RESEARCH ARTICLE

Developing a Transparent Anaemia Prediction Model Empowered With Explainable Artificial Intelligence

MUHAMMAD SAJID FAROOQ[1], MUHAMMAD HASSAN GHULAM MUHAMMAD[2], OUALID ALI[3], ZAHID ZEESHAN[4], MUHAMMAD SALEEM [5], MUNIR AHMAD [6,7], (Senior Member, IEEE), SAGHEER ABBAS[8], (Senior Member, IEEE), MUHAMMAD ADNAN KHAN [9], (Senior Member, IEEE), AND TAHER M. GHAZAL [10], (Senior Member, IEEE)

[1]Department of Cyber Security, NASTP Institute of Information Technology, Lahore 54000, Pakistan
[2]Department of Computer Science, IMS Pak-AIMS, Lahore 54000, Pakistan
[3]College of Arts and Science, Applied Science University, Manama, Bahrain
[4]Department of Computer Science, Lahore Garrison University, Lahore 54000, Pakistan
[5]School of Computer Science, Minhaj University at Lahore, Lahore 54000, Pakistan
[6]College of Informatics, Korea University, Seoul 02841, Republic of Korea
[7]School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan
[8]Department of Computer Science, Prince Mohammad Bin Fahd University, Al Khobar, Dhahran 34754, Saudi Arabia
[9]Department of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si 13557, South Korea
[10]Research Innovation and Entrepreneurship Unit, University of Buraimi, Buraimi 512, Oman

Corresponding authors: Muhammad Adnan Khan (adnan@gachon.ac.kr) and Taher M. Ghazal (taher.ghazal@ieee.org)

ABSTRACT The worldwide health epidemic of anaemia which is a condition with low levels of red blood cells or haemoglobin requires accurate prediction models to act promptly and improve patient outcomes because it is widespread and has different causes. The effective management of anaemia is piled with obstructions, which may include the variability of diagnostic criteria, the resource limitations of healthcare, and the multifactorial nature of the disease including nutritional deficiencies, chronic disease, and genetic factors. Conventional anaemia prediction models, that predominantly rely on statistics and are trained on clinical risk scores, are frequently incapable of providing practical solutions and meaningful insights into anaemia diagnosis. There is a growing interest in focusing on Artificial Intelligence (AI) use for anaemia prediction, however, traditional AI models (black boxes) lack transparency, which causes doctors not to pick them up for practical usage. Actionable insights that are enabled by transparent AI models (white boxes) based on the explainable AI methodologies reveal the rationales of the prediction, clarify the features that are responsible for them, and help clinicians and healthcare providers. In this research work, a transparent anaemia prediction model (white box) empowered by explainable AI techniques is proposed to address the limitations of black boxes in terms of transparency. The proposed model utilizes machine learning algorithms such as Support Vector Machine (SVM), Decision Trees, K-Neighbors Classifier, and Gradient Boosting Classifier, enhanced with Explainable AI (XAI) techniques like SHAP and LIME. With the integration of explainable AI techniques like SHapley Additive exPlanations (SHAP), and Local Interpretable Model-agnostic Explanations (LIME), the proposed model offers insights into the underlying factors influencing anaemia predictions. The proposed model, significantly, represents exceptional growth in the healthcare sector and helps in bridging the gap between predictive performance and clinical interpretability, thus improving patient care and disease management strategies. The model simulation results are showing promising results in terms of the accuracy (98.13%) and the miss-rate (1.87%) which are the superior performance compared to the previous published approaches.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Ni.

INDEX TERMS Anaemia prediction, explainable artificial intelligence, SHAP, LIME.

I. INTRODUCTION

Anaemia is a very common blood disorder that is found in 1.6 million people around the globe [1]. As reported by [2], the worldwide incidence of anaemia in all age groups in 2019 was 22.8% (95% CI: 22.6–23.1). Considering it is a major health burden worldwide, anaemia diagnosis and treatment with short turnaround time should be prioritized since they can make global health status better in the long run. The World Health Organization (WHO) defines anaemia [3] occurs due to a shortage of red blood cells in the human body which also causes a decrease in the ability of the blood to carry oxygen. Gene aspects, for instance, altitude, age, gender, smoking, as well as health conditions like pregnancy are the main causes of the development of this disease. According to the WHO, 42% of children under the age of 6 and 40% of pregnant women universally suffer from anaemia. The iron shortage is a general reason for anaemia, impacting approximately 33% of the world's population [4].

Iron deficiency is a prominent reason for anaemia; however, other expressions of iron deficiency are compounded by folic acid deficiency, inflammation, lack of vitamins A and B12, parasitic, and hereditary infections like thalassemia. The symptoms [5] that are associated with the disease are weakness, fatigue, drowsiness, and dizziness. The majority of anaemia cases are unnoticed due to the painlessness of the signs a general lack of information in poor countries, and incomplete healthcare services.

The ancient technique of diagnosing anaemia is by assessing the haemoglobin concentration in blood using the hematocrit test which involves drawing blood samples for analysis in triage. The invasive nature of this procedure can be time-consuming and puts healthcare employees at risk of blood-borne diseases.

Recently, there has been the widespread use of non-invasive medical devices like smartphone and smartwatches which helps to supervise several health factors like heart rate, oxygen levels, sleep patterns, etc [6], [7], [8]. Thus, there is a need for improved and minimally invasive techniques that can identify anaemia without the complications that come with invasive detection methods. Anaemia, characterized by low red blood cell levels or reduced haemoglobin, demands precise prediction models for effective healthcare management. However, managing anaemia effectively faces several challenges, including diagnostic variability, resource constraints in healthcare, and the multifactorial nature of the disease (encompassing nutritional deficiencies, chronic conditions, and genetic factors).

The complexity of conventional AI models, commonly referred to as "black boxes," curtails their use by medical professionals. Medical practitioners are wary of AI algorithms not being able to explain the decisions made to the patients, which further interferes with proper care given to patients [9]. The black-box nature of AI algorithms renders it challenging to understand how conclusions or outcomes are made, thus lessening the transparency of diagnostic and therapy equipment [10]. In medicine, both AI and ML help in data analysis and diagnosis and are used in the optimization of patient care. However, these traditional AI models may give black-box results which clinicians then may not understand and hence cannot trust the AI system. This has given rise to the concept of XAI to increase the level of interpretability to show that the models are not 'black-box' systems in a healthcare setting.

The artificial intelligence field becomes transparent and interpretable through transparency and interpretability. XAI equips us with deciphering tools that help us use machine learning models [11], [12], [13] to understand how they work and the way they make predictions. Through the demystification of the decision-making process, XAI empowers users through its transparent and auditable feature. It efficiently fills the blank between predictive accuracy and clinical interpretability and that is very crucial in healthcare applications [14].

There are complicated questions, such as high accuracy and in-depth understanding of the causes that stand in the way of developing the anaemia prediction models [15]. Come to the mind of white box models which together with "black boxes" draw the eye and at the same time reveal the hidden systems [16]. This way of modelling might stand as a great source of validation of predictions to be used by clinicians even to find out biases. While exposing the fact, scientific research is tightly tied to public health understanding, decision-making and patient outcomes. In white boxes' concept which emphasizes transparency, revealing their inner workings while predicting Anaemia takes the main focus [17]. The model may use explainable AI, which implies a comprehensible mechanism of decision-making procedure.

SHAP values use the principal technique to attribute the predictions from particular features to the machine learning model. SHAP achieves interpretability by assigning each feature with a value that shows the influence of the feature on the model's outcome [18], [19]. To exemplify this, SHAP reveals which genes lead to anaemia like haemoglobin level, age and dietary habits that are strongly associated with the outcome. Realizing these contributions allows clinicians to grasp the waves that are the basis of these predictions.

