Jurnal Ilmiah
KURSOR
Menuju Solusi Teknologi Informasi
Vol. 13, No. 1, Juli 2025

ISSN 0216 – 0544
e-ISSN 2301– 6914

Hyperparameter optimization of XGBoost using artificial bee colony for predicting medical complications in hemodialysis patients

Rangga Laksana Aryananda [a], **Trimono** [b,*], **Wahyu Syaifullah J** [c], **Wan Suryani Wan Awang** [d]

[a,b,c] *Department of Data Science, Faculty of Computer Science, University of Pembangunan Nasional Veteran Jawa Timur, Surabaya, Indonesia*
[d] *Faculty of Informatics and Computing, University Sultan Zainal Abidin Besut Campus, 22200 Besut, Terengganu, Malaysia*
E-mail: trimono.stat@upnjatim.ac.id*

Abstract

Chronic Kidney Disease (CKD) is a serious global health issue, ranking as the 12th leading cause of death in 2019, with a 31.7% increase since 2010. Many CKD patients require hemodialysis, which poses risks of complications such as hypertension, hypotension, and gastrointestinal disorders, increasing mortality. This study predicts hemodialysis complications using XGBoost optimized with the Artificial Bee Colony (ABC) algorithm. The dataset includes numerical and categorical variables such as blood pressure, hemoglobin levels, gender, and complication history. To improve class distribution, the Synthetic Minority Over-sampling Technique is applied. Five test scenarios with different ABC parameter configurations were conducted to optimize XGBoost hyperparameters. Results indicate that balancing the dataset with SMOTE enhances model accuracy. Among the tested scenarios, Test 3, with ABC parameters n_bees set to 30, max_iter set to 30, and limit set to 10, achieved the highest accuracy, increasing from 89% (unbalanced) to 94% (balanced). Although training time increased, the improved performance highlights the potential of the XGBoost-ABC framework for early complication detection. This approach can enhance patient care, reduce mortality risks, and support clinical decision-making for hemodialysis patients.

Key words: Artificial Bee Colony, Complications, Hemodialysis, Hyperparameter, XGBoost.

INTRODUCTION

Chronic Kidney Disease (CKD) is currently one of the serious health issues that has received global attention. As of 2021, the World Health Organization (WHO) reported that no fewer than 843.6 million individuals worldwide had been affected by this disease. This number is estimated to continue to increase by 41.5% by 2040. The mortality rate recorded by the WHO was 1.2 million deaths in 2019, and this figure continues to increase compared to the previous year, making it one of the diseases with the fastest increase in mortality. In 2015, 500 million people were recorded as suffering from Chronic Kidney Disease (CKD). The high number of deaths from Chronic Kidney Disease (CKD) places this disease as one of the top 12 causes of death in the world [1].

In Indonesia, the number of cases of Chronic Kidney Disease (CKD) continues to show a significant increase. Based on data from Basic Health Research (Riskesdas) conducted by Indonesia’s Ministry of Health (Kemenkes) in 2007, there were 1,885 recorded cases of CKD. This number surged significantly to 11,689 cases in 2013, and by 2018, it had escalated to 713,783 cases [2]. This rising trend highlights CKD as a critical health concern in Indonesia, necessitating focused attention and strategic intervention.

Chronic Kidney Disease (CKD) is a progressive condition that leads to a permanent decline in kidney function, eventually reaching end-stage (stage 5), which requires renal replacement therapy such as hemodialysis, peritoneal dialysis, or kidney transplantation [3]. Hemodialysis is the most commonly used method, as it replaces kidney function by filtering the blood using a machine. The high prevalence of CKD in Indonesia is reflected in the increasing number of hemodialysis patients, as recorded in the 13th Indonesian Renal Registry, with around 60,000 new patients and 130,000 active patients in 2020 [4][5]. This situation highlights the importance of proper CKD management, particularly at Dr. Iskak General Hospital in Tulungagung, which has been designated a Regional Referral Hospital since 2015 and currently treats 60 active hemodialysis patients, with an additional 20–30 new patients each month [6].

Hemodialysis patients are at high risk of various medical complications that can reduce their quality of life. These complications include hypertension, hypotension, nausea, vascular access infections, arrhythmias, headaches, muscle cramps, and other potentially fatal conditions [3]. The 2020 Indonesia Renal Registry (IRR) report from the Indonesian Nephrology Association (PERNEFRI) noted intradialytic hypertension as the most common complication, with an incidence of 30%, while hypotension increased from 14% to 27%. In addition, cardiovascular problems are identified as the primary cause of death in hemodialysis patients, reaching rates of 42% [5].

Given the high risk of complications and mortality in hemodialysis patients, this problem is an urgent health issue to be addressed. Therefore, an early detection system is needed to identify potential complications more quickly and accurately. Currently, doctors still do the detection of complications manually by reviewing patient data directly, a method that has significant limitations [6]. As the number of patients increases and the complexity of clinical variables such as age, gender, blood pressure, hemoglobin levels, serum creatinine, and serum albumin, this manual process becomes increasingly inefficient and prone to human error [7].

To detect potential complications in hemodialysis patients earlier, a machine learning classification model can be applied using patient medical data such as age, gender, blood pressure, hemoglobin levels, and other relevant features. However, implementing this model poses challenges, particularly in handling large and unbalanced datasets that contain both numerical and categorical variables [8]. Data imbalance-such as when patients experiencing complications are significantly fewer than those who do not, or when complication types are unevenly distributed-can reduce model accuracy and performance. Additionally, the complexity and variety of features, time and storage limitations, and the risk of overfitting further complicate the modeling process. To address these issues, this study employs the Extreme Gradient Boosting (XGBoost) classification model, which is known for its efficiency in handling large datasets, strong scalability, resistance to overfitting, and high predictive accuracy [9].

