International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 14, Issue 2, 2026

An Interpretable AI-Based Hybrid Model for Accurate Anaemia Prediction Using Explainable Machine Learning

Mr Dr.D. Durga Prasad [1], Mamidi Jyothika Sai [2] ,Panuganti Manasa [3], Nicholas Sucharitha Mary [4],Katragadda Bhavya Geetika [5],Budati Mohana Aditya [6]
1 Professor & Hod of CSE, Computer Science Engineering Department, Potti Sriramulu Chalavadi Mallikarjuna Rao College Of Engineering, One Town , Vijayawada, India. 2,3,4,5,6 Students; Computer Science Engineering Department, Potti Sriramulu Chalavadi Mallikarjuna Rao College Of Engineering,One Town , Vijayawada, India
Mail Id; ddprasad.fac@gmail.com 1, mamidijyothikasai26@gmail.com 2, manasapanuganti2303@gmail.com 3, nicholasmary283@gmail.com 4, geetikabhavya69@gmail.com 5, mohanaditya333@gmail.com 6

Abstract: Anaemia is a widespread hematological disorder characterized by reduced haemoglobin concentration, leading to decreased oxygen-carrying capacity of blood and significant health complications if not detected early. Traditional diagnostic methods are often invasive, time-consuming, and lack predictive intelligence for proactive healthcare support. To address these limitations, this paper proposes a transparent anaemia prediction framework using hybrid machine learning integrated with Explainable Artificial Intelligence (XAI). The system utilizes a structured anaemia dataset containing demographic and clinical parameters such as gender, haemoglobin level, MCV, MCH, and MCHC to classify individuals as anaemic or non-anaemic. Multiple machine learning algorithms including Decision Tree, K-Nearest Neighbors, Support Vector Machine, and Gradient Boosting are implemented and comparatively evaluated. To enhance predictive performance, a hybrid ensemble model combining Random Forest and XGBoost through a Voting Classifier is proposed, achieving an accuracy of 99.70%. To improve transparency and clinical trust, SHAP and LIME are incorporated to provide interpretable explanations for each prediction by identifying influential features contributing to the model's decisions. The complete framework is deployed through a Flask-based web application for real-time prediction and user interaction. Experimental results demonstrate that the proposed model significantly improves prediction accuracy while maintaining interpretability, making it suitable for practical clinical decision support and early anaemia diagnosis.

***Index terms* - —** *Anaemia Prediction, Explainable Artificial Intelligence (XAI), Machine Learning, Hybrid Ensemble Model, SHAP, LIME, Clinical Decision Support, Healthcare Analytics*

1. INTRODUCTION

Anaemia is one of the most prevalent global health disorders, affecting millions of individuals across different age groups, particularly children, pregnant women, and elderly populations. It is primarily caused by reduced haemoglobin concentration or insufficient healthy red blood cells, which decreases the oxygen-carrying capacity of blood and leads to fatigue, weakness, impaired cognitive function, and severe complications if left untreated. Early detection and timely diagnosis of anaemia are essential for effective treatment and prevention of long-term health risks.

Traditional anaemia diagnosis methods mainly rely on laboratory blood tests and clinical examination, which can be time-consuming, costly, and dependent on expert interpretation. Although machine learning techniques have shown significant potential in automating disease prediction, many existing predictive systems suffer from limited transparency and operate as black-box models, making them difficult for clinicians to trust in real-world healthcare settings. The lack of interpretability remains a major barrier to the adoption of artificial intelligence in medical decision support systems.

To overcome these limitations, this paper proposes a transparent anaemia prediction framework that integrates hybrid machine learning with Explainable Artificial Intelligence (XAI). Multiple classification algorithms including Decision Tree, K-Nearest Neighbors, Support Vector Machine, and Gradient Boosting are evaluated, while a hybrid ensemble model combining Random Forest and XGBoost through a Voting Classifier is employed to enhance predictive performance. Furthermore, SHAP and LIME are incorporated to provide interpretable explanations for prediction outcomes, enabling clinicians to understand the influence of individual features on model decisions. This transparent and accurate framework aims to support early anaemia diagnosis, improve clinician trust, and facilitate reliable AI-assisted healthcare decision-making.