LIME follows up a different way of making complex models fit with simplified and interpretable ones. It focuses on the capabilities of giving locally sufficient explanations based on each prediction [20], [21]. In the instance of anaemia detection, it can be invaluable to doctors. It gives them a picture of this patient's evaluation in the category of anaemia by accurately portraying the characteristics. It gives localized insights to support making the right decisions and raising the quality of healthcare.

To mitigate these issues, both SHAP and LIME now have two important roles being played to improve the transparency of the anaemia prediction model using explainable AI techniques. The use of SHAP and LIME integration in this model may provide insight into the hidden factors contributing to anaemia predictions, allowing for meaningful interpretations and thus improving healthcare service delivery to doctors and patients. This research in healthcare AI is a decisive growth towards the combination of what predicts and how it interprets clinically.

II. LITERATURE REVIEW

In the previous few years, the healthcare predictive model achievements have been notably improved by employing the implementation of artificial intelligence (AI) methods. The major subject to be worked on is the determination of anaemia, a condition with low amounts of red blood cells that can have worse consequences if left untreated. Although the models of AI have demonstrated the capability of managing anaemia, however, the absence of transparency and interpretability in these models leads to their incapability for usage in the healthcare domain at a wide scale. Several research studies assess the effectiveness of different machine learning and AI algorithms [22], [23], notably Random Forest, SVM, and others, for the prediction of anaemia in patients.

In their research [16], the researchers studied the involvement of intelligent software in predicting anaemia as this is a serious health problem in developing countries. Different machine learning algorithms like decision trees, random forests, Naïve Bayes, XGBoost, LightGBM, Lasso, Ridge, and Elasti-net were applied for diagnosing anaemia. The focus of the research was on applying threshold-based feature selection Pearson correlation to pick up feature points with strong associations. The LightGBM algorithm has a better and higher capability in predicting anaemia than other algorithms. The research has employed the Skilcarslan dataset to ascertain the accuracy of the predictions.

In several studies, a diffuse method to standard haemoglobin using blood flow in the palm has been proposed. Santra et al. [24] created a machine vision approach which can automate the acquiring of video of the palm before and after the application of a cuff and they also measure the degree of skin redness to predict blood haemoglobin content with 91% accuracy. Granados-Castro, et al. proposed a method using Multispectral Imaging as well as a variant of linear unmixing which was an alternative to the measurement of blood perfusion in the palm [25]. In this research [26], an approach was given based on multispectral imaging and a variant of linear unmixing, and this is a method of estimation of blood perfusion parameters in the hand palm other than those practices. But with these techniques, contact with patients is required, and the dimensions are only single-point estimation.

Research [27] reported the mathematically based back-propagation rules for the adjustment of the weights in the neural network-based anaemia prediction model. The model

was then able to get the best results of 97.00% accuracy with a sensitivity of 99.21% and a specificity of 95.42% on the dataset created. In [28], the researchers shared Convolutional Neural Networks, Logistic Regression, and Gaussian Blur algorithm to establish a conjunctiva detector and an anaemia detector, based on a server of Fast API linked to a mobile app frontend which was built with React Native. This non-invasive method of detecting anaemia from conjunctival images of the eyes was proposed to build a diagnosis model with 90% sensitivity, 95% specificity, and 92.50% accuracy. The model [29] which was trained on 99 test subjects utilizing k-Folds validation did well and scored 93% accuracy as well as a noninvasive automated technique that is cost-effective was developed.

Anaemia remains one of the key health issues, particularly in settings that have limited medical services and clinical labs. To address the problems of late anaemia diagnosis and intrusive diagnostic procedures, the scientists [30] came up with a non-invasive approach to anaemia detection based on the examination of conjunctival images of the eye. The research evaluated various model methods to eliminate mistakes while classifying anaemic and healthy patients. The Support Vector Machine armed with MobileNetV2 turns out to be the most efficient option, with the performance of 93% accuracy, 91% sensitivity, and 94% specificity. This approach could be able to classify patients well between the healthy group and the anaemic group. This technique offered a very auspicious approach for early anaemia detection, based on the conjunctiva eyes. It can facilitate early identification of patients with anaemia in the healthcare system. These three aspects like accuracy, sensitivity, and specificity of SVM and MobileNetV2 classification technique make it a perfect detector of anaemia.

III. LIMITATIONS OF THE PREVIOUSLY PUBLISHED APPROACHES

Through a comprehensive literature survey depicted in Table 1, this review highlights the benefits and drawbacks of predictive modelling and emphasizes the role of explainability in the acceptance of this innovative tool by healthcare professionals. Table 1 provides an in-depth analysis of different studies that employed AI models for the detection of anaemia based on methods such as image processing, neural networks and support vector machines. These techniques quite often attain high degrees of accuracy, yet still encounter problems of image quality, individual differences, and interpretability.

XAI has emerged as a solution to address the opacity of black-box models and offer transparency in how AI systems make decisions. It enhances transparency (white box), allows clinicians to understand AI decisions (black box) by revealing the reasoning behind predictions, and aids early detection, ultimately improving patient care.

IV. CONTRIBUTION OF THE PROPOSED METHODOLOGY

The significant contribution of this work is the development of an explicit and interpretable anaemia prediction model

TABLE 1. Comprehensive analysis of the previously published approaches.

Reference	Model Used	Predictive model	XAI (white-box)	Outcomes	Accuracy (%)	Limitations
Santra et al., 2017 [24]	Machine vision-based approach	✓	×	Predicts haemoglobin content	91	The non-invasive method may not be as accurate as direct blood sampling, influenced by skin colour and flow variations.
Jain et al., 2020 [27]	Neural network-based model	✓	×	Predict anaemia	97	This method for detecting anaemia may be limited by variations in image quality and individual differences in eye characteristics.
Appiahene et al., 2023 [28]	Convolutional Neural Networks, Logistic Regression, Gaussian Blur algorithm	✓	×	Developed conjunctiva detection model and anaemia detection model.	92.50	the need for image quality control and the reliance on a specific technology stack for deployment.
Bauskar eta l., 2019 [29]	SVM	✓	×	The SVM attained 93% accuracy	93	This method may be limited by variations in image quality and individual differences in eye characteristics.
Wulandari et al., 2024 [30]	Support Vector Machine (SVM), MobileNetV2	✓	×	Identified healthy and anaemic patients.	93	Relies on eye conjunctiva images, limiting generalizability to other data sources, and has a black-box nature.
Agrawal et al., 2021 [31]	ResNet, MobileNetV2 and DenseNet	✓	×	MobileNetV2 had a higher performance	74	Image quality variations
Magdalena et al., 2022 [32]	AlexNet	✓	×	With a 0.0001 learning rate, 94% accuracy using the Adam optimizer achieved	94	Data limitations, generalization challenges, and interpretability concerns.
Peksi et al., 2021 [33]	Naïve Bayes	✓	×	The Naïve Bayes attain an accuracy of 90%	90	Practical implementation
Roychowdhury et al., 2017 [34]	SVM, K-NN, DT, ANN	✓	×	86% classification accuracy was attained	83	This method is not transparent or easily interpretable.
Noor et al., 2019 [35]	MLR, DT, SVR	✓	×	Decision Tree gain a higher accuracy of 88.99%	88.99	Small sample size, reliance on a single imaging modality, and lack of comparative analysis.
Tamir et al., 2017 [36]	RGB Thresholding Algorithm	✓	×	This method achieved an accuracy of 78.90%	78.90	Relies solely on image processing and thresholding, which may not account for variations in lighting conditions and other factors.
Noor et al., 2019 [37]	DT, SVM, K-NN	✓	×	An accuracy of 82.61% was attained by the DT algorithm	82.61	Limited sample size
Rivero-Palacio et al. 2021 [38]	YOLO v5 Neural Network	✓	×	Accuracy of 93% for PC, and 80% for phone apps.	93	Requires a reliable internet connection for real-time processing and analysis
Dimauro et al., 2023 [39]	SVM, random forest, k-NN, MobileNet-V2	✓	×	The MobileNet-V2 achieved an accuracy of 91%.	91	The study has limitations related to dataset bias, non-invasive accuracy, sample size, and external validation.
Dimauro et al., 2023 [40]	Scleral segmentation algorithm	✓	×	Precision 82.53%, 84.10% with recall, and f1, and 86.40% with f2	73.58	inner workings and decision-making process of the method are not transparent or easily interpretable
Muthalagu et al., 2018 [41]	Elman neural system	✓	×	This system achieved 91.3% accuracy.	91.3	inner workings and decision-making process of the method are not transparent or easily interpretable

that complements the existing limitations of the prior models, which include the lack of interpretability in "black box" AI models, a small and specific dataset that would not allow generalizing, as well as challenges with real-life application in the clinics. By integrating both SHAP and LIME, the model offers insightful and helpful interpretations for clinicians, increases the outlook with a diverse pool of data, and operates at a higher level of credibility since the prediction process is transparent and interpretable.