XGBoost has been proven effective in dealing with data imbalance issues in various research fields. For example, in research conducted by Zhao 2018, this algorithm can overcome challenges in unbalanced datasets, even outperforming other techniques in data processing performance [10]. Although XGBoost is known as a very strong and reliable algorithm, it has major challenges, especially the difficulty in choosing optimal parameters [9]. To address this, the Artificial Bee Colony (ABC) optimization method is employed. This method emulates the natural foraging patterns of bees to determine the most suitable parameters for XGBoost.

Research related to the prediction of hemodialysis complications has been conducted in several previous studies. One study by Hsieh, Hwang, et al. 2022 used data from 264 patients collected from May 2019 to March 2021 at Chang-Hua Hospital, Ministry of Health and Welfare (MOHW), Taiwan. This study aims to predict hemodialysis complications, such as hypotension and AV fistula, by applying the XGBoost algorithm. The results showed a precision level ranging from 71% to 90% [11]. In addition, in the study by Othman, Mustafa, et al. 2022 using data from 215 patients at El-Mowasah Hospital, Alexandria University, the study aims to predict seven types of complications that may occur, namely hypotension, hypertension, shortness of breath, nausea, cramps, headaches, and chest pain. The implementation of the Artificial Neural Network (ANN) model in this study resulted in an accuracy of 82%, with specificity ranging from 40% to 95% [12].

Research by Dong, Wang, et al. 2023 used hemodialysis patient data from the Quality Control Center in Shenzhen City, China, to predict intradialytic hypotension. The model applied is the Light Gradient Boosting Machine (LightGBM), which achieves a C-statistics/Area Under Curve (AUC) accuracy of 0.82, while other models show an AUC range between 0.77 and 0.89 [13]. Research on the prediction of hemodialysis complications in Indonesia is still minimal. A study conducted by Suparti and Nurul (2019) analyzed data from 56 hemodialysis patients at Prof. Dr. Margono Soekarjo Purwokerto Regional Hospital. Using a multiple linear regression model, they examined the relationship between oxygen saturation (SpO2), heart rate (HR), and the occurrence of intradialytic complications. The findings indicated that SpO2 and HR could predict 49.9% of complications in the first hour. This predictive power declined to 27.9% in the second hour, remained relatively stable in the third hour (27.0%), and increased slightly to 29.4% in the fourth hour [14].

The study conducted by Nirvan and Rohman (2024) aimed to optimize the performance of the XGBoost algorithm in stunting classification by comparing two hyperparameter optimization methods: Grid Search and Random Search. The results showed that Grid Search improved the model's accuracy from 83.28% to 89.09%, while Random Search achieved an accuracy of 88.71%. Although both methods were effective in enhancing model performance, each has its limitations. Grid Search, despite being systematic, has a major drawback in terms of time and computational efficiency, as it must evaluate all possible parameter combinations-making it impractical when dealing with a large and complex hyperparameter space. On the other hand, Random Search is faster since it evaluates combinations randomly, but lacks thoroughness and may miss the optimal configuration. Both methods tend to be less adaptive and are unable to dynamically guide the search toward more promising solution spaces [15].

Based on previous research, this study proposes a more optimal approach by combining the Artificial Bee Colony (ABC) algorithm to optimize XGBoost parameters to predict complications in hemodialysis patients. The ABC technique is a metaheuristic approach that relies on honeybee foraging behavior. In this procedure, the bees are divided into three groups: employed bees, which investigate initial answers; observer bees, who choose solutions based on likelihood; and scout bees, who hunt for fresh solutions in various places. This concept allows ABC to tackle various optimization issues efficiently and frequently beats other methods like Genetic Algorithms or Particle Swarm Optimization (PSO) [9]. In addition, to overcome the limitations of the unbalanced class distribution in complicated data, this study applies the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE balances the data distribution by synthesizing new samples in minority classes so that the model can learn better to recognize rare complication patterns [16]. This study uses XGBoost, ABC optimization, and SMOTE data balancing approaches to build a prediction model for assessing the risk of hypertension, hypotension, and gastrointestinal problems in hemodialysis patients. The resulting model is expected to help medical personnel detect early risk, enabling more effective interventions to prevent complications.

MATERIAL AND METHODS

Fig. 1 illustrates the research flow designed systematically to predict types of complications in hemodialysis patients. The research began with collecting information on hemodialysis patients and relevant medical records. The next step involved data pre-processing to manage missing values, eliminate duplicate entries, and ensure data consistency. The parameters of the Artificial Bee Colony (ABC) and XGBoost algorithms are then set to ensure optimal model performance. In the modeling stage, XGBoost optimization using ABC is carried out through parameter initialization, division of the worker bee population and foragers, and iterative evaluation to improve prediction accuracy. The best resulting model is used to classify hemodialysis complications, such as hypertension, hypotension, and gastrointestinal disorders. This study's findings are likely to increase the early diagnosis of complications and assist medical personnel in making more accurate and efficient decisions.

Data Collecting

This study utilized a dataset from the medical records of 200 hemodialysis patients treated at Dr. Iskak Tulungagung Regional Hospital from 2022 to 2024. This data includes 15 features used to predict hemodialysis complications, which are classified into four categories: hypertension, hypotension, gastrointestinal disorders, and no complications. The specifics of the utilized features are presented in Table 1.

