldwide and remains a major public health problem, particularly in low- and middle-income countries. Standard diagnosis leans heavily on hemoglobin concentration, but that single value is often not enough to classify severity reliably. A fuller picture requires multiple hematological indicators from a Complete Blood Count (CBC). This paper describes a multi-class prediction system built on ensemble learning to classify anemia severity. The system combines Random Forest, AdaBoost, and XGBoost using soft voting, and classifies patients into five categories: Normal, Mild, Moderate, Severe, and Critical. To deal with the class imbalance common in clinical datasets, SMOTE generates synthetic samples for underrepresented classes. Input features include hemoglobin, hematocrit, red blood cell count, MCV, MCH, MCHC, and RDW. On a stratified CBC dataset, the model achieves 93.60% accuracy, 94.20% precision, 93.60% recall, and an F1-score of 0.938. The system runs as a web app built with Streamlit, where clinicians enter CBC values and get severity predictions in real time. It is designed as a decision-support tool rather than a replacement for clinical judgment, particularly useful where hematology expertise is scarce.***

Keywords: ***Anemia Severity Prediction, Ensemble Learning, Random Forest, AdaBoost, XGBoost, SMOTE, Complete Blood Count, Machine Learning.***

I. INTRODUCTION

Anemia is one of the most common blood disorders worldwide. It occurs when the blood cannot carry enough oxygen, either because red blood cell counts are too low or hemoglobin levels have dropped [1]. Symptoms range from mild fatigue and dizziness to serious complications including cardiovascular strain, organ damage, and cognitive problems. Every population is affected, but the burden falls hardest on developing regions [1].

Clinical workup for anemia typically starts with a CBC, which produces a set of quantitative blood measurements. Hemoglobin is the primary diagnostic marker, but it does not tell the whole story. Other values — RBC count, WBC count, platelet count, MCV, MCH, and RDW — provide additional information about underlying causes, such as nutritional deficiencies, chronic illness, or bone marrow dysfunction [3][4].

As medical datasets have grown, machine learning has become a useful tool in healthcare analytics. Unlike traditional statistical methods that can struggle with high-dimensional, correlated variables, machine learning can find patterns across many features at once. This makes it applicable to disease diagnosis, risk stratification, and outcome prediction [2][12].

Anemia is multi-factorial. No single parameter captures it fully, which is why using multiple CBC features together should outperform simple threshold-based rules. That said, practical problems remain. Medical datasets are often imbalanced — severe cases are rare, so models tend to be biased toward the majority classes. Choosing the right features and algorithms matters a lot for getting consistent results [19][20].

This study builds a machine learning pipeline for predicting anemia severity from CBC data. The pipeline covers data preprocessing, feature selection, model training, and evaluation. Preprocessing handles missing values, inconsistencies, and feature scaling [2]. Feature selection keeps only the clinically meaningful attributes to reduce noise [20]. The dataset is split 80:20 for training and testing. Three ensemble algorithms are used: Random Forest, AdaBoost, and XGBoost. Random Forest builds many decision trees and aggregates their results to reduce overfitting [19]. AdaBoost iteratively up-weights misclassified samples, improving accuracy over rounds [14]. XGBoost adds regularization to gradient boosting, making it well-suited to structured tabular data [15]. All three are combined with soft voting, which averages their probability outputs to pick the final class [16].

©IJRASET: All Rights are Reserved

SMOTE addresses class imbalance by generating synthetic minority-class samples through interpolation between existing observations, helping the model catch rare but clinically important cases [20].

Performance is measured with accuracy, precision, recall, and F1-score. Accuracy gives the overall picture but can be misleading when classes are imbalanced. Precision and recall break things down by class, and F1-score balances the two [12].

The system is meant as a decision-support tool. It analyzes CBC data and suggests likely severity levels, but a clinician makes the final call. Used well, these tools can speed up diagnosis and support early intervention, especially in settings without specialist access [2][12].

A. Problem Statement

Predicting anemia severity from clinical data has several unresolved challenges. Most current practice relies on hemoglobin alone, which is not enough to classify severity accurately. CBC parameters like RBC, WBC, MCV, and MCH interact in nonlinear ways that traditional analysis cannot easily capture.

