Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN INTELLIGENT ADAPTIVE SYSTEM FOR ANEMIA CLASSIFICATION FROM HEMATOLOGICAL PARAMETERS

S.B. SWATHI[1] , DR. N. RAMANA[2]

1Research Scholar, Dept. of CSE, Faculty of Engineering and Technology, Kakatiya University, India, Email: sbswathi870@gmail.com, sbswathikits@gmail.com
2Associate Professor, Dept. of CSE, Faculty of Engineering and Technology, Kakatiya University, India, Email: ramanauce.ku@kakatiya.ac.in

ABSTRACT

Anemia is among the most common blood disorders worldwide, caused by a decrease in red blood cells or hemoglobin concentration, which reduces oxygen transport in the body. Traditional methods of diagnosing anemia usually involve manual interpretation of diagnostic hematological parameters, and these time-consuming processes can still be subject to human error. To mitigate some of these challenges, this study presents a Sharpbelly Fish-based Deep Neural Network Prediction Framework (SbDNNPF) that can classify anemic cases using publicly available hematological data. The SbDNNPF process begins with data pre-processing to eliminate inconsistencies and missing values, followed by feature selection to identify relevant hematological parameters. An optimized feature subset is then used to train a DNN, and the Sharpbelly Fish Optimization algorithm is used to tune network parameters to improve convergence and minimize prediction error. Experimental analysis of the SbDNNPF demonstrates that the proposed DNN-based approach outperforms traditional machine learning (ML) classifiers in terms of performance, accuracy, precision, recall, and F score, thereby effectively enhancing the reliability of anemia diagnosis.
Keywords: *Anemia classification, Hematological Data, Pre-processing, Diagnosis, Feature selection. Model.*

1. INTRODUCTION

Anemia refers to a deficiency of red blood cells (RBCs)/hemoglobin (Hb) below normal limits (1) and is usually indicated by commonly observed signs such as fatigue, palpitations, headaches, shortness of breath, and pallor (2). These signs are not final indicators (3), as they provide only one potential "path," without providing proof of an anemia diagnosis (4). Anemia has multiple causes (etiologies) around the world; hence the term "multifactorial etiology" [5]. One of the major/global causes is iron deficiency anemia, which accounts for approximately 30% of all females and 40% of children <5 years of age in all countries (6). Additionally, the prevalence of anemia within Colombia is correlated with the ongoing food insecurity resulting from the poverty level. The estimated number of Caliburnians living in poverty in Colombia is estimated to be 18.3 million people in 2022, and approximately 6.9 million were estimated to be living in "extreme" poverty [7]. Although a drop of three per cent overall indicates

there was a decline in poverty in 2015, that increase was concerning [8]. Many approaches permit physicians to detect and diagnose anemia [9]. The focus of the majority of these methods is on assessing the makeup and quality of blood elements (mainly red blood cells) and the characteristics of these red blood cells [10]. A few specific examples include a complete blood cell count (CBC), which provides information about various types of blood elements (including the number of red blood cells, white blood cells, and platelets) [11]. The other is a peripheral blood smear, an open-circulation hemodynamic study used to identify abnormalities in the size, shape, and appearance of red blood cells [12]. Through this approach, clinicians can identify patterns (indicating whether or not there is a hemolytic component), which are used in combination with hematological indices to diagnose and explain anemia [13]. Lastly,

reticulocyte counts serve as a valuable diagnostic tool [14]. Reticulocytes, or immature red blood cells, are formed in the bone marrow, and as soon as they mature, they are released into the bloodstream [15]. Iron levels and the body's ability to use iron can be assessed using various serum iron levels, including serum iron, total iron binding capacity (TIBC), and transferrin saturation [16]. Iron deficiency is a prevalent cause of anaemia, and testing these measurements will help assess whether anaemia is caused by inadequate iron availability [17].

Anemia may have many different causes, which leads to the potential for classifying anemia (e.g., microcytic, normocytic, macrocytic) by identifying the underlying cause [18]. Examples of anemia classification include iron-deficiency (microcytic), inflammatory disease (normocytic), and vitamin B12 deficiency (macrocytic) [19]. Additionally, it is possible to classify anemia based on severity (e.g., mild, moderate, severe). There are several barriers in relation to traditionally diagnosing anemia, especially for those in disadvantaged groups who may experience barriers such as pricing and lack of access to quality health care [20]. These challenges lead to delays in diagnosis and treatment for anemia and contribute to a vicious cycle of worsening anemia, leading to additional health complications and decreased quality of life for affected individuals [21]. On the contrary, there is evidence that predictive health models may improve resource distribution within health systems by utilizing machine-learning-based predictive models in healthcare, including deployment of early warning systems that allow for more rapid action taken by providers [22]. Predictive health models utilizing machine learning can also lead to an improved element of objectivity, resulting in more accurate predictions [23]. This study aims to develop an intelligent system using machine learning to classify various types of anemia using hematological data. The key contribution of the research is described as follows;

- Initially, the Complete Blood Counts (CBC) data is gathered and loaded into the Python platform.
- Subsequently, a new Sharpbelly fish-based Deep Neural Network Prediction Framework (SbDNNPF) was established with sufficient characteristics.
- Hence, the collected data undergoes pre-processing to prepare it for further analysis. Pre-processing was performed to eliminate the missing values from the datasets, and relevant features were selected in the feature selection phase.
- Subsequently, the SbDNNPF was used to forecast and classify the type of anaemia based on the selected features.
- Lastly, the suggested method's performance is assessed using a few existing models, such as error rate, f-score, recall, accuracy, and precision.