2. LITERATURE SURVEY

a) Integrating explainable AI (XAI) with machine learning for enhanced disease prediction and decision transparency

Artificial Intelligence (AI) methods for early illness diagnosis and clinical decision assistance have significantly improved medical healthcare. However, because the models are black box, there is a gap in the public's general adoption of these algorithms' outputs. Because medical professionals must comprehend the rationale behind a certain disease's result, the concealed nature of these systems presents fundamental challenges within the medical sectors that manage critical situations. The proposed study investigates a hybrid machine learning (ML) architecture that incorporates Explainable AI (XAI) techniques to enhance predicted performance and interpretability. Within its framework, the system uses Decision Trees, Naive Bayes, Random Forests, and XGBoost algorithms to forecast the risks of diabetes, anemia, thalassemia, heart disease, and thrombocytopenia. By highlighting significant characteristics that contribute to each prediction, SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) enhance the usefulness of the suggested system. In addition to offering comprehensible explanations for the interpretation of model results, the framework maintains an accuracy of 99.2%. Clinical practitioners may make judgments by comprehending AI-generated outputs thanks to the framework's interpretability and performance, which lowers mistrust in AI-driven healthcare.

b) Differential diagnosis of iron deficiency anemia from aplastic anemia using machine learning and explainable Artificial Intelligence utilizing blood attributes:

The World Health Organization lists anemia as the most common blood condition worldwide. Anemia is a reduction of healthy red blood cells. This condition also reduces blood oxygen capacity. This study examines blood test features to establish a reliable approach for identifying Aplastic Anemia and Iron Deficiency Anemia. Currently, no research employ Interpretable AI to make the differential diagnosis. Kasturba Medical College, Manipal provided this study's dataset. The dataset included Red Blood Cell count, Hemoglobin level, Mean Corpuscular Volume, etc. Explainable AI is a Machine Learning trend. They explain machine learning outcomes to stakeholders. XAI tools SHAP, LIME, Eli5, Qlattice, and Anchor are used to understand model predictions. Training and testing data using XAI models uses PLT, PCT, MCV, PDW, HGB, ABS LYMP, WBC, MCH, and MCHC significance features. Data analytics helps medical practitioners make better decisions, manage resources, and improve patient care. XAI argues that medical personnel who rely on AI-assisted diagnosis and therapy should trust model outcomes by considering each patient's unique traits.

c) An Explainable AI and Optimized Multi-Branch Convolutional Neural Network Model for Eye Anemia Diagnosis:

In this study, a novel non-invasive method of detecting ocular anemia using deep learning techniques is proposed. Conventional approaches are expensive and may be uncomfortable for patients since they depend on intrusive procedures like venipuncture. For improved accuracy, our model uses a multi-branch convolutional neural network (CNN) architecture with multiclass support vector machines (SVMs) and the Hippopotamus Optimization (HO) method. We use data augmentation and the Synthetic Minority Oversampling Technique (SMOTE) to overcome data imbalance. A dataset of 211 eye photos is used to train and assess the model. With a Receiver Operating Characteristic (ROC) curve showing an Area Under the Curve (AUC) of 0.973, suggesting great discriminative power, the model achieves an astounding 97.06% accuracy. Training speed and inference time are greatly enhanced by the parallel branch CNN design. Additionally, the model's capacity to differentiate between anemic and non-anemic instances is demonstrated by the t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization, which efficiently clusters data points. We use the SHapley Additive exPlanations (SHAP) approach to determine feature significance in order to guarantee model transparency and dependability. This non-invasive method has great potential for early and effective anemia screening, especially in environments with limited resources.

d) Explainable artificial intelligence for personalized management of inflammatory bowel disease: A minireview of recent advances:

Because the course of inflammatory bowel disease (IBD) is unpredictable, medication response varies, and illness presentation is heterogeneous, personalized care is essential. The "black-box" character of many AI models restricts their clinical use, despite the fact that machine learning algorithms and artificial intelligence (AI) provide potential answers by evaluating complicated, multidimensional patient data. This problem is addressed by Explainable AI (XAI), which increases the transparency and clinical actionability of data-driven forecasts. The latest developments and therapeutic significance of using XAI for individualized IBD treatment are the main topics of this minireview. We discuss the significance of XAI in treatment prioritization and show how XAI methods, such interpretable model topologies and feature-attribution explanations, improve AI model transparency. By giving priority to the predictive indicators for gastrointestinal bleeding and food consumption patterns, XAI models have been used in recent years to identify IBD abnormalities. Additionally, research has shown that the use of XAI improves IBD risk classification and increases the accuracy of treatment effectiveness and patient response prediction. XAI promotes clinician trust, facilitates tailored decision-making, and permits the safe use of AI systems in delicate, customized IBD treatment pathways by converting opaque AI models into comprehensible tools.

