Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 519-527
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

# Comparative Study of Ensemble Learning Models for Blood Cell Anomaly Detection

**Fany Andhina[1,*], Anjar Wanto[2], Agus Perdana Windarto[2], Poningsih[2], Solikhun[2]**

[1] Master's Informatics, Sekolah Tinggi Ilmu Komputer Tunas Bangsa, Pematangsiantar, Indonesia
[2] Department of Information System, Sekolah Tinggi Ilmu Komputer Tunas Bangsa, Pematangsiantar, Indonesia
Jl.Kartini, Proklamasi, 21143, Pematangsiantar, Indonesia
Email: [1,*]fanyandhinaa@gmail.com, [2]anjarwanto@amiktunasbangsa.ac.id, [3]agus.perdana@amiktunasbangsa.ac.id,
[4]poningsih@amiktunasbangsa.ac.id, [5]solikhun@amiktunasbangsa.ac.id
Correspondence Author Email: fanyandhinaa@gmail.com
Submitted: **01/06/2026**; Accepted: **12/06/2026**; Published: **30/06/2026**

**Abstract**−Blood cell anomaly detection plays a crucial role in the early diagnosis of hematological disorders, including anemia, leukemia, and other blood-related diseases. Accurate identification of abnormal blood cells can support timely clinical decision-making and improve patient outcomes. However, selecting the most effective machine learning model for blood cell anomaly detection remains challenging due to variations in data characteristics and model performance. This study aims to compare the effectiveness of three ensemble learning models, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), for blood cell anomaly detection. The study utilized a publicly available blood cell anomaly dataset consisting of 5,880 blood cell records with morphological, staining, and hematological features. Data preprocessing included feature selection, categorical encoding, and train-test splitting using an 80:20 ratio. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC). Experimental results showed that LightGBM achieved the best performance with an Accuracy of 98.13%, Precision of 98.10%, Recall of 96.01%, F1-Score of 97.04%, AUC of 99.83%, and MCC of 0.9569. Feature importance analysis further revealed that granularity score, cell diameter, and cell area were the most influential features. These findings indicate that LightGBM is a highly effective approach for blood cell anomaly detection and can support the development of intelligent hematology diagnostic systems.

**Keywords**: Blood Cell Anomaly Detection; Ensemble Learning; Hematological Disorders; Machine Learning; Feature Importance

## 1. INTRODUCTION

Blood disorders represent a major global health challenge due to their widespread prevalence and significant impact on human health (Miah et al., 2024). Diseases such as anemia, leukemia, sickle cell disease, and other hematological abnormalities affect millions of people worldwide and often require early diagnosis to prevent severe complications and improve treatment outcomes(Mondal et al., 2025). Accurate identification of abnormal blood cells is therefore essential for supporting timely clinical decision-making and effective patient management (Jiao et al., 2023). Traditionally, blood cell examination is performed manually through microscopic observation by hematologists and laboratory specialists (Alhejaily, 2025). Although this approach remains a standard practice in clinical settings, it is often time-consuming, labor-intensive, and susceptible to inter-observer variability, particularly when large numbers of samples must be analyzed (Faiyazuddin et al., 2025).

Recent advances in artificial intelligence (AI) and machine learning (ML) have introduced promising solutions for improving the efficiency and reliability of hematological diagnostics (Rahman et al., 2024). By learning complex relationships from blood cell characteristics, machine learning models can automate classification and anomaly detection tasks while reducing human dependency (Furizal et al., 2023) . The growing availability of digital healthcare data and computational resources has further accelerated the adoption of AI-driven diagnostic systems in various medical domains, including hematology (Wood et al., 2024).

Several studies have demonstrated the effectiveness of machine learning and deep learning techniques for blood disorder diagnosis and blood cell classification (Cakmak & Pacal, 2025). Previous research has explored machine learning algorithms for classifying anemia and blood-related disorders using hematological parameters, while other studies have proposed computer vision and deep learning frameworks for automated blood cell anomaly detection (Dalvi & Gawas, 2025). Furthermore, ensemble deep learning architectures and hybrid feature extraction approaches have achieved promising classification performance in white blood cell and blood cell recognition tasks (Ullah et al., 2025). These findings highlight the growing importance of AI-based methods in supporting hematological diagnosis and medical decision-making (Kholish et al., 2025).

