GAMA CUTE: Development of a Web-based for Gadjah Mada Caring University for Thalassemia Exit Prediction Tool by Applying Machine Learning

Dimas Chaerul Ekty Saputra[1,2,3], Afiahayati[4], Tri Ratnaningsih[5]

[1] Department of Biomedical Engineering, Graduate School, Gadjah Mada University, Yogyakarta 55284, Indonesia
[2] Department of Computer Science, College of Computing, Khon Kaen University, Khon Kaen 40002, Thailand
[3] Department of Informatics, School of Computing, Telkom University Surabaya, Surabaya 60231, Indonesia
[4] Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Yogyakarta 55281, Indonesia
[5] Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Public Health and Nursing, Gadjah Mada University, Yogyakarta 55281, Indonesia

ARTICLE INFO

ABSTRACT

Blood disorders occur in one or several parts of the blood that affect the nature and function, and blood disorders can be acute or chronic. Blood disease consists of several types, such as anemia. Anemia is the most common hematologic disorder associated with a decrease in the number of red blood cells or hemoglobin, causing a decrease in the ability of the blood to carry oxygen throughout the body. Patients with anemia in Indonesia have increased for the age of 15-24 years. This study aimed to conduct a screening test for anemia using machine learning. It is expected to know the process of knowing the type of anemia suffered. The machine learning technique used to identify the cause of anemia is divided into four classes, namely Beta Thalassemia Trait, Iron Deficiency Anemia, Hemoglobin E, and Combination (Beta Thalassemia Trait and Iron Deficiency Anemia or Hemoglobin E and Iron Deficiency Anemia). This study would apply the K-Nearest Neighbor (KNN) and Random Forest (RF) methods to build a model on the data collected. The evaluation results using a confusion matrix in the form of accuracy, precision, recall, and f1-score against the KNN and RF methods are 79.36%, 59.40%, 62.80%, and 62.80%. In comparison, the RF is 87.30%, 90.89%, 78.40%, and 81.00%. From the results of comparing the two methods, the Graphic User Interface (GUI) implementation using python applies the RF method. The classifier that gets the highest value among all these parameters is called the best machine learning algorithm to perform screening tests for anemia.

Corresponding Author:

Afiahayati, Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Yogyakarta 55281, Indonesia
Email: afia@ugm.ac.id

1. INTRODUCTION

Blood cells in the human body influence the health of the human body [1]. Human blood cells are generally divided into three types, namely red blood cells, white blood cells and platelets [2]. The three types certainly have their respective functions and duties in the human body's blood circulation system [3]. A blood disorder is a disorder that occurs in one or several parts of the blood so that it affects the amount and function, and blood disorders can be acute or chronic [3], [4]. There are several types of blood disorders, such as anemia. Anemia is a disease associated with a decrease in the total number of red blood cells or haemoglobin in the blood or a decrease in the ability of the blood to carry oxygen throughout the body [5]. When anemia comes

on slowly, the symptoms are often vague, and you will feel weak, tired, lethargic, limp, and tired [6]. This condition can occur because the blood cells in the body do not get enough oxygen, or the blood cells in the body experience a lack of oxygen supply. Anemia can occur temporarily or in the long term, ranging from mild to severe [7].

Based on research data from the Ministry of Health of the Republic of Indonesia in 2007, 2013, and 2018, there was an increasing trend of anemia prevalence in adolescents [6]–[8]. In 2018, 32% of adolescents in Indonesia experienced anemia. This indicates that approximately 7.5 million Indonesian adolescents are at risk for experiencing barriers to growth and development and cognitive abilities and are susceptible to infection. Artificial intelligence, known as Artificial Intelligence (AI), is a simulation of human intelligence processes by machines, especially computer systems [11]. AI-specific applications include expert systems [12], natural language processing [13], speech recognition [14], and signal processing [15]. As the hype around AI accelerates, vendors have been scrambling to promote how their products and services use AI [16]. They often call AI just one component of AI, such as machine learning [17]. AI requires a base of specialized hardware and software to write and train machine learning algorithms [18]. No programming language is synonymous with AI, but several are popular, including Python, R, and Java [16]. In general, AI systems work by absorbing large amounts of labelled training data, analysing the data for correlations and patterns, and using these patterns to make predictions about future statuses [16]. In this way, chatbots given examples of text chats can learn to generate real exchanges with people, or image recognition tools can learn to identify and describe objects in images by reviewing millions of examples [19].

In the rapid development of artificial intelligence (AI) technology today [20]. AI programming focuses on three cognitive skills: learning [15], reasoning [16] and self-correction [17]. This aspect of AI programming focuses on acquiring data and creating rules about how to turn data into actionable information [21]. Rules, called algorithms, provide computing devices with step-by-step instructions on how to complete a particular task [18]. Not many people know that AI consists of several branches, one of which is Machine Learning (ML) [22]. AI in its application is broadly divided into seven branches, namely Machine Learning, Natural Language Processing, Expert Systems, Vision, Speech, Planning, and Robotics [11]. Artificial intelligence is intended to narrow the scope when developing or learning AI because AI has a very broad scope [17], [18].

AI in human life is expected to accelerate identifying a problem [25]. In research conducted by Laengsri *et al.* [7] using five methods, namely K-Nearest Neighbor (KNN), Decision Tree (DT), Random Forest (RF), Artificial Neural Network (ANN), and Support Vector Machine (SVM) using 6.935 data has resulted in good accuracy, by classifying the data into two classes, namely BTT and IDA, the accuracy is obtained successively equal to 92.36%, 98.03%, 94.17%, 94.11%, and 95.05%. This study uses seven features generated by Haematology Analyzer (HA) to screen patients with anemia. There is difficulty in screening to distinguish the type of anemia using only the patient's blood sample. This screening process will produce an outcome in the form of the type of anemia suffered by the patient, so that doctors do not experience difficulties in carrying out further diagnostic supporting examinations. Thus, the results will be clearer and more accurate for the anemia suffered by the patient [7].

The use of AI in this study was to accelerate the process of identifying the type of anemia suffered by the patient. By using the Random Forest and K-Nearest Neighbor methods to classify anemia into four classes, namely Beta Thalassemia Trait (BTT), Haemoglobin E (Hb E), Iron Deficiency Anemia (IDA), and Combination (BTT and IDA or Hb E and IDA). This research will produce application outputs that can be used at level 1 Health Facilities. Hopefully, the system can help health workers identify the type of anemia in patients.

2. MATERIALS AND METHODS
2.1. Research Description
This study uses the Random Forest and K-Nearest Neighbor methods to classify anemia as a screening tool. The results of this study are the level of accuracy, precision, recall, f1-score, and graphic user interface (GUI) that can be used by health workers to assist the diagnosis process. This study uses a letter of ethics from the Medical and Health Research Ethics Committee (MHREC) Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada – Dr. Sardjito General Hospital with reference number KE/FK/1255/EC/2021.

2.2. Data Acquisition
The data used in this study used complete blood test results at the Laboratory of the Department of Clinical Pathology and the Laboratory of Faculty of Medicine, Public Health and Nursing Universitas Gadjah Mada.