e) Improved CKD classification based on explainable artificial intelligence with extra trees and BBFS:

Chronic renal disease is a long-term condition characterized by a slow deterioration of kidney function. The presence of renal disease and the estimated glomerular filtration rate are the main factors used to classify it. A generally recognized classification system for chronic kidney disease has been devised by the Kidney Disease Improving Global Outcomes Organization. Developing machine learning models that not only correctly forecast results but also provide understandable justifications for their choices is the goal of explainable artificial intelligence for categorization. Because of their complexity and obscurity, traditional machine learning models can make it impossible to understand the complex mechanisms underlying certain categorization decisions. An explainable artificial intelligence-chronic kidney disease model is presented in this paper for the classification procedure. The model uses Shapley additive explanations values and additional trees to construct explainable artificial intelligence. Additionally, the most crucial elements for the suggested explainable artificial intelligence-chronic kidney disease model are chosen using a binary breadth-first search strategy. This approach is intended to provide useful information for improving decision-making techniques in the area of categorizing chronic renal disorders. Accuracy, sensitivity, specificity, F-score, and area under the ROC curve are used to assess how well the suggested model performs in comparison to other machine learning models, including random forest, decision tree, bagging classifier, adaptive boosting, and k-nearest neighbor. The results of the trial showed that the suggested model produced the best outcomes with an accuracy of 99.9%.

3. METHODOLOGY

i) Proposed Work:

The proposed work introduces a transparent and intelligent anaemia prediction framework that combines advanced machine learning algorithms with Explainable Artificial Intelligence (XAI) to support early and reliable diagnosis of anaemia. The system utilizes clinical and demographic input parameters such as gender, haemoglobin level, Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin (MCH), and Mean Corpuscular Hemoglobin Concentration (MCHC) collected from the anaemia dataset. Initially, the dataset undergoes preprocessing operations including missing value handling, duplicate removal, categorical encoding, normalization, and feature scaling to ensure high-quality input for model training. Multiple machine learning algorithms such as Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Gradient Boosting are trained and evaluated to identify the most effective predictive approach.

To further enhance prediction performance, a hybrid ensemble model is developed by combining Random Forest and XGBoost using a Voting Classifier, which leverages the strengths of multiple learners to improve robustness and classification accuracy. The proposed framework integrates SHAP and LIME explainability techniques to provide transparent interpretation of prediction results by identifying feature contributions for each outcome. Additionally, the system includes hereditary condition analysis and personalized food/lifestyle recommendation modules to provide practical healthcare guidance based on prediction results. The entire framework is deployed through a Flask-based web application, enabling real-time user interaction, prediction generation, explanation visualization, and recommendation delivery in an accessible clinical decision support environment.

ii) System Architecture:

The proposed system architecture for the transparent anaemia prediction framework follows a multi-stage pipeline designed to ensure accurate, interpretable, and real-time anaemia diagnosis. The process begins with the Data Collection and Input Layer, where patient clinical parameters such as haemoglobin level, gender, MCV, MCH, and MCHC are collected through the user interface or dataset repository. The input data is then forwarded to the Data Preprocessing Module, where data cleaning, missing value handling, duplicate removal, categorical encoding, normalization, and feature scaling are performed. The processed dataset is split into training and testing subsets to facilitate robust model development and evaluation. This preprocessed data is subsequently passed to the Machine Learning Layer, where multiple classifiers including Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Gradient Boosting are trained to learn predictive patterns associated with anaemia conditions.

To improve predictive performance and robustness, the outputs of the base classifiers are integrated into a Hybrid Ensemble Layer, where Random Forest and XGBoost are combined using a Voting Classifier to

ISSN 2347–3657
Volume 14, Issue 2, 2026

generate the final optimized prediction. The architecture further incorporates an Explainable AI Layer consisting of SHAP and LIME modules, which provide both global and local interpretability by identifying the contribution of each clinical feature to the prediction outcome. These explanations are delivered alongside the final prediction through the Prediction Output Module, enabling clinicians and users to understand the rationale behind model decisions. Finally, the complete system is deployed using a Flask-based Web Application Layer, which provides an interactive interface for data entry, prediction generation, explanation visualization, and recommendation delivery, ensuring practical accessibility and real-time clinical decision support.