Despite these advancements, several challenges remain (Cakmak & Pacal, 2025). First, most existing studies focus on image-based deep learning approaches that require large image datasets, high computational resources, and complex model architectures (Dalvi & Gawas, 2025). Such requirements may limit their practical implementation in resource-constrained environments. Second, comparative evaluations involving modern ensemble learning algorithms, particularly Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), remain relatively limited for blood cell anomaly detection using structured hematological and morphological features (Yıldız & Kalayci, 2025). Third, many studies primarily emphasize classification accuracy while providing limited insight into the contribution of individual features, resulting in reduced model interpretability (Alhejaily, 2025). Consequently, understanding which blood cell characteristics contribute most significantly to anomaly detection remains an important

research challenge. Recent review studies have also emphasized the need for more comprehensive comparative analyses and interpretable machine learning frameworks in hematology(Noor et al., 2025).

Motivated by these research gaps, this study investigates the effectiveness of ensemble learning models for blood cell anomaly detection using a publicly available blood cell dataset consisting of morphological, staining, and hematological features. In addition to model performance evaluation, feature importance analysis is incorporated to improve model interpretability and provide insights into the characteristics associated with blood cell abnormalities.

The objective of this study is to compare the performance of Random Forest, XGBoost, and LightGBM in detecting blood cell anomalies and to identify the most effective model for this task. Model performance is evaluated using multiple metrics, including Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC), to ensure a comprehensive assessment. Experimental results indicate that LightGBM achieves the best overall performance, demonstrating its effectiveness for blood cell anomaly detection.

This study provides three main contributions. First, it presents a comprehensive comparative analysis of three widely used ensemble learning algorithms for blood cell anomaly detection using structured hematological data. Second, it employs multiple evaluation metrics, including MCC, to provide a more robust assessment of classification performance. Third, it performs feature importance analysis on the best-performing model to identify the most influential characteristics associated with blood cell anomalies. The findings are expected to contribute to the development of accurate, interpretable, and computationally efficient intelligent diagnostic systems for hematological applications.

## 2. RESEARCH METHODS

### 2.1 Research Framework

The research framework employed in this study is illustrated in Figure 1. The proposed framework consists of five main stages: dataset collection, data preprocessing, model training, performance evaluation, and feature importance analysis. Initially, a publicly available blood cell anomaly dataset containing morphological, staining, and hematological characteristics was collected and prepared for analysis.

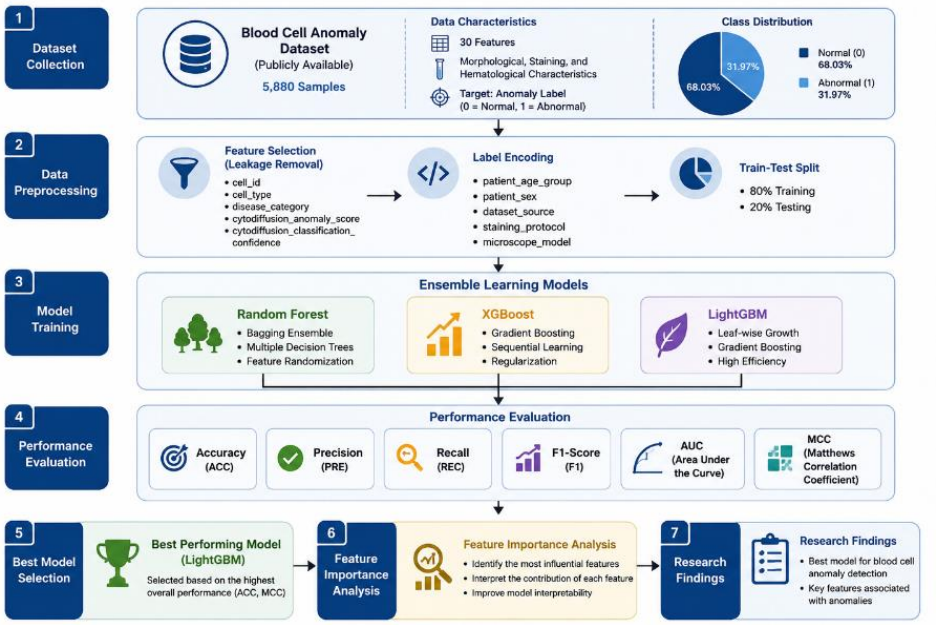


**Figure 1.** Research Framework

During the preprocessing stage, irrelevant and potentially leakage-prone features were removed to ensure model reliability. Categorical attributes were transformed into numerical representations using label encoding, and the dataset was subsequently divided into training and testing subsets using an 80:20 ratio. In the model training stage, three ensemble learning algorithms, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), were trained and evaluated. Model performance was assessed using Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC). Finally, feature importance analysis was conducted on the best-performing model to identify the most influential features contributing to blood cell anomaly detection and to improve model interpretability.

