International Journal of Applied Mathematics and Computing
E-ISSN: 3047-146X
P-ISSN: 3048-1988
IJAMC

Research Article

A Comparison of SVM and ELM Algorithms Based on SMOTE for Anemia Classification Using Hematology Data

Dio Dharmaesa[1,*], Hamdani Hamdani[1], Addy Suyatno[1]

[1] Departement of Informatics, Faculty of Engineering, Universitas Mulawarman, Indonesia.
* Corresponding Author: ddharmaesa@gmail.com

Abstract: Anemia remains a significant global health concern, and its diagnosis through manual interpretation of Complete Blood Count (CBC) results is susceptible to bias and misinterpretation. Machine learning techniques offer a promising solution for identifying complex patterns in medical data; however, their performance is often affected by class imbalance issues commonly found in healthcare datasets. Therefore, this study aims to evaluate and compare the performance of Support Vector Machine (SVM) and Extreme Learning Machine (ELM) algorithms enhanced with the Synthetic Minority Over-sampling Technique (SMOTE) for anemia classification. The proposed approach employs SVM and ELM classifiers with parameter optimization using K-Fold Cross Validation, while SMOTE is applied to address the imbalance in class distribution. The study utilizes a secondary CBC dataset consisting of 364 patient records categorized into Anemia and Non-Anemia classes. Experimental results indicate that the SMOTE-based SVM model achieved an accuracy of 94.52%, precision of 97.14%, recall of 91.89%, and an F1-score of 94.44%, with a computation time of 0.013 seconds. In comparison, the SMOTE-based ELM model attained an accuracy of 91.78%, precision of 89.74%, recall of 94.59%, and an F1-score of 92.11%, while requiring only 0.002 seconds of computation time. The findings suggest that SVM delivers more stable performance and the highest precision, making it highly effective in reducing false positive predictions. On the other hand, ELM demonstrates greater sensitivity to the incorporation of synthetic samples but outperforms SVM in terms of recall and computational efficiency, making it a suitable alternative when rapid processing and higher sensitivity are prioritized.

Keywords: Anemia; Extreme Learning Machine; Hematology Data; SMOTE; Support Vector Machine.

Received: June, 26th 2026
Revised: July, 03th 2026
Accepted: July, 03th 2026
Published: July, 03th 2026
Curr. Ver.: July, 03th 2026

1. Introduction

Anemia continues to be a significant global public health issue that deeply impacts the daily productivity and overall quality of life of those affected. Highlighting the massive scale of this condition, the World Health Organization (WHO) reported that in 2021 alone, approximately 1.92 billion people worldwide suffered from anemia (IHME, 2023). This staggering figure disproportionately affects vulnerable populations, particularly women of reproductive age and young children in developing nations (Shah et al., 2023).

The clinical diagnosis of anemia is primarily performed through a comprehensive blood examination known as a complete blood count (CBC). However, the traditional manual interpretation of this complex hematological data inherently demands a substantial level of clinical expertise and experience from healthcare professionals. Consequently, this conventional approach can potentially lead to human errors or misdiagnoses, particularly in ambiguous cases where the clinical characteristics of the blood parameters closely overlap (Noviandy et al., 2024). With the advancement of artificial intelligence, various machine learning algorithms have been used to analyze hematological parameters to automatically classify anemia conditions (Gómez et al., 2025).

Support Vector Machines (SVMs) are well-known for their ability to handle high-dimensional data and maximize the separation margin between classes (Aditya et al., 2024). On the other hand, Extreme Learning Machines (ELM) are neural network-based algorithms that offer the key advantage of fast training, as the weights in the hidden layers are initialized randomly without an iterative process (Ariyanto et al., 2025; Lin, 2025).

As two reliable algorithms for medical data classification, SVM and ELM have fundamental differences regarding their sensitivity to data distribution. This poses a challenge when dealing with medical datasets, which typically suffer from class imbalance (Martini et al., 2025). To prevent a decline in performance, data balancing techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) are required.

Despite their individual strengths, research directly comparing the performance of these two methods in anemia classification, especially while handling highly imbalanced hematological datasets, remains very limited in the current literature. Therefore, this study aims to systematically compare the classification performance of SMOTE-based SVM and ELM. By doing so, this research seeks to provide practical recommendations for identifying the most optimal and time-efficient algorithm to be deployed in a real-world medical anemia diagnosis system..

2. Literature Review

Several previous studies have explored the use of machine learning for medical data. Research by Wulandari & Badieah (2025) applied an SMOTE-based SVM to imbalanced data and found that this combination was able to achieve an accuracy of 85% with a precision of 87%. In a different classification domain, Ariyanto et al. (2025) used the ELM algorithm combined with a One-Class Support Vector Machine (OCSVM) on a heart disease dataset and achieved an accuracy of 91.27%.

In addition, a study by Aditya et al. (2024) compared Random Forest, SVM, and Logistic Regression on CBC data for anemia and found that Random Forest outperformed SVM (82.10%) and logistic regression. Based on these studies, this research fills a gap by specifically and directly comparing the SVM and ELM algorithms after class balancing using SMOTE on a specialized hematology dataset for anemia.