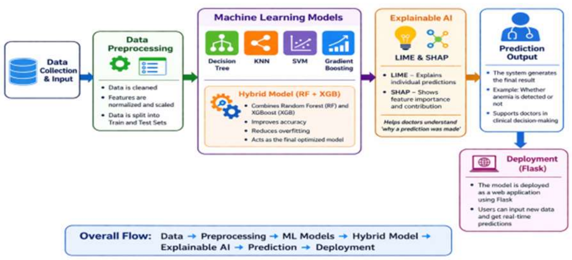

Fig 1: proposed architecture

iii) Modules:

1. Data Collection and Input Module

This module gathers patient clinical and demographic information such as gender, haemoglobin level, MCV, MCH, and MCHC from users or dataset repositories. It validates the input data format and ensures completeness before forwarding it for processing.

2. Data Preprocessing Module

This module performs data cleaning operations including missing value handling, duplicate removal, categorical encoding, normalization, and feature scaling. It prepares the raw data into a structured and machine-readable format suitable for model training.

3. Machine Learning Model Training Module

This module trains multiple machine learning algorithms such as Decision Tree, KNN, SVM, and Gradient Boosting using the preprocessed dataset. It evaluates each model's predictive capability and stores trained classifiers for further analysis.

4. Hybrid Ensemble Model Module

This module combines Random Forest and XGBoost using a Voting Classifier to improve prediction accuracy and robustness. It acts as the final optimized prediction engine by leveraging the strengths of multiple models.

5. Explainable AI Module

This module integrates SHAP and LIME techniques to interpret model predictions. It highlights the importance of individual clinical features and explains why a particular prediction was generated.

6. Prediction and Classification Module

This module processes new patient inputs through the trained hybrid model and classifies the patient as anaemic or non-anaemic. It generates prediction confidence scores for reliability assessment.

7. Recommendation Module

Based on prediction outcomes and hereditary analysis, this module provides personalized food, lifestyle, and healthcare recommendations to support anaemia management and prevention.

8. Web Deployment Module

Implemented using Flask, this module provides a user-friendly web interface for real-time interaction, data submission, prediction visualization, and explanation display.

iv) Algorithms:

1. Decision Tree

Decision Tree is a supervised machine learning algorithm that performs classification by recursively splitting the dataset into branches based on the most informative feature conditions. It creates a tree-like structure of decision rules that maps input clinical parameters to anaemia prediction outcomes. In the proposed system, Decision Tree is used as a baseline interpretable classifier because of its simplicity and ability to clearly represent decision-making logic. Its rule-based structure helps in understanding how medical attributes influence classification.

2. K-Nearest Neighbors (KNN)

K-Nearest Neighbors is a non-parametric instance-based learning algorithm that classifies new samples based on the majority class of their nearest neighboring data points in feature space. It uses distance metrics such as Euclidean distance to identify similarity between patient records. In this system, KNN assists in anaemia prediction by comparing incoming patient data with previously labeled clinical records and assigning the most common neighboring class as the prediction.

3. Support Vector Machine (SVM)

Support Vector Machine is a supervised classification algorithm that identifies an optimal hyperplane to separate data points of different classes with maximum margin. It is particularly effective in high-dimensional feature spaces and for medical datasets with nonlinear boundaries. In the anaemia prediction framework, SVM is employed to distinguish anaemic and non-anaemic patients by learning robust decision boundaries from clinical and demographic parameters.

4. Gradient Boosting

Gradient Boosting is an ensemble learning technique that constructs a strong predictive model by sequentially combining multiple weak learners, typically decision trees. Each subsequent model corrects the errors made by the previous model, thereby improving predictive performance iteratively. In this work, Gradient Boosting is utilized to enhance classification capability and capture complex nonlinear relationships among anaemia-related health features.

5. Random Forest

Random Forest is an ensemble classification algorithm that builds multiple decision trees using random subsets of data and features, then aggregates their predictions to produce a final output. This approach improves model generalization and reduces overfitting compared to a single decision tree. In the proposed system, Random Forest contributes to robust anaemia prediction by effectively handling feature variability and noisy clinical data.