### 2.2 Data Preprocessing

This study utilized a publicly available Blood Cell Anomaly dataset obtained from the Kaggle platform. The dataset was designed to support machine learning research in hematology and contains blood cell records represented by

morphological, staining, and hematological characteristics. Each record corresponds to an individual blood cell observation and is labeled according to its anomaly status. The target variable, referred to as the anomaly label, is a binary classification variable where 0 represents normal blood cells and 1 represents abnormal blood cells.

The dataset consists of 5,880 samples and 30 predictive features describing various characteristics of blood cells. These features include cell morphology measurements, staining properties, structural attributes, and patient-related information. The diversity of the available features enables machine learning models to capture complex patterns associated with blood cell abnormalities. Prior to model development, several features identified as irrelevant or potentially causing data leakage were excluded during the preprocessing stage to ensure reliable model evaluation. Regarding class distribution, the dataset contains 4,000 normal blood cell samples (68.03%) and 1,880 abnormal blood cell samples (31.97%). Although the dataset exhibits a moderate class imbalance, both classes remain sufficiently represented for supervised machine learning tasks. Table 1 summarizes the main characteristics of the dataset used in this study.

**Table 1**. Dataset Characteristics

| Item | Value |
|---|---|
| Total Samples | 5,880 |
| Total Features | 30 |
| Target Variable | Anomaly Label |
| Classification Type | Binary |
| Normal Samples | 4,000 (68.03%) |
| Abnormal Samples | 1,880 (31.97%) |
| Features Categories | Morphological, Staining, and Hematological |

### 2.3 Data Preprocessing

Data preprocessing was conducted to improve data quality and ensure that the machine learning models were trained using relevant and reliable features. This stage consisted of feature selection, categorical data encoding, and dataset partitioning.This stage consisted of three main processes: feature selection, categorical data encoding, and dataset partitioning. The preprocessing workflow is illustrated in Figure 1 and was applied consistently across all machine learning models used in this study. The first preprocessing step involved feature selection through the removal of irrelevant and leakage-prone attributes. Several features, including *cell_id*, *cell_type*, *disease_category*, *cytodiffusion_anomaly_score*, and *cytodiffusion_classification_confidence*, were excluded from the dataset. These attributes were removed because they either served as identifiers, contained information directly related to the target variable, or had the potential to introduce data leakage, which could lead to overly optimistic model performance and compromise the validity of the experimental results.

After feature selection, categorical variables were transformed into numerical representations using the Label Encoding technique. The encoded attributes included *patient_age_group*, *patient_sex*, *dataset_source*, *staining_protocol*, and *microscope_model*. This transformation was necessary because machine learning algorithms require numerical input features for model training and prediction. Finally, the processed dataset was divided into training and testing subsets using an 80:20 ratio. The training set was used to develop the machine learning models, while the testing set was reserved for performance evaluation. This data partitioning strategy was adopted to ensure an unbiased assessment of model generalization capability on unseen data. The main data preprocessing procedures applied in this study are summarized in Table 2.

**Table 2**. Data Preprocessing Procedures

| Procedure | Purpose |
|---|---|
| Feature Selection | Removed irrelevant and leakage-prone attributes to improve model reliability and prevent biased predictions. |
| Label Encoding | Converted categorical variables into numerical representations suitable for machine learning algorithms. |
| Train-Test Split (80:20) | Divided the dataset into training and testing subsets for model development and performance evaluation. |

### 2.4 Ensemble Learning Models

Data Ensemble learning has become one of the most effective machine learning approaches for classification tasks due to its ability to combine multiple learners and improve predictive performance (Mahajan et al., 2023). By integrating the strengths of several decision-making processes, ensemble methods can reduce variance, minimize overfitting, and enhance model generalization (Saha et al., 2026). In this study, three widely used ensemble learning algorithms, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), were employed to detect blood cell anomalies. Random Forest is a bagging-based ensemble algorithm that constructs multiple decision trees using randomly selected subsets of training samples and features (Abdulatif & Ali, 2025). The final

prediction is obtained through majority voting among individual trees, enabling the model to achieve high robustness and resistance to overfitting . Due to its simplicity and strong baseline performance, Random Forest has been widely adopted in various medical classification applications (Vlachas et al., 2022).

XGBoost is a gradient boosting algorithm that builds decision trees sequentially, where each new tree is trained to correct the prediction errors of previous trees (Chen & Guestrin, 2016). The algorithm incorporates regularization mechanisms to improve model generalization and prevent overfitting. Owing to its high predictive capability and computational efficiency, XGBoost has become one of the most popular machine learning algorithms in healthcare analytics and biomedical research. LightGBM is an advanced gradient boosting framework that utilizes a leaf-wise tree growth strategy to improve training efficiency and predictive accuracy. Compared to conventional boosting algorithms, LightGBM is capable of handling large datasets with lower computational costs while maintaining high classification performance (Ismanto et al., 2024) . Its ability to efficiently capture complex relationships among features makes it particularly suitable for blood cell anomaly detection. These three ensemble learning algorithms were selected because they represent different ensemble learning strategies, including bagging and boosting, and have demonstrated strong performance in various classification tasks. Their comparative evaluation enables a comprehensive assessment of their effectiveness in detecting blood cell anomalies using structured hematological data.