V. PROPOSED METHODOLOGY

In past years, professionals in the field of healthcare and the healthcare industry, in general, have had to deal with a large number of concerns related to anaemia, a disease that is characterized by either a low volume of red blood cells or low hemoglobin. Anaemia represents a severe condition with the potential for serious effects such as fatigue, weakness, decreased cognitive performance, and pregnancy complications. Detecting promptly and managing the anaemia effectively are considered crucial to avoid bad health outcomes and giving patients the best quality of life. However, conventional approaches to anaemia detection are characterized by their limited interpretability and transparency, which ultimately hinder healthcare professionals' comprehension of the driving factors behind the diagnosis. The incorporation of XAI in healthcare promotes the interpretability of predictive models for anaemia diagnosis, which helps clinicians make judgments and strengthens the trust that patients have in their healthcare providers.

The proposed XAI-based anaemia prediction model in Figure 1 portrays such a novel concept in healthcare, revolutionizing anaemia detection as well as addressing the healthcare specialists' concerns.
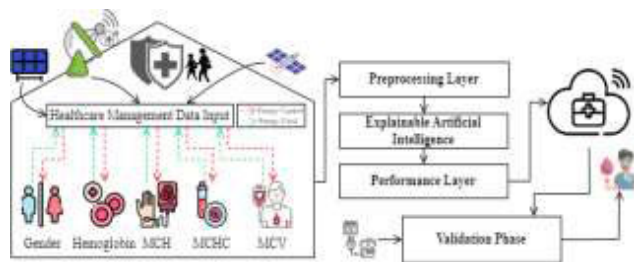

FIGURE 1. Abstraction of the proposed model.

Figure 1 shows the abstraction of the proposed model for anaemia prediction, which consists of the input layer, a preprocessing layer, a training layer where explainable AI is applied, a performance layer, and a validation phase. The input layer empowers healthcare devices to increase efficiency by exchanging sensory data wirelessly and getting rid of cellular communication costs. This layer with anaemia parameters presents an opportunity for data collection and transition, which is efficient and effective for healthcare management. The sensed data that are received by the input layer are then passed through the preprocessing layer, training layer

(XAI) and performance layer. The output of the performance layer is saved on cloud data storage. In the validation phase, the learned data is imported from the cloud storage that is used further to predict the development of anaemia.
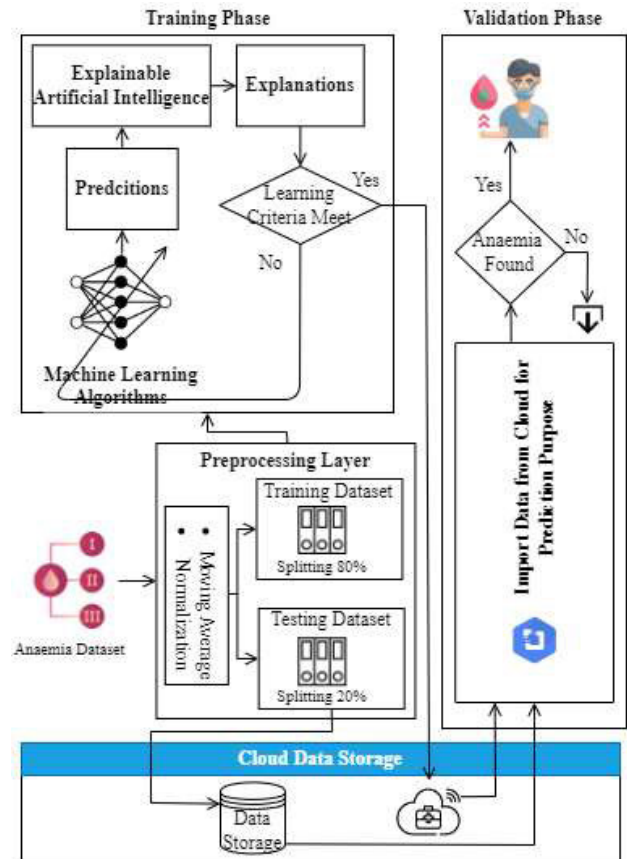

FIGURE 2. Proposed model for anaemia prediction using XAI.

Figure 2 represents the proposed model for anaemia prediction which is comprised of the training and validation phases to predict eye disease with the help of the XAI method. Firstly, data collection of the patient's entire record is the crucial operation for the preprocessing layer. The collection of patient records from anaemia parameters is the pivotal link between the patient information and the analysis stage to integrate insights into the predictive model. The data is then sent to the preprocessing supporting the healthcare sector by using normalization and moving average techniques to get rid of the noise. The preprocessed data is then constantly split into training (80%) and testing (20%) datasets that aim at supporting the prediction process so that it can have a better performance/accuracy result. The dataset features are given in Table 2.

Subsequently, this data is directed to the training layer, whereas the testing (20%) dataset is stored in the cloud data storage. In the training layer, several machine-learning algorithms (Decision Tree, K Neighbors Classifier, SVM, and Multilayer Perceptron) are applied for improved prediction accuracy. These algorithms are responsible for predicting

Sr. No.	Feature	Data type
1	Gender	Integer
2	Haemoglobin	Float
3	MCH	Float
4	MCHC	Float
5	MCV	Float
6	Result	Integer

patterns based on multiple machine learning algorithms like Decision Tree, K Neighbors Classifier, SVM, Multilayer Perceptron etc. In this proposed model, the SVM algorithm is used for prediction. As we know the line equation is:

$$k = mx + b \qquad (1)$$

In Equation 1, m represents the slope of the line and b is the intercept. Therefore, for the equation of a hyperplane in higher dimensions, we can write:

$$w.x + b = 0 \qquad (2)$$

where w is the weight vector and x is the input feature vector. The direction of the vector $x = (x_1, x_2, \ldots, x_n)$ is determined by the weight vector w. The vector is normalized using the Euclidean norm:

$$|| w || = \sqrt{w_1^2 + w_1^2 + \ldots + w_n^2} \qquad (3)$$

We can also express the direction cosines for each dimension as:

$$cos(\theta_1) = \frac{w_1}{||w||}, \quad cos(\theta_1) = \frac{w_1}{||w||}, \ldots$$

The dot product of vectors w and x can be written as:

$$w.x = \sum_{i=1}^{n} w_i x_i \qquad (4)$$

The SVM algorithm aims to find a hyperplane that maximizes the margin between two classes of data points. The functional margin of a dataset is given by:

$$\gamma = \min_{i=1,\ldots,n} (y_i(w.x_i + b))$$

where y_i is the label of the i-th data point. The goal of SVM is to maximize the geometric margin γ, which represents the distance between the support vectors and the hyperplane.