2.1. Support Vector Machine

Support Vector Machine (SVM) is a supervised learning algorithm widely used for classification and regression tasks (Fawwaz et al., 2023). SVMs work by finding the optimal hyperplane that clearly separates data classes and maximizes the margin between each class (Avci et al., 2023). The Radial Basis Function (RBF) kernel was used in this study due to its ability to handle nonlinear separation. The regularization parameter C controls the balance between the margin width and the classification error rate (Guido et al., 2024).

2.2. Extreme Learning Machine

The Extreme Learning Machine (ELM) is a machine learning algorithm based on a Single-Layer Feedforward Neural Network (SLFN) in which the input weights in the hidden layer are initialized randomly and are not updated during training (Yustisio et al., 2024). The output weights are calculated analytically using the Moore-Penrose pseudoinverse, as shown in Eq. (1) (Permana et al., 2025):

$$\beta = H \dagger \ . \ T \tag{1}$$

where β is the output weight, $H\dagger$ is the Moore-Penrose generalized inverse of the hidden layer output matrix, and T is the target output. This approach allows the ELM to train in a much shorter time compared to traditional gradient-based algorithms (Nugraha et al., 2025).

2.3. Synthetic Minority Over-sampling Technique

SMOTE is an oversampling technique that addresses class imbalance by generating new synthetic samples in the minority class based on the k-nearest neighbors (KNN) principle (Rahayu et al., 2024). For each minority sample, SMOTE selects k nearest neighbors and generates a new sample from among them using the formula: new_sample = minority_sample + $\lambda \times$ (neighbor - minority_sample), where λ is a random number between 0 and 1. This technique is applied exclusively to the training data to avoid data leakage (Hairani et al., 2024; Joloudari et al., 2023).

2.4. K-Fold Cross Validation

K-Fold Cross-Validation is a validation technique that divides a dataset into k subsets (folds). In each iteration, one fold is used as validation data, while the remaining k-1 folds are used as training data (Rahmawati et al., 2026). This process is repeated k times until all folds have had a chance to serve as validation data, The average performance metric across all iterations is used as a more objective and less biased estimate of the model's performance (Shiddiq et al., 2023).

2.5. Model Performance Evaluation

The model's performance is evaluated using a confusion matrix, which is an evaluation method that compares the model's predictions with the actual labels to determine the model's classification ability. The confusion matrix consists of four main components: True Positive (TP), which indicates the number of positive data points correctly classified; True Negative (TN), which indicates the number of negative data points correctly classified; False Positive (FP), which refers to negative data points incorrectly classified as positive; and False Negative (FN), which refers to positive data points incorrectly classified as negative (Tharwat, 2021). Based on these four components, several evaluation metrics are calculated to provide a more comprehensive picture of the classification model's performance. The metrics used in this study include:

- Accuracy, which measures the overall percentage of correct predictions, Eq. (2) is used to calculate Accuracy (Hinojosa Lee et al., 2024).

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{2}$$

- Precision measures the accuracy of positive-class predictions, Eq. (3) is used to calculate Precision (Silmina et al., 2025).

$$\text{Presicion} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{3}$$

- Recall measures the model's ability to identify all true positive-class data, Eq. (4) is used to calculate Recall (Alimi et al., 2025).

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{4}$$

- The F1-score is the harmonic mean of precision and recall, Eq. (5) is used to calculate F1-score (Widyawati & Faradibah, 2023).

$$\text{F1} - \text{score} = 2 \times \frac{(\text{Presicion x Recall})}{(\text{Presicion} + \text{Recall})} \tag{5}$$

3. Materials and Method

This study follows a systematic research methodology to develop and evaluate machine learning models for anemia classification using Complete Blood Count (CBC) data. The proposed methodology consists of several sequential stages, including data collection, data preprocessing and SMOTE implementation, SVM and ELM model implementation, model evaluation, and results analysis. Each stage is designed to ensure an effective data processing workflow, optimal model development, and a comprehensive performance comparison between the two classification algorithms. The overall research workflow is illustrated in Figure 1.

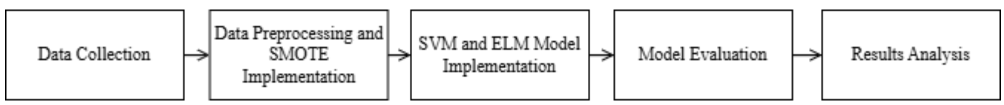

Figure 2. Research Flowchart

3.1. Data Collection

The data used in this study were secondary publicly available data obtained online from the Mendeley Data repository (data.mendeley.com), which was published on January 31, 2026. The dataset was accessed and downloaded on May 8, 2026. The collected data consist of Complete Blood Count (CBC) examination records. The dataset contains a total of 364 patient samples, along with the corresponding target class labels, namely Anemic and Normal. All data were downloaded and stored in CSV (.csv) format for subsequent analysis (Ahmad et al., 2026).