Table 1. Hemodialysis dataset feature

Feature	Description
Gender	Gender (male/female)
Age	Patient's age in years
Blood pressure	Systolic and diastolic blood pressure (mmHg)
Hemoglobin	Hemoglobin level in the blood (g/dL)
SI-TIBC	Transferrin capacity filled by iron in the blood (%)
Urea	Urea level in the blood (mg/dl)
Creatinine	Creatinine level in the blood (mg/dl)
EPO_3000	Administration of erythropoietin 3000 IU to treat anemia
EPO_2000	Administration of erythropoietin 2000 IU to stimulate red blood cell production.
Tron_Sucrose	Intravenous iron supplements to treat iron deficiency.
Hemodialysis Duration	Duration of hemodialysis per month
Medical history	Patient and family medical history
Type of Complications	Type of patient complications (hypertension, hypotension, and gastrointestinal)

Data Preprocessing

The preprocessing stage in this study includes data understanding, handling missing values, encoding data, and applying test scenarios to different data parameters and distributions.

Fig. 2. Preprocessing data

Data Understanding and Exploration

The results of data exploration show the distribution of complication classes in the dataset as follows: without complications (142 cases), hypertension (93 cases), gastrointestinal complications (93 cases), and hypotension (58 cases).

Check and Handling Missing Value

This study imputed blank data in numeric columns, such as systolic, diastolic, hemoglobin, and urea, using the mean value. In contrast, categorical data, such as drug type (EPO 2000 and 3000), were filled with the mode value.

Data Encoding

Before modeling, categorical variables such as gender, EPO 2000, EPO 3000, medical history, and type of complications are converted to numerical format using label encoding. This transformation ensures that all features can be processed by the XGBoost algorithm, including the target variable representing the type of patient complications.

ABC-XGBOOST METHOD PRINCIPLE

Tianqi Chen invented Extreme Gradient Boosting (XGBoost), a boosting-based ensemble technique, in 2014. This algorithm is designed to improve model accuracy by correcting classification errors from previous iterations and optimizing decision tree learning in parallel to speed up execution [17]. XGBoost has scalability, efficiency, and flexibility advantages, making it popular in various machine learning tasks, including classification and regression. As a Gradient Boosting Decision Tree (GBDT) development, this algorithm can significantly speed up the computing process [18]. However, the primary disadvantage of this approach is the difficulty of establishing hyperparameters, which takes time and experience in optimal adjustment [9].

To overcome the challenge of determining the best hyperparameters in XGBoost, the Artificial Bee Colony (ABC) algorithm is applied to optimize parameter selection and enhance model performance efficiently. This metaheuristic optimization technique simulates the foraging behavior of bees in identifying and utilizing optimal resources. This algorithm divides bees into three leading roles: worker bees that explore food sources, observer bees that choose the best source based on the information obtained, and explorer bees that look for new alternatives to maintain the diversity of solutions. This process continues until it meets specific criteria, such as the maximum limit of iterations [9].

The process of combining the ABC and XGBoost algorithms in predicting complications in hemodialysis patients based on medical record data is carried out the stages shown in Fig. 3.

The conventional Gradient Boosting Decision Tree (GBDT) approach only uses first-order derivatives and has limitations in parallel training due to the dependence between weak learner models. XGBoost overcomes this by incorporating a second-order Taylor expansion, as shown in Equation (1), which refines the approximation of the loss function. Moreover, XGBoost incorporates a regularization component to optimize both the objective function and model complexity, effectively minimizing the potential for overfitting [19][20]. The XGBOOST formula is defined as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F \tag{1}$$

In this context, K signifies the total count of decision trees, while $f_k(x_i)$ represents the input function associated with the k-th decision tree. The predicted outcome is denoted $\hat{y}_i$, whereas F encompasses all potential Classification and Regression Trees (CART).