6. XGBoost

XGBoost is an advanced gradient boosting framework optimized for speed, regularization, and predictive performance. It enhances traditional boosting by using parallel tree construction, shrinkage, and regularization techniques to minimize overfitting and improve generalization. Within the proposed framework, XGBoost is employed as a high-performance classifier capable of learning complex feature interactions in anaemia diagnosis.

7. Voting Classifier (Hybrid Ensemble Model)

The Voting Classifier is a hybrid ensemble approach that combines the predictions of multiple classifiers and determines the final output based on majority voting. In this project, Random Forest and XGBoost are integrated into a Voting Classifier to leverage the strengths of both models. This hybrid ensemble improves predictive robustness, reduces the limitations of individual algorithms, and delivers the highest classification accuracy among all evaluated models.

8. SHAP (SHapley Additive Explanations)

SHAP is an Explainable Artificial Intelligence technique based on cooperative game theory that assigns contribution values to each feature involved in a prediction. It explains both global model behavior and local individual predictions by quantifying the impact of each clinical parameter on the final result. In this system, SHAP is used to provide transparent insights into the anaemia prediction process, enabling clinicians to understand feature importance comprehensively.

9. LIME (Local Interpretable Model-Agnostic Explanations)

LIME is a local explanation technique that approximates complex machine learning models with simple interpretable surrogate models around individual predictions. It explains why a specific prediction was made by highlighting the contribution of nearby feature values. In the proposed anaemia prediction framework, LIME is integrated to provide case-specific explanations for each patient prediction, thereby increasing model transparency and clinical trust.

4. EXPERIMENTAL RESULTS

The proposed transparent anaemia prediction framework was evaluated using the Kaggle Anaemia dataset containing clinical and demographic parameters such as gender, haemoglobin level, MCV, MCH, and MCHC. The dataset was preprocessed through missing value handling, categorical encoding, normalization, and feature scaling before being divided into training and testing subsets for model development. Multiple machine learning algorithms including Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Gradient Boosting, and the proposed Hybrid Voting Ensemble Model were trained and evaluated using standard classification metrics such as Accuracy, Precision, Recall, and F1-Score. Experimental analysis demonstrates that the proposed hybrid model outperformed all baseline classifiers, confirming the effectiveness of ensemble learning in improving anaemia prediction reliability.

The comparative performance evaluation indicates that the Hybrid Ensemble Model combining Random Forest and XGBoost through a Voting Classifier achieved the highest classification performance with 99.649% accuracy, 99.699% precision, 99.583% recall, and 99.640% F1-score, significantly surpassing individual machine learning models. In addition to prediction performance, SHAP and LIME visualizations were generated to explain the contribution of clinical features toward each prediction, thereby improving interpretability and trustworthiness of the system. The Flask-based deployment further validated the practical usability of the framework by enabling real-time anaemia prediction, explanation visualization, and recommendation generation through a user-friendly web interface. These results demonstrate that the proposed model is highly effective, accurate, and suitable for transparent clinical decision support in anaemia diagnosis.

Accuracy: The ability of a test to differentiate between healthy and sick instances is a measure of its accuracy. Find the proportion of analysed cases with true positives and true negatives to get a sense of the test's accuracy. Based on the calculations:

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: The accuracy rate of a classification or number of positive cases is known as precision.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 14, Issue 2, 2026

Accuracy is determined by applying the following formula:
Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$Precision = \frac{TP}{(TP + FP)}$$

Recall: The recall of a model is a measure of its capacity to identify all occurrences of a relevant machine learning class. A model's ability to detect class instances is shown by the ratio of correctly predicted positive observations to the total number of positives.

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

mAP: One ranking quality statistic is Mean Average Precision (MAP). It takes into account the quantity of pertinent suggestions and where they are on the list. The arithmetic mean of the Average Precision (AP) at K for each user or query is used to compute MAP at K.

$$\boldsymbol{mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k}$$

$$\boldsymbol{AP_k = \text{ the AP of class } k}$$

$$\boldsymbol{n = \text{ the number of classes}}$$

F1-Score: A high F1 score indicates that a machine learning model is accurate. Improving model accuracy by integrating recall and precision. How often a model gets a dataset prediction right is measured by the accuracy statistic..