The data used are 190 data with 127 training data and 63 test data. The data used are haematological parameters from patients who have been diagnosed with BTT, IDA, Haemoglobin E, and Combination. This study uses seven features, namely Red Blood Count (RBC), Haemoglobin (Hb), Haematocrit (HCT), Mean Corpuscular Volume (MCV), Mean Corpuscular Haemoglobin (MCH), Mean Corpuscular Haemoglobin Concentration (MCHC), Red-Cell Distribution Width (RDW). The selection of these seven features is because these features are routinely used in all HA in clinical pathology laboratories. So that by choosing these seven features, it will be more applicable to be used in the anemia screening process. In this study, the number of datasets used was 190 data sets by dividing them into four different classes, namely BTT, Hb E, IDA, and Combination (BTT and IDA or Hb E and IDA). The combination class is a combination of BTT and IDA or Hb E and IDA. The combination class does not consist of BTT and Hb E because the two classes are both in the same clump or hemoglobinopathy group. Of the 190 data used, they were divided into 127 training data and 63 test data. Based on these 190 data, there were 24 data including BTT class, 41 data including Hb E, 104 data including IDA, and 21 data including Combination. Fig. 1 relates to the number of datasets used in this study.
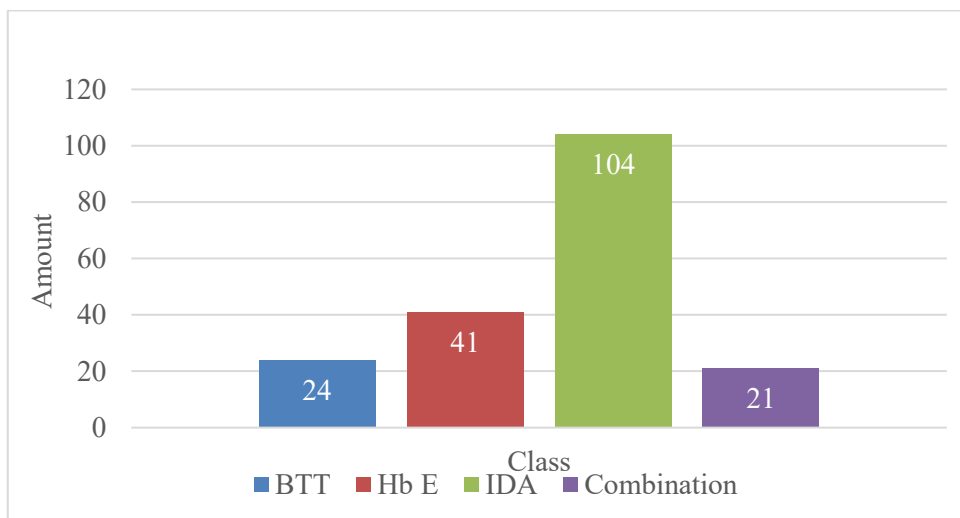

Fig. 1. Research Dataset

2.3. Research Flowchart

The course of the research conducted in this study can be seen in Fig. 2. The first stage that will be carried out is a library study. The second stage was followed by data collection at the Laboratory of the Department of Clinical Pathology and Medicine and the Laboratory of Faculty of Medicine, Public Health and Nursing Universitas Gadjah Mada. In the third stage, it was pre-processing to normalize data with StandardScaler and LabelEncoder to change labels from characters to numeric. The classification uses several algorithms in the fourth stage, namely Random Forest and k-Nearest Neighbor. After getting the best accuracy from one of the methods used in this research, the next step will be a GUI implementation using Python-based Flask to create a user interface. Making GUI is inseparable from testing to ensure that the system is ready for use by users, the test is divided into two, namely Black-Box Testing for testing the system in appearance and function then SUS testing to assess the usability of the system from the user's side. Furthermore, in the last process, the evaluation using the Confusion Matrix to get the results of the algorithm's accuracy was used in this study.
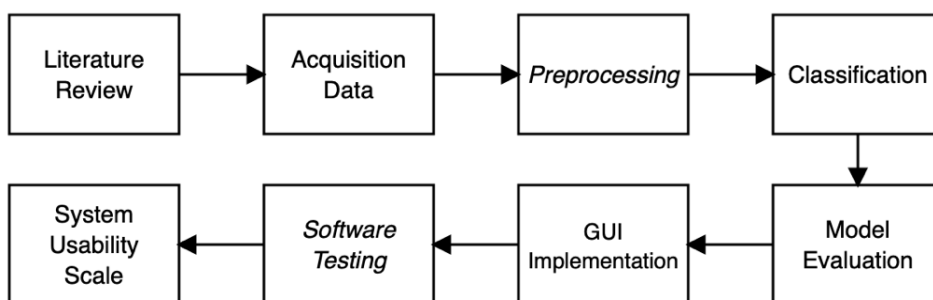

Fig. 2. Research Flowchart

2.4. K-Nearest Neighbor Flowchart

K-Nearest Neighbor (KNN) is a classification method for a data set based on learning the previously classified data. The selection of the KNN method is because KNN is the most straightforward method for the classification process, the development of this system is expected to be used for the classification process properly [26]. The KNN process uses Euclidean Distance calculations to calculate the distance according to the number of K (neighbors) entered by the user. Based on Fig. 3, this KNN method is included in supervised learning. The new data results will be classified based on most of the distance proximity of the categories in the KNN.
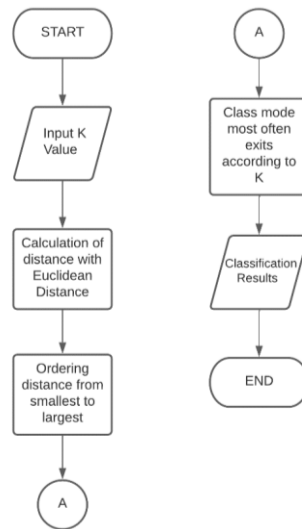
Fig. 3. Flowchart KNN

2.5. Random Forest Flowchart

The Random Forest Algorithm is an algorithm whose classification is supervised. This method creates a forest with a predefined number of trees. This method creates a tree that has multiple decisions. Because all node selection calculations will be the same for the same dataset. The choice of the RF method is because RF will create a forest randomly to get predictions from each tree and choose the best solution from each branch. In Fig. 4, the following Random Forest Flowchart.
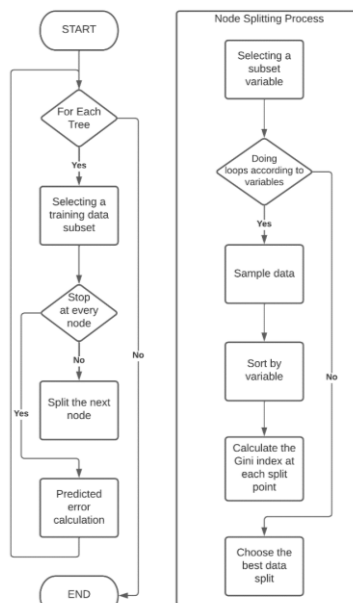
Fig. 4. Flowchart Random Forest

The following is an explanation of Fig. 4:

1. Suppose N is the number of training data and M is the number of features. M is the number of features used to determine the sorting decision at a node, where m is smaller than M;
2. Select a random training data set by selecting N times with replacement from all N data, and this is also called a bootstrap sample, only $\frac{2}{3}$ of the original data is used. The rest of the dataset is used for Out-of-Bag purposeful testing, which estimates the OOB error for classification;
3. For each tree node, the node is split with the best feature among the randomly selected features using the Gain Impurity separator criterion, then calculates the best split based on m features in the training data.

Making a decision tree is maximal if there is no root pruning. For prediction, the new sample is inserted into the tree, and then the labelled/train sample class is at the node where the new sample is located. This procedure is repeated for all trees in the forest collection. The predicted results of all trees are applied by majority vote to be taken as prediction results.

