

Article

Anemia Classification System Using Machine Learning

Jorge Gómez Gómez ^{1,*}, Camilo Parra Urueta ¹, Daniel Salas Álvarez ¹, Velssy Hernández Riaño ¹
and Gustavo Ramirez-Gonzalez ²

¹ Departamento de Ingeniería de Sistemas y Telecomunicaciones, Universidad de Córdoba, Montería 230002, Colombia; cparraurueta63@correo.unicordoba.edu.co (C.P.U.); danielsalas@correo.unicordoba.edu.co (D.S.Á.); vlhernandez@correo.unicordoba.edu.co (V.H.R.)

² Departamento de Telemática, Universidad del Cauca, Popayán 190001, Colombia; gramirez@unicauca.edu.co

* Correspondence: jelienergomez@correo.unicordoba.edu.co

Abstract: In this study, a system was developed to predict anemia using blood count data and supervised learning algorithms. Anemia, a common condition characterized by low levels of red blood cells or hemoglobin, affects oxygenation and often causes symptoms, such as fatigue and shortness of breath. The diagnosis of anemia often requires laboratory tests, which can be challenging in low-resource areas where anemia is common. We built a supervised learning approach and trained three models (Linear Discriminant Analysis, Decision Trees, and Random Forest) using an anemia dataset from a previous study by Sabatini in 2022. The Random Forest model achieved an accuracy of 99.82%, highlighting its capability to subclassify anemia types (microcytic, normocytic, and macrocytic) with high precision, which is a novel advancement compared to prior studies limited to binary classification (presence/absence of anemia) of the same dataset.

Keywords: anemia; machine learning; e-health



Received: 12 November 2024

Revised: 2 February 2025

Accepted: 7 February 2025

Published: 11 February 2025

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

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Anemia is a disease characterized by a concentration below the normal standards of the number of red blood cells (RBCs) or hemoglobin (Hb). The most frequent symptoms are fatigue, palpitations, headache, shortness of breath, and conjunctival and palm pallor. Although these symptoms are an initial guide, they are not definitive. They act as compasses indicating a possible direction, but do not guarantee an accurate diagnosis [1].

Anemia worldwide has a multifactorial etiology; that is, there are many causes by which the disease can occur, with iron deficiency being the most common cause in developing countries. It is estimated that 30% of women and 40% of children under five years of age worldwide have this disease. In Colombia, anemia is associated with poor nutrition due to poverty [2]. In 2022, 18.3 million people were economically impoverished and 6.9 million people were extremely impoverished. Although there was a 3.1% reduction in poverty rates that year, this is still an alarming statistic [3].

Anemia can be classified according to its cause. For example, anemia can be microcytic, owing to iron deficiency, normocytic, owing to inflammation, or macrocytic, owing to vitamin B12 deficiency. In addition, it can be classified according to severity as mild, moderate, or severe. Conventional methods for diagnosing anemia present significant barriers for the most disadvantaged populations. The lack of access to quality health services, coupled with the costs associated with laboratory tests, prevents many people from receiving timely diagnoses [4]. This situation generates a vicious circle in which anemia evolves out of control, aggravating complications and deteriorating the quality of life of those who have the disease.

On the other hand, it has been shown that predictive models of health based on machine learning manage to improve or optimize the allocation of resources in healthcare, for example, in early warning systems which allow professionals to act in advance. In addition, they reduce subjectivity, which translates into more precise and accurate predictions [5].

By using machine learning techniques, this study aims to develop an intelligent system capable of accurately and efficiently classifying the various types of anemia based on the analysis of hematological data, such as blood counts. This tool will not only speed up the diagnostic process, but will also allow for treatment that is tailored to the needs of each individual, improving clinical outcomes and quality of life.

The World Health Organization estimates that 29.9% of women between 15 and 49 years of age have anemia. In children aged 6 to 59 months, the prevalence of this disease reached 39.8%, while in Africa, within the same age group, 60.2% of children were affected [6]. This condition can lead to serious complications if not properly diagnosed and treated.

Traditional diagnosis of anemia often requires complex laboratory tests and specialized personnel, which can be a challenge in rural and low-resource areas [4]. In addition, identifying the specific type of anemia is crucial for proper treatment, as causes can range from iron deficiency to genetic disorders such as thalassemia and sickle cell anemia.

In recent years, machine learning and deep learning have shown their potential to improve the diagnosis and classification of different types of anemia using blood count data [7].

The use of machine learning and data analysis techniques offers a promising solution to improve the accuracy and efficiency of anemia diagnosis, building advanced AI models, such as those described by Prajapati et al. [8]. Machine learning models can analyze large volumes of blood count data and provide rapid and accurate diagnoses. These models not only identify the presence of anemia, but also classify its different types based on specific hematological parameters.

However, much work remains to be done to integrate these technologies into clinical practice. Further research is needed to improve the robustness, interpretability, and scalability of these AI-based models [7].

The use of machine learning for the classification of anemia types using blood count data represents a significant advancement in the field of hematology. By addressing the limitations of traditional methods and improving diagnostic accuracy, these technologies have the potential to transform healthcare, especially in economically vulnerable areas.

This article is organized as follows: it starts with an introduction, then there is the Related Works section, then the background section, then comes the Materials and Methods section, then the Results section, then there is the Discussion section, and finally Conclusions and Future Works.

2. Related Works

Our consideration of the present state of the art is organized into two main sections. The first is a bibliometric analysis that highlights a network diagram visualizing the current state of research. The second section is dedicated to related research, where studies that are directly related to the objective of classifying types of anemia using blood count data and machine learning are presented, as well as those that, although not directly related, have a more tangential relationship with the objective.

2.1. Bibliometric Analysis

It is crucial to understand the current state of research on anemia and to evaluate the impact of artificial intelligence and machine learning technologies on the diagnostic

support of this condition. For this purpose, we defined and analyzed the following key variables in detail:

- Anemia: types of anemia, such as microcytic, normocytic, macrocytic, and hemolytic anemia.
- Blood count: complete blood count (CBC), Mean corpuscular volume (MCV), hemoglobin, etc.
- Deep learning: applications involving neural networks, machine learning, predictive, and deep learning techniques.