2. RELATED WORK

Qasrawi et al. [21] made use of data mining and ML methods to investigate nutrient intake and anemia among university students. Data from 755 females participated, and K-means clustering and Decision Tree (DT) were the ML methods used to determine associations. There was a specific focus on vitamins and minerals affecting anemia prevalence, and the results indicated that inadequate intake led to an increased prevalence of anemia. The DT model was applied and produced 82.1% accuracy, clearly identifying key diet contributors to anemia risk.

Asare et al. [22] constructed anaemia detection models using ML algorithms and assessed anaemia by analyzing conjunctiva pictures of children. The digital images were processed using the CIELAB colour space, and regions of interest were extracted for feature analysis. The dataset was partitioned into three sections with a ratio of 70:10:20 for training, validation, and test sets. The accuracy of the models was reported, with the CNN achieving the highest accuracy of 98.45%. The findings indicate this method can provide non-invasive, affordable anaemia detection in poor communities.

Abdullah et al. [23] explored the use of Gene Expression Programming (GEP) for forecasting anemia using a publicly available Kaggle data containing clinical features such as hemoglobin (HGB) and red blood cell (RBC) indices. The optimized GEP model achieved an impressive accuracy of 99.30%. Explainable AI techniques, including LIME and SHAP, were employed to enhance model transparency, identifying HGB levels and gender as key

predictors. The study highlights the potential of interpretable ML models for reliable and automated anemia diagnosis in clinical practice.

Inoue et al. [24] applied ML strategies to enhance the care of anemia among patients with end-stage kidney disease (ESKD) on hemodialysis. They developed models through PyCaret using 67 patients with anemia, incorporating clinical variables related to laboratory and treatment-related hemoglobin levels, iron levels, and ESA doses. Both the LightGBM and XGBoost models had superior performance compared to logistic regression. The analysis showed that previous ESA, iron doses, and ferritin levels were significant predictors. The authors conclude that machine learning has the potential to improve clinical decision-making and the precision of drug dosing for treating anemia.

Peng et al. [25] have proposed a new NLP hybrid method for diagnosing anemia in Traditional Chinese Medicine (TCM). The methodology integrates a pre-trained LERT model with a GRU network, which improves text feature extraction and minimizes the model's computational complexity. An Adaptive Gate Network (AGN) module was included to address the long-term dependencies, and the new FCS loss function was added to tackle data imbalance. Their findings demonstrate the method's efficacy by using a dataset of TCM anemia and achieving an accuracy value of 94.3%. The methodology outperformed other methods currently in use and suggests promising capabilities for future intelligent clinical decision support.

3. PROBLEM STATEMENT

Millions of people suffer from anemia, a common illness. Individuals experience symptoms such as fatigue, weakness, shortness of breath, or organ damage in extreme cases. Recognizing and classifying an anemia diagnosis is necessary for the urgent and effective management of the condition. Diagnostics are currently based on practitioners examining laboratory blood tests. These investigations are subjective, time-consuming, and can expose health professionals to human error. The challenge lies in the fact that different types of anemia have considerable overlap in clinical presentation and minimal variance in blood tests, which adds complexity and makes differentiating between anemias difficult even for seasoned clinicians. In addition, the increasing number of patients and the volume of blood test results in clinical practice make it difficult to provide timely and accurate diagnoses. This situation creates a need for a systematic, robust, and reliable method to evaluate hematology data, extract features, and support clinicians in accurately classifying anemia to improve care.

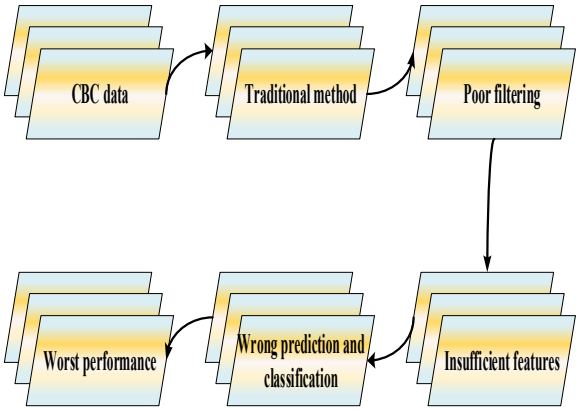

Fig. 1 System model with problem

4. PROPOSED METHODOLOGY

At first, a CBC dataset was collected from publicly available hematology sources and loaded into the Python environment for evaluation. The final aim is to develop a Deep Neural Network Prediction Framework based on the novel Sharpbelly fish (SbDNNPF), which has high aptitudes for learning and optimizing data features to enhance prediction accuracy and reliability. The data processing on the dataset included handling missing values and selecting the most relevant hematologic features in an efficient feature selection phase. The dataset with all processed features was transferred into the SbDNNPF model to predict and classify multiple types of anaemia. The proposed framework was evaluated using standard evaluation metrics, including accuracy, precision, recall, F-score, and error rate, and compared to the performance of existing methods.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