To solve the optimization problem, SVM uses Lagrange multipliers. The Lagrangian function is given by:

$$L(w, b, a) = \frac{1}{2}||w||^2 - \sum_{i=1}^{n} \alpha_i(y_i(w.x_i + b) - 1) \qquad (5)$$

where α_i are the Lagrange multipliers. By taking the gradient of the Lagrangian concerning w and b, we obtain the following conditions:

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i \qquad (6)$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \qquad (7)$$

Substituting these into the Lagrangian gives the dual form of the optimization problem:

$$\max_a \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i.x_j)$$

Subject to the conditions $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i y_i = 0$ (8)

Once the optimal $\boldsymbol{\alpha}$ values are obtained, we can compute the intercept b using the support vectors:

$$b = y_i - w.x_i \qquad (9)$$

Finally, the SVM model makes predictions based on the following decision function:

$$f(x) = sign(w.x + b) \qquad (10)$$

where $f(x) = 1$ predicts the presence of anaemia, and $f(x) = 0$ predicts its absence.

The predictions from the SVM algorithm are then passed to XAI techniques such as SHAP and LIME to provide human-friendly interpretations, ensuring transparency (white box) and interpretability of the results.

The primary goal of XAI is to build trust in machine learning (black box) through the importance of local and global variables using post-hoc explanations. It focuses on the trial of deciphering the mysteries of the black boxes, but it also suggests accountable AI because it can aid in creating transparent models. Enabling explainability in ML aims to make it simpler for end-users and other stakeholders to identify the reasoning behind algorithmic decisions.

The model-agnostic interpretation approaches, the current advancements in ML, are used to explain the multi-faceted models while retaining a good prediction performance. The model-agnostic interpretation is more flexible than the model-specific explanation method as it differentiates the model from the explanations. There are two types of model-agnostic interpretation methods: local and global explanations.

LIME as a local LIME explanation approach is most often used. Although PDP and SHAP are the most prevalent interpretation methods on a global-scale level. Employing LIME as a local explanation model, the local surrogate models are designed to be capable of explaining the complex models. It presents the new dataset by manipulating the original data first. Using the new dataset to train, it then builds an interpretable model, such as an SVM. Finally, the black box model's prediction accuracy is associated with the interpretable model's accuracy. LIME is distinct as trails:

$$\gamma(x) = argminL(f, g \atop g \in G} \pi_x) + \Omega(g) \qquad (11)$$

The loss function L is used to determine how closely the interpretable model g matches the predictions of the original complicated model f. f is the original complicated model. g represents the interpretable model for the instance x (e.g., logistic regression). G represents the family of the interpretable models. π_x represents the closeness of the sampled

cases to the instance x. $\Omega(g)$ is the complexity of the model g.

PDP shows that a single feature has only a minor impact on the predicted outcomes for a complex ML model. The PDP is a representation of the relationship joining the output and input. The partial dependence function $\widehat{f_{x_s}}$ distinct as:

$$\widehat{f_{x_s}}(x_s) = \frac{1}{n}\sum_{i=1}^{n}\widehat{f_{x_s}}\left(x_s, x_c^i\right) \qquad (12)$$

where $\widehat{f_{x_s}}(x_s)$ is the partial function showing the global association of an input feature through the predicted result. s is a feature set comprising only one or two features, x_s represents the set of features is to be plotted by $\widehat{f_{x_s}}(x_s)$, x_c states the real feature standards from the dataset for the features we are not involved in. ns the number of instances of the dataset.

The Shapley values approach, which is a feature impact assessment technique using SHAP, is used to measure the impact of a feature on a complicated model. Shapley value is precisely the arithmetic mean of the marginal contributions. Concerning any possible coalition, the feature value is influential on prediction. The Shapley value formula is shown as follows:

$$\varphi_j(x) = \sum_{s\subseteq\{x_1,x_2,...,x_m\}\setminus\{x_j\}} \frac{|s|!\,(m-|s|-1)!}{m!} \\ \times (val(s\cup\{x_j\}) - val(s))) \qquad (13)$$

where $\varphi_j(x)$ is the Shapley value of x_j, x_j signifies a feature value, s is a feature subclass of the model, m shows the number of features, val is the prediction for feature standards in the set s.

After obtaining the explanations from predicted outcomes, then it is further checked whether the learning rate is found or not. If found, the trained explainable patterns are stored on the cloud, whereas in the case of No, the fine-tuned model is retained, and so on.

Then, in the validation phase, the trained explainable model is stored on the cloud, and the testing dataset is imported for post-training purposes, ensuring performance and robustness. Subsequently, the testing data is utilized to evaluate the model's performance objectively, delivering a final assessment of its effectiveness before deployment in real-world scenarios. Then it is checked whether the anaemia disease is found or not. If "Yes," the message is shown that the disease is found, whereas in the case of No, the process is discarded.

VI. SIMULATION RESULTS

Anaemia is a prevalent haematological disorder marked by a shortage of red blood cells or haemoglobin concentration in the blood and presents complex clinical symptoms that require early prediction. Despite encountering various obstacles, such as the intricacy of anaemia's clinical indication and the challenge of integrating diverse data sources, this research proposed an XAI-based anaemia prediction model to revolutionize the healthcare sector. This proposed model aims to validate and showcase the effectiveness and transparency of

the predictive model for anaemia which is applied to a dataset (80% for training, 20% for testing), which is augmented with XAI.
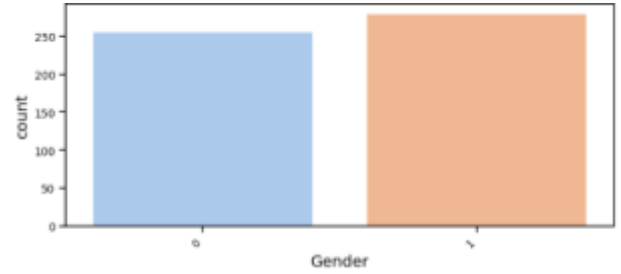

FIGURE 3. Distribution of gender over the dataset.

Figure 3 illustrates the gender distribution within a dataset using a count plot, set in a visually appealing pastel palette. This figure displays the frequency of each gender category from the "Gender" column, with gender labels on the x-axis for improved readability, and is crucial for evaluating gender diversity within the dataset.
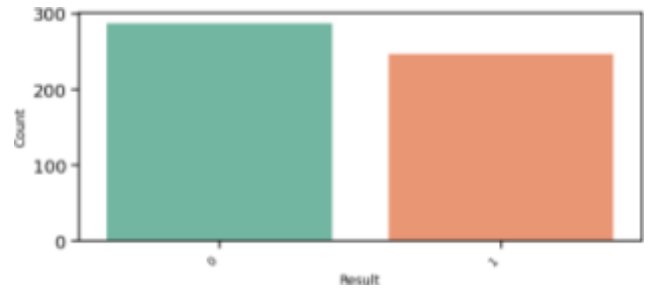

FIGURE 4. Distribution of MCHC over the dataset.

Figure 4 displays the distribution of "Result" data across a dataset, using a palette count plot for vibrant visualization, the x-axis labels are rotated for readability, and to enhance clarity.

Figure 5 displays a seaborn pairplot, clearly differentiating data categories "0" and "1" with blue and orange colors, respectively. It uses kernel density estimates to show variable distributions and sets plot height to 2 units for concise visualization. This offers an insightful view into the dataset's variable interactions and correlations based on the "Result" categories, making patterns and relationships easily discernible.

Figure 5 displays a seaborn pairplot, clearly differentiating data categories "0" and "1" with blue and orange colors, respectively. It uses kernel density estimates to show variable distributions and sets plot height to 2 units for concise visualization. This offers an insightful view into the dataset's variable interactions and correlations based on the "Result" categories, making patterns and relationships easily discernible.