2.6. Classification Process

When carrying out a classification process, you will usually deal with an abnormal distribution of data. The purpose of data normalization is to eliminate and reduce data redundancy and ensure data dependencies [27]. In the first pre-processing process, the use of LabelEncoder is used to change the class in the form of characters to numeric. The use of LabelEncoder here is to simplify the calculation process during classification [28]. The next pre-processing is using StandardScaler. StandardScaler standardizes features by subtracting the mean and then scaling to the unit variance. Unit variance means dividing all values by the standard deviation. StandardScaler returns a distribution with a standard deviation equal to 1. The variance equals 1 too, because variance = standard deviation squared (1 squared = 1) [29]. The classification process in Fig. 5 uses data from a complete blood check. Complete blood data obtained after laboratory tests using HA is then processed in the system for classification/screening. After the classification process is complete, it displays the accuracy, precision, sensitivity, and f1-score of the classification results that have been obtained.
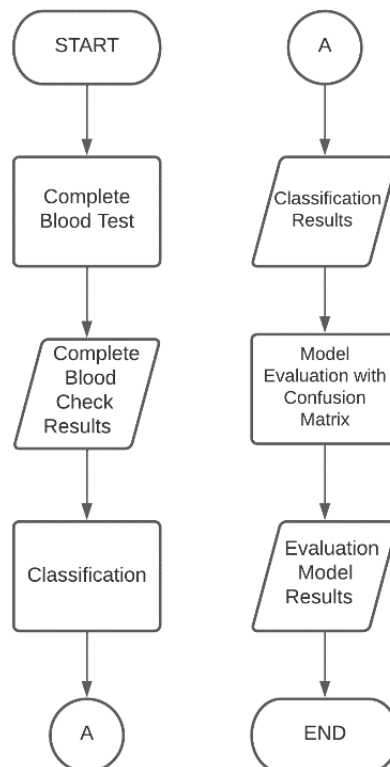

Fig. 5. Classification Flowchart

2.7. Evaluation Model

At the testing stage, the researcher divided the data into training and test data. Testing of training data and test data was carried out using the Confusion Matrix model. Confusion Matrix is a method usually used to

calculate the accuracy of the methods used in the classification process. The value-sharing process is described in Table 1 Confusion Matrix.

Table 1. Confusion Matrix

Prediction Class		Real Class	
		True	**False**
	True	True Positive	False Positive
	False	False Negative	True Negative

From Table 1, it can be explained that:
1. TP, namely True Positive, is the number of positive data classified correctly by the system.
2. TN, True Negative, is the number of negative data correctly classified by the system.
3. FN, namely False Negative, is the number of negative data classified incorrectly by the system.
4. FP, namely False Positive, is the number of positive data classified as wrong by the system.

Based on the values above, the formulas for calculating the value can be explained as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{9}$$

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{10}$$

3. RESULTS AND DISCUSSION

In this section, the results of the research are explained and, at the same time is given a comprehensive discussion. Results can be presented in figures, graphs, tables, and others that make the reader understand easily. The discussion can be made in several sub-chapters. It is strongly suggested that comparison with results from other published articles are provided to give more context and to strengthen the claim of novelty.

3.1. Performance Evaluation Methods

This research uses 2 (two) different methods, namely KNN and Random Forest. In KNN using Euclidean Distance and in Random Forest using entropy. The difference in the results of the evaluation of the two methods is very different, the results of the performance of the method can be seen in Table 2 and Table 3.

The determination of the number of n_estimators in the random forest method is chosen based on Fig. 6 of the n_estimators graph based on new predicted data and previously trained data. In the precision section the highest score is 100% in the Combination class, in the recall section the highest score is 100% in the BTT class and the best F1-Score in the BTT class with a score of 92%. The optimal value of n_estimators used in this system is n_estimators = 120, with an accuracy rate of 87.3%.

Table 2. Performance Evaluation Random Forest

	Precision	Recall	F1-Score
0 (BTT)	0.86	1.00	0.92
1 (Hb E)	0.93	0.76	0.84
2 (IDA)	0.85	0.97	0.91
3 (Combination)	1.00	0.40	0.57
Accuracy			0.87
Macro avg	0.91	0.78	0.81
Weighted avg	0.88	0.87	0.86
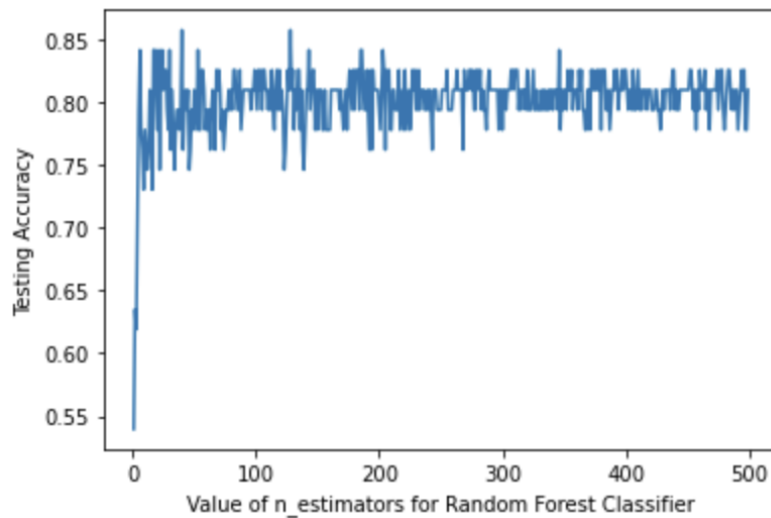
Fig. 6. Determination Graph of n_estimators Random Forest

Table 3. Performance Evaluation KNN

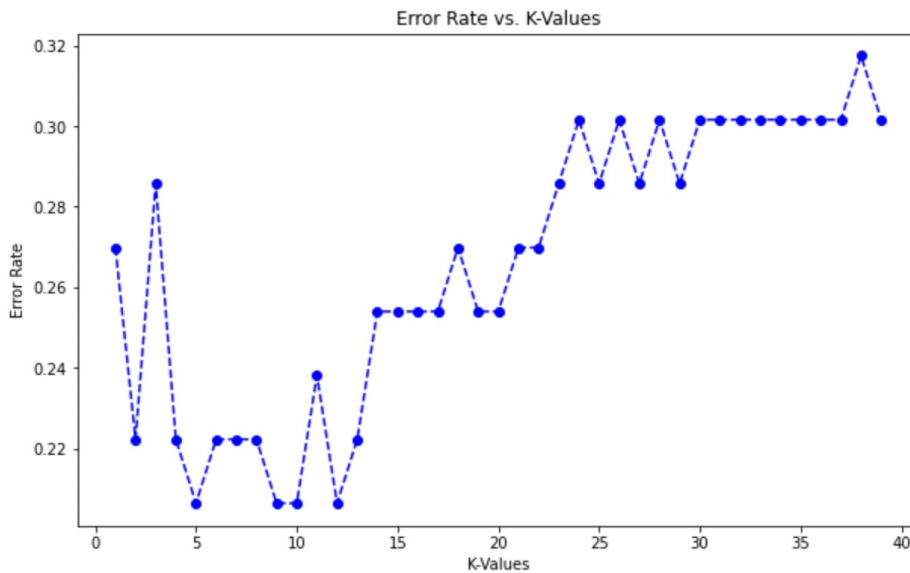|	Precision	Recall	F1-Score
0 (BTT)	0.83	0.83	0.83
1 (Hb E)	0.72	0.76	0.74
2 (IDA)	0.82	0.91	0.86
3 (Combination)	0.00	0.00	0.00
Accuracy			0.79
Macro avg	0.59	0.63	0.61
Weighted avg	0.73	0.79	0.76

Fig. 7. Determination of K-Values KNN

As known in the random forest method, the K in the KNN is determined by the user according to the best error rate. The number of K in the KNN is determined based on the Fig. 7 error rate graph based on new data that is predicted to be not the same as the data that has been trained previously. The optimal K value used in this system is K = 5, 9, 10, and 12, with an error rate of less than 22% and an accuracy of 79.36%. Based on table 3 regarding the results of the method performance for each class, the precision section has the highest score of 83% in the BTT class. In the recall section, the highest score is 91% in the IDA class and the best F1-Score in the IDA class with 86%.

3.2. System Requirements Analysis

The system requirements that have been built have been described in the system requirements analysis stage, of course, so that the system can be used optimally in identifying the type of anemia suffered by the patient. System requirements include:

3.2.1. Input Needs Analysis

    3.2.1.1. The system is capable of inputting data in *.xlsx format.

    3.2.1.2. The system can perform a single data input without files.

3.2.2. Process Needs Analysis

    3.2.2.1. The system can perform RF calculations for data classification.

    3.2.2.2. The system can diagnose data into predetermined categories.

3.2.3. Output Need Analysis

The system can display the results of the classification diagnosis into predetermined categories.