Table 1. Hematology Data

Age	Sex	RBC	PCV	MCV	MCH	MCHC	RDW	TLC	PLT	HGB	Label
28	Female	5.66	34.0	60.1	17.0	28.2	20.0	11.1	128.3	9.6	Anemic
41	Female	4.78	44.5	93.1	28.9	31.0	13.0	7.02	419.0	13.8	Normal
40	Male	4.65	41.6	89.5	28.8	32.2	13.0	8.09	325.0	13.4	Normal
76	Female	4.24	36.7	86.6	26.7	30.8	14.9	13.4	264.0	11.3	Anemic
20	Male	4.14	36.9	89.1	27.8	31.2	13.2	4.75	196	11.5	Anemic
24	Female	4.29	40.1	93.5	29.6	31.7	14.5	13.96	233	12.7	Anemic
28	Male	4.98	42.3	84.9	24.9	29.3	16.2	9.33	213	12.4	Normal
...	....	...	...	...	...	...	...	...	...	...	...
14	Female	4.97	43.8	88.1	28	31.7	15.2	3.92	229	13.9	Normal
16	Female	4.16	38.7	93	28.8	31	17.9	5.77	211	12	Anemic
62	Female	5.25	45.6	86.9	25.3	29.2	15.6	10.68	151	13.3	Normal

Following the sample dataset presented in Table 1, a detailed description of each hematological attribute is provided in Table 2. The table includes the attribute name, data type, description, and unit of measurement, which serve as the input features for the machine learning models developed in this study.

Table 2. Variable of Hematology Data

Attribute Name	Data Type	Description	Unit
Age	Integer	Patient's age at the time of the blood examination.	Years
Gender	Categorical	Patient's gender (Male/Female).	-
(RBC) Red Blood Cell	Float	Number of red blood cells in the blood.	million/µL
(PCV) Packed Cell Volume	Float	Percentage of red blood cell volume relative to the total blood volume	%
(MCV) Mean Corpuscular Volume	Float	Average volume of red blood cell	fL
(MCH) Mean Corpuscular Hemoglobin	Float	Average amount of hemoglobin contained in a single red blood cell	pg/cell
(MCHC) Mean Corpuscular Hemoglobin Consentration	Float	Average concentration of hemoglobin within red blood cells.	g/dL
(RDW) Redcell Distribution Width	Float	Measure of the variation in the size of red blood cells	%
(TCL) T-Cell Lymphoma	Float	Total number of white blood cells in the blood	cells/µL
(PLT/mm3) Platelet	Float	Number of platelets in the blood	cells/µL
(HGB) Hemoglobin	Float	Hemoglobin concentration in the blood	g/dL
Label	String	Target class for classification	-

Figure 2 presents the class distribution of the dataset before the implementation of SMOTE. The visualization highlights the proportion of Anemia and Non-Anemia samples, serving as the basis for evaluating the need for class balancing in the subsequent preprocessing stage.

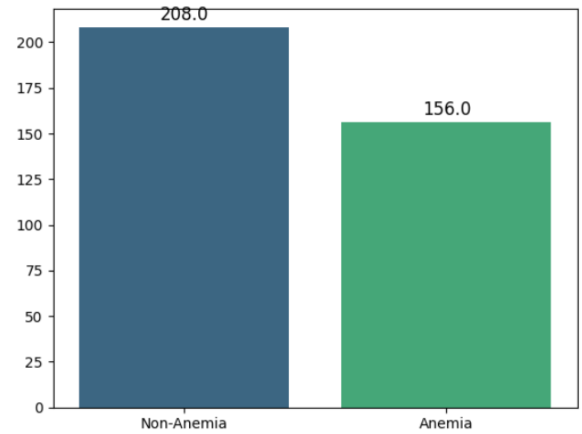

Figure 2. Class Distribusion.

3.2. Data Preprocessing and SMOTE Implementation

The data preprocessing stage was carried out to improve data quality and prepare the dataset for machine learning classification. This process included several steps. First, missing values were handled using mean imputation for numerical attributes and mode imputation for categorical attributes. Second, the Gender attribute was encoded into a numerical format to enable processing by the classification algorithms. The dataset was then split into training and testing sets using an 80:20 ratio, resulting in 291 training samples and 73 testing samples. Finally, all numerical features were normalized using the Min-Max Scaling technique to ensure that each feature contributed equally during model training. The Min-Max normalization formula is presented in Eq. (6):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

The initial class distribution analysis revealed an imbalance between the Anemia and Non-Anemia classes. To address this issue, the Synthetic Minority Over-sampling Technique (SMOTE) was applied only to the training dataset to generate synthetic samples for the minority class, resulting in a more balanced class distribution while preserving the original testing data. The SMOTE-balanced training dataset was then used as an additional experimental scenario and compared with the original training dataset to evaluate its impact on the classification performance of the SVM and ELM models.