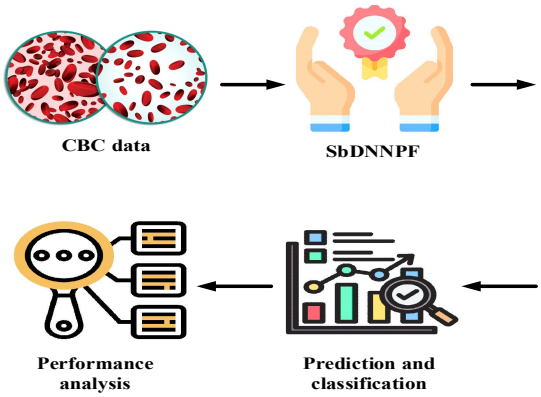

Fig. 2 Proposed architecture of SbDNNPF

4.1 Process of suggested SbDNNPF

Initially, the complete blood counts (CBCs) dataset was collected from publicly accessible haematological databases and loaded into the Python platform. The dataset includes the essential blood parameters that are often used in the diagnosis of anemia. The collected data is equated in Eqn. (1).

$$D(F_x) = Cs\ \{1, 2, 3, \ldots, n\} \quad (1)$$

Here, $D(Fx)$ represents the collected dataset, Cs denotes the CBC data, and n denotes the amount of data. Hematological data may contain inconsistencies, missing values, and erroneous values that may hinder model performance. Therefore, data processing is conducted to prepare the database for analysis.

4.1.1 Data processing

Data processing is an essential stage in preparing the collected hematological dataset for the SbDNNPF, ensuring accurate and practical analysis. For example, raw medical dataset accounts may contain missing values, noise, or inconsistencies that could compromise the performance of the Hb classification model. First, the missing values in the dataset were examined and filled in. Second, data cleaning was employed to remove duplicate accounts and correct any inconsistencies with the hematological parameters. The filtration process can be represented in Eqn. (2).

$$J(F_x) = [F_x(g, v, m) - F_x(g, v)] \quad (2)$$

Here J is the filtration variable, g which stands for missing variables, m denotes the standard feature, and v represents the noise data. Data processing ensures that the dataset provided to the SbDNNPF model has been cleaned and optimized, benefiting the analysis of relevant features and classifications of anemia types.

4.1.2 Feature selection

After data processing, feature selection involves identifying the most relevant hematology features, eliminating redundant or irrelevant ones, and retaining those that significantly predict anemia. The importance of every feature is assessed through a Sharpbelly Fish Optimization (SFO) function, which quantifies the correlation and statistical relationship between hematology features and anemia classes. The feature selection procedure can be formalized mathematically as Eqn. (3).

$$B[F_x] = S_f\ [J(F_x) - L_d] \quad (3)$$

Here, $B[F_x]$ denotes the feature selection variable, S_f stands for the Sharpbelly Fish-based Optimization feature tracking variable, L_d denotes the selected features of anemia classification, such as HGB, RBC, mean corpuscular volume (MCV), mean corpuscular Haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), blood platelets (PIT), white blood cells (WBC), and variability of blood platelet size distribution (PDW). The algorithm rapidly searches the search space to identify the optimal subset of attributes that maximizes classification performance, drawing inspiration from the collective behaviour and adaptive hunting movement of Sharpbelly fish in nature.

4.1.3 Prediction and classification

Within the predicting and classifying phase, this study provides reliable predictions and classifications of the type of anemia from hematological data by using a DNN as part of the SbDNNPF. The optimized dataset available from the feature selection phase will

be input into the DNN. The hidden layers use nonlinear activation functions to discover intricate correlations between the features. The network parameters are fine-tuned using the SFO algorithm during training to enhance the network's overall learning efficiency and reduce errors in predicting the class label. The prediction phase is exposed in Eqn. (4).

$$G_a = Kn\left[\, B(F_x)\,\right] * \alpha \quad (4)$$

Where G_a signifies the prediction variable, F_x denotes the selected features, Kn represents the neural network function, and α indicates the weight parameters. The model's predictions are then categorised into relevant classifications. The final stage is to classify the anemia types, such as healthy, iron deficiency anemia, leukaemia, macrocytic anemia, leukaemia with thrombocytopenia, normocytic hypochromic anemia, other microcytic anemia, and thrombocytopenia. The classification phase is equated in Eqn. (5).

$$R = \begin{cases} if \;\; G_a = 0; & Healthy \\ G_a = 1; & Iron \;\; deficiency \;\; syndrome \\ G_a = 2; & Leukemia \\ G_a = 3; & Macrocytic \;\; anemia \\ G_a = 4; & Leukemia \;\; with \;\; thrombocyt \;\; openia \\ G_a = 5; & Normocytic \;\; hypochromi \;\; c \;\; anemia \\ G_a = 6; & Normocytic \;\; normochrom \;\; ic \;\; anemia \\ G_a = 7; & Thrombocyt \;\; ophenia \\ G_a = 8; & Other \;\; microtican \;\; emia \end{cases} \quad (5)$$

Where R indicates the classification variable, G_a and represents the anticipated outcome. The class with the highest probability value is the final prediction. The DNN-based prediction and classification provide reliable anemia diagnoses from public hematological data with high accuracy and robustness. The present work's algorithm is described as follows.