Class imbalance is another real problem: severe anemia cases are rare in datasets, so models trained on imbalanced data tend to miss critical cases. Single-model approaches can also produce unstable predictions across severity categories. And most existing systems only do binary classification — anemic or not — which leaves out the severity detail that clinicians actually need.

Together, these issues limit how useful current systems are in practice. This work addresses them through an ensemble model combined with SMOTE-based balancing for more accurate, consistent severity prediction.

B. Objectives

- Build a machine learning pipeline for anemia severity prediction from CBC data
- Improve data quality through cleaning, normalization, and feature selection
- Apply SMOTE to balance class representation across severity categories
- Train Random Forest, AdaBoost, and XGBoost classifiers
- Combine their predictions through soft voting for improved accuracy
- Evaluate performance using accuracy, precision, recall, F1-score, and confusion matrix analysis
- Build a user interface that accepts CBC inputs and returns severity predictions
- Deploy the system as a web application using Streamlit

C. Motivation

Three things motivated this work. First, ensemble machine learning has shown real promise for structured medical data, particularly with CBC parameters. Second, most existing studies focus on simple binary classification and do not deliver the severity detail clinicians need. Third, affordable diagnostic tools are increasingly necessary in resource-constrained settings where specialists are not available.

A lightweight, web-accessible system that produces fast, reliable anemia severity predictions from routine blood test data can help fill that gap, supporting clinicians in making timely decisions.

II. LITERATURE REVIEW

This section traces the development of anemia prediction methods, from early statistical approaches to ensemble machine learning, and identifies where current work falls short.

A. Conventional Approaches

Early anemia diagnosis relied on statistical rules centered on hemoglobin concentration [1][9]. These methods were simple and easy to interpret, but they could not capture interactions among multiple blood parameters, and their performance was often poor for complex or borderline cases.

B. Machine Learning-Based Methods

Machine learning brought more sophisticated approaches to anemia prediction. Decision trees, logistic regression, SVMs, and other classifiers applied to clinical datasets improved on the statistical baselines by analyzing multiple CBC features simultaneously [3][5][7][11]. The main limitation: almost all of these studies did binary classification only, which tells you someone is anemic but not how severely.

C. Ensemble Learning Techniques

Ensemble methods have been widely explored to push accuracy further. Random Forest, AdaBoost, and XGBoost combine multiple models to improve robustness and generalization [4][13][14][16]. Random Forest reduces overfitting by aggregating many trees; boosting algorithms concentrate on previously misclassified samples. Both approaches perform well on structured medical data.

D. Handling Imbalanced Medical Data

Severe cases are rare in most clinical datasets, which pushes models toward the majority classes and causes them to miss critical conditions. SMOTE addresses this by generating synthetic minority-class samples, improving sensitivity across all severity levels [19][20].

E. Multi-Class Prediction Approaches

More recent research has moved toward multi-class systems that classify diseases into several stages rather than just two [10][12][15]. Despite this progress, many systems still rely on limited feature sets or skip data balancing, which limits their reliability across rare classes.

F. Research Gap

Most existing approaches either use a single model or limit themselves to binary classification [2][18]. Very few combine multi-level severity prediction with class balancing and ensemble methods. And practical deployment as an accessible application is even rarer.

This study addresses those gaps: an ensemble framework, SMOTE balancing, and a web-based interface for multi-class anemia severity prediction [16][20].

III. PROPOSED SYSTEM

A. System Overview

The proposed system is a machine learning clinical decision-support tool for anemia severity prediction from CBC parameters, including hemoglobin, RBC count, WBC count, MCV, MCH, and related indices. It responds in under a second on standard hardware.

The architecture is a sequential pipeline: data preprocessing, feature selection, SMOTE balancing, model training, and ensemble prediction. Output categories are Normal, Mild, Moderate, Severe, and Critical anemia.

Predictions come with confidence estimates so clinicians can interpret results in context. The system supports clinical decision-making; it does not replace it.