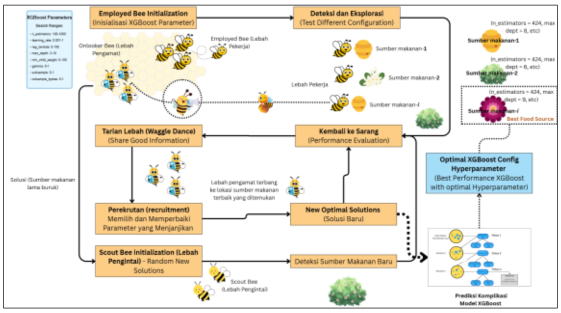

Fig. 3. Ilustration model XGBoost and ABC

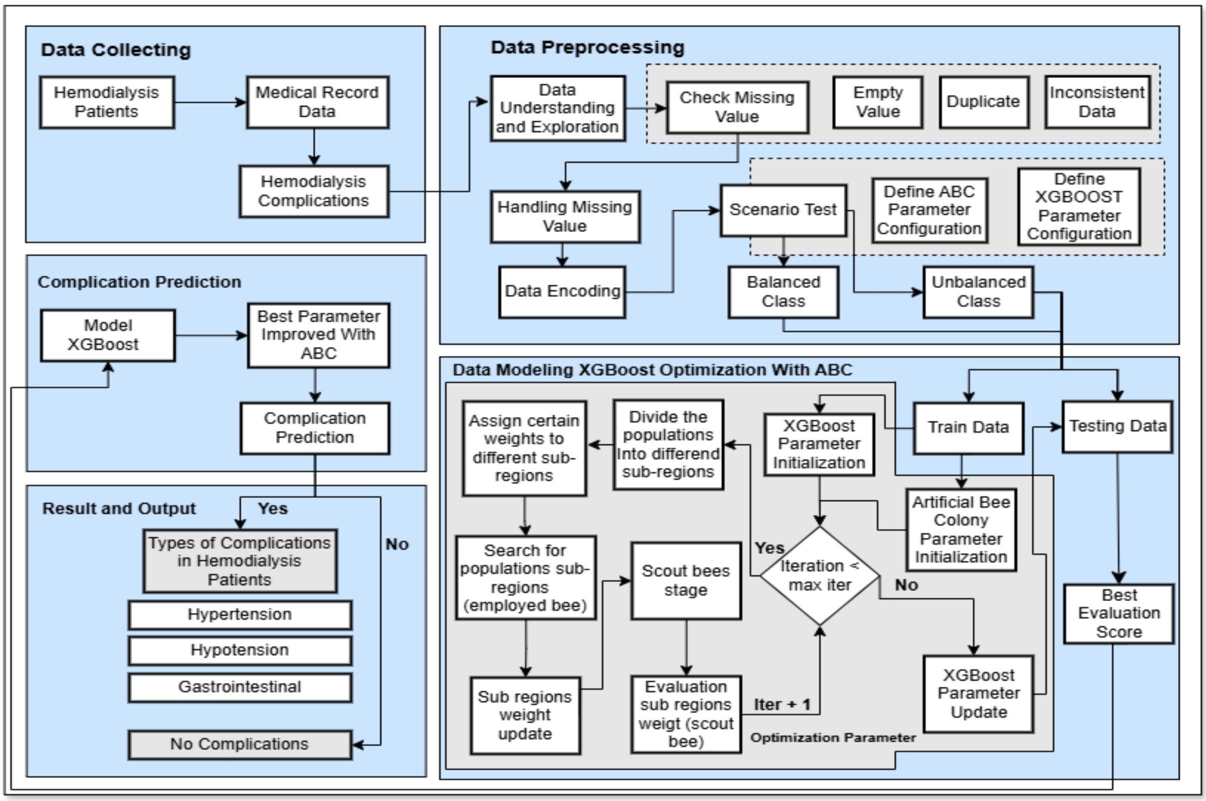

Fig. 1. Systematic research flow for predicting complication types in hemodialysis patients

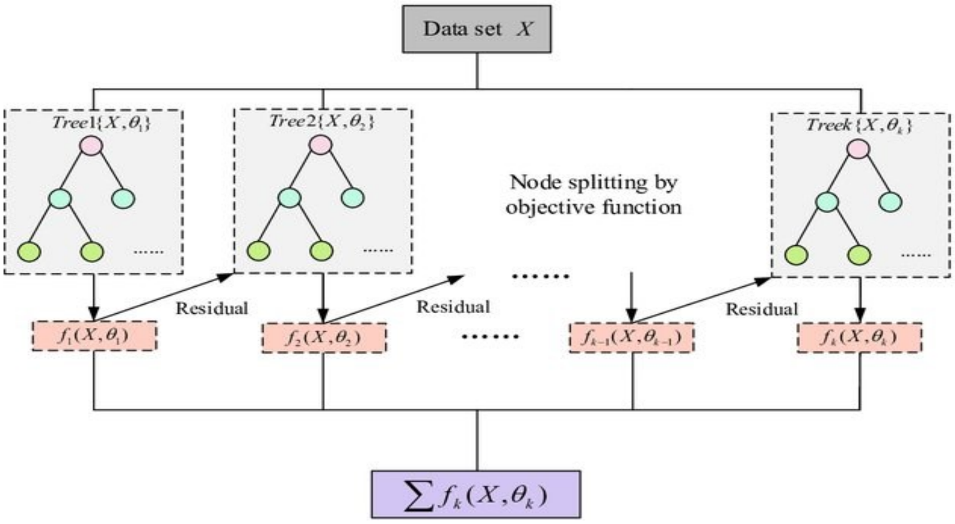

Fig. 4. Flowchart of XGBoost model (adapted from [21])

The XGBoost equation iteratively constructs new trees that adjust to the residuals of the previous trees, gradually improving model accuracy. This tree-adding process is governed by the objective function, which evaluates and optimizes model performance at each iteration [19]. The objective function for the at-th iteration, as defined in Equation (2), is expressed as follows:

$$\boldsymbol{Obj}^{(t)} = \sum_{i=1}^{n} \boldsymbol{L}\,[(\boldsymbol{y}_i, \hat{\mathrm{y}}_i^{(t-1)}) + \boldsymbol{f}_t(\boldsymbol{x}_i)] + \boldsymbol{\Omega}(\boldsymbol{f}_t) \qquad (2)$$

Where $L\ (y_i\ , \hat{y}_i)$ represents the training loss for a given sample, with $\hat{y}_i$ as the predicted value and y_i as the actual classification label. Meanwhile, $\Omega(f_t)$ serves as the regularization component, refining the model structure while minimizing the risk of overfitting, as defined in Equation (3):

$$\Omega\ (f) = \gamma\ T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \qquad (3)$$

In this context, T signifies the total count of leaf nodes, while ω_j represents the weight assigned to each leaf node. The constants γ and λ function as penalty coefficients. The objective function is estimated through a Taylor series expansion, utilizing a second-order approximation derived from Equation (2), as formulated in Equation (4).

$$\boldsymbol{Obj}^{(t)} = \sum_{i=1}^{n} [\ \boldsymbol{L}\ (\boldsymbol{y}_i, \hat{\mathrm{y}}_i^{(t-1)}) + \boldsymbol{g}_i\, \boldsymbol{f}_t(\boldsymbol{x}_i) + \frac{\mathbf{1}}{\mathbf{2}} \boldsymbol{h}_i\, \boldsymbol{f}_t^2(\boldsymbol{x}_i)\ ] + \boldsymbol{\Omega}(\boldsymbol{f}_t) \qquad (4)$$