The following search equation was designed to ensure topical relevance and precision, which are crucial elements in any study. We included specific terms such as “anemia”, “blood count”, and “deep learning”, interconnected using the AND operator to ensure that the retrieved documents comprehensively covered key aspects of the research. In addition, we used the OR operator to encompass a variety of literature that may employ different vocabulary when describing similar concepts, which is essential in fields such as medicine and technology, where terminology can vary significantly. For example, terms such as “CBC” and “neural networks” help capture variants in the description of complete blood count and neural networks, respectively. The NOT operator was used to exclude documents that, although related, often lacked the necessary depth of research.

(“anemia” OR “anemic conditions”) AND (“blood count” OR “CBC” OR “hemogram”) AND (“deep learning” OR “neural networks” OR “machine learning”) AND NOT (“conference review” OR “book”) (1).

Three academic databases were selected for their broad coverage and relevance to medicine and technology. These included PubMed, which is essential for accessing biomedical literature; IEEE Xplore, which provides a technical perspective with its focus on engineering and computational sciences; and Science Direct, known for its interdisciplinary indexing and analytical tools. A four-year inspection window, from 2021 to 2024, was implemented to ensure the inclusion of the latest research.

Figure 1 shows a network that explores the relationships between several key concepts related to anemia and machine learning. Each node in the diagram represents a key concept or a topic. The nodes are colored differently to group related topics, making it easier to visualize the areas of research that are most closely connected to each other, and the lines connecting the nodes represent the relationships or links between different concepts. These may indicate conceptual relationships, such as topics that are often studied together, or may reflect more direct relationships, such as techniques applied to a specific problem. The main thematic groups were as follows. Machine learning: The center of the diagram shows how machine learning is applied to the diagnosis and classification of anemia. This is a central topic that appears to connect related subtopics. Anemia types: On the right, the nodes indicate different types of anemia, such as “iron deficiency anemia” and “hereditary anemias”, showing how these specific categories are being studied or diagnosed using machine learning technologies. Diagnostic technology: On the left, nodes such as “algorithms”, “non-invasive”, and “invasive” refer to the methods used to diagnose anemia. This includes invasive and noninvasive techniques that can be optimized using machine learning algorithms. Specific applications: on the periphery of the diagram, specific applications such as “digital holography” and “point-of-care diagnostics” are observed, indicating emerging technologies and their implementation in the field of anemia diagnostics.

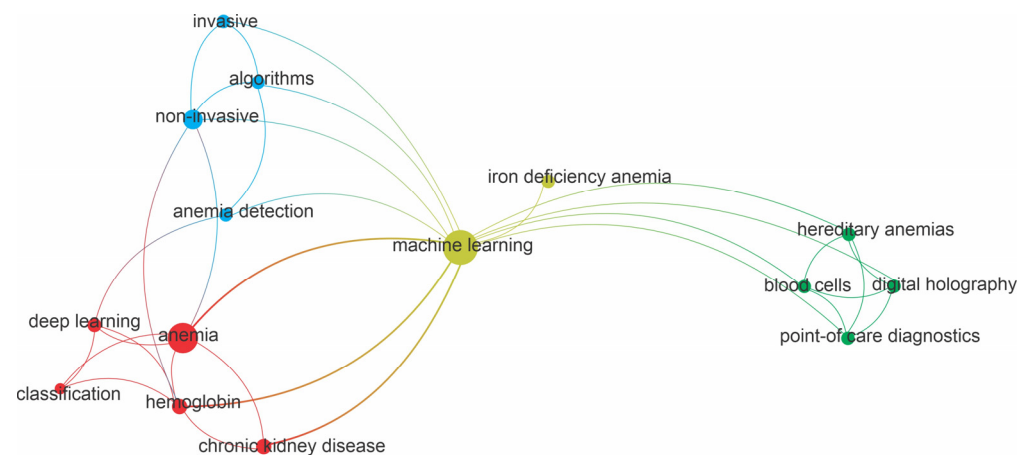


Figure 1. Network diagram of current state of research, generated in VOSviewer software, version 1.6.20.

2.2. Related Research

In ref. [8], the authors propose an explainable model (XAIA) for the automated classification of anemia types, using the XGBoost classifier and SHAP (Shapley Additive Explanations) values for the interpretation and causality of decisions. The model, trained on data from the All India Institute of Medical Sciences, Raipur, India, achieved an accuracy of 96.95%. The XAIA not only classifies eight different types of anemia but also provides a clear explanation of the model predictions, helping patients and healthcare workers to understand the underlying causes of the diagnosed anemia. The inclusion of interpretability in the anemia classification improves the clarity of predictions and promotes the uptake of AI-based automated diagnostics in resource-limited settings.

In ref. [9], the authors addressed anemia prediction using various machine learning techniques including logistic regression, K-Nearest Neighbors (KNN), Decision Trees, Naive Bayes, Random Forest, Stochastic Gradient Descent (SGD), Ridge Classifier, Extreme Gradient Boosting (XGBoost), and Bagging Classifier. Using a dataset from a local pathology center with 1000 instances and 8 attributes, the study achieved 95% accuracy with the logistic regression technique, highlighting the effectiveness of machine learning algorithms in the early and accurate detection of anemia from CBC reports.

In ref. [10], an anemia prediction model was proposed using data analytics and explainable AI to provide detailed clinical assistance. Using data preprocessing techniques, normalization, and class imbalance approaches, the model, based on lightweight gradient boosting learning (LGM Boost), achieved an accuracy of 91%. The implementation of explainable AI allows for a detailed interpretation of the most influential attributes in the prediction, facilitating healthcare professionals to make informed decisions regarding the optimal treatment for anemia.

In ref. [11], the authors investigated and validated two multiclass classification strategies, One-Vs-All (OVA) and One-Vs-One (OVO), using logistic regression (LR) models and Support Vector Machines (SVMs) to predict anemia levels. Using a public dataset, the authors achieved outstanding accuracy, precision and recall of up to 95.05%, 0.951, and an area under the curve (AUC) of 0.990 with the OVO strategy using the LR model. This superior performance demonstrates the viability of advanced machine learning techniques for the accurate and differential diagnosis of anemia levels, which could significantly improve the identification and treatment of this condition in clinical settings. Furthermore, the study provides a detailed analysis of how each model configuration and strategy affects the classification performance, providing valuable insights for future research and practical applications in personalized medicine.