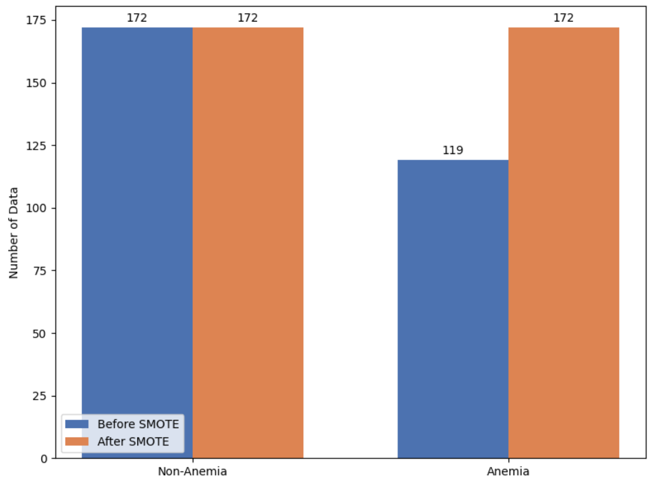

Figure 3. Result SMOTE

Figure 3 shows the class distribution of the training dataset before and after SMOTE implementation. Before oversampling, the dataset contained 172 Non-Anemia and 119 Anemia samples. After applying SMOTE, the minority class was balanced to 172 samples, resulting in an equal class distribution. The balanced training dataset was then used for model training.

3.3. SVM and ELM Model Implementation

The SVM model was implemented using the Radial Basis Function (RBF) kernel with regularization parameter values of 0.1, 1, 10, and 100, while probability=True and random_state=42 were applied to enable probability estimation and ensure reproducible results. The optimal parameter was determined using 5-Fold Cross Validation, where the training data were randomly divided into five folds (shuffle=True), with four folds used for training and one fold for validation in each iteration. The average performance across all folds was calculated for each parameter configuration, and the parameter yielding the highest average accuracy was selected as the optimal configuration. Finally, the selected parameter was used to train the final SVM model before evaluation on the testing dataset.

Algorithm 1. Pseudocode Train SVM

```
INPUT    : X_train_smote, y_train_smote, C_options : {0.1, 1, 10, 100}, k = 5
OUTPUT : best_C, final_svm_model, performance_time.
INITIALIZE best_accuracy = 0
INITIALIZE best_C = 1
SPLIT X_train_smote AND y_train_smote INTO k folds RANDOMLY

FOR EACH c_val IN C_options DO:
    INITIALIZE empty arrays: fold_acc, fold_prec, fold_rec, fold_f1

    FOR i FROM 1 TO k DO:
        SET X_tr, y_tr AS training data FOR fold i
        SET X_val, y_val AS validation data FOR fold i

        DEFINE model svm_cv WITH C = c_val AND selected kernel
        TRAIN svm_cv USING X_tr AND y_tr
        PREDICT X_val USING svm_cv TO GET preds

        COMPUTE accuracy, precision, recall, AND f1-score BETWEEN y_val AND preds
        APPEND metric results TO fold_acc, fold_prec, fold_rec, AND fold_f1
    ENDFOR

    COMPUTE average accuracy (avg_acc) FROM fold_acc
    COMPUTE average precision, recall, AND f1-score FROM respective arrays

    IF avg_acc > best_accuracy THEN
        best_accuracy = avg_acc
        best_C = c_val
    ENDIF
ENDFOR

RECORD start_time
DEFINE final_svm_model WITH C = best_C AND selected kernel
TRAIN final_svm_model USING ALL X_train_smote AND y_train_smote
RECORD end_time
COMPUTE performance_time = end_time - start_time
RETURN best_C, final_svm_model, performance_time
```

The ELM model was implemented using the Sigmoid activation function with the number of hidden neurons set to 10, 20, 30, and 50, while the input weights and biases were randomly initialized following a normal distribution. The optimal number of hidden neurons was determined using 5-Fold Cross Validation, following the same procedure as the SVM model. The average performance across all folds was calculated for each hidden neuron configuration, and the configuration yielding the highest average accuracy was selected as the optimal parameter. Finally, the selected hidden neuron configuration was used to train the final ELM model before evaluation on the testing dataset.