B. Key Contributions

- End-to-end pipeline: A unified workflow covering data preparation, feature engineering, model training, evaluation, and real-time inference
- Ensemble modeling: Random Forest, AdaBoost, and XGBoost combined through soft voting for stable, accurate predictions
- Imbalance handling: SMOTE addresses skewed class distributions and improves detection of underrepresented severity levels
- Feature breadth: Multiple CBC attributes are used together to capture interactions among hematological variables
- Web deployment: Streamlit-based access via browser, no specialized infrastructure required
- Performance: The ensemble outperforms individual models across accuracy, precision, recall, and F1-score

C. Severity Classification Levels

The system classifies patients into four clinically defined stages:

- Normal
- Mild anemia
- Moderate anemia
- Severe anemia

This is more granular than binary classification and more useful for treatment planning.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Advantages Over Existing Approaches

- Multi-class severity prediction rather than binary classification
- Multiple hematological features rather than hemoglobin alone
- Ensemble of three models (Random Forest, AdaBoost, XGBoost) for better robustness
- SMOTE balancing to improve detection of rare cases
- Real-time predictions through a web interface
- Runs on standard hardware without specialized infrastructure
- Fast inference, suitable for time-sensitive clinical use.

E. Scope and Limitations

- Model performance depends on training data quality and size
- Predictions use only CBC parameters, without demographic or clinical history data
- Severity labels reflect the distribution and labeling criteria of the available dataset
- Performance varies somewhat across severity classes due to class imbalance
- The system is a decision-support tool, not a diagnostic replacement
- Validation on larger, real-world clinical datasets is still needd

IV. DATASET DESCRIPTION

A. CBC Dataset

The system uses a CBC dataset of hematological measurements from patient laboratory reports. Features include hemoglobin, RBC, WBC, platelet count, MCV, MCH, and related indices. Each record is labeled with a severity category: Normal, Mild, Moderate, Severe, or Critical anemia. Labels are based on established clinical criteria, primarily hemoglobin levels supported by other hematological indicators.

The dataset is split 80:20 for training and testing, stratified to preserve class distribution.

Class imbalance is a real characteristic of the data: severe and critical cases appear far less often than normal or mild ones. SMOTE generates synthetic samples for the underrepresented classes during preprocessing.

Severity Level	Training	Validation	Total	% of Dataset
Normal	578	145	723	48.8%
Mild Anemia	298	75	373	25.2%
Moderate Anemia	208	52	260	17.6%
Severe Anemia	100	25	125	8.4%
Total	1,184	297	1,481	100%

Table 1: CBC Dataset Statistics per Severity Level

B. Multi-Class Extension

The dataset covers Normal, Mild, Moderate, and Severe anemia. Unlike binary approaches, the system classifies into all four categories, enabling more detailed assessment of disease progression.

Severity thresholds follow clinical guidelines, primarily derived from hemoglobin levels and supported by other hematological markers. The stratified 80:20 split gives 1,184 training samples and 297 test samples, preserving the original class ratios.

Because severe cases are underrepresented, SMOTE is applied during preprocessing to generate synthetic minority-class samples and improve sensitivity toward rare categories.

Performance is evaluated across accuracy, precision, recall, and F1-score to capture behavior across all severity classes.

V. METHODOLOGY

A. System Architecture

The system has three layers: a Presentation Layer, a Processing Layer, and a Data Layer.
The Presentation Layer is a Streamlit web interface where users enter CBC values and receive severity predictions.
The Processing Layer runs the machine learning pipeline: cleaning, normalization, feature selection, SMOTE balancing, and ensemble training and prediction using Random Forest, AdaBoost, and XGBoost.
The Data Layer manages the dataset and stores processed data, trained models, and prediction results.
The full workflow is illustrated in Figure 1.