### 2.5 Performance Evaluation

The performance of the proposed models was evaluated using six widely adopted classification metrics, namely Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC). These metrics were selected to provide a comprehensive assessment of classification performance from different perspectives, including overall classification accuracy, positive prediction reliability, sensitivity to abnormal cases, discrimination capability, and classification quality under class imbalance conditions. Accuracy measures the proportion of correctly classified instances among all observations and is calculated using Equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision evaluates the proportion of correctly predicted positive instances among all instances predicted as positive and is defined in Equation (2).

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall measures the ability of a model to correctly identify positive instances and is expressed in Equation (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

F1-Score represents the harmonic mean of Precision and Recall and is calculated using Equation (4).

$$F1 - Score = 2x\frac{Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

Matthews Correlation Coefficient (MCC) provides a balanced evaluation by considering all elements of the confusion matrix simultaneously and is defined in Equation (5).

$$MCC = \frac{(TP\ x\ TN)-(FP\ x\ FN)}{\sqrt{(TP+FP)\ (TP+FN)\ (TN+FP)\ (TN+FN}} \quad (5)$$

Where *TP* denotes True Positive, *TN* denotes True Negative, *FP* denotes False Positive, and *FN* denotes False Negative. In addition, the Area Under the Curve (AUC) metric was employed to evaluate the discriminative capability of each model across different classification thresholds. Higher AUC values indicate better separation between normal and abnormal blood cell classes. Among these evaluation metrics, MCC received particular attention because it provides a more balanced assessment by considering all components of the confusion matrix simultaneously. Therefore, MCC was used together with Accuracy, Precision, Recall, F1-Score, and AUC to comprehensively evaluate the performance of the proposed ensemble learning models.

## 3. RESULTS AND DISCUSSION

This section presents the experimental results obtained from the implementation of three ensemble learning models, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), for blood cell anomaly detection. The evaluation was conducted using multiple performance metrics, including Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC), to provide a comprehensive assessment of model effectiveness. In addition to quantitative performance evaluation, graphical analyses using Receiver Operating Characteristic (ROC) curves, Precision–Recall (PR) curves, and MCC comparisons were performed to further investigate the classification capabilities of the evaluated models. The experimental findings are discussed to identify the most effective ensemble learning approach for blood cell anomaly detection.

Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 519-527
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

### 3.1 Classification Performance Result

Table 3 presents the classification performance of the evaluated ensemble learning models. Overall, all models achieved excellent predictive performance, indicating that the selected hematological and morphological features contained substantial information for distinguishing between normal and abnormal blood cells. However, performance differences were observed among the evaluated algorithms.

**Table 3.** Performance Comparison of Ensemble Learning Models

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) | MCC |
|---|---|---|---|---|---|---|
| Random Forest | 97.11 | 97.77 | 93.09 | 95.37 | 99.65 | 0.933 |
| XGBoost | 98.04 | 98.09 | 95.74 | 96.90 | 99.75 | 0.955 |
| LightGBM | 98.13 | 98.10 | 96.01 | 97.04 | 99.83 | 0.957 |

As shown in Table 3, LightGBM achieved the highest performance across all evaluation metrics, obtaining an Accuracy of 98.13%, Precision of 98.10%, Recall of 96.01%, F1-Score of 97.04%, AUC of 99.83%, and MCC of 0.957. XGBoost ranked second with highly competitive results, while Random Forest demonstrated the lowest performance among the evaluated models. Although the differences between LightGBM and XGBoost were relatively small, LightGBM consistently outperformed the other models across all metrics, suggesting superior capability in identifying blood cell anomalies.

The Receiver Operating Characteristic (ROC) curves generated by the evaluated models are presented in Figure 2. ROC analysis was conducted to examine the discriminative capability of each model across different classification thresholds.