Algorithm 2. Pseudocode Train ELM

```
INPUT     : X_train_smote, y_train_smote, neuron_options : {10, 20, 30, 50}, k = 5
OUTPUT : best_neuron, final_elm_model, performance_time.
INITIALIZE best_accuracy = 0
INITIALIZE best_neuron = 0
SPLIT X_train_smote AND y_train_smote INTO k folds RANDOMLY

FOR EACH n IN neuron_options DO:
    INITIALIZE empty arrays: fold_acc, fold_prec, fold_rec, fold_f1

    FOR i FROM 1 TO k DO:
        SET X_tr, y_tr AS training data FOR fold i
        SET X_val, y_val AS validation data FOR fold i

        GENERATE random input_weights AND biases FOR n neurons
        COMPUTE hidden layer output matrix H = sigmoid(X_tr * input_weights + biases)
        COMPUTE Moore-Penrose pseudoinverse H_pinv OF matrix H
        COMPUTE output_weights = H_pinv * y_tr

        COMPUTE validation matrix H_val = sigmoid(X_val * input_weights + biases)
        COMPUTE raw_predictions = H_val * output_weights
        SET preds = 1 IF raw_predictions >= 0.5 ELSE 0

        COMPUTE accuracy, precision, recall, AND f1-score BETWEEN y_val AND preds
        APPEND metric results TO fold_acc, fold_prec, fold_rec, AND fold_f1
    ENDFOR

    COMPUTE average accuracy (avg_acc) FROM fold_acc
    COMPUTE average precision, recall, AND f1-score FROM respective arrays

    IF avg_acc > best_accuracy THEN
        best_accuracy = avg_acc
        best_neuron = n
    ENDIF
ENDFOR

RECORD start_time
GENERATE random input_weights AND biases FOR best_neuron
COMPUTE final H = sigmoid(X_train_smote * input_weights + biases)
COMPUTE final H_pinv OF matrix H
COMPUTE final output_weights = H_pinv * y_train_smote
RECORD end_time
COMPUTE performance_time = end_time - start_time
RETURN best_neuron, final_elm_model (input_weights, biases, output_weights), performance_time
```

3.4. Model Evaluation

The trained SVM and ELM models were evaluated using the testing dataset, comprising 20% of the total data, to compare their classification performance under three experimental scenarios: the baseline model, the model trained with SMOTE, and the model trained without SMOTE. The prediction results from each scenario were compared using a confusion matrix and evaluated by calculating the Accuracy based on Eq. (2), Precision based on Eq. (3), Recall based on Eq. (4), and F1-Score based on Eq. (5). In addition, the training time of each model was recorded to assess its computational efficiency.

3.5. Results Analysis

The evaluation results were analyzed to compare the performance of the SVM and ELM models under different experimental scenarios. Based on the 5-Fold Cross Validation, the optimal parameters were C = 100 for SVM and 30 hidden neurons for ELM. The analysis compared Accuracy, Precision, Recall, F1-Score, and training time, as well as the effect of applying SMOTE on the classification performance of both models.

4. Results and Discussion

The experiments were conducted using Google Colaboratory (Google Colab) with the Python programming language and supporting libraries. A publicly available Complete Blood Count (CBC) dataset obtained from the Mendeley Data repository was used in this study. The performance of the SVM and ELM models was evaluated three experimental scenarios: the baseline model, the model with SMOTE, and the model without SMOTE, to analyze the impact of SMOTE on the classification performance.

4.1. Baseline Model Testing

Table 3. Baseline Model Testing

Algorithm	Accruacy	Presision	Recall	F1-score	Time Performace
SVM	0.5753	0.00	0.00	0.00	0.079085s
ELM	0.4521	0.4366	0.1	0.6078	0.002236s

The baseline evaluation was conducted using an 80:20 train-test split to establish the initial performance of the SVM and ELM models before applying SMOTE and hyperparameter tuning. As shown in Table 3, the baseline SVM model achieved an accuracy of 57.53%, while failing to correctly identify the minority class, resulting in 0.00 precision, recall, and F1-score. In contrast, the baseline ELM model achieved an accuracy of 45.21%, with a precision of 43.66%, recall of 100%, and F1-score of 60.78%. These baseline results serve as a reference for evaluating the performance improvements after applying SMOTE.

4.2. Testing the Model Using SMOTE

Table 4. Testing the Model Using SMOTE

Algorithm	Accruacy	Presision	Recall	F1-score	Time Performace
SVM	0.9452	0.9714	0.9189	0.9444	0.013257s
ELM	0.9178	0.8974	0.9459	0.9211	0.002415s

The evaluation with SMOTE demonstrated a substantial improvement in the performance of both classification models. As shown in Table 4, the SVM model achieved an accuracy of 94.52%, precision of 97.14%, recall of 91.89%, and an F1-score of 94.44%, with a training time of 0.013257 s. Meanwhile, the ELM model achieved an accuracy of 91.78%, precision of 89.74%, recall of 94.59%, and an F1-score of 92.11%, requiring only 0.002415 s for training. These results indicate that SMOTE effectively improved the classification performance of both models, with SVM providing the highest overall performance and ELM offering faster computation and higher recall.

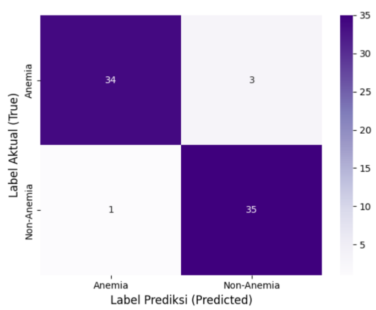

Figure 4. Confusion Matrix SVM using SMOTE

Figure 4 presents the confusion matrices of the SVM models after applying SMOTE. For the SVM model, 34 Anemia and 35 Non-Anemia samples were correctly classified, with 3 false negatives and 1 false positive, indicating strong performance in reducing incorrect positive predictions.