Algorithm1:SbDNNPF

Start

{

Data initialization()

{

$$D(F_x) = C_S\{1, 2, 3, \ldots\ldots, n\}$$

//The CBC data is collected

}

Pre-processing()

{

int $J(F_x), g, v, m$

//initializing the pre-processing variables

$\Pr e - process \sin g \leftarrow J(F_x)$

// the missing and erroneous values are removed

}

Feature extraction()

{

int $B, S_f, L_d, J(F_x)$

//initializing the feature extraction variables

$Feature\ extraction \leftarrow B(F_x)$

//The most influential haematological parameters were selected using the Sharpbelly Fish optimization.

}

Prediction and classification

{

int G_a, Kn, α

//initialize the prediction variables

$\Pr ediction \leftarrow G_a$

//The anemia types were predicted and classified using the deep neural network function.

}

}

Stop

used to execute the Python code and validate the results. For all mathematical function factors, the approach produces a pseudo-code representation. The SbDNNPF flowchart is displayed in Fig. 3.

The step-by-step algorithm offers a comprehensive explanation of the process and approaches outlined in the suggested framework. These procedures are

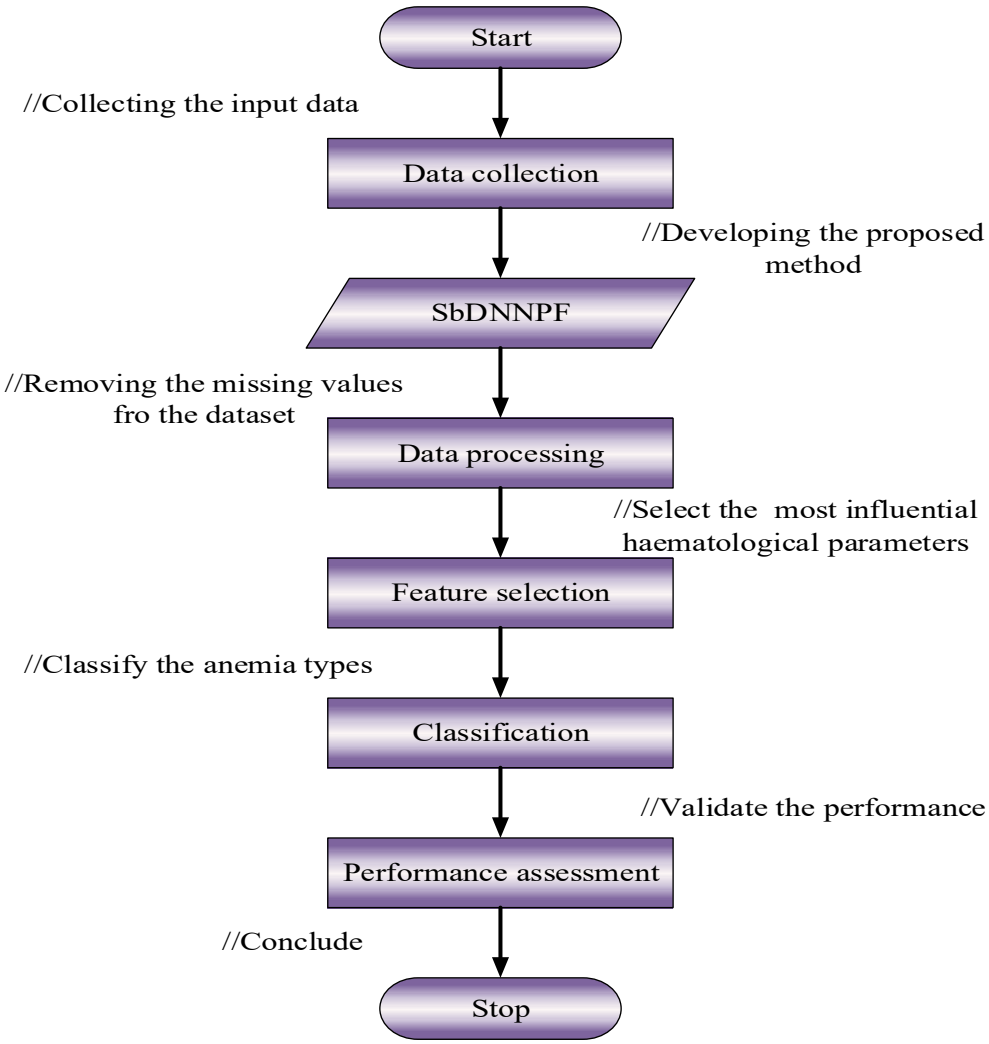

Fig. 3 Flowchart of SbDNNPF

5 RESULT AND DISCUSSION

The Python environment version and Windows 10 operating system are used to validate the suggested framework. Additionally, the model's effectiveness is thoroughly examined using the provided data. The results produced by the selected model are contrasted with those made by the current model to illustrate the extent of improved efficiency and the improvement percentage. Table 1 displays the parameters for the implementation.