In ref. [12], the authors develop an alpha thalassemia classifier using machine learning techniques to categorize patients based on specific genetic mutations. Using algorithms such as Decision Trees, Artificial Neural Networks, Naive Bayes Algorithms, and Support Vector Machines, the study achieves an accuracy rate of 95% and Kappa statistics of 0.947. The classifier was evaluated using a patient dataset, demonstrating a significant ability to accurately and efficiently identify alpha thalassemia, which could revolutionize the diagnosis and treatment of this genetic disease.

In ref. [13], the authors explored the use of explainable artificial intelligence (XAI) to improve transparency and understanding in the diagnosis of iron deficiency anemia using blood parameters. The authors applied machine learning models to classify anemia states and used the SHAP (Shapley Additive Explanations) model to explain the importance of different blood parameters in predicting anemia. This not only improved diagnostic accuracy, but also increased physicians' confidence in automated decisions, facilitating a more informed and fair approach to the treatment of anemia.

In ref. [14], the authors presented an alternative method for diagnosing and differentiating malaria from several types of anemia, including sickle cell anemia, megaloblastic anemia, and thalassemia, using a convolutional neural network (CNN). The authors applied the CNN technique to process high-resolution images of blood samples, without the need for a standard complete blood count (CBC) protocol. This approach provides a fast and low-cost method for diagnosing these medical conditions, achieving an overall accuracy of 93.4% for testing.

In ref. [15], the study focuses on the use of advanced artificial intelligence techniques for disease estimation, specifically employing the YOLO algorithm for the detection and counting of blood cells in blood smear images. The modified YOLO algorithm was trained on a custom dataset to detect white blood cells (WBCs), red blood cells (RBCs), and platelets, providing a powerful and efficient tool for performing complete blood cell counts (CBCs), which are critical for diagnosing diseases such as leukemia, thrombocytopenia, and anemia.

In ref. [16], the authors addressed hemoglobin estimation and anemia severity prediction using machine learning techniques by applying regression and classification models based on complete blood count (CBC) data collected from Bangabandhu Sheikh Mujib Medical University. They used advanced algorithms, such as Random Forest and neural networks, to accurately predict the severity of anemia and improve diagnostic capability in clinical settings.

In ref. [17], the authors present a prediction model to determine the curability of anemia using machine learning techniques, including Naive Bayes, linear regression (LR), LASSO, and Exponential Smoothing (ES) algorithms. Using complete blood count (CBC) data collected from pathology centers, this model not only predicts the presence of anemia, but also estimates the probability of patients being cured within 90 days. The Naive Bayes algorithm showed superior performance in terms of accuracy compared to the other methods, offering a valuable tool for clinical decision-making in anemia management.

In ref. [18], the authors developed an ensemble classifier for the identification of β -thalassemia carriers using red blood cell indices. By implementing a combined model of three machine learning algorithms (Support Vector Machine, Gradient Boosting Machine, and Random Forest), this method achieved an accuracy of 93%. This approach provides a fast and inexpensive solution for screening β -thalassemia carriers using accessible complete blood count (CBC) tests.

In ref. [19], the authors studied the classification of 12 different types of anemia using artificial learning methods, including Artificial Neural Networks, Support Vector Machines, Naïve Bayes, and Ensemble Decision Trees. A dataset of 1663 samples with 25 attributes obtained from patient files at a university hospital in Turkey was used. The models achieved

high accuracy, with the Bagged Decision Tree method achieving the highest accuracy of 85.6%. This approach provides significant support for accurately and quickly classifying anemia under general clinical conditions, potentially improving treatment management.

In ref. [20], the study presents an innovative multi-level Deep Convolution Encode–Decode Network (DCED-Net) for the semantic segmentation of anemic red blood cells (RBCs). The proposed method addresses the pixel-level segmentation challenges that are crucial for diagnosing blood-related diseases. Two state-of-the-art datasets were developed: one of healthy RBCs and one of anemic RBCs, each containing 1000 images and ground-truth masks for cross-correspondence analysis. The DCED-Net model achieved training, validation, and testing accuracies of 0.9856, 0.9760, and 0.9720, respectively, for the healthy RBC dataset, and 0.9736, 0.9696, and 0.9591, respectively, for the anemic RBC dataset. Furthermore, IoU and BScore values of 0.9311 and 0.9138 were achieved for the healthy RBC dataset, and 0.9032 and 0.8978 for the anemic RBC dataset, respectively. This approach ensures the accurate and efficient segmentation of blood cells and improves the identification of morphological features relevant to clinical diagnosis.

In ref. [21], the study focuses on exploratory data analysis (EDA) and feature selection to achieve maximum accuracy in anemia classification using Random Forest, which allows 100% accuracy. The study does not explicitly mention the application of the model in low-resource settings or its potential use in clinical practice.

As can be seen in Table 1, most studies exploring anemia prediction using machine learning algorithms achieved very high performance, with accuracies ranging from 91% to 100%. This indicates that AI techniques can not only match but also outperform traditional medical diagnostic approaches. However, in our proposal, we not only identified the presence of anemia, but also subclassified the types of anemia, such as microcytic, normocytic, and macrocytic anemia. As can be seen in the results obtained, our system achieved an accuracy of 99.82% in prediction, outperforming the methods described above, considering these subclassifications.

Table 1. A comparative analysis of the different techniques used for the detection of anemia.