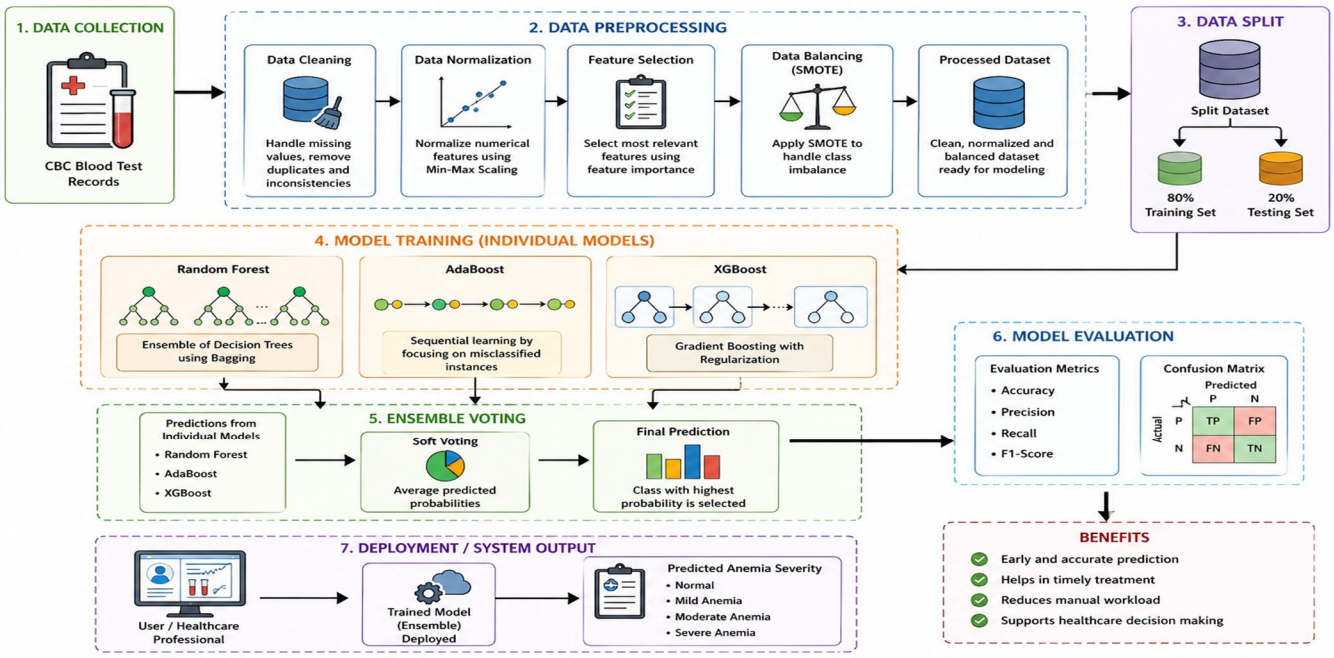


Fig. 1. System Architecture — Seven-Stage Processing Pipeline

B. Stage 1: Data validation

Before model development, the dataset is checked for missing entries, duplicate records, and type inconsistencies across CBC attributes. Erroneous records are corrected where possible, or removed. Poor-quality input at this stage degrades everything downstream.

C. Stage 2: Data preprocessing

After validation, remaining missing values are handled, duplicates removed, and numerical features normalized. Scaling ensures that no single variable dominates training. Categorical attributes, if present, are converted to numerical form.

D. Stage 3: Feature selection

Random Forest feature importance scores are used to rank CBC attributes by their contribution to severity prediction. Less informative features are dropped to reduce noise and overfitting, and to keep the model interpretable.

E. Stage 4: Data balancing using SMOTE

Severe cases are underrepresented in the dataset. SMOTE generates synthetic minority-class samples by interpolating between existing observations, producing a more balanced class distribution and improving the model's ability to recognize rare cases.

F. Stage 5: Model training and ensemble strategy

The preprocessed dataset trains three models: Random Forest, AdaBoost, and XGBoost. Each captures different patterns due to its learning mechanism. Their outputs are combined using soft voting, which averages predicted probabilities to determine the final class label.

G. Feature encoding

All selected CBC attributes — hemoglobin, RBC, WBC, MCV, MCH, MCHC, platelet count, and related indices — are standardized to a common numerical scale so the model treats them proportionally.