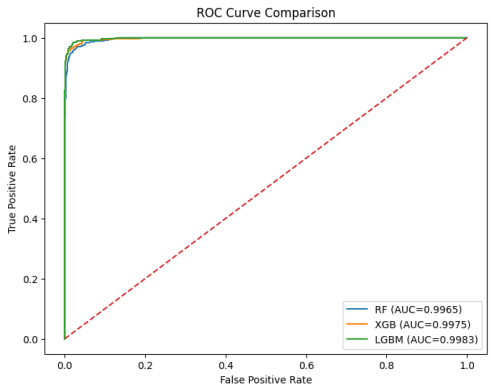


**Figure 2.** ROC Curve Comparison of Ensemble Learning Models

Figure 2 shows that all evaluated models achieved near-perfect ROC curves, with AUC values exceeding 99%. These results indicate that the selected features possess strong discriminative characteristics for blood cell anomaly detection. Among the evaluated models, LightGBM obtained the highest AUC value of 99.83%, followed by XGBoost with 99.75% and Random Forest with 99.65%. The superior ROC performance achieved by LightGBM demonstrates its ability to distinguish between normal and abnormal blood cells more effectively across varying classification thresholds.

To further assess classification performance under class distribution differences, Precision–Recall (PR) curve analysis was conducted. The resulting curves are presented in Figure 3.

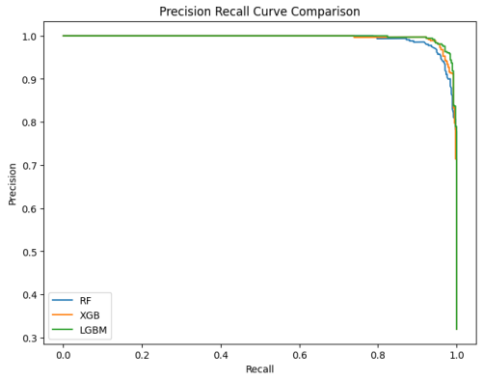


**Figure 3.** ROC Curve Comparison of Ensemble Learning Models

As illustrated in Figure 3, all models maintained high precision and recall values throughout the evaluation process. However, LightGBM demonstrated a more favorable balance between correctly identifying abnormal blood cells and minimizing false positive predictions. This finding is consistent with the superior Recall and F1-Score values reported in

Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 519-527
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

Table 3. The results suggest that LightGBM provides a more reliable classification performance, which is particularly important in medical diagnostic applications where missed abnormal cases may lead to serious clinical consequences.

In addition to conventional evaluation metrics, this study employed Matthews Correlation Coefficient (MCC) to provide a balanced assessment of classification quality. The MCC comparison among the evaluated models is shown in Figure 4.

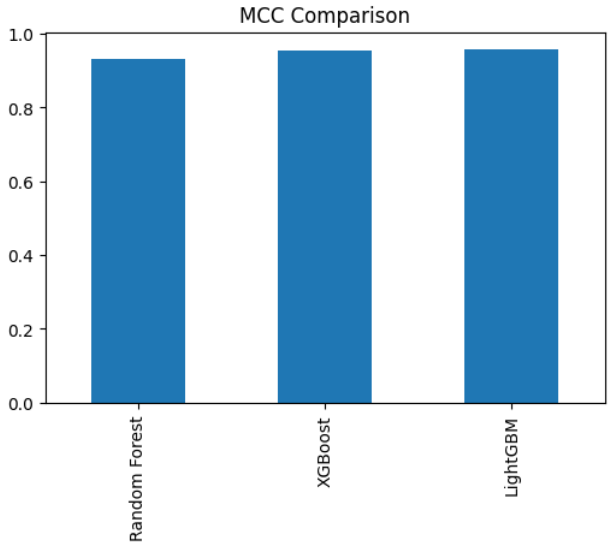


**Figure 4.** MCC Comparison of Ensemble Learning Models

Figure 4 indicates that LightGBM achieved the highest MCC value of 0.957, followed by XGBoost with 0.955 and Random Forest with 0.933. Since MCC simultaneously considers true positives, true negatives, false positives, and false negatives, it provides a more comprehensive evaluation of classification performance. The consistently superior MCC obtained by LightGBM further confirms its robustness and reliability in blood cell anomaly detection. These findings collectively demonstrate that boosting-based ensemble learning approaches, particularly LightGBM, are highly effective for classification tasks involving structured hematological data.

### 3.2 Feature Importance Analysis

To improve model interpretability and provide insights into the factors contributing to blood cell anomaly detection, feature importance analysis was conducted using the best-performing model, namely LightGBM. Feature importance analysis enables the identification of variables that have the greatest influence on the classification process and helps explain how the model distinguishes between normal and abnormal blood cells.

The feature importance results obtained from LightGBM are illustrated in Figure 5, while the ranking of the ten most influential features is presented in Table 4. The analysis revealed that *granularity_score*, *cell_diameter_um*, and *cell_area_px* were the three most important features contributing to blood cell anomaly detection. These features obtained importance scores of 578, 556, and 460, respectively, indicating that morphological and structural characteristics play a dominant role in the classification process.