Figure 6 provides a clear comparison of standardized features within a dataset through boxplots. Graphically displaying provides an efficient understanding of how features
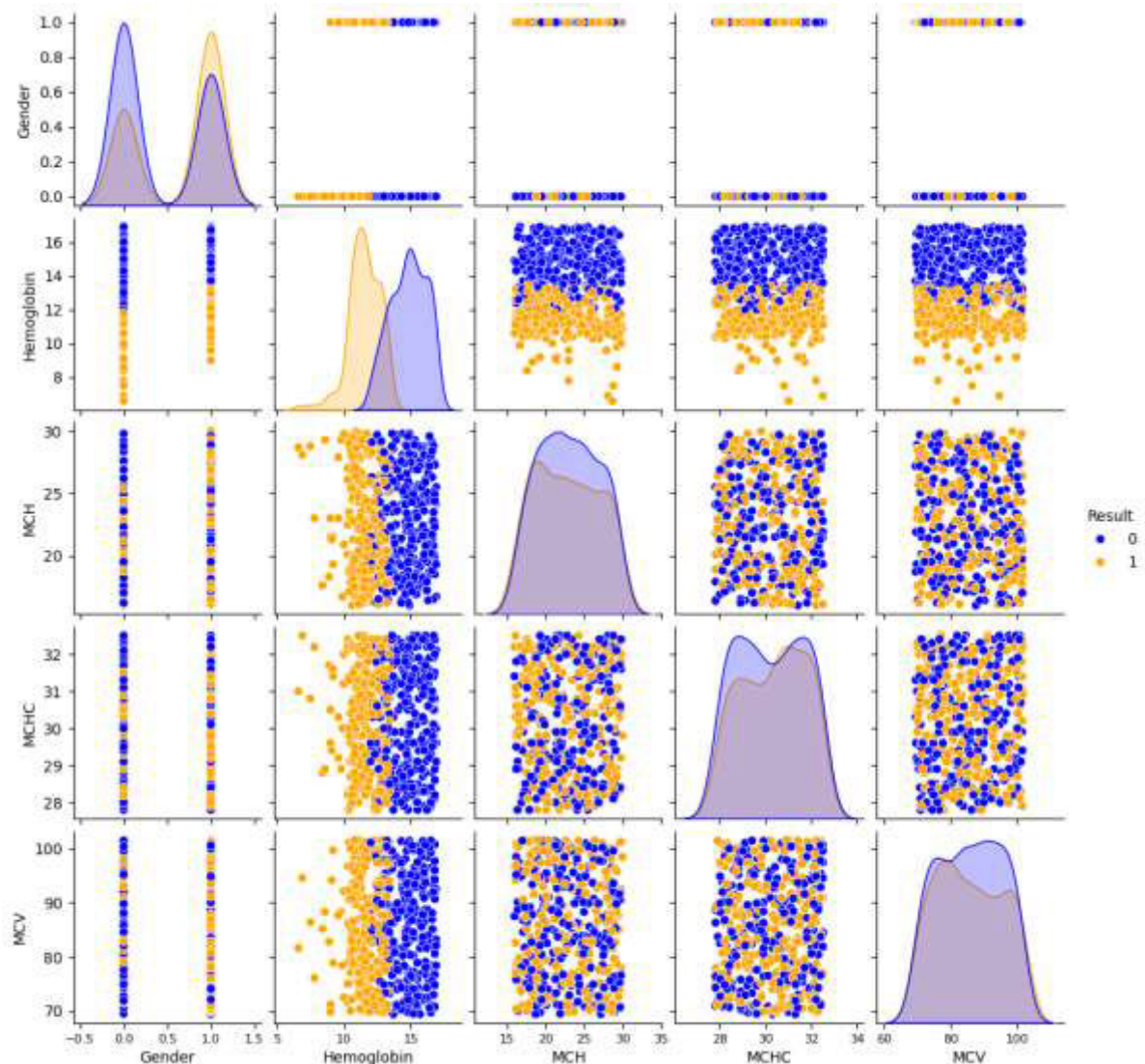
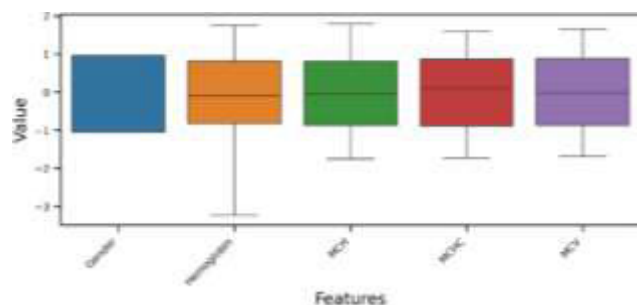**FIGURE 5.** Variable interactions and distributions by result category.

FIGURE 6. Comparative distribution of standardized features in the dataset.

FIGURE 7. Comparative analysis of feature distributions in the dataset.

with their central tendencies vary, and as an analytical tool, it helps to notice different patterns, anomalies, and outliers making the data analysis and pattern recognition task easier.

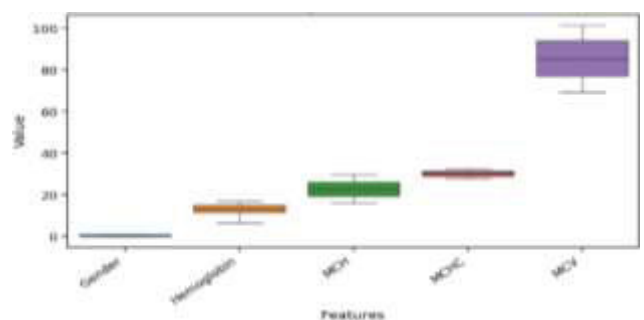Figure 7 provides a representative diagram for the various distribution and diversity of values across different features of the dataset. Colors are used to differentiate the boxes depicting each variable. The key goal of this method is to visualize the mixture of the range, median and outlier data, which enables a fast comparison between various attributes. Color-coded differentiation emphasizes visual clarity, enabling

TABLE 3. Performance evaluation of the proposed model using multiple machine learning algorithms with statistical measures.

|		Accuracy (%)	Sensitivity (%)	Specificity (%)	Miss-Rate (%)	Fallout (%)	LR+ve	LR-ve	Precision (%)	NPV (%)
Decisoin Tree	Training	98.36	97.40	99.49	1.64	0.51	190.98	0.016	99.56	97.01
	Validation	89.72	94.83	83.67	10.28	16.33	5.80	0.122	87.30	93.18
KNeighbors	Training	97.89	99.54	96.15	2.11	3.85	25.85	0.021	96.46	99.50
Classifier	Validation	73.83	72	78.13	26.17	21.88	3.29	0.334	88.52	54.35
SVM	Training	99.07	100	98.06	0.93	1.94	51.54	0.009	98.23	100
	Validation	98.13	100	95.83	1.87	4.17	23.98	0.019	96.72	100
Gradient Boosting	Training	95.55	94.04	97.40	4.45	2.60	36.17	0.045	97.79	93.03
Classifier	Validation	84.11	78.21	100	15.89	0	-	0.158	100	63.04
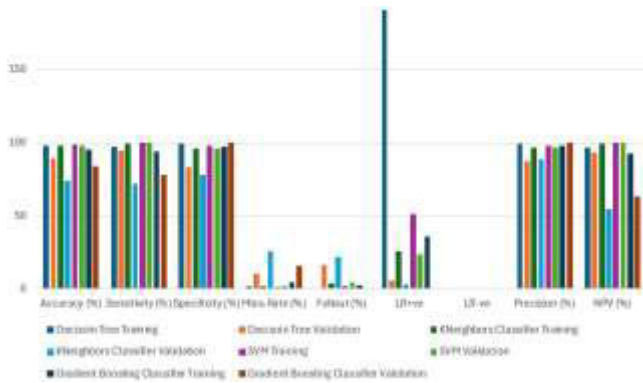
FIGURE 8. Graphical representation of the performance evaluation of the proposed model using multiple machine learning algorithms with statistical measures.

quick spotting of patterns, deviations, and possible areas of interest within the data. This visualization is meant to assist users in having a better appreciation of the data structure which, in turn, allows them to make informed data-driven decisions and analysis.