3.3. Implementation of Graphic User Interface

This research has produced a product in a Flask Python-based User Interface (UI). From the results of comparing the evaluation models of the two methods, it was decided to use Random Forest to classify patients with anemia into four classes. Fig. 8-Fig. 13 related to the implementation of the UI in this GAMA CUTE application.
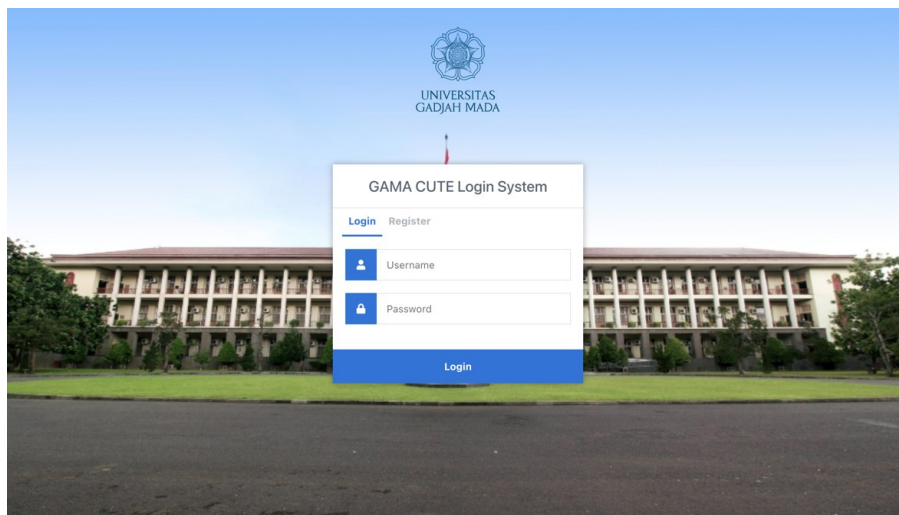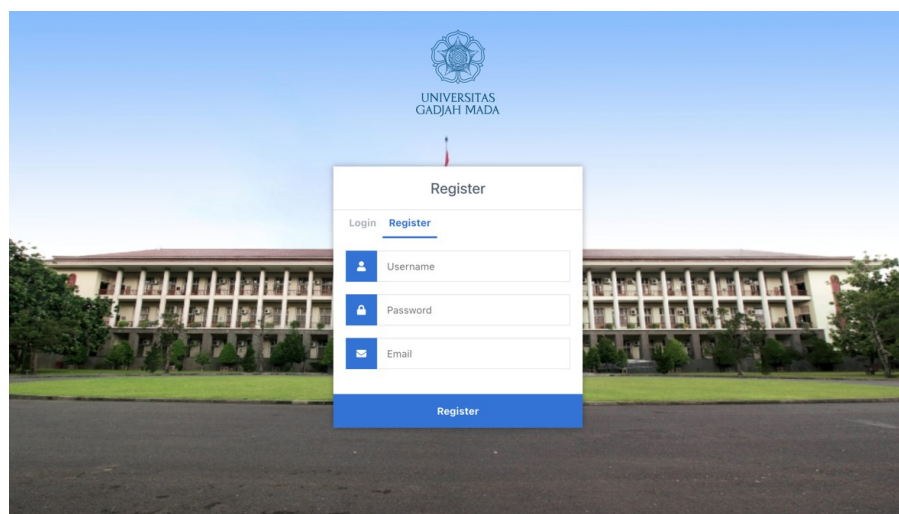

Fig. 8. UI Login System

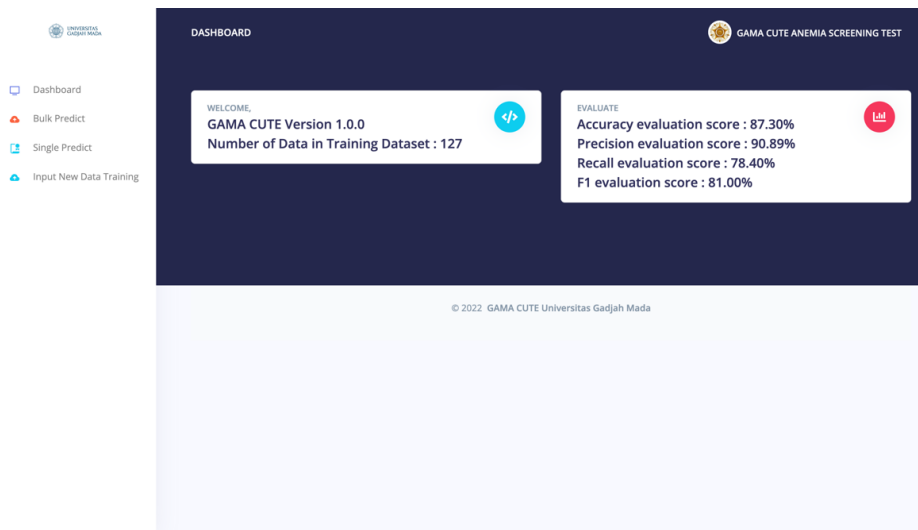**Fig. 9**. UI Register for New Account
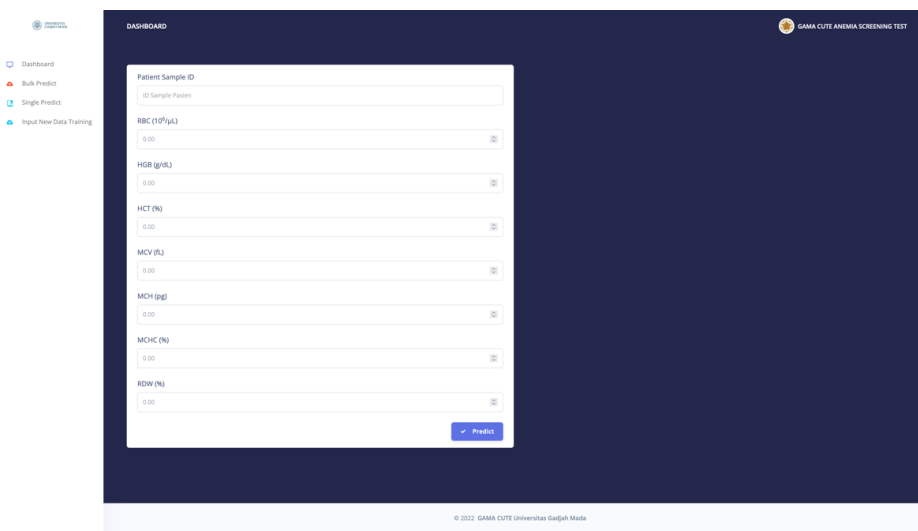
Fig. 10. UI Dashboard
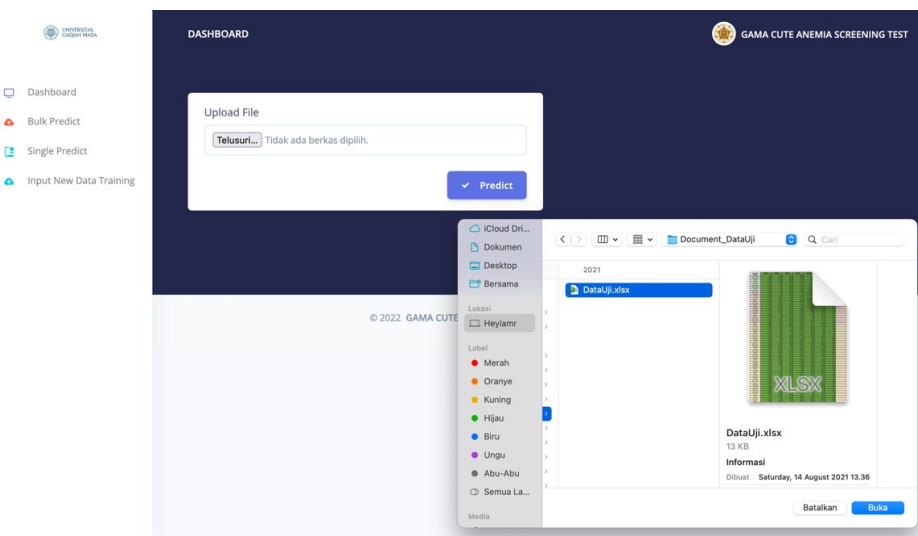