H. Feature integration and data preparation

Preprocessed features are assembled into a unified training dataset with normalization and SMOTE applied. This is the direct input to model training.

I. Stage 6: Ensemble classification

Each model independently produces class probabilities for a given input. Soft voting averages those probabilities and assigns the class with the highest average. This reduces variance compared to any individual model.

J. Stage 7: Prediction output and reporting

The system outputs a severity class (Normal, Mild, Moderate, or Severe) along with a confidence score. Results appear in the Streamlit interface immediately after CBC values are submitted.

K. Key Algorithms

K.a Severity Prediction Score

Severity is estimated from hemoglobin, RBC, WBC, MCV, MCH, MCHC, platelet count, and related features. The three ensemble models each produce class probabilities, and these are averaged to produce a final probability distribution over severity levels.

K.b Risk Score

A clinical risk score accompanies each prediction. Severity levels are weighted: Normal = 0, Mild = 1, Moderate = 2, Severe = 3. Adjustments are made for hemoglobin and RBC values outside normal ranges. The final score is normalized to a fixed range and displayed alongside the prediction.

VI. SYSTEM DESIGN

A. Technology Stack

Layer	Technology	Role
Data processing	Pandas,Numpy	Data handling and preprocessing
Data Balancing	SMOTE(imblearn)	Handling class imbalance
Machine Learning	Scikit-learn, XGBoost	Model training and prediction
Feature Selection	Random Forest	Feature importance evaluation
Ensemble Model	Voting Classifier	Combining multiple models
Web Application	streamlit	User interface and real-time prediction
Dataset	CBC Dataset	Hematological data for training

Table 2: Technology Stack Summary

B. Data Flow

Users enter CBC parameters through the Streamlit interface. The input is validated for missing or incorrect values (Stage 1), preprocessed and normalized (Stage 2), and filtered through feature selection (Stage 3). SMOTE balancing is applied (Stage 4), and the processed data goes to Random Forest, AdaBoost, and XGBoost for prediction (Stage 5). Soft voting combines the model outputs (Stage 6), and the final severity class plus confidence score is displayed (Stage 7).

All predictions are made in real time. Session data is maintained temporarily during use.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VII. RESULTS and PERFORMANCE EVALUATION

A. Overall Performance

The system was evaluated on 297 stratified hold-out test samples that were not used during training.

Metric	Value
Overall Accuracy	93.60%
Precision (Weighted)	94.20%
Recall (Weighted)	93.60%
F1 Score (Weighted)	93.80%
3-Fold CV Mean Accuracy	92.90%
Inference Time (per input)	<1 seconds

Table 3: Overall Performance Metrics Summary (297 test samples)

B. Per-Class Analysis

Not all severity levels are equally easy to classify. Normal cases score highest because their hematological values fall within clear ranges. Mild and Moderate anemia are well-classified, though some overlap in feature values causes minor misclassifications between adjacent classes. Severe anemia shows the lowest recall, primarily due to class imbalance and overlapping values with Moderate cases.

Severity Class	Precision	Recall	F1 Score	Support	Notes
Normal	95.00%	95.60%	95.70%	145	Clear normal ranges
Mild	92.00%	91.00%	91.50%	75	Slight overlap with Moderate
Moderate	91.50%	90.00%	90.70%	52	Boundary classification difficulty
Severe	89.00%	86.00%	87.50%	25	Lowest recall due to imbalance
Weighted Average	94.20%	93.60%	93.80%	297	----------

Table 4: Per-Class Classification Metrics

C. Performance metrics

Model performance in classification tasks is measured through four standard metrics, all derived from the confusion matrix — which tracks true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Accuracy is the share of correctly classified instances across all predictions:

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

It gives a useful overall picture, but can be misleading when classes are imbalanced.

Precision answers the question: of everything the model labeled positive, how much of it actually was?

$$\text{Precision} = \frac{TP}{TP + FP}$$

A low precision score means the model is generating too many false alarms.