Each term corresponds to the first and second derivatives of a given pair, allowing Equation (4) to be rewritten in a simplified form as follows:

$$Obj^{(t)} = \sum_{j=1}^{T} [\ (\sum_{i \in I\ j} g_i\ )\omega_j + \frac{1}{2}(\sum_{i \in I\ j} h_i\ + \lambda\ )\omega_j^2\ ] + \gamma\ T \qquad (5)$$

The optimal weight that minimizes the objective function can be determined by deriving its partial derivative and solving for zero, as shown in Equation (6).

$$\omega_j^* = -\frac{\sum_{i \in I\ j} (\ g_i\ )}{\sum_{i \in I\ j}\ (h_i\ + \lambda)} \qquad (6)$$

By replacing Equation (6) into Equation (5), the best objective function value is determined, as shown in Equation (7).

$$\boldsymbol{Obj}^{\ *} = -\frac{1}{2} \frac{\sum_{i \in I\ j}(\ g_i\ )}{\sum_{i \in I\ j}(h_i\ + \lambda)} +\ \gamma\ T \qquad (7)$$

Throughout the training phase, the model incrementally computes the loss at every node to identify the leaf node with the most significant loss increase. XGBoost sequentially adds new trees by selecting optimal feature splits. Each new tree learns a function $f_k(x, \theta_k)$ to refine the residuals from previous sample is mapped to a particular leaf node in every tree, where each node is assigned a corresponding score [19]. The scores across all trees are then aggregated to generate the final prediction for the sample. A flowchart of the XGBoost process is shown in Fig. 4.

In XGBoost modeling, it is necessary to determine various hyperparameters, such as the number of trees *(T)*, learning rate (η), and regularization coefficients (λ), to optimize performance. The large number of parameters makes the tuning process quite complex. To overcome this challenge, XGBoost optimization can be done using the Artificial Bee Colony (ABC) algorithm with the following stages:

(1) In the first step, the ABC algorithm creates the initial population by determining the position of the food supply, which represents the first solution set. Furthermore, the iterative procedure is repeated till it reaches the maximum number of cycles. Three sorts of bees play a role in this process: employed bees, who seek solutions surrounding their current position; observer bees, which pick options based on likelihood; and scout bees, who hunt for new solutions when the previous ones aren't ideal [19]. The following equation (8) is used to generate a population of solutions (location of food sources):

$$x_{ij} = l_i + rand\,(0,1) * (u_i - u_j) \tag{8}$$

Where $i = 1,2,\ldots,SN$ and $j = 1,2,..,D$; SN indicate the total number of food sources, while D represents the number of parameters or problem dimensions, the variable u_i and u_j is the upper and lower bound of the solution space for the objective function. The function generates random numbers that are usually distributed in the range [0, 1].

(2) Furthermore, in the employed bee phase, worker bees are tasked with finding new food sources (alternative solutions) by comparing the amount of nectar (fitness value) of existing solutions. Solutions with better fitness values will be retained, while less optimal solutions will be ignored [19]. The acquisition of new solutions from old solutions is calculated using Equation (9), which determines the mechanism for updating the bee's position in the search space to improve the exploration of optimal solutions.

$$v_{ij} = z_{ij} + \varphi\,(z_{ij} - z_{kj}) \tag{9}$$

Where and represent two different food sources, indicate a certain dimension, and are random values that are uniformly distributed in the range [-1, 1].

(3) During the onlooker bee phase, bees determine food sources based on the probability associated with the fitness of employed bees. As in the previous phase, alternatives with higher fitness values will be preferred, while less optimum options will be ignored [19]. The chance of picking a food source is computed using the following formula:

$$P_i = 1 - \frac{E_i}{\sum_{j=1}^{SN}(1 - E_j)} \tag{10}$$

(4) In the scout bee phase, ineffective food sources are replaced with new ones. If a source cannot be increased in a certain number of cycles (called a "limit"), it is considered obsolete and ignored [19]. Scout bees then generate new food sources using the same approach as in the initial phase. Once a new source is found, its fitness value is assessed against the prior one through a greedy selection process. The source exhibiting the highest fitness value is retained, while the previous one is discarded.

(5) Next, the ABC algorithm will select the best food source. The processes in the employee bee, onlooker bee, and scout bee phases take place repeatedly until a predetermined iteration limit is reached. In each iteration, the results of each phase are compared, and the final solution is chosen based on the highest fitness value [19].

(6) The result of selecting the best food source represents the optimal combination of XGBoost parameters. The parameters obtained from this best solution are used to train the XGBoost model to improve prediction accuracy. Thus, the ABC algorithm automatically adjusts the XGBoost hyperparameters, include the number of trees, learning rate, and regularization coefficient, resulting in

the best-performing model. This optimization process continues until the model reaches the desired level of convergence or performance. The full stages of the ABC algorithm in the optimization process are illustrated in Fig. 5.

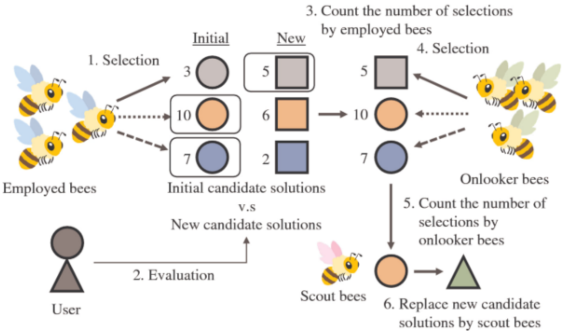

Fig. 5. Artificial bee colony algorithm process

Define XGBOOST and ABC Parameter Configuration

At this stage, various parameter configuration settings were made for the Artificial Bee Colony (ABC) and XGBoost algorithms. In optimization using ABC, the three primary parameters adjusted are n_bee (number of bees), max_iter (maximum number of iterations), and limit (limit for searching for new solutions). These parameters were tested using five different testing schemes to evaluate their effect on optimization performance.