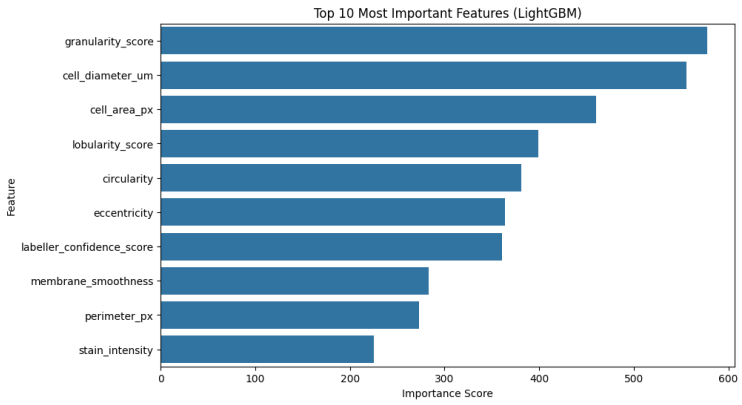


**Figure 5.** Top 10 Most Important Features Identified by LightGBM

The results shown in Figure 5 demonstrate that the importance values gradually decrease after the top three features, suggesting that multiple attributes contribute to model predictions rather than relying on a single dominant variable. In addition to *granularity_score*, several shape-related features, including *lobularity_score*, *circularity*, and

*eccentricity*, were also ranked among the most influential predictors. This finding highlights the significance of cell morphology in distinguishing abnormal blood cells from normal cells.

**Table 4.** Top 10 Feature Importance Ranking

| Rank | Feature | Importance |
|---|---|---|
| 1 | granularity_score | 578 |
| 2 | cell_diameter_um | 556 |
| 3 | cell_area_px | 460 |
| 4 | lobularity_score | 399 |
| 5 | circularity | 381 |
| 6 | eccentricity | 364 |
| 7 | labeller_confidence_score | 361 |
| 8 | membrane_smoothness | 283 |
| 9 | perimeter_px | 273 |
| 10 | stain_intensity | 225 |

As presented in Table 4, the majority of highly ranked features are associated with morphological characteristics, including cell size, shape, texture, and structural properties. The prominence of *cell_diameter_um*, *cell_area_px*, *perimeter_px*, and *membrane_smoothness* indicates that abnormal blood cells exhibit distinguishable morphological patterns that can be effectively captured by machine learning models. Furthermore, the importance of *granularity_score* suggests that cellular texture and granule distribution are critical indicators for anomaly identification. From a hematological perspective, these findings are consistent with clinical observations, where abnormalities in cell morphology frequently serve as important indicators of blood-related disorders. Variations in cell size, texture, and shape are commonly associated with pathological conditions and are routinely evaluated during laboratory examinations. Therefore, the feature importance analysis not only enhances model transparency but also demonstrates that the predictions generated by LightGBM are supported by biologically meaningful characteristics.

Overall, the feature importance results provide valuable insights into the decision-making mechanism of the model and strengthen the practical applicability of the proposed approach. By identifying the most influential blood cell characteristics, the developed system can support more interpretable and trustworthy machine learning-assisted diagnostic tools for hematological applications.

### 3.3 Discussion

The experimental results demonstrate that ensemble learning algorithms are highly effective for blood cell anomaly detection using structured hematological data. Among the evaluated models, LightGBM consistently achieved the best performance across all evaluation metrics, including Accuracy, F1-Score, AUC, and MCC. This finding indicates that the leaf-wise learning strategy employed by LightGBM is capable of capturing complex relationships among blood cell characteristics more effectively than Random Forest and XGBoost (Deng et al., 2025). The results obtained in this study are consistent with previous studies that reported the effectiveness of artificial intelligence and machine learning techniques in hematological diagnosis (Cakmak & Pacal, 2025; Naouali & El Othmani, 2025). Research on blood disorder classification and blood cell anomaly detection has shown promising predictive performance using both machine learning and deep learning approaches (Ullah et al., 2025). However, most previous studies primarily focused on image-based methods, whereas this study demonstrates that structured hematological and morphological features can also achieve excellent classification performance when combined with ensemble learning algorithms (Vičaitė et al., 2024).

Feature importance analysis revealed that *granularity_score*, *cell_diameter_um*, and *cell_area_px* were the most influential features contributing to anomaly detection. These findings indicate that cell morphology and structural characteristics play a crucial role in distinguishing abnormal blood cells from normal cells (Mărcău et al., 2025). The identified features are also consistent with hematological knowledge, where variations in cell size, texture, and shape are commonly associated with pathological conditions (Ennab & Mcheick, 2024).