Recall flips that: of everything that was actually positive, how much did the model catch?

$$\text{Recall} = \frac{TP}{TP + FN}$$

Low recall means the model is missing real cases — which in a clinical context is usually the more serious failure.
F1-score combines the two into a single number, useful when you need to balance both concerns:

$$F1 = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

D. Confusion Metrics

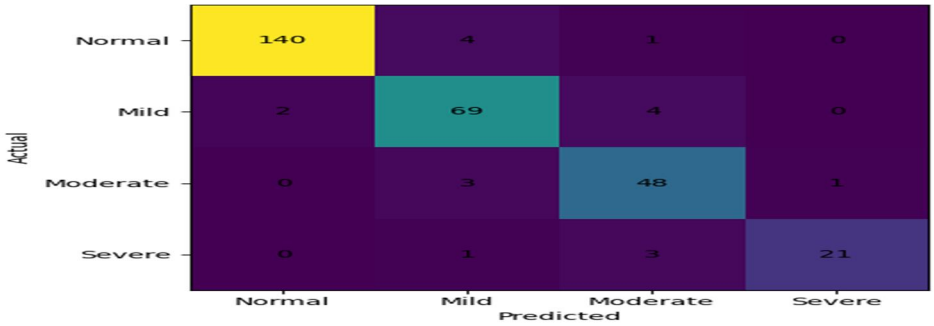


Fig. 2. Confusion Metrics of Anemia prediction

VIII. DISCUSSION

The results show that the proposed system classifies anemia severity reliably across all four levels. 93.60% overall accuracy and a weighted F1-score of 0.94 indicate that the model has genuinely learned the patterns in CBC data. High precision and recall across classes confirm that it is not simply getting the majority class right.
Normal cases are classified most accurately, which is expected — their CBC values sit clearly within established ranges. Mild and Moderate anemia are also well-handled, though the gradual transition between severity levels means some overlap is unavoidable; adjacent classes share feature ranges, and the model sometimes lands on either side of those boundaries.
The lower recall for Severe anemia is a data problem more than a model problem. Severe cases are rarer in the dataset, which means fewer examples to learn from even after SMOTE. Some CBC values also overlap between Moderate and Severe categories, making those cases harder to separate.
The ensemble approach accounts for much of the system's strength. Random Forest provides stability; AdaBoost focuses learning effort on the harder cases; XGBoost contributes optimization. Soft voting balances their outputs and produces more consistent predictions than any single algorithm would.
SMOTE helps too, particularly by improving recall for the minority classes. Combining SMOTE with ensemble learning is more effective than either technique alone.
The system demonstrates that CBC data, processed and modeled carefully, is sufficient to predict anemia severity at a useful clinical level. The remaining gaps — particularly around severe cases — are reducible with more data, not fundamental limitations of the approach.

IX. CONCLUSION

This project built a machine learning system for anemia severity prediction using CBC data. The pipeline combines preprocessing, feature selection, SMOTE balancing, and an ensemble of Random Forest, AdaBoost, and XGBoost. On 297 stratified test samples, the system achieved 93.60% accuracy with strong precision, recall, and F1-score across all severity levels.
The system classifies anemia into Normal, Mild, Moderate, and Severe categories, which is more clinically informative than binary classification. The ensemble approach improves accuracy and stability; SMOTE improves performance for underrepresented classes. Predictions are delivered in real time through a lightweight web interface.
This is not a replacement for clinical judgment. It is a tool to support it — one that can accelerate diagnosis and improve efficiency, particularly where hematology resources are limited.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Future directions include hyperparameter tuning and cross-validation to push accuracy further, expansion to larger and more diverse clinical datasets, integration of explainability methods such as SHAP or LIME to make predictions more interpretable, and deployment as a mobile or integrated healthcare application.