Study	Year	AI/ML Techniques Used	Accuracy
[8]	2024	XGBoost, SHAP	96.95%
[9]	2024	Logistic regression, KNN, Decision Trees, Naive Bayes, Random Forest, SGD, XGBoost, Bagging Classifier	95%
[10]	2024	LightGBM, Explainable AI (XAI)	91%
[11]	2023	Logistic regression, SVM	95.05%
[12]	2024	Decision Trees, RNA, Naive Bayes, SVM	95%
[13]	2024	ML models, SHAP	N/A
[14]	2024	Convolutional Neural Networks (CNNs)	93.4%
[15]	2024	YOLO	N/A
[16]	2024	Random Forest, Redes Neuronales	N/A
[17]	2024	Naive Bayes, LR, LASSO, ES	N/A
[18]	2024	SVM, Gradient Boosting, Random Forest	93%
[19]	2024	Artificial Neural Networks, SVM, Naive Bayes, Decision Trees	85.6%
[20]	2021	DCED-Net	97.36%
[21]	2022	EDA, Random Forest	100%
Current Study		Random Forest, Decision Trees, Linear Discriminant Analysis	99.82%

3. Materials and Methods

3.1. Framework and Workflow

The framework and workflow of the proposed model is presented below, Figure 2 shows the main elements that form the workflow of the model.

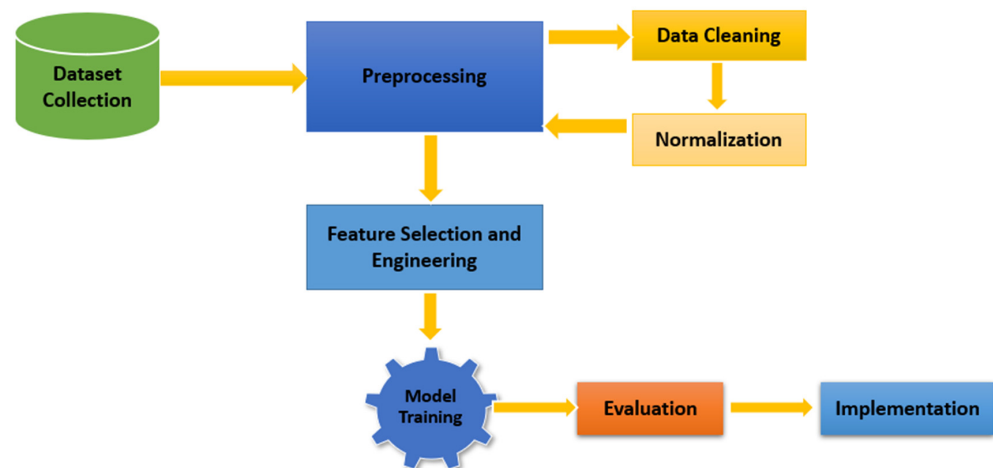


Figure 2. Framework and workflow of proposed model.

An extended description of the framework is included in the following steps:

1. Dataset Collection and Preprocessing:
 - Source: Kaggle dataset with 1421 instances.
 - Data cleaning: removal of incomplete or inconsistent records.
 - Normalization: ensuring values are scaled to eliminate bias due to variable magnitudes.
2. Feature Selection and Engineering:
 - Parameters like HGB, MCV, MCHC, and RDW were selected based on their diagnostic significance.
 - Thresholds were defined for anemia classification (e.g., MCV for microcytic: <80).
3. Model Training:
 - Algorithms: Random Forest, Decision Trees, and Linear Discriminant Analysis.
 - Dataset split into training (80%) and testing (20%) sets.
4. Evaluation:
 - Metrics: accuracy, precision, recall, F1 score, and confusion matrix.
5. Implementation:
 - Integration into a diagnostic system for clinical use.
 - Designed for regions with limited access to laboratory tests.

3.2. DataSet Definition

The dataset used in this study was sourced from Kaggle and was previously analyzed by [21], who achieved 100% accuracy using a Random Forest. It consisted of 1421 instances, each representing a patient's hematological parameters, including hemoglobin (HGB), mean cell volume (MCV), mean cell hemoglobin (MCH), and mean cell hemoglobin concentration (MCHC). Table 2 presents the clinical variables considered.

Table 2. Description of the dataset.

Variable	Abbreviation
Gender	Gender
Hemoglobin	Hemoglobin
Mean cell volume	MCV
Mean cell hemoglobin	MCH
Mean cell hemoglobin concentration	MCHC
Target type of anemia	Target

The dataset consisted of hematological parameters such as hemoglobin (HGB) and mean cell volume (MCV), among others. Each value contributed differently to the diagnosis and classification of anemia, as follows:

- HGB: determined the presence of anemia based on threshold levels (e.g., <13.6 g/dL for men).
- MCV: differentiated anemia types (microcytic, normocytic, macrocytic) based on red blood cell size.
- MCH and MCHC: indicated the content and concentration of hemoglobin within cells, aiding in subclassification.

Once these metrics were understood, the algorithms could set thresholds for classification.

3.3. Training the Data

To train the data, a classification algorithm was developed by defining thresholds for the variables HGB, MCH, MCHC, and RDW. The algorithm initially discriminates according to sex and the presence of anemia. For men, the HGB threshold must be below 13.6 g/dL (136 g/L) to be considered anemic. In women, the HGB threshold must be below 12 g/dL (120 g/L). Once it was identified that the person has anemia, the type of anemia was labeled. According to the medical literature [22], there are three main types of anemia that are described below in Table 3.

Table 3. Types of anemia according to MVC.

Type of Anemia	Ranges or Threshold
Microcytic	<80
Normocytic	80–100
Macrocytic	>100

Figure 3 shows the flow chart of the rules for classifying the three types of anemia (microcytic, normocytic, macrocytic); then, depending on the type of anemia, a subclassification was made for each one.

The function of the anemia-type classification algorithm is briefly described below:

As can be seen in Algorithm 1, it consisted of two parts: the first part was related to the type of anemia according to the mean cell volume (MVC), that is, the algorithm identified whether the patient had one of the three types of anemia (microcytic, normocytic, and macrocytic). Thus, they were classified by class, including healthy patients, that is, those who did not have anemia. The second step was to classify anemia according to the cell type and other parameters that were defined as a threshold or range, such as hemoglobin (HGB) if it was microcytic, to determine if the patient had chronic anemia disease (ACD, moderate or mild). If the cell type was normocytic, the MCHC (mean cellular hemoglobin concentration) was examined, as well as the HGB again and the RDW (red blood cell distribution width) to assess whether the patient had thalassemia (mild or moderate) or, failing that, whether the patient had severe or moderate iron deficiency. Finally, the cells were assessed to determine whether they were macrocytic, and HGB was considered again. Depending on the threshold value, whether the patient had severe or moderate aplastic anemia was determined. Although our algorithm could classify variations in the different types of anemia given the cell type, we only classified the three types of anemia, as described in the first step.