Table 1: Configuration parameters

Parameter	Specification	Optimization	Sharpbelly fish
		Network	Deep neural network
Platform	Python	Dataset	CBC data
Version	3.7.14	Total samples	1281
OS	Windows 10		

5.1 Case study

The efficacy of the proposed SbDNNPF was evaluated using a case study with a publicly available CBC hematological dataset. The dataset collected key clinical variables that serve as predictors for the diagnosis of the different types of anemia. The primary variables used in this study were HGB, platelet count, WBC, RBC count, MCV, MCH, MCHC, and PDW. Each of

these parameters provides essential information about the makeup of blood and indicates pathologies associated with different types of anemia. For example, low HGB and RBC values generally indicate the possibility of iron deficiency anemia. Elevated MCV and MCH are good indicators of megaloblastic anemia, and variations in both PLT and PDW reflect disorders of hematology that affect the clotting mechanisms of blood. Table 2 displays the dataset's details.

Table 2. Dataset details

Total samples: 1281	
Training samples	1024
Testing samples	257
Training samples (80%): 1024	
Healthy	269
Iron deficiency anemia	151
Leukaemia	38
Macrocytic anemia	14
Leukaemia with thrombocytopenia	9
Normocytic hypochromic anemia	223
Normocytic normochromic anemia	215
Thrombocytopenia.	58
other microcytic anemia	47
Testing samples (20%): 257	
Healthy	67
Iron deficiency anemia	38
Leukaemia	9
Macrocytic anemia	4
Leukaemia with thrombocytopenia	2
Normocytic hypochromic anemia	56
Normocytic normochromic anemia	54
Thrombocytopenia.	15
other microcytic anemia	12

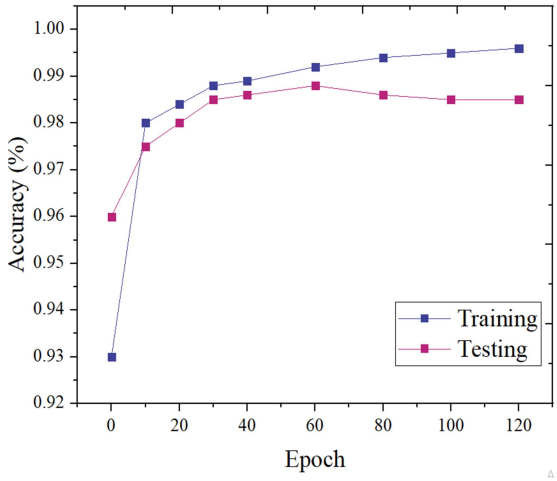

Fig. 5 Accuracy curve graph

Fig. 5 illustrates the model's training and testing accuracies following multiple epochs of training on both datasets. The training accuracy begins at roughly 92% and consistently rises across all epochs, ultimately reaching approximately 99%. Testing accuracy behaves similarly, starting at nearly 96% and plateauing near 98% after about 40 epochs. This suggests that the model is learning well, with high training and testing accuracy but minimal overfitting and/or excessive overlearning.

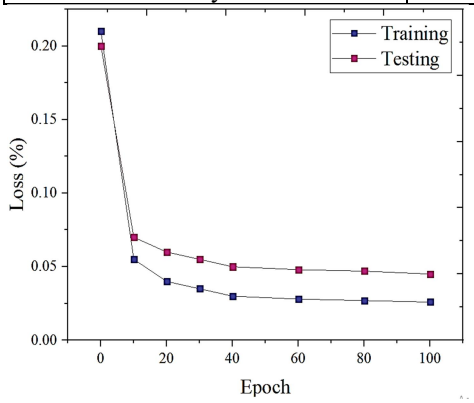

Fig. 6 Loss curve graph

Fig. 6 shows the training and testing loss values over the trajectory of epochs. For both training and testing, the initial loss values were somewhat high (around 0.20), but they dropped significantly within the first 20 epochs. Although the training loss did fall, it is now stabilised at about 0.02, and the last testing loss at epoch 40 was about 0.04. The dropping and stabilising traces from the loss curves indicate that the model is effectively reducing errors and generalising well to the unrecognised data.

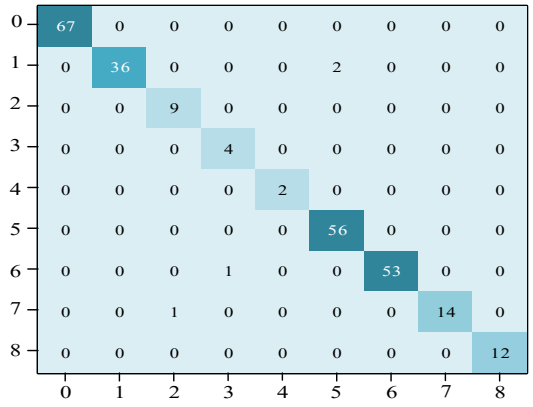

Fig. 7 Confusion matrix

The confusion matrix shown above illustrates the classification performance of the proposed SbDNNPF for anemia categorization. Each diagonal element represents the number of correctly classified samples for each anemia category, while the off-diagonal elements indicate that no samples were misclassified. The matrix demonstrates excellent predictive accuracy, with perfect class separation across all nine categories — for example, Class 0 achieved 67 correct predictions, Class 1 had 36, Class 5 had 56, and Class 6 had 53. This result confirms that the proposed framework effectively distinguishes between different anemia types based on hematological parameters, achieving near-perfect classification performance without confusion between categories.