The main contribution of this study lies in the comprehensive comparison of Random Forest, XGBoost, and LightGBM for blood cell anomaly detection, combined with feature importance analysis to improve model interpretability. The findings demonstrate that LightGBM provides superior predictive performance while maintaining meaningful feature interpretation, making it a promising approach for intelligent hematological diagnostic systems (Deng et al., 2025).

## 4. CONCLUSION

This study investigated the effectiveness of ensemble learning algorithms for blood cell anomaly detection using a publicly available hematological dataset consisting of 5,880 samples and 30 features. Three ensemble learning models, namely Random Forest, XGBoost, and LightGBM, were evaluated using multiple performance metrics, including Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Matthews Correlation Coefficient (MCC). The experimental results demonstrated that all evaluated models achieved excellent classification performance, indicating that

structured hematological and morphological features contain substantial information for anomaly detection. Among the evaluated algorithms, LightGBM consistently achieved the best performance, obtaining an Accuracy of 98.13%, F1-Score of 97.04%, AUC of 99.83%, and MCC of 0.957. Furthermore, feature importance analysis revealed that *granularity_score*, *cell_diameter_um*, and *cell_area_px* were the most influential features contributing to the classification process, highlighting the importance of morphological and structural characteristics in distinguishing abnormal blood cells. These findings confirm that ensemble learning, particularly LightGBM, is a promising approach for developing accurate, reliable, and interpretable blood cell anomaly detection systems. Despite these encouraging results, this study is limited to a single publicly available dataset and focuses exclusively on structured tabular features. Therefore, future research should investigate the generalizability of the proposed approach using larger and more diverse datasets, explore hybrid approaches combining tabular and image-based information, and incorporate advanced Explainable Artificial Intelligence (XAI) techniques to further enhance model transparency and clinical applicability in hematological diagnostic systems.

## ACKNOWLEDGMENT

Thank you to those who have supported the implementation of this research.

## REFERENCES

Abdulatif, F., & Ali, A. (2025). Review of Classification Medical Images Using Ensemble Learning. *NTU Journal of Pure Sciences*, *3*(4). https://doi.org/10.56286/10q1f229

Alhejaily, A. M. G. (2025). Artificial intelligence in healthcare (Review). In *Biomedical Reports* (Vol. 22, Number 1). Spandidos Publications. https://doi.org/10.3892/br.2024.1889

Cakmak, Y., & Pacal, I. (2025). AI-Driven Classification of Anemia and Blood Disorders Using Machine Learning Models. *Computers and Electronics in Medicine*, *2*(2), 43–52. https://doi.org/10.69882/adba.cem.2025073

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. https://doi.org/10.1145/2939672.2939785

Dalvi, P. T., & Gawas, M. A. (2025). A comprehensive review of machine learning approaches for detecting anemia and abnormal red blood cells using RBC indices and medical imaging. *Discover Artificial Intelligence*. https://doi.org/10.1007/s44163-025-00698-8

Deng, L., Lu, K., & Hu, H. (2025). An interpretable LightGBM model for predicting coronary heart disease: Enhancing clinical decision-making with machine learning. *PLOS ONE*, *20*(9 September). https://doi.org/10.1371/journal.pone.0330377

Ennab, M., & Mcheick, H. (2024). Enhancing interpretability and accuracy of AI models in healthcare: a comprehensive review on challenges and future directions. In *Frontiers in Robotics and AI* (Vol. 11). Frontiers Media SA. https://doi.org/10.3389/frobt.2024.1444763

Faiyazuddin, M., Rahman, S. J. Q., Anand, G., Siddiqui, R. K., Mehta, R., Khatib, M. N., Gaidhane, S., Zahiruddin, Q. S., Hussain, A., & Sah, R. (2025). The Impact of Artificial Intelligence on Healthcare: A Comprehensive Review of Advancements in Diagnostics, Treatment, and Operational Efficiency. In *Health Science Reports* (Vol. 8, Number 1). John Wiley and Sons Inc. https://doi.org/10.1002/hsr2.70312

Furizal, Ma'arif, A., & Rifaldi, D. (2023). Application of Machine Learning in Healthcare and Medicine: A Review. In *Journal of Robotics and Control (JRC)* (Vol. 4, Number 5, pp. 621–631). Department of Agribusiness, Universitas Muhammadiyah Yogyakarta. https://doi.org/10.18196/jrc.v4i5.19640

Ismanto, E., Fadlil, A., Yudhana, A., & Kitagawa, K. (2024). A Comparative Study of Improved Ensemble Learning Algorithms for Patient Severity Condition Classification. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *6*(3), 312–321. https://doi.org/10.35882/jeeemi.v6i3.452