Fig. 11. UI Single Predict

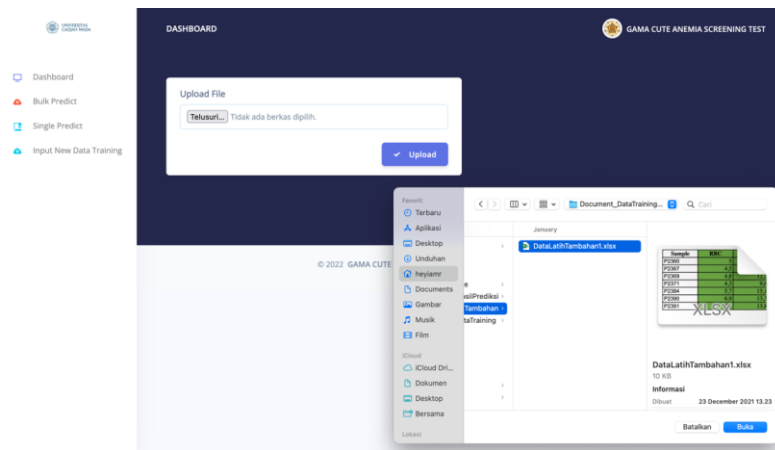**Fig. 12**. UI Bulk Predict using File Excel (*.xlsx)

Fig. 13. UI Input New Training Data using File Excel (*.xlsx)

3.4. Black-Box Testing

Black-box is a software testing method of the application instead of the work structure. This test aims to find some errors such as missing functions and errors in the interface. Testing is carried out to ensure an application runs appropriately according to the proper process and produces the desired output. This black-box test was carried out on the main menu program of this web-based GAMA CUTE Version 1.0.0 application. Table 4 is the result of testing using Black-Box.

Table 4. Black-Box Testing

Input	Process	Output	Test Results
Click menu → Dashboard	Displays the evaluation results of the method used and the amount of training data used in the system.	The name of the application 'GAMA CUTE Version 1.0.0' appears, the amount of training data and the evaluation results are: Accuracy, Precision, Recall, F1-Score.	Succeed
Click menu → Bulk Predict	Displays columns for data input with buttons for processing data.	Shows columns for uploading excel files.	Succeed
Click menu → Single Predict	Displays columns of each feature/attribute used to predict single data (without files).	Displays columns of seven attributes used: RBC, HGB, HCT, MCV, MCH, MCHC, and RDW.	Succeed
Click menu → Input New Data Training	Displays columns for data input with buttons for processing data.	Shows columns for uploading excel files.	Succeed
Button → Predict on menu Bulk Predict	Displays a button to carry out the data prediction process.	Processing data entered in the Upload File column and displays prediction results on the data in the form of a complete table with the diagnosis.	Succeed
Button → Upload on menu Input New Data Training	Displays a button to add training data.	Process the data that has been inputted in the File Upload column and display the prediction results of the data in the form of a complete table with the diagnosis.	Succeed
Button → Predict on menu Single Predict	Displays a button to perform the data prediction process.	Process the data entered in each attribute column and display the prediction results on the data in the form of a pop-up complete with the diagnosis.	Succeed

After testing with the black-box in Table 4, GAMA CUTE is an application to screening for anemia. In this application, user can screen for the type of anemia suffered by the patient. The results of this test are to make it easier for GAMA CUTE users to improve the user experience.

3.5. System Usability Scale

The SUS evaluation consisted of 10 questions using a Likert scale of 1 to 5. Questions were asked to 70 respondents. Respondents filling out the SUS evaluation consisted of general practitioners, clinical pathology

specialists, internal medicine specialists, pediatricians, software engineers, informatics engineering students, computer science students, electrical engineering students, biomedical engineering students, and Medical Laboratory Technologists. Fig. 15 shows a graph of the Average SUS Score. Based on the results of the data recap of respondent's SUS scores in Fig. 15, the average value of SUS is 76.64. The results of the average SUS value taken and matched based on the SUS score are shown in Fig. 14. Based on the SUS score matching, the SUS average score is good because the standard SUS average is 68, in Fig. 16 with details of 70 respondents: one respondent assessed that the application was unacceptable, 21 respondents rated it marginal, and 48 respondents stated the application was acceptable. On the Grade Scale, the rating for the application category gets a C value and the Adjective Ratings are in the 'Good' category because SUS has an average SUS score of 76.64.

List of Questions:
1.   I think I will use this system again;
2.   I find this system complicated to use;
3.   I find this system easy to use;
4.   I need help from other people or technicians in using this system;
5.   I feel the features of this system are working properly;
6.   I feel many things are inconsistent (incompatible in this system);
7.   I feel other people will understand how to use this system quickly;
8.   I find this system confusing;
9.   I feel there are no obstacles in using this system;
10.   I need to get used to it before using this system.
Answer List:
1.   Strongly Disagree.
2.   Disagree.
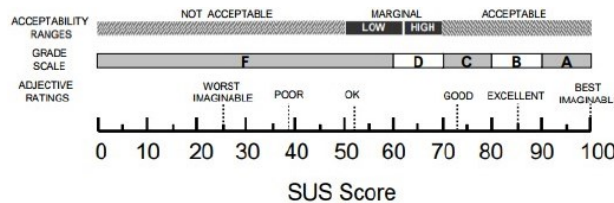3.   Doubtful.
4.   Agree.
5.   Strongly Agree.

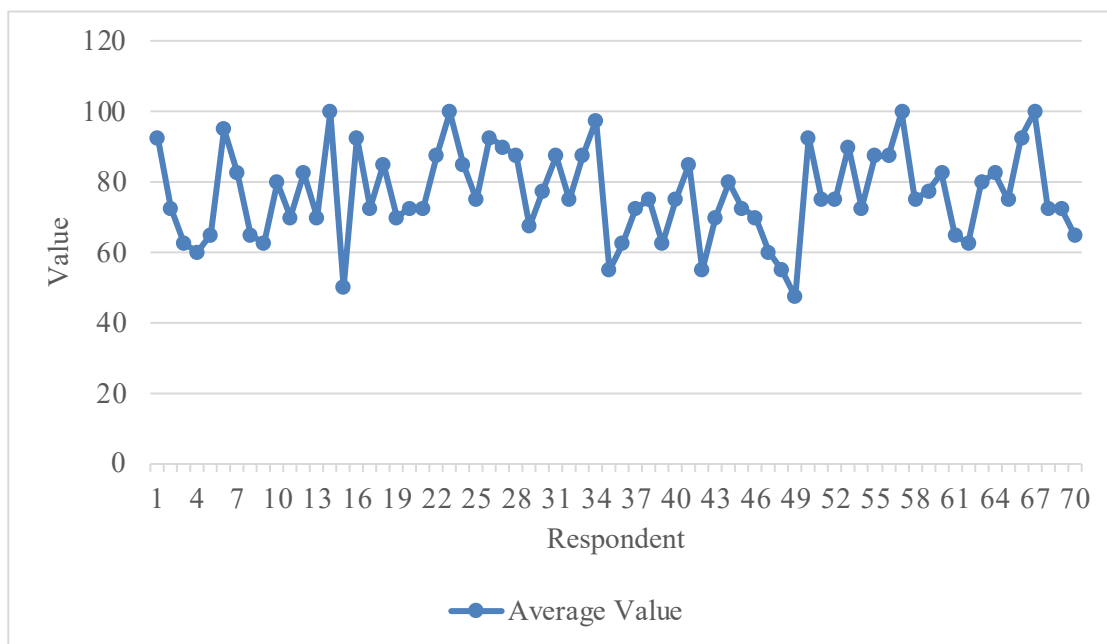**Fig. 14.** Acceptable Score SUS [30]
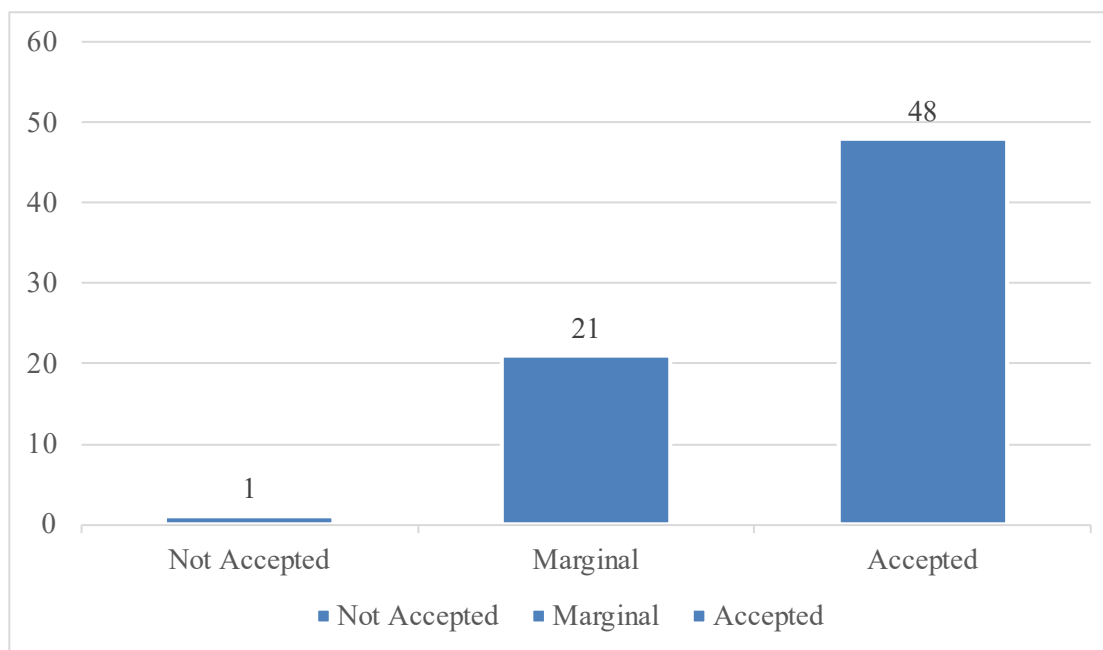