Meanwhile, for the XGBoost algorithm, an optimal parameter range search was carried out for eight parameters that affect model performance. The range of values used in this parameter search is shown in more detail in Table 2.

Metrics Evaluating Model Performance

A confusion matrix is essential for assessing classification performance by comparing predicted labels with actual outcomes. It represents classification results using four key indicators: TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), and FN (*False Negative*). These form the foundation for calculating key metrics like accuracy, precision, recall, and the F1 score, offering insights into the model's ability to differentiate categories while minimizing error [21]. The confusion matrix allows for a more exact assessment of the model's prediction performance, revealing its strengths and limitations [22]. The mathematical formulations for these metrics are presented in Equations (11)-(14) [23].

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{11}$$

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

$$Recall = \frac{TP}{TP+FN} \tag{13}$$

$$F1-score = \frac{2 \times precision \times recall}{precision+recall} \tag{14}$$

Table 2. XGBoost parameter range

Parameter	Search Range
n estimators (T)	100-1000
learning rate (η)	0,001 - 1
regular coefficient (λ)	0-100
max depth	3-10
min child weight	0-100
gamma (γ)	0-1
subsample (s)	0-1
colcample bytree	0-1

RESULT AND DISCUSSION

The implementation of medical complication prediction using XGBoost optimized with Artificial Bee Colony (ABC) is carried out through a series of systematic experiments. Initial testing focuses on selecting the optimal ABC parameter configuration. In this stage, five test scenarios were conducted by varying ABC parameters, including the count of worker bees (*n_bees*), the upper limit of iterations (*max_iter*), and the exploitation threshold (*limit*), to assess their impact on XGBoost's effectiveness, as shown in Table 3.

Table 3. Testing Scenario with Artificial Bee Colony Parameter Variation

Testing Scheme	Set Parameter ABC		
	n_bees	max_iter	limit
1	10	10	5
2	20	10	5
3	30	30	10
4	20	10	10
5	20	20	10

In addition, the model was tested on data with balanced and unbalanced class distributions to assess the robustness of its performance in various conditions. SMOTE (Synthetic Minority Over-sampling Technique) was employed to handle dataset imbalance. The dataset was subsequently divided, with 80% used for training and 20% for testing, ensuring robust evaluation and generalization.

Evaluation of the Best XGBoost Model with Parameter ABC in the Unbalanced Dataset

Fig. 6 presents the results of the model evaluation based on five test scenarios performed on an unbalanced dataset. Of the five scenarios, Test Scenario 1 shows the best performance with the highest accuracy, which is 91%. This scenario uses ABC parameters with the configuration *n_bee* = 20, *max_iter* = 10, and *limit* = 5. In addition, a comparison of iteration times shows that Scenario 1 has the fastest training time, which is 6 minutes, while Scenario 3 takes the longest time, which is 52 minutes. The results of the model training time comparison are presented in more detail in Fig. 7.

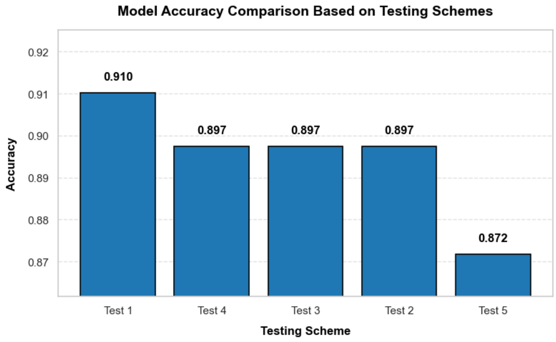

Fig. 6. Accuracy comparison of testing schemes in the unbalanced dataset

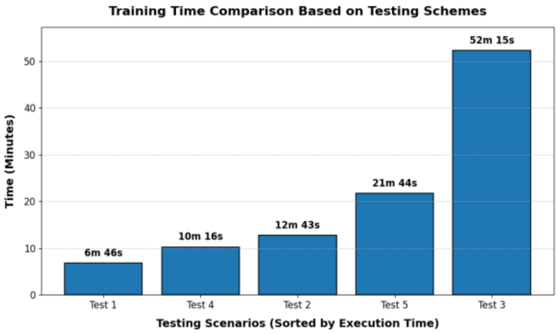

Fig. 7. Iteration time for each testing scheme in the unbalanced dataset

Test Scenario 1 performs best in generating the ideal XGBoost model, according to the findings of the model evaluation of the accuracy in the imbalanced dataset. The optimization's optimal hyperparameters are displayed in Table 4.

Table 4. Optimal hyperparameters for xgboost in test 1

Parameter	Best Parameter Value
n estimators (T)	476
learning rate (η)	0.7130940260014383
reg lambda (λ)	19.83568411211097
max depth	5
min child weight	0
gamma (γ)	0.6076823245551388
subsample (s)	0.7941295826978276
colsample bytree	0.5119665847979727

The performance evaluation of the best test scenario (Test 1) is illustrated in Fig. 8, which presents the learning curve based on Stratified K-Fold Cross-Validation. The graph compares the training score and cross-validation score as the training data increases, ensuring a balanced class distribution across folds for stable evaluation results [24]. As the amount of data increases, the model's accuracy improves, achieving a training score of 0.93 and a cross-validation score of 0.87. This result suggests strong generalization capabilities with no significant overfitting.