Once the types of anemia were classified for training, we defined four classes, as shown in Table 4. For people without anemia, the class was 0; for patients with microcytic anemia 1, normocytic anemia 2, and macrocytic anemia 3.

Algorithm 1. Classification of anemia

```

1  function classify_anemia(MCV, MCHC, HGB, RDW)
2  # First Step Classification of anemia type according to MVC
3  if MCV < 80 Then
4      cell_type = "Microcytic"
5  elif MCV >= 80 and MCV <= 100 Then
6      cell_type = "Normocytic"
7  elif MCV > 100 Then
8      cell_type = "Macrocytic"
9  else
10     cell_type = "Unknown"
11 # Second Step Classification of anemia type according cell type and other parameters
12 if cell_type == "Microcytic" Then
13     if HGB < 10 Then
14         return "ACD_Severe"
15     else
16         return "ACD_Moderate"
17 elif cell_type == "Normocytic" Then
18     if MCHC < 32 Then
19         if HGB < 10 Then
20             if RDW < 14.16 Then
21                 return "Severe_thalassemia"
22             else do
23                 return "Severe_iron_deficiency anemia"
24         else do
25             if RDW < 14.16 Then
26                 return "Moderate_thalassemia"
27             else do
28                 return "Moderate_iron_deficiency_anemia"
29     else:
30         return "Normocytic anemia (Unknown)"
31
32 elif cell_type == "Macrocytic" Then
33     if HGB < 10:
34         return "Severe_aplastic_anemia"
35     else do
36         return "Moderate_aplastic_anemia"
37 else do
38     return "Unknown anemia type"

```

Table 4. Classification of types of anemia.

Type of Anemia	Class
No anemia	0
Microcytic	1
Normocytic	2
Macrocytic	3

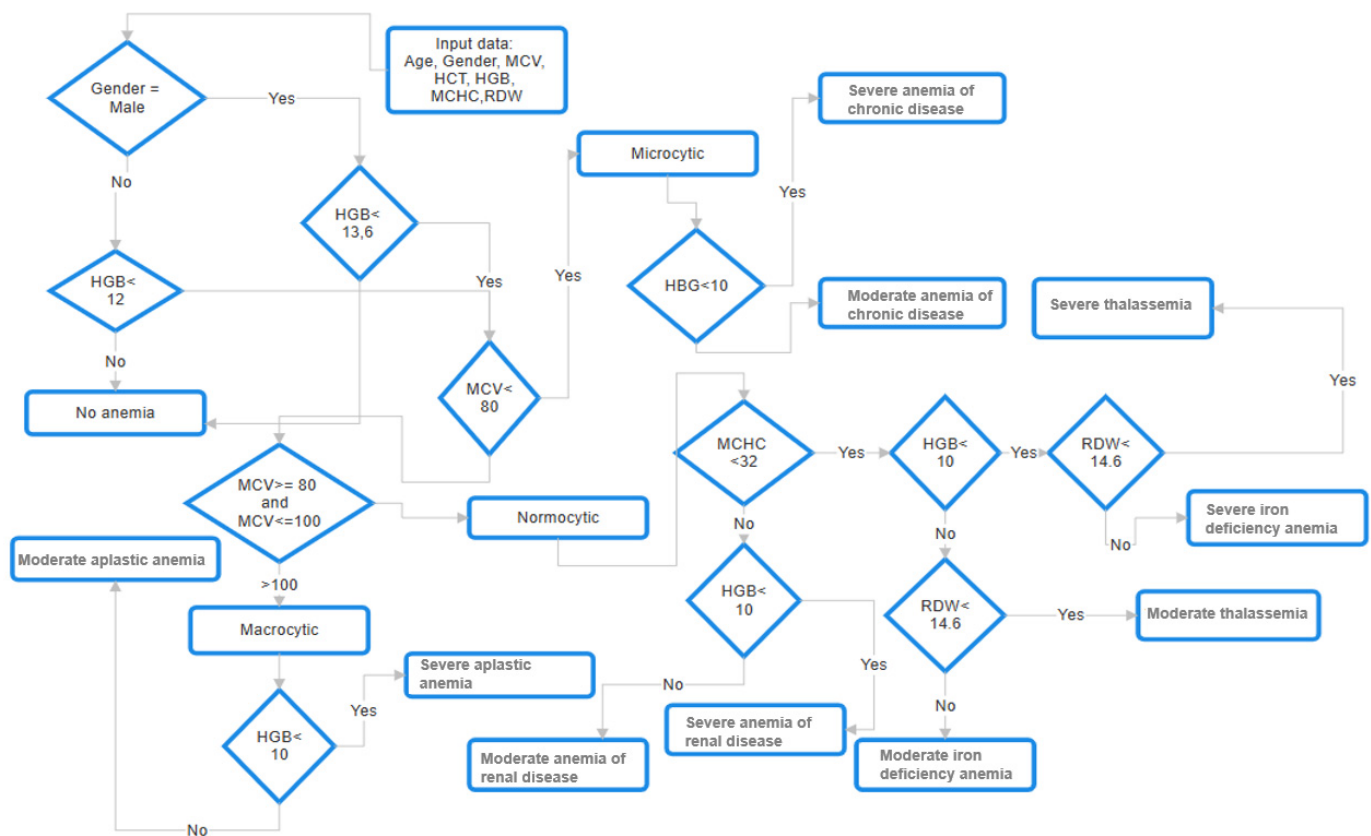


Figure 3. Classification of anemias.

3.4. Metrics

To select the best supervised learning model, we used a confusion matrix, which allowed us to evaluate the models' performance in making predictions. The confusion matrix is also known as the error matrix, which seeks to evaluate the number of correct and incorrect predictions with count values derived from each class [23]. To carry out this process, it were necessary to observe the four quadrants of the confusion matrix, the elements of which are described below:

- Positive (P): the observation were positive (example: it is anemia)
- Negative (N): the observation were not positive (example: it is not anemia)
- True positive (TP): the model correctly predicted the positive class
- True negative (TN): the model correctly predicted the negative class
- False positive (FP), which is also known as a type 1 error, that is, the model incorrectly predicted the positive class when in reality it was negative.
- False negative (FN) is also known as a type 2 error, that is, the model incorrectly predicted the negative class when in practice it was positive.