Multiple machine learning algorithms have been applied to this dataset, with the training and validation of their performance metrics: Accuracy, Sensitivity, Specificity, miss - rate, fallout, Positive Likelihood Ratio (LR+ve), Negative Likelihood Ratio (LR-ve), precision, and Negative Predictive Value (NPV) are provided in Table 3 below.

Table 3 and Figure 8 provide the performance of the mentioned machine learning algorithms (Decision Tree, K-Neighbors Classifier, SVM, and Gradient Boosting Classifier) on training and validation datasets across multiple metrics. The SVM model achieves a very high accuracy, reaching 99.07% on training and 98.13% on validation, demonstrating the robustness. The table reflects each model's strengths and areas for improvement, highlighting the importance of balancing various metrics to achieve optimal model performance.

SHAP and LIME are interpretative tools used to demystify predictions made by complex ''black box'' machine learning models, particularly useful in healthcare for conditions like anaemia. SHAP assigns importance values to features for each prediction, using game theory to quantify each feature's contribution. LIME explains individual predictions by creating simpler, local models, offering insights into

specific patient outcomes. Together, they enable healthcare professionals to understand the factors influencing a model's prediction of a patient's disease progression, facilitating personalized and informed treatment decisions.
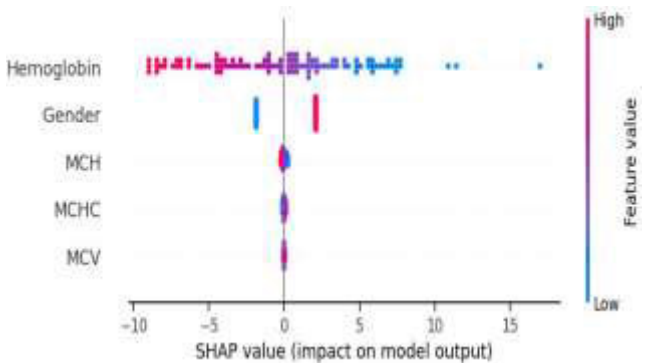

FIGURE 9. Feature impact analysis with SHAP on SVM predictions.

Figure 9 displays a SHAP summary plot that visualizes the impact of each feature on the predictions of an SVM classifier for test data. This plot presents a summary of features, arranged in the order of their importance and illustrating the spread of their impact on model predictions. The plot is used to depict both the direction and magnitude of each feature's influence, and it clearly shows which features are the strongest drivers toward a specific outcome. This phenomenon is particularly important for understanding the factors that the model uses as input, for a transparent visualization of the model's decision-making process, and for identifying features necessitating improvement and for the model domain.

Figure 10 is an illustration of a LIME explanation for the SVM classifier's prediction on an experiment instance about the difference between anaemia and not-anaemia diagnosis. This visualization illustrates how each feature impacts the model's decision-making process, including positive or negative contribution factors. It offers a clear, interpretable breakdown of the model's reasoning, enhancing transparency and understanding of its predictive behaviour on individual cases.

Table 4 presents the comparison of the performance of the proposed model with previously published approaches [24], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37],
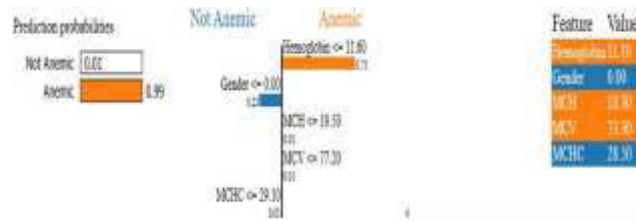
FIGURE 10. LIME explanation of SVM prediction for anaemia diagnosis.

TABLE 4. Comparison of the proposed anaemia prediction model with previously published approaches.

Reference	Model Used	Accuracy (%)	Miss-rate (%)
Santra eta l., 2017 [24]	Machine vision-based approach	91	9
Jain et al., 2020 [27]	Neural network-based model	97.00	3
Appiahene et al., 2023 [28]	Convolutional Neural Networks, Logistic Regression, Gaussian Blur algorithm	92.50	7.5
Bauskar eta l., 2019 [29]	SVM	93	7
Wulandari et al., 2024 [30]	Support Vector Machine (SVM), MobileNetV2	93	7
Agrawal et al., 2021 [31]	ResNet, MobileNetV2 and DenseNet	74	26
Magdalena et al., 2022 [32]	AlexNet	94	6
Peksi et al., 2021 [33]	Naïve Bayes	90	10
Roychowdhury et al., 2017 [34]	SVM, K-NN, DT, ANN	83	17
Noor et al., 2019 [35]	MLR, DT, SVR	88.99	11.01
Tamir et al., 2017 [36]	RGB Thresholding Algorithm	78.90	21.1
Noor et al., 2019 [37]	DT, SVM, K-NN	82.61	17.39
Rivero-Palacio et al. 2021 [38]	YOLO v5 Neural Network	93	7
Dimauro et al., 2023 [39]	SVM, random forest, k-NN, MobileNet-V2	91	9
Dimauro et al., 2023 [40]	Scleral segmentation algorithm	73.58	26.42
Muthalagu et al., 2018 [41]	Elman neural system	91.3	8.7
Proposed anaemia prediction model using XAI	Machine learning-based XAI	98.13	1.87

[38], [39], [40], [41] and clearly shows that the proposed model's accuracy and miss rate 98.13% and 1.87%, respectively, are significantly more reliable than the alternative algorithm. This superior performance is primarily attributed to the integration of XAI techniques, such as SHAP and LIME, which enhance interpretability by providing actionable insights into the factors influencing predictions. These insights improve the practical usability of the model in clinical settings, fostering trust among healthcare professionals.

VII. CONCLUSION

During the past years, researchers have proven that anaemia is the most spread threat around the world. This disease is defined by low red blood cells of haemoglobin counts with various causes, and the latest step in its tackling is using precise predicting models to get better patient results. The traditional anaemia prediction models that are based on statistical and clinical indicators are not suitable for practical purposes in the real world and are also not as insightful because of the disease's complexity and healthcare restrictions. This has fostered the look for AI solutions, however, the opacity in standard AI models (black boxes) has slowed down their acceptance in clinical usage.

This research work proposed a transparent, explainable AI (white box) model for anaemia prediction, employing techniques like SHAP and LIME to demystify the predictive process and offer actionable insights to clinicians. This innovative model not only improves predictive accuracy and clinical interpretability, evidenced by its 98.13% accuracy and 1.87% miss rate. but also represents a significant development in patient care and management, overcoming the transparency issues associated with previous approaches.

VIII. FUTURE WORK

The concept of the proposed model is supported by promising findings; however, further studies should consider broadening the model's application across different populations and healthcare contexts. Adding data from other sources like genotyping, continuous monitoring through wearable gadgets can improve its precision and reliability. Expanding these uses into real-world applications and investigating the ability to integrate into Electronic Health record (HER) more easily will also contribute to wider clinical applicability.

REFERENCES

[1] E. McLean, M. E. Cogswell, I. Egli, D. Wojdyla, and B. D. Benoist, "Worldwide prevalence of anaemia, WHO vitamin and mineral nutrition information system, 1993–2005," *Public Health Nutrition*, vol. 12, no. 4, pp. 444–454, May 2008.