$$F1 = 2 \cdot \frac{(Recall \cdot Precision)}{(Recall + Precision)}$$

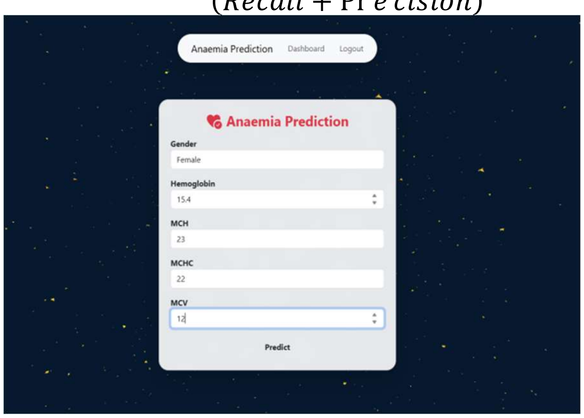

Fig 2 Anaemia Prediction Input Screen

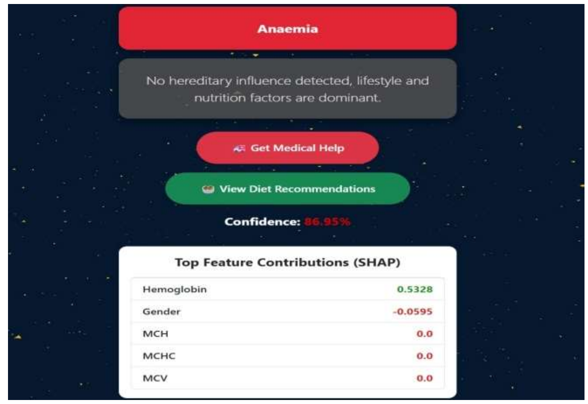

Fig 3 Prediction Result with Explainable AI, Recommendation

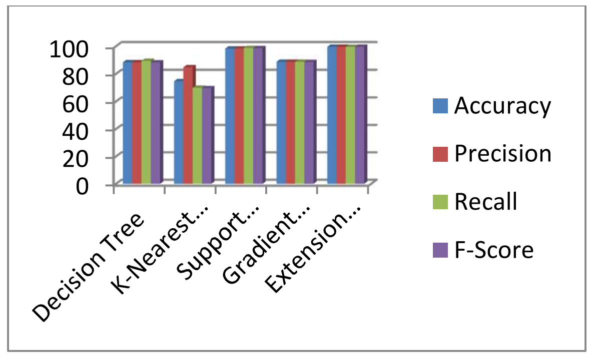

Fig 4 Accuracy graph

Algorithm Name	Accuracy	Precision	Recall	F-Score
Decision Tree	88.421	88.449	89.432	88.352
K-Nearest Neighbors (KNN)	74.737	84.810	70.000	69.616
Support Vector Machine (SVM)	98.387	98.387	98.788	98.567
Gradient Boosting	88.772	88.745	88.735	88.698
Extension Hybrid Model	99.649	99.699	99.583	99.640

Table 1 Performance Evaluation

5. CONCLUSION

This paper presented a transparent and intelligent anaemia prediction framework that integrates hybrid machine learning with Explainable Artificial Intelligence to support accurate and interpretable clinical diagnosis. Multiple machine learning algorithms were evaluated for anaemia classification, and a hybrid ensemble model combining Random Forest and XGBoost through a Voting Classifier demonstrated superior performance over individual classifiers. The proposed framework achieved a maximum accuracy of 99.649%, indicating its effectiveness in identifying anaemic and non-anaemic cases with high reliability.

To enhance transparency and clinical trust, SHAP and LIME were incorporated to explain prediction outcomes by identifying the contribution of individual clinical features influencing each decision. The deployment of the model through a Flask-based web application further enables real-time prediction, explanation visualization, and recommendation generation in a user-friendly environment. Overall, the proposed system successfully bridges the gap between predictive performance and interpretability, making it a practical and trustworthy decision support tool for early anaemia diagnosis in healthcare applications.

6. FUTURE SCOPE

Future enhancements of the proposed anaemia prediction system can focus on expanding the dataset with additional medical, nutritional, and lifestyle-related parameters such as age, dietary habits, chronic disease history, and environmental factors to improve prediction generalization and personalization. Incorporating larger real-world clinical datasets from hospitals can further enhance model robustness and applicability across diverse patient populations. Advanced deep learning and automated feature selection techniques may also be explored to improve predictive performance while maintaining interpretability.