Next, we describe the metrics of the confusion matrix:

- Accuracy was equal to the proportion of predictions that the model classified correctly.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision is known as positive predictive value and corresponded to the proportion of relevant instances among the recovered instances.

$$Precision = \frac{TP}{TF + FP}$$

- Sensitivity corresponded to the hit rate or real positive rate, that is, it is the proportion of the number of instances that were recovered.

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

- Specificity is known as the true negative rate and measured the proportion of true negatives that were correctly identified; in other words, it was the opposite of sensitivity.

$$Specificity = \frac{TN}{TN + FP}$$

- F1 Score is known as a measure of the accuracy of a test. It could have a maximum accuracy of 1 and a minimum of 0.

$$\text{F1 Score} = \frac{2 * \text{Precision} * \text{Recall}}{2 * TP + FP + FN}$$

4. Results

To perform training, three supervised learning algorithms were selected (Linear Discriminant Analysis, Decision Trees, and Random Forests), as shown in Table 5.

Table 5. Comparison of supervised learning algorithms.

	Score 1	Score 2	Score 3	Score 4	Score 5	Mean	Std
Random Forest	1	1	1	0.99118943	1	0.99823789	0.00394021
Linear Discriminant	0.9122807	0.90350877	0.91189427	0.94273128	0.95594714	0.92527243	0.02273305
Decision Tree	1	1	1	0.98678414	1	0.99735683	0.00591031

The results shown in Table 5 indicate that the algorithm with the best performance for predicting the classification of anemia types is a Random Forest, with an accuracy of 99.82%, followed by the Decision Tree algorithm, with an accuracy of 99.73%. However, the algorithm that had the lowest performance for making predictions was Linear Discriminant Analysis, with an accuracy of 92.52%. Figure 4 shows the performance of each supervised learning algorithm used in this study.

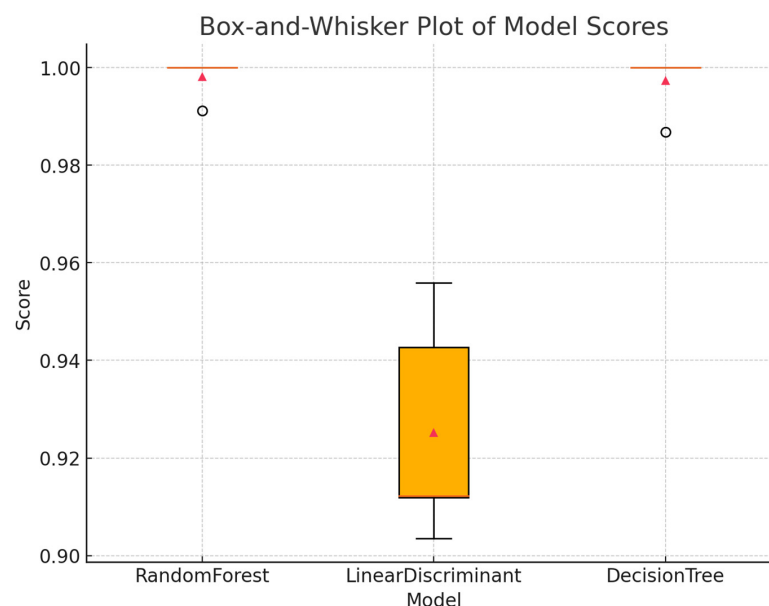


Figure 4. Performance of supervised learning algorithms.

As shown in Figure 4, the Random Forest algorithm has an accuracy of 99.82%, which indicates that its predictions are highly reliable and can be used by primary care medical services to generate recommendations for patients based on clinical data obtained from a basic blood count test, which is routinely performed on patients on a regular basis.

Analysis of the Metrics for Random Forest

As the results of the Random Forest algorithm are promising for predictions, given its 99.82% performance, the interpretation of the four classes is described below:

- Class 1 (non-anemia)

Correct predictions (diagonal): 162, with no false positives or false negatives, indicating that all instances of Class 1 were correctly identified, and there were no misclassifications.

- Class 2 (microcytic anemia)

Correct predictions: 37, with no false positives or negatives, indicating 100% classification for this class.

- Class 3 (normocytic anemia)

Correct predictions: 76, with no false positives or negatives, indicating 100% classification for this class.

- Class 4 (macrocytic anemia)

Correct predictions: 9, with no false positives or false negatives, also suggesting 100% classification for this class.

Figure 5 shows the confusion matrix for the four classes used for anemia prediction using the Random Forest algorithm.

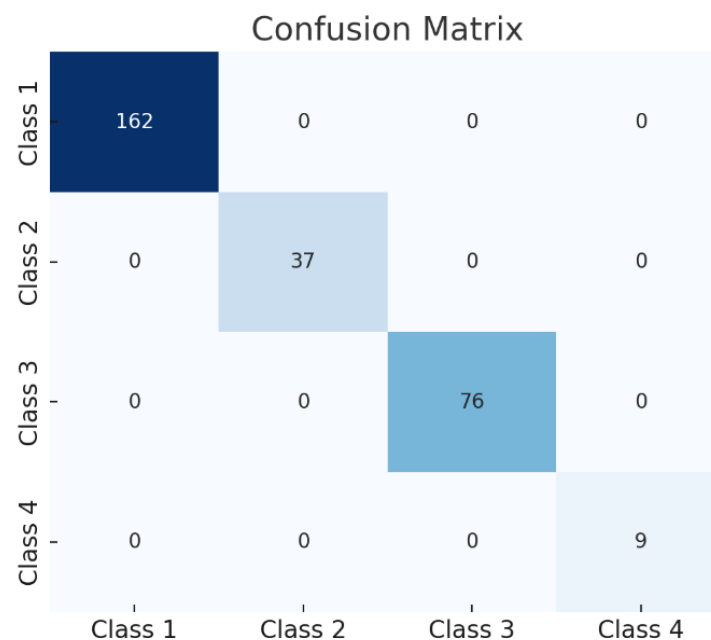


Figure 5. Confusion matrix.