5.2 Performance analysis

The implementation of the validation process involves assessing several standard performance measures of the models for predicting various aspects of cardiovascular disease, specifically analyzing error rate, precision, f-score, and recall accuracy. Then, the approach is compared with some existing algorithms, including Gaussian Naive Bayes (GNB), DT, K-Nearest Neighbours (KNN), XGBoost, Random Forest (RF), and Multilayer Perceptron-neural Network (MLP-NN), to validate its performance improvement [26].

5.2.1 Accuracy

Accuracy indicates the percentage of accurate predictions among all predictions. Accuracy shows that a model predicts both positive and negative predictions. Accuracy can be expressed as in Eqn. (6).

$$Accuracy = \frac{P_R + N_R}{P_R + N_R + P_W + N_W} \quad (6)$$

Where P_R signifies the true positive, N_R represents the true negative, P_W means the false positive, and N_W specifies the false negative.

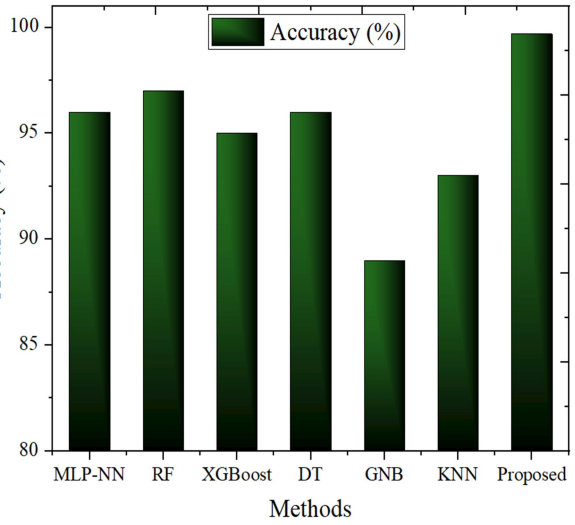

Fig. 8 Accuracy assessment

The results of accuracy validation are presented in Fig. 8. The efficiency improvement of the developed framework is demonstrated by comparing it with other existing methods in terms of accuracy in dataset evaluation. In SbDNNPF, the suggested model had a 99.7% accuracy rate. Compared to other accuracy values like KNN, DT, GNB, MLP-NN, RF, and XGBoost, which were 93%, 96%, 89%, 96%, 97%, and 95%, respectively.

5.2.2 precision

Precision can be determined by calculating the probability of actual cases of disease for anticipated cases. Precision is utilized to ascertain the reliability of predictions of positive symptoms. Precision can be expressed as in Eqn. (7).

$$\text{Pr}\,ecision = \frac{P_R}{P_R + P_W} \quad (7)$$

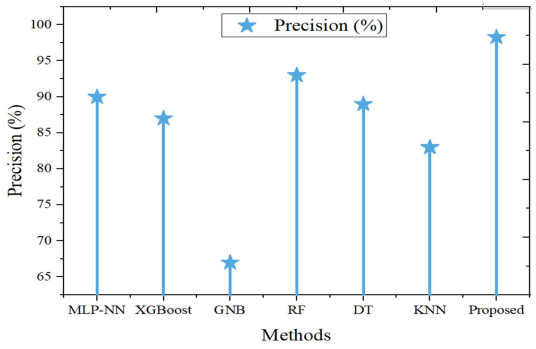

Fig. 9 Precision assessment

Fig. 9 shows the precision measurement. The precision value of the KNN method was 83%, the DT received 89%, the MLP-NN gained 90%, the RF method had a score of 93%, GNB obtained 67% and the XGBoost method had 87%. The proposed SbDNNPF method achieved 98.3% precision compared to the mechanism.

5.2.3 recall

Recall, defined as the percentage of actual disease cases that the model accurately predicts, is higher if the model can identify optimistic scenarios. Recall can be expressed as in Eqn. (8).

$$\text{Re}call = \frac{P_R}{P_R + N_W} \quad (8)$$

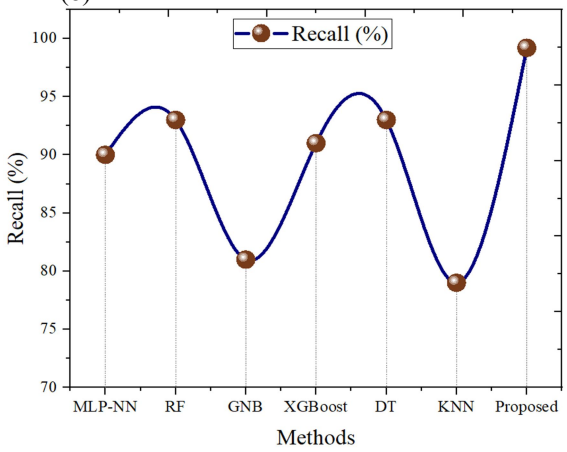

Fig. 10 Recall assessment

Fig. 10 displays the comparison of recall. The proposed method has a recall rate of 99.2%, and the recall rates of DT, XCBoost, GNB, KNN, RF, and MLP-NN are 93%, 91%, 81%, 79%, 93%, and 90%, respectively. The model's high rate of recall proved its efficacy.