The framework can be extended by integrating Optical Character Recognition (OCR) and image-processing techniques to automatically extract blood test parameters from uploaded medical reports, thereby reducing manual input effort and improving usability. Future work may also include developing a mobile application for remote healthcare accessibility, integrating wearable health monitoring devices for continuous anaemia assessment, and strengthening compliance with healthcare security standards to support large-scale clinical deployment. These enhancements can transform the system into a comprehensive intelligent healthcare assistant for proactive anaemia management.

REFERENCES

[1] Agrawal, R., Gupta, T., Gupta, S., Chauhan, S., Patel, P., & Hamdare, S. (2025). Fostering trust and interpretability: integrating explainable AI (XAI) with machine learning for enhanced disease prediction and decision transparency. *Diagnostic Pathology*, *20*(1), 105.

[2] Darshan, B. D., Sampathila, N., Bairy, G. M., Prabhu, S., Belurkar, S., Chadaga, K., & Nandish, S. (2025). Differential diagnosis of iron deficiency anemia from aplastic anemia using machine learning and explainable Artificial Intelligence utilizing blood attributes. *Scientific Reports*, *15*(1), 505.

[3] Mohammed, K. K., Dahmani, N., Ahmed, R., Darwish, A., & Hassanien, A. E. (2025). An Explainable AI and Optimized Multi-Branch Convolutional Neural Network Model for Eye Anemia Diagnosis. *IEEE Access*.

[4] Okpete, U. E., & Byeon, H. (2025). Explainable artificial intelligence for personalized management of inflammatory bowel disease: A minireview of recent advances. *World Journal of Gastroenterology*, *31*(35), 111033.

[5] Elshewey, A. M., Selem, E., & Abed, A. H. (2025). Improved CKD classification based on explainable artificial intelligence with extra trees and BBFS. *Scientific Reports*, *15*(1), 17861.

[6] S. Sadiq, M. U. Khalid, S. Ullah, W. Aslam, A. Mehmood, G. S. Choi, and B.-W. On, ''Classification of β-thalassemia carriers from red blood cell indices using ensemble classifier,'' IEEE Access, vol. 9, pp. 45528–45538, 2021.

[7] J.W.Asare,P.Appiahene,E.T.Donkoh,andG.Dimauro,' 'Irondeficiency anaemia detection using machine learning models: A comparative study of fingernails, palm and conjunctiva of the eye images,'' Eng. Rep., vol. 5, no. 11, 2023, Art. no. e12667.

[8] M. K. Hasan, M. H. Aziz, M. I. I. Zarif, M. Hasan, M. Hashem, S. Guha, R. R. Love, and S. Ahamed, ''Noninvasive hemoglobin level prediction in a mobile phone environment: State of the art review and recommenda tions,'' JMIR mHealth uHealth, vol. 9, no. 4, Apr. 2021, Art. no. e16806.

[9] S. M. A. Iqbal, I. Mahgoub, E. Du, M. A. Leavitt, and W. Asghar, ''Advances in healthcare wearable devices,'' Npj Flexible Electron., vol. 5, no. 1, p. 9, Apr. 2021.

[10] C. Kim, S. U. Gadgil, A. J. DeGrave, J. A. Omiye, Z. R. Cai, R. Daneshjou, and S.-I. Lee, ''Fostering transparent medical image AI via an image-text foundation model grounded in medical literature,'' Nature Med., vol. 30, pp. 1154–1165, 2024.

[11] S. Abbas, A. Qaisar, M. S. Farooq, M. Saleem, M. Ahmad, and M. A. Khan, ''Smart vision transparency: Efficient ocular disease prediction model using explainable artificial intelligence,'' Sensors, vol. 24, no. 20, p. 6618, Oct. 2024.

[12] T. Shahzad, M. Saleem, M. S. Farooq, S. Abbas, M. A. Khan, and K.Ouahada, ''Developing a transparent diagnosis model for diabetic retinopathy using explainable AI,'' IEEE Access, vol. 12, pp. 149700–149709, 2024.