Fig. 15. Average Score SUS

Fig. 16. Result Details SUS

4. DISCUSSIONS

4.1. METHOD PERFORMANCE

In the research that researcher have done using 127 training data and 63 test data, the performance level of the RF method is far more optimal than that of the KNN. Accuracy on RF is much higher with 87.30% accuracy, 90.89% precision, 78.40% recall, and 81.00% F1-Score compared to using KNN with accuracy 79.36%, precision 62.61%, recall 62.80%, and f1-score 62.80%. The difference between the two methods is that the RF method classifies data by dividing it into several tree nodes to be much easier to classify data from each existing feature. In contrast to the KNN classification according to the most classes according to the K that the user has determined. The software has been developed to analyse blood results for anemia. However, the software only divides into two types of anemia: Iron Deficiency Anemia (IDA) and Thalassemia Trait (TT). Previous research that inspired this research was related to a website to predict the discriminatory nature of thalassemia and iron deficiency anemia. This study used data from the Centre of Medical Laboratory Service, Mahidol University, from 2014 to September 2016. Complete blood test results collected were 237 sets of haematological data from Thai subjects aged 18 to 60 years. This study divides the classification results into two classes: Iron deficiency anemia and Thalassemia trait. The method used is based on machine learning, such as K-Nearest Neighbor (KNN), Decision Tree (DT), Random Forest (RF), Artificial Neural Network (ANN), and Support Vector Machine (SVM). To evaluate the method's performance, K-Fold Cross Validation was used with K=5, with sequential accuracy of 92.36%, 98.03%, 94.17%, 94.11%, 95.05%. This study was limited to two classes of anemia, so further research is needed to classify anemia more broadly [7]. In contrast to previous studies that used two classes of IDA and TT, this study will use four classes for the type of anemia suffered by the patient, namely Iron Deficiency Anemia, Beta Thalassemia Trait, Haemoglobin E, and Mixed. To get maximum results, it is necessary to use the proper data processing method. The use of machine learning to process data aims to speed up the screening process for each patient.

Several studies address the same issue regarding machine learning to classify anemia. The research conducted by [31]-[35] has the same goal, namely to predict data in the form of complete blood count (CBC) results and build a model used to identify anemia. This study [31] used three machine learning algorithms, namely Decision Tree, Naïve Bayes, Neural Networks, and a hybrid algorithm, which combines the three methods called Hybrid Classifier. To evaluate the performance of the method, MAE and RMSE are used. The MAE results for case two are 4.440, 3.042, 0.139, and 0.083, respectively. Meanwhile, the RMSE was 0.867, 0.601, 0.022, and 0.015. For the results, the score for the two cases is 0.715, 0.831, 0.995, and 0.996, respectively. This study [33] used three machine learning methods: Random Forest, C4.5, and Naïve Bayes. To evaluate the model, K-Fold Cross-Validation and MAE were used. The accuracy results obtained a score

for Random Forest 95.3241 with an MAE of 0.0332, for an accuracy of C4.5 96.0909 with an MAE of 0.0333 and the Naïve Bayes model produces an accuracy score of 95.4602 with an MAE of 0.0347.

The development of models for anemia prediction also experienced development in 2020. Research conducted by [34] used five machine learning methods as ANN, SVM, KNN, RF, and Naïve Bayes (NB). To evaluate the machine learning model, a confusion matrix is used. The results of the evaluation of the method obtained the highest evaluation results in the ANN method of 84.17%, ANN sensitivity of 99.5%, and specificity of 27.27%. The most recent research in 2021 was conducted by [32] using six methods such as LDA, CART, KNN, SVM, RF, and LR. The evaluation of the machine learning model uses a confusion matrix by taking accuracy, sensitivity, and specificity. The highest evaluation results were the CART method 62.35%, 71.54%, and 53.52, respectively. The study conducted by [35] seud COVID-19 patients to identify whether the CBC results of COVID-19 patients changed or not. This study uses five machine learning models as MLP, SVM, Random Forest (RF), Average Ensemble (AE), and KNN Ensemble. It uses RMSE, MAE, and MBE to evaluate the machine learning model. The lowest error results were obtained using the KNN Ensemble RMSE 0.007, KNN Ensemble MAE 0.010, KNN Ensemble MBE -0.012.

Recent studies conducted by [36], [37], and [38] used Deep Learning to deepen the process of identifying anemia in patients. In 2020, a study was conducted by [36] using patient blood image data. This research uses the deep learning AlexNet model. A study conducted by [36] used a combined machine learning and deep learning method to further deepen the identification process. This study uses three models, namely Genetic Algorithm (GA), Deep Learning Algorithm of Stacked Autoencoder (SAE) and Convolutional Neural Network (CNN). A confusion matrix was used to evaluate the model with the results of GA-CNN accuracy 98.5%, GA-CNN F-Score 98.8%, GA-CNN sensitivity 98.7%, and GA-CNN precision 99.01%. The confusion matrix to evaluate the method resulted in 95.92% accuracy, 77% sensitivity, 98.82% specificity, and 90% precision. While the research conducted by [38] used the Multi-Layer Perceptron (MLP) model with the confusion matrix to evaluate the MLP model. A score of 96.04% was obtained for 1.387 datasets (August 2017 to August 2019). Then the data was tested with the 100 most recent datasets (September 2019 to August 2020), obtaining 99% accuracy.

Based on this research, the SVM and Deep Learning method can make good predictions of the data, so it is hoped that using the SVM and Deep Learning method in the development of further research on GAMA CUTE can get more optimal and faster results.