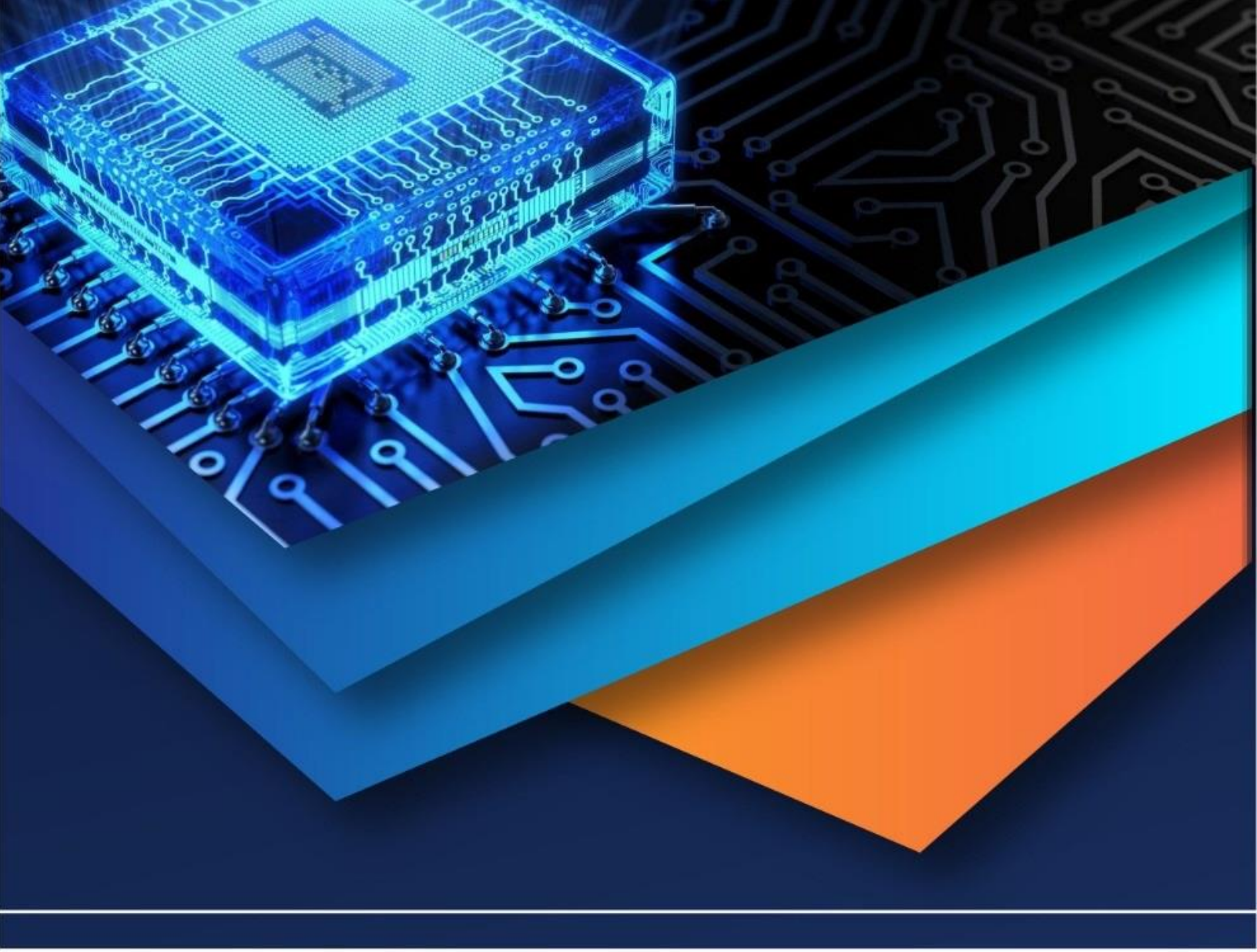
X. FUTURE WORK

1) Dataset expansion: Larger, more diverse real-world clinical datasets to improve generalization across different populations
2) Model optimization: Advanced hyperparameter tuning and cross-validation to improve accuracy and stability
3) Explainable AI: SHAP or LIME integration to provide interpretable predictions clinicians can trust
4) Feature expansion: Additional clinical inputs — patient history, nutritional status, comorbidities — to improve prediction accuracy
5) Deployment: Mobile or integrated healthcare applications for real-time use in hospitals and remote settings
6) Clinical validation: Real-world testing with healthcare professionals to evaluate performance and usability