5.2.4 F-score

F1-score evaluates the mean average of sensitivity and precision. The two metrics accurately balance false positives and negatives. The f-score can be expressed as in Eqn. (9).

$$F - score = 2 * \frac{\text{Re}call * \text{Pr}ecision}{\text{Re}call + \text{Pr}ecision} \quad (9)$$

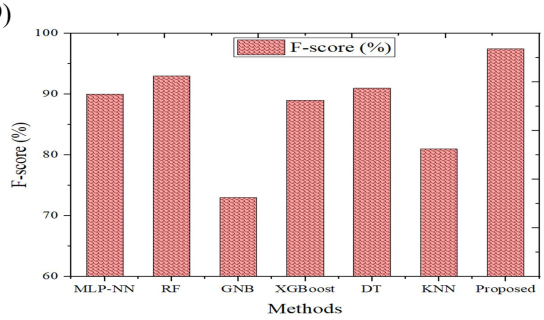

Fig. 11 F-score assessment

The f-score validation results are shown in Fig. 11. RF has an f-score of 93%, GNB has an f-score of 73%, DT has an f-score of 91%, MLP-NN has an f-score of 90%, XGBoost has an f-score of 89%, and KNN has an f-score of 81%. The f-score of the suggested approach is 97.5%. The SbDNNPF model, therefore, has a high f-score that suggests excellent performance.

5.2.5 Error rate

The error rate is the percentage of inaccurate forecasts of all predictive events. The error rate can be expressed as in Eqn. (10).

$$Error\ \ rate = \frac{P_W + N_W}{P_R + N_R + P_W + N_W} \quad (10)$$

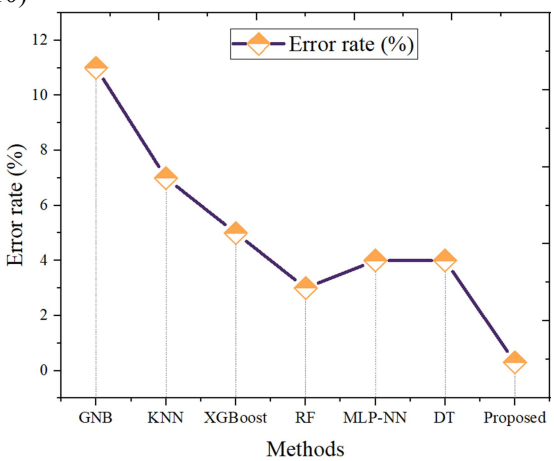

Fig. 12 Error rate assessment

Fig. 12 shows the error rate assessment. The error rates are as follows: the RF method is 3%, the DT method is 4.2%, the MLP-NN method is 4%, the GNB method is 11%, the

XGBoost method is 5%, and the KNN method is 7%. The error rate of the proposed model is 0.3%. As a result, an extremely low error rate is achieved, proving the suggested model's excellent efficacy.

Table 3. Comparison of existing methods

Methods	Recall (%)	Accuracy (%)	F-score (%)	Precision (%)	Error rate (%)
MLP-NN	90	96	90	90	4
XGBoost	91	95	89	87	5
GNB	81	89	73	67	11
RF	93	97	93	93	3
DT	93	96	91	89	4
KNN	79	93	81	83	7
Proposed	**99.2**	**99.7**	**97.5**	**98.3**	**0.3**

5.3 Discussion

The primary goal of the proposed model was to analyze the methods of predicting the risk of intracranial aneurysm rupture. The last element of the study ensured the increased performance through comparison analysis. Its stability score is computed and compared with other models currently in use to further ensure the framework's efficiency. Table 5 displays the effectiveness of the suggested method.

Table 5: SbDNNPF's performance

Metrics	Performance (%)
Accuracy	99.7
F-score	97.5
Error rate	0.3
Recall	99.2
Precision	98.3

The exceptional performance of the suggested model is demonstrated by comparing it with other approaches. The metrics are used to assess the results of the proposed method. The recommended approach produces good results when compared to the available data. With a 99% f-score, 99.2% recall, 99.5% precision, 0.3% error rate, and 99.7% accuracy, the proposed SbDNNPF model demonstrated exemplary performance.

6. CONCLUSION

This research introduced an innovative approach for classifying anemia using the SbDNNPF, which can classify various types of anemia through hematological data. This framework utilizes systematic data processing, feature selection, and an optimized deep-learning architecture to classify anemia populations accurately and reliably. Ultimately, the optimization of the DNN model through the SFOA enhances parameter tuning and learning efficiency, which improves predictive accuracy beyond traditional models. According to the experimental results, the suggested system exhibits a high-performance profile for recall, accuracy, precision, and F-score metrics, indicating the framework's resilience and effectiveness for medical data processing. In essence, the SbDNNPF model offers researchers and clinicians an effective decision support tool for diagnosing and classifying anemia in patients while reducing manual workload and uncertainty. Future work may involve extending the model through larger and more diverse hematological datasets and utilizing additional clinical parameters to improve overall generalization and diagnostic performance.

DECLARATIONS

Ethical Approval: Not applicable. This study does not involve any human or animal subjects.

Author Contributions:
S.B. Swathi: Conceptualization, Algorithm Design, Simulation, Manuscript Drafting, Editing,Simulation Validation, Literature Review, Visualization.

Dr. N. Ramana: Mathematical Modeling, Data Analysis, Review &amp.

Data Availability: Python models and optimization scripts used in this study are available from the corresponding author up.

REFERENCES:

[1] Gómez, Jorge Gómez, et al. "Anemia Classification System Using Machine Learning." *Informatics*. Vol. 12. No. 1. MDPI, 2025.