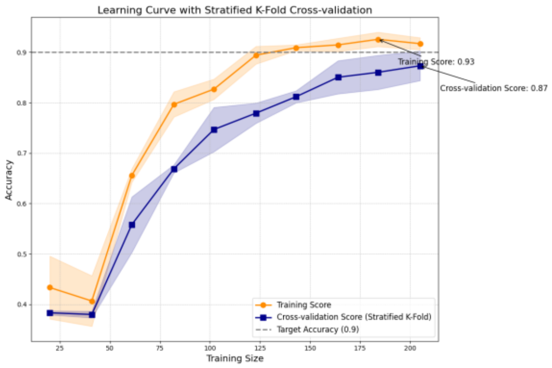

Fig. 8. Learning curve of test 1

To analyze the performance of Test Scenario 1 in multiclass classification, Fig. 9 illustrates the ROC curve. The *Area Under the Curve* (AUC) is derived by evaluating the relationship between the *True Positive Rate* (TPR) and *False Positive Rate* (FPR), providing insights into the model's capability to distinguish between multiple classes. The results indicate that all classes achieve a remarkably high AUC score (≥ 0.97), demonstrating the model's strong classification performance even in the presence of class imbalance.

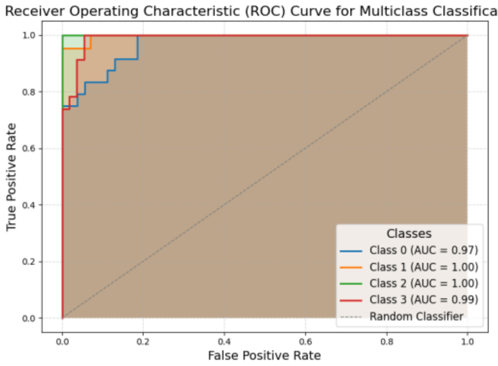

Fig. 9. Receiver operating characteristic curve of test 1

Based on the model obtained, Fig. 10 displays the feature importance analysis using XGBoost to identify the most influential features in the classification. The results show that Systolic has the most significant contribution to the model prediction, followed by SI-TIBC (%), Urea, and Diastolic. These findings confirm that these factors play a significant role in determining the classification results by XGBoost.

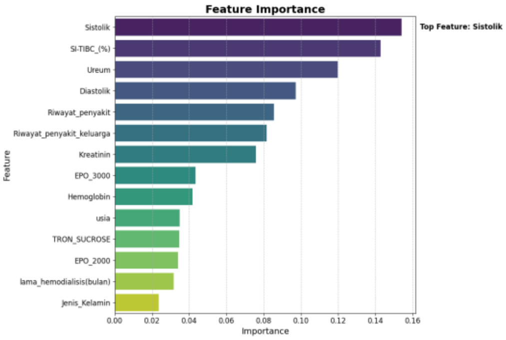

Fig. 10. Feature importance of the XGBoost model in test 1

Evaluation of the Best XGBoost Model with Parameter ABC in a Balanced Dataset

Testing was also carried out on a dataset with a balanced class distribution. Fig. 11 presents the results of the model evaluation based on five testing scenarios carried out on data that has been balanced using SMOTE. Of the five testing scenarios, Scenario 3 shows the best performance with the highest accuracy, which is 94%. This scenario uses ABC parameters with the configuration n_bee = 30, max_iter = 30, and limit = 10. In terms of time efficiency, Scenario 1 has the fastest training time, which is 10 minutes, while Scenario 3 requires the longest training time, which is 104 minutes. The results of the model training time comparison are presented in more detail in Fig. 12.

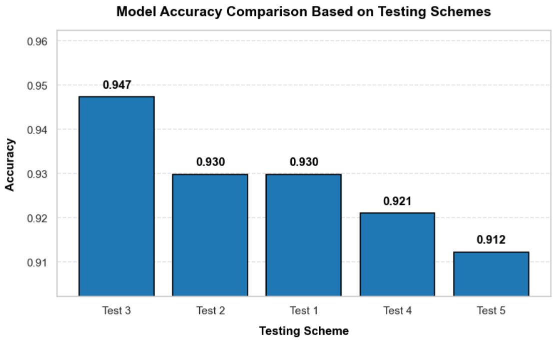

Fig. 11. Accuracy comparison of testing schemes in the balanced dataset

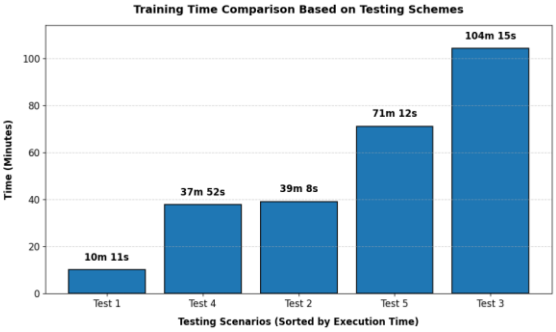

Fig. 12. Iteration time for each testing scheme in the balanced dataset

Based on the model evaluation results, the accuracy obtained on the balanced dataset indicates that Test 3 is the best test scenario for the XGBoost model. The optimal XGBoost hyperparameters were successfully identified, as presented in Table 5.

Table 5. Optimal Hyperparameters for XGBoost in Test 3

Parameter	Best Parameter Value
n estimators (T)	332
learning rate (η)	0.8321564289807684
reg lambda (λ)	46.07085476078473
max depth	5
min child weight	0
gamma (γ)	0.0117333768716177
Subsample (s)	0.1970623288575786
colsample_bytree	0.9239102632618799

The performance evaluation of the best test scenario (Test 3) is illustrated in Fig. 13, which presents the learning curve based on Stratified K-Fold Cross-Validation. The graph compares the training score and cross-validation score as the training data increases, ensuring a balanced class distribution across folds for stable evaluation results. As the amount of data increases, the model's accuracy improves, achieving a training score of 0.97 and a cross-validation score of 0.90. This result suggests strong generalization capabilities with no significant overfitting.