4.2. TIME COMPLEXITY

Based on the research that has been done using 190 data, the distribution of 127 training data and 63 test data and using the Random Forest and k-Nearest Neighbor methods has got good accuracy. Based on Table 2 and Table 3 regarding the results of the confusion matrix of the two methods, it was found that the evaluation of the performance of the random forest method was more optimal than the KNN. The computational complexity of machine learning is the mathematical study of computers' possibilities for efficient learning. It works in a model for machine inference based on computational complexity theory and places an explicit emphasis on efficient and generalized algorithms for learning. Theorems are presented, which help explain the limits of what can be learned efficiently from examples. These results are in the form of two algorithms with evidence of their performance, and the evaluation results show the toughness of learning. Time complexity increases with the increasing number of training points. So, if the data set is very large with many dimensions, the algorithm becomes very slow. Time complexity can be seen as a measure of how fast or slow an algorithm will perform for an input size. Testing the length of time for conducting training and predicting data using the KNN and Random Forest methods have a much different time difference. The KNN method trains 127 data for 19 milliseconds and predicts 63 data for 19 milliseconds.

Meanwhile, in the random forest method, the time required to train 127 data for 29 seconds is 71 milliseconds. Predictions for 63 data are 29 seconds and 48 milliseconds. The difference in time complexity here affects the length of the machine learning process on the results to be obtained. Research [39] uses 1.011 features, requiring feature reduction to reduce the number of features used. The number of training data examples is so small due to time constraints in data collection and the sheer number of features that when we do live training and testing, it creates the problem of over-customization and increases computational complexity.

4.3. IMBALANCED DATASET

The classification problem in this study has four classes with unbalanced datasets presenting different challenges from binary classification problems. The skewed distribution makes machine learning algorithms less effective, especially in predicting minority class examples. Based on Fig. 3 regarding training data, the

most data distribution is in the IDA class with 70 samples, while the least distribution is in the Combination class with 14 samples. Based on the data distribution, there are majority and minority classes with different numbers of data samples. Based on Table 2 and Table 3, the data distribution is still the majority in the IDA class and the minority in the Combination class. With an unbalanced data distribution, less than optimal results are obtained for each data class. Research conducted by [30,31] regarding unbalanced datasets shows that the most popular voting method is primary voting, based on the RF algorithm. The RF classifier can achieve high accuracy in data classification compared to many standard classification methods. It can minimize the overall rate of misclassification and can handle unbalanced class data.

5. CONCLUSION

In this article, an analysis of the health system and the issues and problems faced by the world's health systems is summarized. This is an attempt to conduct a comprehensive review of the health system and highlight potential avenues for quantitative research questions. This article provides the following contributions to support knowledge. The use of machine learning methods in this study speeds up the screening process for the type of anemia suffered by the patient. The performance of Accuracy, Precision, Recall, and F1-Score, the results for KNN, respectively 79.36%, 59.40%, 62.80%, and 62.80%. While in the Random Forest, respectively 87.30%, 90.89%, 78.40%, and 81.00%. From the two methods used, KNN and Random Forest, it was found that the Random Forest method has an excellent method performance compared to KNN. Thus, it can be concluded that the Random Forest method has optimal results compared to KNN. So the GAMA CUTE system was created using the Random Forest method with Python Flask for its Graphic User Interface. This study has shortcomings in terms of time complexity and the balance of the number of datasets in each class. So it is hoped that further development will be carried out in future research.

REFERENCES

[1] X. Han, C. Wang, and Z. Liu, "Red Blood Cells as Smart Delivery Systems," *Bioconjugate Chem.*, vol. 29, no. 4, pp. 852–860, Apr. 2018, https://doi.org/10.1021/acs.bioconjchem.7b00758.

[2] J. Amin *et al.*, "An Integrated Design Based on Dual Thresholding and Features Optimization for White Blood Cells Detection," *IEEE Access*, vol. 9, pp. 151421–151433, 2021, https://doi.org/10.1109/ACCESS.2021.3123256.

[3] H. Li *et al.*, "Mechanics of diseased red blood cells in human spleen and consequences for hereditary blood disorders," *Proc Natl Acad Sci USA*, vol. 115, no. 38, pp. 9574–9579, Sep. 2018, https://doi.org/10.1073/pnas.1806501115.

[4] C. Hézode *et al.*, "Elbasvir/Grazoprevir for Patients With Hepatitis C Virus Infection and Inherited Blood Disorders: A Phase III Study: Hézode *et al.*," *Hepatology*, vol. 66, no. 3, pp. 736–745, Sep. 2017, https://doi.org/10.1002/hep.29139.

[5] M. A. Ehsani, E. Shahgholi, M. S. Rahiminejad, F. Seighali, and A. Rashidi, "A new index for discrimination between iron deficiency anemia and beta-thalassemia minor: results in 284 patients.," *Pakistan journal of biological sciences : PJBS*, vol. 12, no. 5, pp. 473–5, Mar. 2009, https://doi.org/10.3923/PJBS.2009.473.475.

[6] Y. Xu, T. Hu, H. Ding, and R. Chen, "Effects of anemia on the survival of patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis," *Expert Review of Respiratory Medicine*, vol. 14, no. 12, pp. 1267–1277, 2020, https://doi.org/10.1080/17476348.2020.1816468.

[7] V. Laengsri, W. Shoombuatong, W. Adirojananon, C. Nantasenamart, V. Prachayasittikul, and P. Nuchnoi, "ThalPred: A web-based prediction tool for discriminating thalassemia trait and iron deficiency anemia," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–14, 2019, https://doi.org/10.1186/s12911-019-0929-2.

[8] Badan Penelitian dan Pengembangan Kesehatan, "Laporan Nasional Riskesdas 2007," *Laporan Nasional 2007*, pp. 1–384, 2007.

[9] Depkes RI., "Riset Kesehatan Dasar 2013," 2013.

[10] Laporan Nasional Riskesdas, "Laporan_Nasional_RKD2018_FINAL.pdf," 2018.

[11] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, pp. 1–13, 2019, https://doi.org/10.1002/widm.1312.

[12] Y. A. Jasim, "High-Performance Deep learning to Detection and Tracking Tomato Plant Leaf Predict Disease and Expert Systems," *ADCAIJ*, vol. 10, no. 2, Feb. 2021, https://doi.org/10.14201/ADCAIJ202110297122.

[13] Y. Juhn and H. Liu, "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research," *Journal of Allergy and Clinical Immunology*, vol. 145, no. 2, pp. 463–469, Feb. 2020, https://doi.org/10.1016/j.jaci.2019.12.897.

[14] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech Recognition Using Deep Neural Networks: A Systematic Review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019, https://doi.org/10.1109/ACCESS.2019.2896880.

[15] D. C. E. Saputra, A. Azhari, and A. Ma"arif, "K-Nearest Neighbor of Beta Signal Brainwave to Accelerate Detection of Concentration on Student Learning Outcomes," *Engineering Letters*, vol. 30, no. 1, pp. 318–234, 2022, [Online]. Available: http://www.engineeringletters.com/issues_v30/issue_1/EL_30_1_38.pdf

[16] G. Briganti and O. Le Moine, "Artificial Intelligence in Medicine: Today and Tomorrow," *Frontiers in Medicine*, vol. 7, no. February, pp. 1–6, 2020, https://doi.org/10.3389/fmed.2020.00027.

[17] B. Rao, "Machine Learning Algorithms: A Review," *International Journal of Computer Science and Information Technologies*, vol. 7, no. 3, pp. 1174–1179, 2016, https://doi.org/10.21275/ART20203995.

[18] R. Cioffi, M. Travaglioni, G. Piscitelli, A. Petrillo, and F. De Felice, "Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions," *Sustainability*, vol. 12, no. 2, p. 492, Jan. 2020, https://doi.org/10.3390/su12020492.

[19] H. Amakdouf, M. El Mallahi, A. Zouhri, A. Tahiri, and H. Qjidaa, "Classification and recognition of 3d image of charlier moments using a multilayer perceptron architecture," *Procedia Computer Science*, vol. 127, pp. 226–235, 2018, https://doi.org/10.1016/j.procs.2018.01.118.