[2] Kasthuri, E., S. Subbulakshmi, and Rajasree Sreedharan. "Insightful clinical assistance for anemia prediction with data analysis and explainable AI." *Procedia computer science* 233 (2024): 45-55.

[3] Darshan, BS Dhruva, et al. "Differential diagnosis of iron deficiency anemia from aplastic anemia using machine learning and explainable Artificial Intelligence utilizing blood attributes." *Scientific Reports* 15.1 (2025): 505.

[4] Göl, Mehmet, et al. "Predicting malnutrition-based anemia in geriatric patients using machine learning methods." *Journal of Evaluation in Clinical Practice* 31.2 (2025): e14142.

[5] Sharif, Nowaj, et al. "Anemia in Asian children: trends in the prevalence, etiology, and analysis of geographic inequalities." *World Journal of Pediatrics* (2025): 1-10.

[6] Ramzan, Muhammad, et al. "Revolutionizing anemia detection: integrative machine learning models and advanced attention mechanisms." *Visual Computing for Industry, Biomedicine, and Art* 7.1 (2024): 18.

[7] Afrifa, Stephen, et al. "Machine learning algorithms for iron deficiency anemia detection in children using palm images." *International Journal of Education and Management Engineering* 14.1 (2024): 1.

[8] Darshan, BS Dhruva, et al. "Differential diagnosis of iron deficiency anemia from aplastic anemia using machine learning and explainable Artificial Intelligence utilizing blood attributes." *Scientific Reports* 15.1 (2025): 505.

[9] Pullakhandam, Siddartha, and Susan McRoy. "Classification and explanation of iron deficiency anemia from complete blood count data using machine learning." *BioMedInformatics* 4.1 (2024): 661-672.

[10] Cheng, Longcan, et al. "An ensemble machine learning model for predicting one-year mortality in elderly coronary heart disease patients with anemia." *Journal of Big Data* 11.1 (2024): 99.

[11] Mohammed, Kamel K., et al. "An Explainable AI and Optimized Multi-Branch Convolutional Neural Network Model for Eye Anemia Diagnosis." *IEEE Access* (2025).

[12] Darshan, BS Dhruva, et al. "Detection of anemic condition in patients from clinical markers and explainable artificial intelligence." *Technology and Health Care* 32.4 (2024): 2431-2444.

[13] Skorupski, Clarissa P., et al. "Preoperative anemia and iron deficiency in elective gastrointestinal cancer surgery patients." *Journal of Surgical Oncology* 131.4 (2025): 614-623.

[14] Berenji, Fariba, et al. "Assessment of Blastocystis hominis as a risk factor for iron deficiency anemia in pregnant women." *Iranian Journal of Parasitology* 20.1 (2025): 75.

[15] Suva, Manoj A., and Pravin R. Tirgar. "Comparative evaluation of different oral iron salts in the management of iron deficiency anemia." *DARU Journal of Pharmaceutical Sciences* 32.2 (2024): 485-494.

[16] Cui, Xiuxiu, et al. "Evaluation of left ventricular systolic function in patients with iron deficiency anemia based on non-invasive left ventricular pressure–strain loops." *BioMedical Engineering Online* 23.1 (2024): 82.

[17] Gupta, Arun. "Iron deficiency anemia." *Decision making through problem-based learning in hematology: a step-by-step approach in patients with anemia*. Singapore: Springer Nature Singapore, 2024. 1-16.

[18] Koh, Jeong Suk, and Ik-Chan Song. "Functional iron deficiency anemia in patients with cancer." *Blood research* 59.1 (2024): 26.

[19] Pullakhandam, Siddartha, and Susan McRoy. "Classification and explanation of iron deficiency anemia from complete blood count data using machine learning." *BioMedInformatics* 4.1 (2024): 661-672.

[20] Chang, Yu-Hsin, et al. "Machine learning for detecting iron deficiency through comprehensive blood analysis." *Clinical Chemistry* 71.9 (2025): 949-961.

[21] Qasrawi, Radwan, et al. "Identification and prediction of association patterns between nutrient intake and anemia using machine learning techniques: results from a cross-sectional study with university female students from Palestine." *European Journal of Nutrition* 63.5 (2024): 1635-1649.

[22] Asare, Justice Williams, et al. "Application of machine learning approach for iron deficiency anaemia detection in children using conjunctiva images." *Informatics in medicine unlocked* 45 (2024): 101451.

[23] Abdullah, Abdullah, et al. "Anemia prediction using gene expression programming (GEP) and explainable artificial intelligence approaches." *Computers in Biology and Medicine* 196 (2025): 110856.

[24] Inoue, Takahiro, et al. "Predicting anemia management in dialysis patients using open-source machine learning libraries." *Renal Replacement Therapy* 11.1 (2025): 47.

[25] Peng, Chongxiao, et al. "Research on a multi-classification diagnosis method for anemia text medical records in Traditional Chinese Medicine based on an improved pre-trained LERT model." *International Journal of Machine Learning and Cybernetics* (2025): 1-15.

[26] Kitaw, Bekan, Chera Asefa, and Firew Legese. "Leveraging machine learning models for anemia severity detection among pregnant women following ANC: Ethiopian context." *BMC Public Health* 24.1 (2024): 3500.