[13] M. Saleem, M. Sajid Farooq, T. Shahzad, A. Hassan, S. Abbas, T. Ali, E.-H.-M. Aggoune, and M. A. Khan, ''Secure and transparent mobility in smart cities: Revolutionizing AVNs to predict traffic congestion using MapReduce, private blockchain,

and XAI,'' IEEE Access, vol. 12, pp. 131541–131555, 2024.
[14] P. Dhakal, S. Khanal, and R. Bista, ''Prediction of anaemia using machine learning algorithms,'' AIRCC's Int. J. Comput. Sci. Inf. Technol., vol. 15, no. 1, pp. 15–30, 2023.
[15] K.Aishwarya,''Threshold based feature selection for anaemia prediction,'' in Proc. Int. Conf. Comput. Commun. Inform. (ICCCI), 2023, pp. 1–5.
[16] S. A. Wulandari, H. A. Azies, M. Naufal, W. A. Prasetyanto, and F. A. Zahra, ''Breaking boundaries in diagnosis: Non-invasive anemia detection empowered by AI,'' IEEE Access, vol. 12, pp. 9292–9307, 2024.
[17] Kasthuri, E., Subbulakshmi, S., & Sreedharan, R. (2024). Insightful clinical assistance for anemia prediction with data analysis and explainable ai. *Procedia computer science*, *233*, 45-55.
[18] Prajapati, J., Uduthalapally, V., Das, D., Mahapatra, R., & Wasnik, P. N. (2023, December). XAIA: An Explainable AI approach for classification and analysis of blood anemia. In *2023 OITS International Conference on Information Technology (OCIT)* (pp. 88-93). IEEE.
[19] Abdullah, A., Saddiqi, H. A., Qasim, M., Khitab, A., Khan, M., & Ahmad, S. (2025). Anemia prediction using gene expression programming (GEP) and explainable artificial intelligence approaches. *Computers in Biology and Medicine*, *196*, 110856.
[20] Darshan, B. D., Sampathila, N., Bairy, M. G., Belurkar, S., Prabhu, S., & Chadaga, K. (2024). Detection of anemic condition in patients from clinical markers and explainable artificial intelligence. *Technology and Health Care*, *32*(4), 2431-2444.
[21] Végh, L., Takáč, O., Czakóová, K., Dancsa, D., & Nagy, M. (2024). Evaluating Optimizable Machine Learning Models for Anemia Type Prediction from Complete Blood Count Data. *International Journal of Advanced Natural Sciences and Engineering Researches*, *8*(7), 108-119.
[22] Gómez, J. G., Parra Urueta, C., Álvarez, D. S., Hernández Riaño, V., & Ramirez-Gonzalez, G. (2025, February). Anemia Classification System Using Machine Learning. In *Informatics* (Vol. 12, No. 1, p. 19). MDPI.
[23] Mugdha, A. G., Pinki, F. T., & Talukdhar, S. K. (2023, December). Hemoglobin estimation and anemia severity prediction using machine learning algorithms. In *2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI)* (pp. 1-6). IEEE.
[24] Noviandy, T. R., Idroes, G. M., Suhendra, R., Bakri, T. K., & Idroes, R. (2024). Advanced Anemia Classification Using Comprehensive Hematological Profiles and Explainable Machine Learning Approaches. *Infolitika Journal of Data Science*, *2*(2), 72-81.
[25] Ponnusamy, U., Darshan, D., & Sampathila, N. (2023, November). Approaching explainable artificial intelligence methods in the diagnosis of iron deficiency anemia using blood parameters. In *2023 International Conference on Recent Advances in Information Technology for Sustainable Development (ICRAIS)* (pp. 201-206). IEEE.
[26] Qadah, T., & Munshi, A. (2024). Synthesis and Prediction of Anemia from Multi-Data Attribute Co-Existence. *IEEE Access*.
[27] Hasan, M., Tahosin, M. S., Farjana, A., Sheakh, M. A., & Hasan, M. M. (2023, May). A harmful disorder: Predictive and comparative analysis for fetal Anemia disease by using different machine learning approaches. In *2023 11th International Symposium on Digital Forensics and Security (ISDFS)* (pp. 1-6). IEEE.
[28] Zhang, A., Lou, J., Pan, Z., Luo, J., Zhang, X., Zhang, H., ... & Chen, L. (2022). Prediction of anemia using facial images and deep learning technology in the emergency department. *Frontiers in Public Health*, *10*, 964385.
[29] Yang, J. Y., Lee, T. C., Liao, W. T., & Hsu, C. C. (2023). Multi-head self-attention mechanism enabled individualized hemoglobin prediction and treatment recommendation systems in anemia management for hemodialysis patients. *Heliyon*, *9*(2).
[30] Sasikala, C., Ashwin, M. R., Dharanessh, M. D., & Dhanabalan, M. (2022, April). Curability prediction model for anemia using machine learning. In *2022 8th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-7). IEEE.