[2] W. Gardner and N. Kassebaum, "Global, regional, and national prevalence of anemia and its causes in 204 countries and territories, 1990–2019," *Current Develop. Nutrition*, vol. 4, p. 830, Jun. 2020.

[3] 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.

[4] J. W. Asare, P. Appiahene, E. T. Donkoh, and G. Dimauro, "Iron deficiency 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.

[5] D. Newhall, R. Oliver, and S. Lugthart, "Anaemia: A disease or symptom," *Neth. J. Med.*, vol. 78, no. 3, pp. 104–110, Apr. 2020.

[6] G. Dimauro, D. Caivano, P. Di Pilato, A. Dipalma, and M. G. Camporeale, "A systematic mapping study on research in anemia assessment with non-invasive devices," *Appl. Sci.*, vol. 10, no. 14, p. 4804, Jul. 2020.

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

[8] 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.

[9] 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.

[10] B. Chan, "Black-box assisted medical decisions: AI power vs. ethical physician care," *Med., Health Care Philosophy*, vol. 26, no. 3, pp. 285–292, Sep. 2023.

[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] S. K. Mandala, "XAI renaissance: Redefining interpretability in medical diagnostic models," 2023, *arXiv:2306.01668*.

[15] 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.

[16] K. Aishwarya, "Threshold based feature selection for anaemia prediction," in *Proc. Int. Conf. Comput. Commun. Inform. (ICCCI)*, 2023, pp. 1–5.

[17] S. Casper, C. Ezell, C. Siegmann, N. Kolt, T. L. Curtis, B. Bucknall, A. Haupt, K. Wei, J. Scheurer, M. Hobbhahn, and L. Sharkey, "Black-box access is insufficient for rigorous AI audits," 2024, *arXiv:2401.14446*.

[18] R. Richman and M. V. Wüthrich, "Conditional expectation network for SHAP," 2023, *arXiv:2307.10654*.

[19] J. Sun, C. Sun, Y.-X. Tang, T.-C. Liu, and C.-J. Lu, "Application of SHAP for explainable machine learning on age-based subgrouping mammography questionnaire data for positive mammography prediction and risk factor identification," in *Proc. MDPI*, Jul. 2023, vol. 11, no. 14, p. 2000.

[20] M. R. Zafar and N. Khan, "Deterministic local interpretable model-agnostic explanations for stable explainability," *Mach. Learn. Knowl. Extraction*, vol. 3, no. 3, pp. 525–541, Jun. 2021.

[21] M. Rehman Zafar and N. Mefraz Khan, "DLIME: A deterministic local interpretable model-agnostic explanations approach for computer-aided diagnosis systems," 2019, *arXiv:1906.10263*.

[22] M. A. Aslam, M. Zonain, M. Suneer, O. Sattar, M. Salahat, and M. Saleem, "Neurological disorder detection using OCT scan image of eye," in *Proc. Int. Conf. Cyber Resilience (ICCR)*, Oct. 2022, pp. 1–13.

[23] M. Saleem, M. S. Khan, G. F. Issa, A. Khadim, M. Asif, A. S. Akram, and H. K. G. Nair, "Smart spaces: Occupancy detection using adaptive back-propagation neural network," in *Proc. Int. Conf. Bus. Analytics Technol. Secur. (ICBATS)*, Mar. 2023, pp. 1–6.

[24] B. Santra, D. P. Mukherjee, and D. Chakrabarti, "A non-invasive approach for estimation of hemoglobin analyzing blood flow in palm," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1100–1103.

[25] L. Granados-Castro, O. Gutiérrez-Navarro, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, E. R. Zavala-Sánchez, and D. U. Campos-Delgado, "Estimation of deoxygenated and oxygenated hemoglobin by multispectral blind linear unmixing," in *Proc. 25th Euromicro Conf. Digit. Syst. Design (DSD)*, 2022, pp. 849–854.

[26] A. Kesarwani, S. Das, M. Dalui, and D. R. Kisku, "Non-invasive haemoglobin estimation using different colour and texture features of palm," in *Proc. Int. Conf. Comput. Vis. Image Process.* Cham, Switzerland: Springer, Jan. 2023, pp. 176–189.

[27] P. Jain, S. Bauskar, and M. Gyanchandani, "Neural network based non-invasive method to detect anemia from images of eye conjunctiva," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 1, pp. 112–125, Mar. 2020.

[28] P. Appiahene, E. J. Arthur, S. Korankye, S. Afrifa, J. W. Asare, and E. T. Donkoh, "Detection of anemia using conjunctiva images: A smartphone application approach," *Med. Novel Technol. Devices*, vol. 18, Jun. 2023, Art. no. 100237.

[29] S. Bauskar, P. Jain, and M. Gyanchandani, "A noninvasive computerized technique to detect anemia using images of eye conjunctiva," *Pattern Recognit. Image Anal.*, vol. 29, no. 3, pp. 438–446, Jul. 2019.

[30] 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.

[31] A. Agrawal, "Detecting anemia from retinal images using deep learning," Indian Stat. Inst. (ISI), Kolkata, India, Tech. Rep. ISI/2023/7292, 2021.

[32] R. Magdalena, S. Saidah, I. D. S. Ubaidah, Y. N. Fuadah, N. Herman, and N. Ibrahim, "Convolutional neural network for anemia detection based on conjunctiva palpebral images," *J. Teknik Informatika (Jutif)*, vol. 3, no. 2, pp. 349–354, 2022.

[33] N. J. Peksi, B. Yuwono, and M. Y. Florestiyanto, "Classification of anemia with digital images of nails and palms using the Naive Bayes method," *Telematika, J. Informatika dan Teknologi Informasi*, vol. 18, no. 1, pp. 118–130, 2021.

[34] S. Roychowdhury, D. Sun, M. Bihis, J. Ren, P. Hage, and H. H. Rahman, "Computer aided detection of anaemia-like pallor," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Inform. (BHI)*, Feb. 2017, pp. 461–464.

[35] N. B. Noor, M. S. Anwar, and M. Dey, "An efficient technique of hemoglobin level screening using machine learning algorithms," in *Proc. 4th Int. Conf. Electr. Inf. Commun. Technol. (EICT)*, Dec. 2019, pp. 1–6.

[36] A. Tamir, C. S. Jahan, M. S. Saif, S. U. Zaman, M. M. Islam, A. I. Khan, S. A. Fattah, and C. Shahnaz, "Detection of anemia from image of the anterior conjunctiva of the eye by image processing and thresholding," in *Proc. IEEE Region 10 Humanitarian Technol. Conf.*, Dec. 2017, pp. 697–701.

[37] N. B. Noor, M. S. Anwar, and M. Dey, "Comparative study between decision tree, SVM and KNN to predict anaemic condition," in *Proc. IEEE Int. Conf. Biomed. Eng., Comput. Inf. Technol. Health (BECITHCON)*, Nov. 2019, pp. 24–28.

[38] M. Rivero-Palacio, W. Alfonso-Morales, and E. Caicedo-Bravo, "Anemia detection using a full embedded mobile application with YOLO algorithm," in *Proc. IEEE Colombian Conf. Appl. Comput. Intell.* Cham, Switzerland: Springer, May 2021, pp. 3–17.

[39] G. Dimauro, M. E. Griseta, M. G. Camporeale, F. Clemente, A. Guarini, and R. Maglietta, "An intelligent non-invasive system for automated diagnosis of anemia exploiting a novel dataset," *Artif. Intell. Med.*, vol. 136, Feb. 2023, Art. no. 102477.

[40] G. Dimauro, M. G. Camporeale, A. Dipalma, A. Guarini, and R. Maglietta, "Anaemia detection based on sclera and blood vessel colour estimation," *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104489.

[41] R. Muthalagu, V. T. Bai, D. Gracias, and S. John, "Developmental screening tool: Accuracy and feasibility of non-invasive anaemia estimation," *Technol. Health Care*, vol. 26, no. 4, pp. 723–727, Sep. 2018.