Once the behavior of the confusion matrix has been observed, as shown in Table 6, the results obtained from the metrics are as follows:

- True positive (TP) = 162, which corresponds to correctly classified cases of class 1.
- True negatives (TN) = 37, which corresponds to correctly classified cases of class 2.
- False positives (FP) = 0, indicating that there were no instances of other classes incorrectly classified as Class 1.

- False negatives (FN) = 0, which means that all instances of Class 1 were correctly identified without omissions.

Table 6. Confusion matrix metrics result.

Metrics	Value
True Positives (TPs)	162
True Negatives (TNs)	37
False Positives (FPs)	0
False Negatives (FNs)	0

In Table 7, you can see the classification report according to the confusion matrix:

Table 7. Classification report.

Class	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	1.00	1.00	1.00
2	1.00	1.00	1.00
3	1.00	1.00	1.00

Overall, as can be seen in Table 7, the precision and accuracy are perfect because there were no false positives or false negatives in any of the classes, and the algorithm correctly classified all the instances in their respective classes. This indicates that this dataset has perfect performance. This model can be used by doctors in primary health units to make preliminary diagnostic judgments for the detection of anemia.

Within the dataset, there is an imbalance in the classes, especially in the dominant classes, which in this case are people who do not have anemia (Class 0), followed by patients who have Class 2 normocytic anemia, Class 1 microcytic anemia, and Class 3 macrocytic anemia.

According to this, the model can be affected by the following factors:

- Prediction bias: models can favor dominant classes.
- Reduced performance metrics, such as precision, recall, and F1 scores, for minority classes can be poor.
- Overfitting to common classes: models can have difficulty generalizing minority class patterns.

However, once the model that performed best was defined, which for our problem was the Random Forest algorithm, to solve this imbalance, we used a multiclass logarithmic loss function. Considering that this function is valid if and only if the number of classes is greater than or equal to 3 ($n_classes \geq 3$), for this method, the Gradient Boosting Classifier function was used.

According to the Gradient Boosting Classifier function, the predicted probability p determines the value of loss [24]. In this case, if the value of p was high (i.e., $p = 1$), the model was rewarded for making a correct prediction. Otherwise, if the value of $p < 1$, this indicated a low value of loss, that is, a bad prediction.

The result of the model score is 1.0, which means that the model is highly reliable for multiclass predictions. That is, the diagnosis of the presence or absence of anemia and the type of anemia can be predicted with a high level of confidence.

5. Discussion

This study highlights the importance of artificial intelligence for early and accurate diagnosis in low-resource regions. The main objective of our study was to implement

supervised learning algorithms to identify anemia types (microcytic, normocytic, and macrocytic). Linear Discriminant Analysis, Decision Trees, and Random Forest were used to train the model. Our subclassification was based on predefined clinical thresholds (e.g., $MCV < 80$ for microcytic anemia) derived from the medical literature. The Random Forest algorithm learns these thresholds algorithmically, which speeds up classification, but does not represent a clinical breakthrough. In future studies, these thresholds will be validated using dynamic clinical criteria.

While [21] achieved 100% accuracy in binary classification (presence/absence of anemia) using the same dataset, our study extended this by subclassifying anemia into microcytic, normocytic, and macrocytic types with 99.82% accuracy, demonstrating the feasibility of multiclass classification.

6. Conclusions

Supervised learning algorithms, especially Random Forest, proved to be effective for classifying types of anemia using blood count data, achieving very high prediction accuracy. Implementing this system in areas with limited medical resources could significantly improve the diagnosis of anemia and optimize treatment. The results highlight the great potential of implementing artificial intelligence in medical diagnosis, highlighting the accuracy and reliability of these models in clinical practice. It is important to note that this tool must be used only by medical personnel to accelerate the diagnosis of patients.

Despite the results obtained, the external validity of the model requires evaluation in diverse contexts and populations.

Future Research Directions Include the Following

Data expansion: it is suggested to use additional and more varied datasets to verify the accuracy of the model in different demographic groups and regions, which will allow the generalization of the results and improve the robustness of the model.

Incorporating explainable artificial intelligence: implementing explainable artificial intelligence techniques would help medical professionals to better understand the model's predictions, increasing its reliability and transparency.

Optimization of mobile devices: because many rural areas have limited access to complete laboratories, it would be interesting to adapt this system to mobile devices to facilitate their application in low-resource regions.

Real-time assessment: implementing a real-time assessment system would allow this model to be integrated into clinical settings where rapid diagnoses are required, such as in rural clinics or during public health campaigns.

These future directions could contribute to improving medical care in areas where access to diagnostic services is limited, as well as optimizing the use of resources in the health sector.