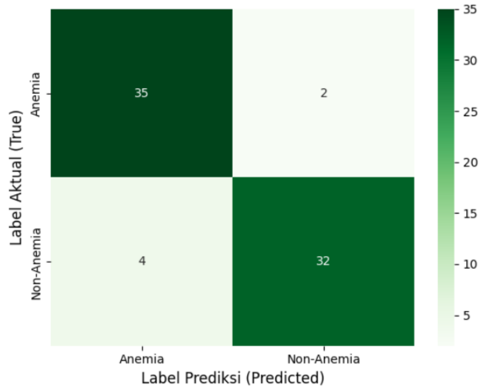

Figure 5. Confusion Matrix ELM using SMOTE

Figure 5 presents the confusion matrices of the ELM models after applying SMOTE. For the ELM model, 35 Anemia and 32 Non-Anemia samples were correctly classified, with 2 false negatives and 4 false positives. These results indicate that ELM achieved slightly higher sensitivity in detecting Anemia cases but produced more false positive predictions than SVM.

4.3. Testing the Model Without SMOTE

Table 5. Testing the Model Without SMOTE

Algorithm	Accruacy	Presision	Recall	F1-score	Time Performace
SVM	0.9452	0.9714	0.9189	0.9444	0.010997s
ELM	0.9452	0.9459	0.9459	0.9459	0.019888s

The evaluation without SMOTE showed that both models achieved high classification performance. As presented in Tables 5, the SVM model obtained an accuracy of 94.52%, precision of 97.14%, recall of 91.89%, and an F1-score of 94.44%, with a training time of 0.010997 s. Meanwhile, the ELM model achieved the same accuracy of 94.52%, with precision, recall, and F1-score of 94.59%, requiring 0.019888 s for training. Although both models achieved identical accuracy, SVM produced higher precision, whereas ELM demonstrated a more balanced performance across the evaluation metrics.

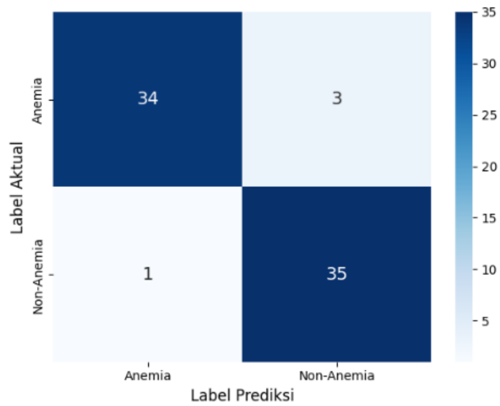

Figure 6. Confusion Matrix SVM using SMOTE

Figure 6 presents the confusion matrices of the SVM models without applying SMOTE. For the SVM model, 34 Anemia and 35 Non-Anemia samples were correctly classified, with 3 false negatives and 1 false positive, indicating strong classification performance and a low number of misclassified samples.

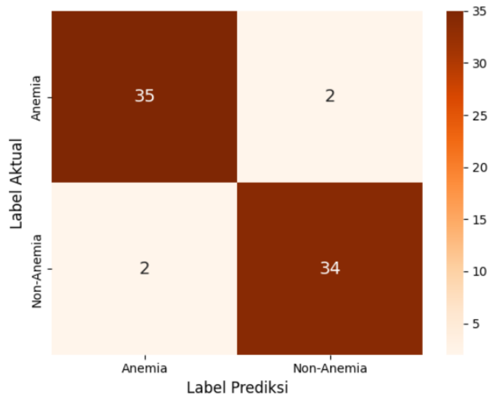

Figure 7. Confusion Matrix SVM using SMOTE

Figure 7 presents the confusion matrices of the ELM models without applying SMOTE For the ELM model, 35 Anemia and 34 Non-Anemia samples were correctly classified, with 2 false negatives and 2 false positives. Compared with SVM, the ELM model correctly identified more Anemia cases while maintaining a relatively low number of false positive predictions, resulting in a balanced classification performance.

4.4. Model Performance Analysis

The experimental results demonstrate that both SVM and ELM are capable of effectively classifying anemia based on Complete Blood Count (CBC) data. Although both algorithms achieved satisfactory performance, they exhibited different classification characteristics. SVM tended to produce more stable predictions and fewer false positive classifications, indicating its ability to construct a more reliable decision boundary. In contrast, ELM showed greater sensitivity in identifying anemia cases, making it suitable for applications where minimizing missed positive cases is more important.

The implementation of SMOTE also influenced the performance of both algorithms. For SVM, the application of SMOTE maintained stable classification performance, suggesting that the algorithm is relatively robust to synthetic samples generated through oversampling. On the other hand, ELM was more sensitive to the addition of synthetic data, which affected its classification behavior. Nevertheless, SMOTE contributed to improving the representation of the minority class, allowing both models to learn more balanced decision boundaries compared with the baseline model.