[20] N. Haefner, J. Wincent, V. Parida, and O. Gassmann, "Artificial intelligence and innovation management: A review, framework, and research agenda☆," *Technological Forecasting and Social Change*, vol. 162, p. 120392, Jan. 2021, https://doi.org/10.1016/j.techfore.2020.120392.

[21] N. Mehta and M. V. Devarakonda, "Machine learning, natural language programming, and electronic health records: The next step in the artificial intelligence journey?," *Journal of Allergy and Clinical Immunology*, vol. 141, no. 6, pp. 2019-2021.e1, Jun. 2018, https://doi.org/10.1016/j.jaci.2018.02.025.

[22] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, Oct. 2020, https://doi.org/10.1016/j.chaos.2020.110059.

[23] C. González García, E. Núñez-Valdez, V. García-Díaz, C. Pelayo G-Bustelo, and J. M. Cueva-Lovelle, "A Review of Artificial Intelligence in the Internet of Things," *IJIMAI*, vol. 5, no. 4, p. 9, 2019, https://doi.org/10.9781/ijimai.2018.03.004.

[24] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, Sep. 2021, https://doi.org/10.1016/j.jii.2021.100224.

[25] P. Vamplew, R. Dazeley, C. Foale, S. Firmin, and J. Mummery, "Human-aligned artificial intelligence is a multiobjective problem," *Ethics Inf Technol*, vol. 20, no. 1, pp. 27–40, Mar. 2018, https://doi.org/10.1007/s10676-017-9440-6.

[26] J. Bartlett, "Generalized Information: A Straightforward Method for Judging Machine Learning Models," *cbi*, vol. 1, no. 2, pp. 13–22, Jun. 2019, https://doi.org/10.33014/issn.2640-5652.1.2.bartlett.1.

[27] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Applied Soft Computing*, vol. 97, p. 105524, 2020, https://doi.org/10.1016/j.asoc.2019.105524.

[28] C. Lin, D. Wu, H. Liu, X. Xia, and N. Bhattarai, "Factor identification and prediction for teen driver crash severity using machine learning: A case study," *Applied Sciences (Switzerland)*, vol. 10, no. 5, 2020, https://doi.org/10.3390/app10051675.

[29] T. R. Gadekallu *et al.*, "Early detection of diabetic retinopathy using PCA-firefly based deep learning model," *Electronics (Switzerland)*, vol. 9, no. 2, pp. 1–16, 2020, https://doi.org/10.3390/electronics9020274.

[30] N. Z. A. Halim, S. A. Sulaiman, K. Talib, and M. N. Isa, "Assessing the usability of the NDCDB checklist with Systematic Usability Scale (SUS)," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 169, p. 012099, Jul. 2018, https://doi.org/10.1088/1755-1315/169/1/012099.

[31] E. M. T. El-kenawy, "A Machine Learning Model for Hemoglobin Estimation and Anemia Classification Anemia Classification Module Hemoglobin Estimation Module Data Cleaning Data Preprocessing," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 17, no. 2, pp. 100–108, 2019,

[32] J. R. Khan, S. Chowdhury, H. Islam, and E. Raheem, "Machine Learning Algorithms To Predict The Childhood Anemia In Bangladesh," *Journal of Data Science*, vol. 17, no. 1, pp. 195–218, 2021, https://doi.org/10.6339/jds.201901_17(1).0009.

[33] M. Jaiswal, A. Srivastava, and T. J. Siddiqui, *Machine learning algorithms for anemia disease prediction*, vol. 524. Springer Singapore, 2019, https://doi.org/10.1007/978-981-13-2685-1_44.

[34] B. Sow, H. Mukhtar, H. F. Ahmad, and H. Suguri, "Assessing the relative importance of social determinants of health in malaria and anemia classification based on machine learning techniques," *Informatics for Health and Social Care*, vol. 45, no. 3, pp. 229–241, 2020, https://doi.org/10.1080/17538157.2019.1582056.

[35] M. M. Eid and A. Ibrahim, "Anemia Estimation for COVID-19 Patients Using A Machine Learning Model," *Journal of Computer Science and Information Systems*, vol. 17, no. 11, pp. 2535-1451, 2021, https://www.semanticscholar.org/paper/Anemia-Estimation-for-COVID-19-Patients-Using-A-El-kenawy-Eid/e1a71f96005c7790d6112d5933b81369b6676559.

[36] H. A. Aliyu, M. A. A. Razak, R. Sudirman, and N. Ramli, "A deep learning alexnet model for classification of red blood cells in sickle cell anemia," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 2, pp. 221–228, 2020, https://doi.org/10.11591/ijai.v9.i2.pp221-228.

[37] S. Kilicarslan, M. Celik, and Ş. Sahin, "Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification," *Biomedical Signal Processing and Control*, vol. 63, 2021, https://doi.org/10.1016/j.bspc.2020.102231.

[38] S. Yeruva, M. Sharada Varalakshmi, B. Pavan Gowtham, Y. Hari Chandana, and P. E. S. N. Krishna Prasad, "Identification of sickle cell anemia using deep neural networks," *Emerging Science Journal*, vol. 5, no. 2, pp. 200–210, 2021, https://doi.org/10.28991/esj-2021-01270.

[39] S. Purwar, R. Tripathi, R. Ranjan, and R. Saxena, "Classification of Thalassemia Patients Using a Fusion of Deep Image and Clinical Features," in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 410–415, 2021. https://doi.org/10.1109/Confluence51648.2021.9377054.

[40] H. Luo, X. Pan, Q. Wang, S. Ye, and Y. Qian, "Logistic Regression and Random Forest for Effective Imbalanced Classification," in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, pp. 916–917, 2019, https://doi.org/10.1109/COMPSAC.2019.00139.

[41] M. Zhu *et al.*, "Class Weights Random Forest Algorithm for Processing Class Imbalanced Medical Data," *IEEE Access*, vol. 6, pp. 4641–4652, 2018, https://doi.org/10.1109/ACCESS.2018.2789428.

BIOGRAPHY OF AUTHORS

Dimas Chaerul Ekty Saputra received his bachelor"s degree from the Department of Informatics, Faculty of Industrial Technology, Ahmad Dahlan University, Yogyakarta, Indonesia in 2020, an M.Sc. degree in Department of Biomedical Engineering, Graduate School, Universitas Gadjah Mada, Yogyakarta, Indonesia in 2022, and now he is a PhD candidate in the Department of Computer Science, College of Computing, Khon Kaen University, Khon Kaen, Thailand. He is currently a lecturer at Telkom University Surabaya, Indonesia. He is also a member of the Association for Scientific Computing and Electronics, Engineering (ASCEE) Student Branch Indonesia. His research interests include artificial Intelligence, pattern recognition, machine learning, signal processing, and bioinformatics. Email: dimaschaerulekty@telkomuniversity.ac.id, Orcid: 0000-0001-6978-2846.

Afiahayati is a computer scientist with a PhD in Science, currently she is a faculty member at the Dept. of Computer Science and Electronics, FMIPA UGM, Indonesia. She got her PhD from Keio University, Japan. Her passion in bioinformatics led her to join Sakakibara Lab, Keio University, Japan as a PhD student. In her PhD, she devotes her time to genome assembly research, especially metagenomic assembler, how to assemble multiple genomes from mix of multiple species reads. She has developed two metagenomic assemblers, MetaVelvet-SL and Xgenovo. Email: afia@ugm.ac.id, Orcid: 0000-0003-4647-5493.

Tri Ratnaningsih received the Medical Doctor degree from the Faculty of Medicine, Gadjah Mada University, Indonesia, in 1995, and the master"s degree in tropical medical science and the Ph.D. degree in health science from Gadjah Mada University. Subsequently, she continued her career as Clinical Pathologist and became a Hematology and Transfusion Medicine consultant, in 2014. She is currently a Lecturer and holds the position of Head of the Department of Clinical Pathology and the Laboratory Medicine, Universitas Gadjah Mada, Indonesia. Email: triratnaningsih@ugm.ac.id, Orcid: 0000-0001-9847-4770.