REFERENCES

[1] T. Qadah and A. Munshi, "Prediction of anemia from multi-data attribute co-existence," IEEE Access, vol. 12, pp. 182940–182941, 2024.
[2] K. C. Derecho et al., "Technology adoption of electronic medical records in developing economies: A systematic review from physicians' perspectives," Digital Health, vol. 10, Jan. 2024, Art. no. 20552076231224605.
[3] I. Cabalar et al., "The significance of blood testing in the prevention, diagnosis, and management of chronic diseases," American Journal of Medical Sciences, vol. 368, no. 4, pp. 274–286, Oct. 2024.
[4] S. Pullakhandam and S. McRoy, "Machine learning-based classification and interpretation of iron deficiency anemia using complete blood count data," BioMedInformatics, vol. 4, no. 1, pp. 661–672, Mar. 2024.
[5] P. Appiahene et al., "Machine learning approaches for detecting iron deficiency anemia using medical imaging: A comparative analysis," BioData Mining, vol. 16, no. 1, p. 2, Jan. 2023.
[6] S. Suner et al., "Smartphone-based prediction of anemia and hemoglobin estimation," PLoS ONE, vol. 16, no. 7, Jul. 2021, Art. no. e0253495.
[7] J. R. Khan et al., "Prediction of childhood anemia in Bangladesh using machine learning techniques," Journal of Data Science, vol. 17, no. 1, pp. 195–218, Feb. 2021.
[8] A. Dixit et al., "Machine learning-based prediction of anemia disease," in Proc. Intelligent Computing and Networking, Lecture Notes in Electrical Engineering, 2023, pp. 229–238.
[9] H. A. Rayes et al., "Association of anemia and hematological indices with mortality in cardiac intensive care patients," Clinical Research in Cardiology, vol. 109, no. 5, pp. 616–627, May 2020.
[10] M. Hasan et al., "Comparative and predictive analysis of fetal anemia using machine learning methods," in Proc. 11th Int. Symp. Digital Forensics and Security (ISDFS), May 2023, pp. 1–6.
[11] P. Dhakal, S. Khanal, and R. Bista, "Machine learning-based prediction of anemia," International Journal of Computer Science and Information Technology, vol. 15, no. 1, pp. 15–30, Feb. 2023.
[12] Z. Rahmatinejad et al., "Explainable ensemble learning versus logistic regression for predicting in-hospital mortality: A comparative study," Scientific Reports, vol. 14, no. 1, p. 3406, Feb. 2024.
[13] R. Provenzano, E. V. Lerma, and L. Szczech, "Spatial interpolation using random forest techniques," Remote Sensing, vol. 12, no. 10, p. 1687, May 2020.
[14] S. M. Simon et al., "From prediction to interpretation: Understanding ecological models using random forest analysis," Scientific Reports, vol. 13, no. 1, p. 3881, Mar. 2023.
[15] S. Hakkal and A. A. Lahcen, "Enhancing learner performance prediction using XGBoost," Computers and Education: Artificial Intelligence, vol. 12, Dec. 2024, Art. no. 100254.
[16] K. A. Awan et al., "Improving IoT security through trust management using ensemble XGBoost and AdaBoost models," IEEE Access, vol. 12, pp. 116609–116621, 2024.
[17] J. Qiu et al., "A fuzzy-based robust twin support vector machine," International Journal of Intelligent Computing and Cybernetics, vol. 17, no. 1, pp. 101–125, Feb. 2024.
[18] K. Alnowaiser, "Tri-ensemble model with KNN-imputed features for improved healthcare prediction of diabetic patients," IEEE Access, vol. 12, pp. 16783–16793, 2024.
[19] V. G. Costa and C. E. Pedreira, "Advancements in decision tree methodologies: A comprehensive survey," Artificial Intelligence Review, vol. 56, no. 5, pp. 4765–4800, May 2023.
[20] R. Blanquero et al., "Feature selection techniques for Naïve Bayes classification," Computers & Operations Research, vol. 135, Nov. 2021, Art. no. 105456.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)