Overall, the findings indicate that handling class imbalance plays an important role in anemia classification. The combination of SVM and SMOTE provides more stable and reliable classification performance, while ELM remains an attractive alternative due to its high sensitivity and computational efficiency. These results demonstrate that the selection of a classification algorithm should consider not only predictive performance but also the specific requirements of the intended medical application.

5. Comparison

To assess the contribution of this study in a more quantifiable manner, the results obtained were compared with several relevant previous studies, as shown in Table 6.

Table 6. Comparison with Previous Research

Research	Algorithm	Accuracy	Presision	F1-score
Wulandari & Badieah (2025)	SVM + SMOTE (OvO)	85,00%	87,00%	87,00%
Aditya et al. (2022)	SVM	82,10%	-	-
Ariyanto et al. (2025)	ELM + OCSVM	91,27%	-	-
This Research	SVM + SMOTE (K-Fold)	94,52%	97,14%	94,44%
This Research	ELM + SMOTE (K-Fold)	91,78%	89,74%	92,11%

Based on Table 6, the SMOTE-based SVM model in this study achieved the highest precision (97.14%) compared to all previous studies that used SVM. Overall, this study contributes by providing the first head-to-head comparison between SMOTE-based SVM and ELM on a hematology dataset focused on anemia, along with an analysis of the impact of SMOTE on each algorithm across various evaluation metrics.

6. Conclusion

This study compared the performance of Support Vector Machine (SVM) and Extreme Learning Machine (ELM) for anemia classification using Complete Blood Count (CBC) data and evaluated the impact of the Synthetic Minority Over-sampling Technique (SMOTE) on model performance. The experimental results showed that the SMOTE-based SVM model achieved the best overall performance, with an accuracy of 94.52%, precision of 97.14%, recall of 91.89%, and an F1-score of 94.44%. Meanwhile, the SMOTE-based ELM model achieved an accuracy of 91.78%, precision of 89.74%, recall of 94.59%, and an F1-score of 92.11%, while requiring a shorter training time. Without SMOTE, both algorithms achieved the same accuracy of 94.52%, although SVM demonstrated higher precision and ELM provided a more balanced performance in terms of recall and F1-score.

These findings indicate that SMOTE effectively improves the classification performance by addressing class imbalance, while the combination of SVM, SMOTE, and 5-Fold Cross Validation provides the most stable and reliable results for anemia classification. This study contributes by presenting a direct comparison between SVM and ELM under the same experimental conditions and demonstrates that appropriate class balancing can enhance machine learning performance for hematological data analysis.

This study is limited to a single publicly available CBC dataset and two machine learning algorithms. Future research should investigate larger and more diverse datasets, explore alternative data balancing techniques, feature selection methods, and additional machine learning or deep learning models to further improve the robustness and generalizability of anemia classification systems.

References

Aditya, M. R., Sutanto, T., Budiman, H., Noor Ridha, M. R., Syapotro, U., & Azijah, N. (2024). Machine Learning Models for Classification of Anemia from CBC Results:Random Forest, SVM, and Logistic Regression. *INTI Journal*, *2022*(1). https://doi.org/10.61453/jods.v2023no49

Ahmad, M. H., Bormon, S., & Hossain, Md. A. (2026). *Expert-Labeled Clinical Dataset for Anemia Detection* [Dataset]. Mendeley Data. https://doi.org/10.17632/8SYCPBHZN7.2

Alimi, F., Khan, A., Ali, H., Bouzidi, M., Alshammari, A. O., & Ahmad, B. (2025). M-estimation activation functions for high-performance extreme learning machine ensemble classification. *Scientific Reports*, *15*(1), 32154. https://doi.org/10.1038/s41598-025-16798-5

Ariyanto, D., Novitasari, D. C. R., & Hamid, A. (2025). *Heart Disease Classification Using Extreme Learning Machine (ELM) Method With Outlier Handling One-Class Support Vector Machine*. *9*(5). https://doi.org/https://doi.org/10.30871/jaic.v9i5.9763

Avci, C., Budak, M., Yağmur, N., & Balçik, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, *8*(1), 1–10. https://doi.org/10.26833/ijeg.987605

Fawwaz, Mazdadi, Muliadi, Saragih, & Friska Abadi. (2023). Implementation of Information Gain Ratio and Particle Swarm Optimization in the Sentiment Analysis Classification of Covid-19 Vaccine Using Support Vector Machine. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *5*(4), 261–270. https://doi.org/10.35882/jeeemi.v5i4.328

Gómez, J. G., Parra Urueta, C., Álvarez, D. S., Hernández Riaño, V., & Ramirez-Gonzalez, G. (2025). Anemia Classification System Using Machine Learning. *Informatics*, *12*(1), 19. https://doi.org/10.3390/informatics12010019

Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. *Information*, *15*(4), 235. https://doi.org/10.3390/info15040235

Hairani, H., Widiyaningtyas, T., & Dwi Prasetya, D. (2024). Addressing Class Imbalance of Health Data: A Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies. *JOIV : International Journal on Informatics Visualization*, *8*(3), 1310. https://doi.org/10.62527/joiv.8.3.2283