[42] B. Rao. *Anemia Dataset*. Accessed: Sep. 2024. [Online]. Available: https://www.kaggle.com/datasets/biswaranjanrao/anaemia-dataset/data

MUHAMMAD SAJID FAROOQ received the Ph.D. degree in computer science from NCBA&E, Lahore, Pakistan. He is currently working as the Head of Department Cyber Security, NASTP Institute of Information Technology, Lahore. His research expertise includes machine learning, deep learning, federated learning, and cloud computing. He has contributed significantly to these fields through his publications, with many research works published in international conferences and journals. His work continues to inspire and shape the future of computing and data sciences.

MUHAMMAD HASSAN GHULAM MUHAMMAD received the Ph.D. degree in computer science. He has extensive experience in both academic and industry settings. He is currently an Assistant Professor with the Department of Computer Science, IMS Pak-AIMS. He is a highly skilled computer science professional. His expertise spans cloud computing, intelligent system design, high-performance computing, the IoT, and big data. Over the past several years, he has held various roles, including a MacOS Synchronization Specialist, an Academic Coordinator, and an IT Consultant, where he successfully managed synchronization protocols, led teams, developed marketing strategies, and ensured security compliance. His academic achievements include a Gold Medal in his master's program, and he has a strong publication record in reputable journals. He is passionate about integrating real-world applications into his teaching, contributing to groundbreaking research, and leveraging his management and technical skills to enhance organizational performance. His goal is to inspire future leaders in the field of computer science by fostering critical thinking, problem-solving skills, and a deep understanding of technological advancements.

OUALID ALI is a distinguished and recognized as an international expert in the field of quality assurance and accreditation in the sector of higher education, renowned for his contributions to enhancing academic standards and institutional effectiveness. With over than 18 years of experience, he has become a leading authority on quality management, accreditation processes, and educational reform. He has a wide experience in quality assurance and accreditation, on a national level for GCC countries (UAE, Oman, Bahrain, Saudi Arabia, and Qatar) and international level (AACSB, ABET, and IET). He held several positions, such as Quality Officer and Quality and Accreditation Department Director in colleges/universities in the Sultanate of Oman. In these colleges/universities, he played a pivotal role in achieving and maintaining the local accreditation status (ISA). His expertise in navigating complex accreditation requirements and his ability to foster collaborative relationships with stakeholders were instrumental in the institutions' successful accreditation outcomes. Moreover, he holds several international certificates related to quality such as ISO 9001 certification (Internal Auditor), Black Belt Lean Six Sigma Certification, KPIs Professional Certification, ISO 31000 Certification, and Strategic Planning Expert Certification. He has performed several reviews across the GCC and he is recognized as an international expert/reviewer in several GCC countries in Institutional Standard Assessment (ISA), National Quality Frameworks (NQF), Program Standards Assessment (PSA), as well as a reviewer for AACSB and ABET International accreditations. He is an active member of several professional organizations working in the field of quality assurance and accreditation in the sector of higher education. He is an advocate for continuous improvement in higher education. He is involved in several projects aiming at enhancing QA practices globally.

ZAHID ZEESHAN is a scholar having the best research background expertise including machine learning, deep learning, and cloud computing. He has contributed significantly to these fields through his publications, with many research works published in international conferences and journals. His work continues to inspire and shape the future of computing and data sciences.

MUHAMMAD SALEEM received the Ph.D. degree in computer science. He is currently an Assistant Professor with the School of Computer Science, Minhaj University Lahore, Pakistan. He is also a Dedicated Researcher. He has specialized in artificial intelligence, machine learning, deep learning, cloud computing, and explainable AI. His research portfolio interests include advancements in energy optimization, intelligent traffic control, and predictive healthcare models. He is committed to developing AI-driven solutions to address digital challenges, contributing significantly to both academic and industry objectives. His work is highly appreciated, indicating that he is experienced and contributes to computer science in terms of advancing its frontiers and encouraging teamwork.

MUNIR AHMAD (Senior Member, IEEE) received the master's degree in computer science from the Virtual University of Pakistan, Pakistan, and the Ph.D. degree in computer science from the School of Computer Science, National College of Business Administration and Economics. He is a distinguished professional with over 16 years of experience. As the Executive Director/CIO at United International Group, Lahore, Pakistan, he has excelled in data management and resource optimization within multinational organizations. He is renowned for his extensive research in sentiment analysis, AI applications in healthcare, and animal facial identification. His expertise lies in data mining, big data, and artificial intelligence.

SAGHEER ABBAS (Senior Member, IEEE) received the M.Phil. and Ph.D. degrees in computer science from the School of Computer Science, NCBA&E, Lahore, Pakistan, in 2016. He is currently a Research Professor with Prince Mohammad Bin Fahad University, Saudi Arabia. He has been teaching graduate and undergraduate students in computer science and engineering for the past 15 years. He has published about 180 research papers in international journals and reputed international conferences. His research interests include cloud computing, the IoT, intelligent agents, image processing, and cognitive machines, with various publications in international journals and conferences.

MUHAMMAD ADNAN KHAN (Senior Member, IEEE) received the B.S. and M.S. degrees from International Islamic University, Islamabad, Pakistan, and the Ph.D. degree from ISRA University, Islamabad. He is currently an Assistant Professor with the School of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam, Republic of Korea. Before joining Gachon University, he worked in various academic and industrial roles in Pakistan and Republic of Korea. He has been teaching graduate and undergraduate students in computer science and engineering for the past 16 years. He received a scholarship from the Higher Education Commission, Islamabad, in 2016, for the Ph.D. study; and a scholarship from Punjab Information and Technology Board, Government of Punjab, Pakistan, for the B.S. and M.S. studies. He is guiding five Ph.D. scholars and six M.Phil. scholars. He has published more than 300 research articles with Cumulative JCR-IF is more than 980 in reputed international journals and international conferences. According to studies conducted by Stanford University, in 2023 and 2024, released by Elsevier, he was listed among the top 2% of impactful researchers in the world. His research interests primarily include computational intelligence, machine learning, MUD, image processing, medical diagnosis, and channel estimation in multi-carrier communication systems using soft computing. He received the University Best Researcher Award, in 2018, 2019, 2020, 2022, and 2023. He is the general chair, the co-chair, and a speaker of conferences and workshops to promote knowledge sharing and learning within academic and societal realms. He serves as an Editorial Member of high-impact journals, such as IEEE Transactions on Neural Networks and Learning Systems and *Computational Intelligence and Neurosciences*.

TAHER M. GHAZAL (Senior Member, IEEE) received the B.Sc. degree in software engineering from Al Ain University, in 2011, the M.Sc. degree in information technology management from The British University in Dubai, in 2013, associated with The University of Manchester and The University of Edinburgh, and the Ph.D. degree in information science and technology from Universiti Kebangsaan Malaysia, in 2023. He is a seasoned academician with a comprehensive educational background. He possesses over a decade of multifaceted expertise, he has fulfilled various roles, including a Lecturer, an Instructor, a Tutor, a Researcher, a Teacher, an IT Support/Specialist Engineer, and a Business/Systems Analyst. He has contributed significantly across diverse departments, such as Engineering, Computer Science, and ICT, and the Head of STEM and Innovation. His professional engagements have extended to governmental and private educational institutions under the purview of KHDA, the Ministry of Education, and the Ministry of Higher Education and Scientific Research, United Arab Emirates. His scholarly pursuits encompass a wide array of interests, including the IoT, artificial intelligence, cybersecurity, information systems, software engineering, web development, building information modeling, quality of education, management, big data, quality of software, and project management. He is actively engaged in community service through his involvement in impactful projects and research endeavors.

● ● ●