Jiao, J., Yin, X., Ma, J., Xia, Y., Xu, J., Zhao, S., & Liu, J. (2023). Comparison of an automated digital cell morphology analysis system with manual counting. *ILABMED*, *1*(1), 22–28. https://doi.org/10.1002/ila2.10

Kholish, A., Shobiry, F., & Puspitasari, R. (2025). Hybrid Feature Benchmark for Blood Cell Classification Using ResNet50 and EfficientNetV2 Features with SVM and ANN Classifiers via Unsupervised Segmentation. *International Journal of Artificial Intelligence in Medical Issues*, *3*. https://doi.org/10.56705/ijaimi.v3i2.364

Mahajan, P., Uddin, S., Hajati, F., & Moni, M. A. (2023). Ensemble Learning for Disease Prediction: A Review. In *Healthcare (Switzerland)* (Vol. 11, Number 12). MDPI. https://doi.org/10.3390/healthcare11121808

Mărcău, F. C., Peptan, C., Iliuta, F. P., Cojoaca, M. E., Musetescu, A. M., Holt, A. G., Tomescu, I. R., Căruntu, G. A., & Gheorman, V. (2025). The Impact of the Ukraine Conflict on the Quality of Life of the Young Population in Romania from a Societal Security Perspective. *Healthcare (Switzerland)*, *13*(2). https://doi.org/10.3390/healthcare13020156

Miah, H., Kollias, D., Pedone, G. L., Provan, D., & Chen, F. (2024). *Can Machine Learning Assist in Diagnosis of Primary Immune Thrombocytopenia? A feasibility study*. http://arxiv.org/abs/2405.20562

Mondal, S. K., Talukder, M. S. H., Aljaidi, M., Sulaiman, R. Bin, Tushar, M. M. S., & Alsuwaylimi, A. A. (2025). Efficient blood cell classification from microscopic smear images using U-Net segmentation and a lightweight CNN. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-26947-5

Naouali, S., & El Othmani, O. (2025). AI-Driven Automated Blood Cell Anomaly Detection: Enhancing Diagnostics and Telehealth in Hematology. *Journal of Imaging*, *11*(5). https://doi.org/10.3390/jimaging11050157

Noor, A. A., Manzoor, A., Mazhar Qureshi, M. D., Qureshi, M. A., & Rashwan, W. (2025). Unveiling Explainable AI in Healthcare: Current Trends, Challenges, and Future Directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *15*(2). https://doi.org/10.1002/widm.70018

Rahman, A., Debnath, T., Kundu, D., Khan, M. S. I., Aishi, A. A., Sazzad, S., Sayduzzaman, M., & Band, S. S. (2024). Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities. In

*AIMS Public Health* (Vol. 11, Number 1, pp. 58–109). American Institute of Mathematical Sciences. https://doi.org/10.3934/publichealth.2024004

Saha, R., Saif, S., Khan, T., & Alabdultif, A. (2026). A critical systematic review of ensemble learning methods for predictive healthcare. In *Discover Computing* (Vol. 29, Number 1). Springer Science and Business Media B.V. https://doi.org/10.1007/s10791-026-10005-3

Ullah, T., Ullah, K., Khan, P., Brikova, O., Voznesensky, A., Eroshkin, A., & Hussain, I. (2025). Fuzzy logic and deep learning approach for automated white blood cell detection and classification via multi-CNN feature fusion. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-28354-2

Vičaitė, G., Barišauskaitė, L., Bakstytė, V., Siesky, B., Verticchio Vercellin, A., & Janulevičienė, I. (2024). Cardiac Surgery Patients Have Reduced Vascularity and Structural Defects of the Retina Similar to Persons with Open-Angle Glaucoma. *Diagnostics*, *14*(5). https://doi.org/10.3390/diagnostics14050515

Vlachas, C., Damianos, L., Gousetis, N., Mouratidis, I., Kelepouris, D., Kollias, K.-F., Asimopoulos, N., & Fragulis, G. F. (2022). Random forest classification algorithm for medical industry data. *SHS Web of Conferences*, *139*, 03008. https://doi.org/10.1051/shsconf/202213903008

Wood, D., Mu, T., Webb, A., Reeve, H., Luján, M., & Brown, G. (2024). *A Unified Theory of Diversity in Ensemble Learning*. http://arxiv.org/abs/2301.03962

Yıldız, A. Y., & Kalayci, A. (2025). *Gradient Boosting Decision Trees on Medical Diagnosis over Tabular Data*. https://doi.org/10.1109/ICAD65464.2025.11114069