Hinojosa Lee, M. C., Braet, J., & Springael, J. (2024). Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores. *Applied Sciences*, *14*(21), 9863. https://doi.org/10.3390/app14219863

IHME, I. (2023). *The Lancet: New study reveals global anemia cases remain persistently high among women and children. Anemia rates decline for men. | Institute for Health Metrics and Evaluation*. https://www.healthdata.org/news-events/newsroom/news-releases/lancet-new-study-reveals-global-anemia-cases-remain-persistently

Joloudari, J. H., Marefat, A., Nematollahi, M. A., Oyelere, S. S., & Hussain, S. (2023). Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks. *Applied Sciences*, *13*(6), 4006. https://doi.org/10.3390/app13064006

Lin, Z. (2025). Optimizing Kernel Extreme Learning Machine based on a Enhanced Adaptive Whale Optimization Algorithm for classification task. *PLOS ONE*, *20*(1), e0309741. https://doi.org/10.1371/journal.pone.0309741

Martini, Pradipta, G. A., & Huizen, R. R. (2025). Analysis of the Impact of Data Oversampling on the Support Vector Machine Method for Stroke Disease Classification. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(2), 404–421. https://doi.org/10.35882/jeeemi.v7i2.698

Noviandy, T. R., Idroes, G. M., Suhendra, R., Bakri, T. K., & Idroes, R. (2024). Advanced Anemia Classification Using Comprehensive Hematological Profiles and Explainable Machine Learning Approaches. *Infolitika Journal of Data Science*, *2*(2), 72–81. https://doi.org/10.60084/ijds.v2i2.237

Nugraha, I. D., Saragih, T. H., Budiman, I., Kartini, D., Indriani, F., & Caesarendra, W. (2025). Performance Comparison of Extreme Learning Machine (ELM) and Hierarchical Extreme Learning Machine (H-ELM) Methods for Heart Failure Classification on Clinical Health Datasets. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(3), 713–728. https://doi.org/10.35882/jeeemi.v7i3.904

Permana, I., Shir Li Wang, Febi Nur Salisah, Sanusi, & Febi Yanto. (2025). A Benchmark of Activation Functions in Extreme Learning Machine for High-Dimensional Low-Sample-Size Microarray Classification. *Jurnal Inotera*, *10*(2), 502–515. https://doi.org/10.31572/inotera.Vol10.Iss2.2025.ID593

Rahayu, W., Deny Jollyta, Hajjah, A., Johan, Gusrianty, Gustientiedina, Marlim, Y. N., & Desnelita, Y. (2024). Synthetic Minority Oversampling Technique (SMOTE) for Boosting the Accuracy of C4.5 Algorithm Model. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, *3*(3), 624–630. https://doi.org/10.59934/jaiea.v3i3.469

Rahmawati, N. P., Irwan Budiman, Muhammad Itqan Mazdadi, Andi Farmadi, & Friska Abadi. (2026). Comparison Between K-Fold Cross Validation And Percentage Split In Decision Tree Algorithms For Anemia Classification. *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *8*(1), 105–114. https://doi.org/10.35882/ijeeemi.v8i1.315

Shah, S. A., Soomro, U., Ali, O., Tariq, Y., Waleed, M. S., Guntipalli, P., & Younus, N. (2023). The Prevalence of Anemia in Working Women. *Cureus*. https://doi.org/10.7759/cureus.44104

Shiddiq, M., Candra, F., Anand, B., & Rabin, M. F. (2023). Neural network with k-fold cross validation for oil palm fruit ripeness prediction. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *22*(1), 164. https://doi.org/10.12928/telkomnika.v22i1.24845

Silmina, E. P., Hardiani, T., & Mahfida, S. L. (2025). The Effect of the Number of Classes on the Values Resulting from Evaluation Metrics in the YOLOv5 Model. *International Journal on Advanced Science, Engineering and Information Technology*, *15*(2), 419–425. https://doi.org/10.18517/ijaseit.15.2.20495

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

Widyawati & Faradibah. (2023). Comparison Analysis of Classification Model Performance in Lung Cancer Prediction Using Decision Tree, Naive Bayes, and Support Vector Machine. *Indonesian Journal of Data and Science*, *4*(2), 80–89. https://doi.org/10.56705/ijodas.v4i2.76

Wulandari, N., & Badieah, B. (2025). Implementasi Teknik Resampling Untuk Mengatasi Ketidakseimbangan Data Terhadap Klasifikasi Anemia Menggunakan Support Vector Machine. *Jurnal Rekayasa Sistem Informasi dan Teknologi*, *2*(3), 942–951. https://doi.org/10.70248/jrsit.v2i3.1856

Yustisio, D., Siswanah, E., & Tafrikan, M. (2024). Extreme Learning Machine and Multilayer Perceptron Methods for Predicting COVID-19. *Sinkron*, *8*(4), 2298–2308. https://doi.org/10.33395/sinkron.v8i4.14029