Author Contributions: All authors (J.G.G., C.P.U., D.S.Á., V.H.R. and G.R.-G.) declare that they have worked collaboratively in its development, from the formulation phase to the experimentation and writing of the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the University of Córdoba for financing this research project according to the internal call with project code FI-03-23.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: We also thank the SOCRATES research group of the Systems Engineering and Telecommunications program for supporting the development of this project.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Garcia-Casal, M.N.; Dary, O.; Jefferds, M.E.; Pasricha, S. Diagnosing anemia: Challenges selecting methods, addressing underlying causes, and implementing actions at the public health level. *Ann. N. Y. Acad. Sci.* **2023**, *1524*, 37–50. [CrossRef] [PubMed]
2. Del Castillo, L.; Cardona-Castro, N.; Whelan, D.R.; Builes, J.P.; Serrano-Coll, H.; Arboleda, M.; Leon, J.S. Prevalence and risk factors of anemia in the mother–child population from a region of the Colombian Caribbean. *BMC Public Health* **2023**, *23*, 1533. [CrossRef] [PubMed]
3. Pobreza Monetaria en Colombia Durante 2022: Efecto Compensado y sus Implicaciones en un Panorama Inflacionario | Programa De Las Naciones Unidas Para El Desarrollo. Available online: <https://www.undp.org/es/colombia/discursos/pobreza-monetaria-colombia-efecto-compensado-implicaciones-panorama-inflacionario> (accessed on 20 July 2024).
4. Cappellini, M.D.; Motta, I. Anemia in Clinical Practice—Definition and Classification: Does Hemoglobin Change with Aging. *Semin. Hematol.* **2015**, *52*, 261–269. [CrossRef]
5. Pineda, J.M. Predictive models in health based on machine learning. *Rev. Medica Clin. Las Condes* **2022**, *33*, 583–590. [CrossRef]
6. Anaemia in Women and Children. Available online: https://www.who.int/data/gho/data/themes/topics/anaemia_in_women_and_children (accessed on 20 July 2024).
7. Calle-Pesántez, S.E.; Pallo-Chiguano, J.M. Inteligencia Artificial en la comunicación científica. *Espejo Monogr. Común. Soc.* **2024**, *23*, 59–81. [CrossRef]
8. Prajapati, J.; Uduthalapally, V.; Das, D.; Mahapatra, R.; Wasnik, P.N. XAIA: An Explainable AI Approach for Classification and Analysis of Blood Anemia. In Proceedings of the OCIT 2023-21st International Conference on Information Technology, Raipur, India, 13–15 December 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023; pp. 88–93. [CrossRef]
9. Rahman, M.; Mojumdar, M.U.; Shifa, H.A.; Chakraborty, N.R.; Stenin, N.P.; Hasan, A. Anemia Disease Prediction using Machine Learning Techniques and Performance Analysis. In Proceedings of the 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 28 February–1 March 2024.
10. Kasthuri, E.; Subbulakshmi, S.; Sreedharan, R. Insightful Clinical Assistance for Anemia Prediction with Data Analysis and Explainable AI. In *Procedia Computer Science*; Elsevier B.V.: Amsterdam, The Netherlands, 2024; pp. 45–55. [CrossRef]
11. Trigka, M.; Dritsas, E.; Mylonas, P. A Multi-Class Classification Approach for Anemia Level Prediction with Machine Learning Models. In Proceedings of the 2023 8th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference, Piraeus, Greece, 10–12 November 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023. [CrossRef]
12. Tressa, N.; Asha, V.; Singh, S.K.; Subhash, J. Alpha Thalassemia Classifier Using Machine Learning Techniques Based on Genetic Mutations. In Proceedings of the 2023 3rd International Conference on Ubiquitous Computing and Intelligent Information Systems, ICUIS 2023, Gobichettipalayam, India, 1–2 September 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023; pp. 118–122. [CrossRef]
13. Ponnusamy, U.; Darshan, D.; Sampathila, N. Approaching Explainable Artificial Intelligence Methods in the Diagnosis of Iron Deficiency Anemia Using Blood Parameters. In Proceedings of the 2023 International Conference on Recent Advances in Information Technology for Sustainable Development, ICRAIS 2023-Proceedings, Dhaka, Bangladesh, 21–23 September 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023; pp. 201–206. [CrossRef]
14. Abdulhay, E.W.; Allow, A.G.; Al-Jalouly, M.E. Detection of Sick Cell, Megaloblastic Anemia, Thalassemia and Malaria through Convolutional Neural Network. In Proceedings of the 2021 Global Congress on Electrical Engineering, GC-ElecEng 2021, Valencia, Spain, 10–12 December 2021; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2021; pp. 21–25. [CrossRef]
15. Shah, A.R.; Javed, I.; Shams, U.A.; Naveed, M.A.; Mehmood, M.Q. Disease estimation using robust AI methods. In Proceedings of the 2023 4th International Conference on Computing, Mathematics and Engineering Technologies: Sustainable Technologies for Socio-Economic Development, iCoMET 2023, Sukkur, Pakistan, 17–18 March 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023. [CrossRef]
16. Mugdha, A.G.; Pinki, F.T.; Talukdhar, S.K. Hemoglobin Estimation and Anemia Severity Prediction Using Machine Learning Algorithms. In Proceedings of the 2023 5th International Conference on Sustainable Technologies for Industry 5.0, STI 2023, Dhaka, Bangladesh, 9–10 December 2023; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023. [CrossRef]

17. Sasikala, C.; Ashwin, R.M.; Dharanesh, D.M.; Dhanabalan, M. Curability Prediction Model for Anemia Using Machine Learning. In Proceedings of the 8th International Conference on Smart Structures and Systems, ICSSS 2022, Chennai, India, 21–22 April 2022; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2022. [\[CrossRef\]](#)
18. Sadiq, S.; Khalid, M.U.; Din, M.Z.; Ullah, S.; Aslam, W.; Mehmood, A.; Choi, G.S.; On, B.-W. Classification of β -Thalassemia Carriers from Red Blood Cell Indices Using Ensemble Classifier. *IEEE Access* **2021**, *9*, 45528–45538. [\[CrossRef\]](#)
19. Yıldız, T.K.; Yurtay, N.; Öneç, B. Classifying anemia types using artificial learning methods. *Eng. Sci. Technol. Int. J.* **2021**, *24*, 50–70. [\[CrossRef\]](#)
20. Shahzad, M.; Umar, A.I.; Shirazi, S.H.; Shaikh, I.A. Semantic Segmentation of Anaemic RBCs Using Multilevel Deep Convolutional Encoder-Decoder Network. *IEEE Access* **2021**, *9*, 161326–161341. [\[CrossRef\]](#)
21. Sabatini, A. Anemia EDA: Max Accuracy & Features Selection. Kaggle. 2022. Available online: <https://www.kaggle.com/code/antoniosabatini/anemia-eda-max-accuracy-features-selection> (accessed on 10 February 2024).
22. Green, R. Anemias beyond B12 and iron deficiency: The buzz about other B's, elementary, and nonelementary problems. *Hematol. Am. Soc. Hematol. Educ. Program* **2012**, *2012*, 492–498. [\[CrossRef\]](#)
23. Vujović, Ž. Classification model evaluation metrics. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 599–606. [\[CrossRef\]](#)
24. Sun, J.J.; Li, J.; Fujita, H. Multi-class imbalanced enterprise credit evaluation based on asymmetric bagging combined with light gradient boosting machine. *Appl. Soft Comput.* **2022**, *130*, 109637. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.