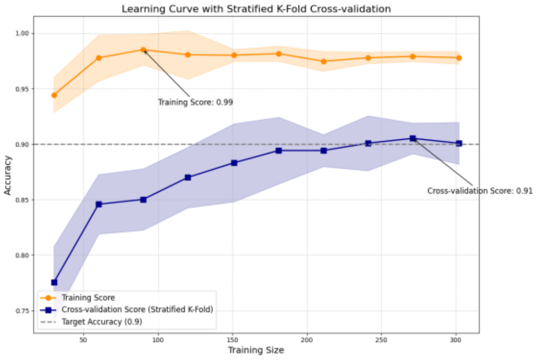

Fig. 13. Learning curve of test 3

The Operating Characteristic (ROC) Curve evaluation test was also carried out on the test with balanced data, as illustrated in Fig. 14, to evaluate the model's performance on balanced data with various testing schemes. The best test result (Test 3) shows that each class has a very high AUC value (≥ 0.99), indicating that the model performs very well in distinguishing the classes in the balanced data.

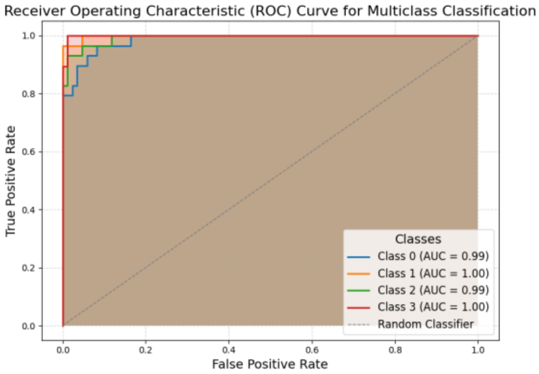

Fig. 14. Receiver operating characteristic curve of test 3

The balanced data test with the best test (Test 3) shows that Systolic is the feature with the greatest contribution to model prediction, as illustrated in Fig. 15, followed by SI-TIBC (%), Urea, and Diastolic. These findings indicate that these factors have a significant influence on determining the classification results produced by the XGBoost model for balanced data.

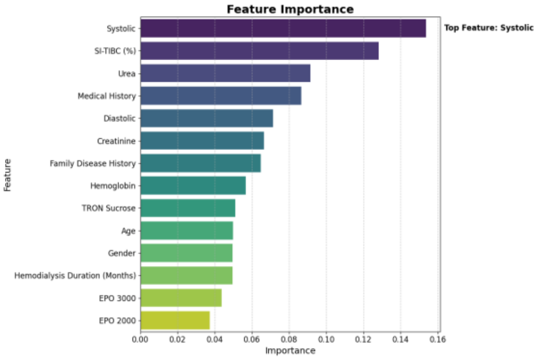

Fig. 15. Feature importance of the xgboost model in test 3

Performance Comparison on Balanced and Unbalanced Data

Testing on both unbalanced and balanced datasets reveals a notable disparity in model effectiveness. The assessment outcomes demonstrate that the model attains greater accuracy when dealing with datasets where hemodialysis complication classes are more evenly distributed. From Table 6, training on a balanced dataset improves accuracy compared to an unbalanced dataset. Test 3 recorded the highest increase in accuracy, from 89% (unbalanced) to 94% (balanced), albeit with a longer training time. A similar thing happened in Test 1, where accuracy increased from 91% to 93%, with a little extra training time. A comprehensive evaluation of XGBoost performance, enhanced through the Artificial Bee Colony (ABC) algorithm, is displayed in Table 6.

The evaluation of model performance on unbalanced and balanced data indicates that the Test 3 scenario, with Artificial Bee Colony parameters (n_bees = 30, max_iter = 30, and limit = 10), attains the highest accuracy, reaching 94%. This model is proven to be more optimal in detecting hemodialysis complications than other scenarios. The confusion matrix results for the Test 3 model are presented in Table 7 to provide a more detailed overview of its classification performance.

Table 6. Comparison of ABC-optimized XGBoost across test scenarios on unbalanced and balanced datasets

Test	Unbalanced Dataset		Balanced Dataset	
	Accuracy (%)	Time (minute)	Accuracy (%)	Time (minute)
1	91	6	93	10
2	89	12	93	39
3	89	52	94	104
4	89	10	92	37
5	87	21	91	71

Table 7. Confusion matrix of the best test scenario for ABC-Optimized XGBoost in classifying hemodialysis complications (test 3)

Types of Complications	Precision (%)	Recall (%)	F1-score (%)
No complication	90	90	90
Hipertensi	100	96	98
Hipotensi	96	93	95
Gastrointestinal	93	100	97

CONCLUSION

This study aims to develop a predictive model for medical complications in hemodialysis patients—including hypertension, hypotension, and gastrointestinal disorders—using the XGBoost algorithm optimized with the Artificial Bee Colony (ABC) method, while also addressing data imbalance using the Synthetic Minority Over-sampling Technique (SMOTE).

The results from five testing scenarios indicate that the application of the ABC method significantly contributed to finding the optimal parameter combinations for the XGBoost model. The best scenario (Test 3) achieved the highest accuracy of 94% on balanced data, an improvement from 89% on imbalanced data. The SMOTE technique proved effective in enhancing model performance by ensuring a more even distribution of data across classes, thereby making the model more reliable in identifying each type of complication.

Although this study does not directly compare the ABC method with conventional optimization approaches such as Grid Search or Random Search, previous literature has highlighted limitations of those methods, particularly in terms of time efficiency and inefficiency in exploring large solution spaces. In contrast, the ABC method is adaptive and capable of navigating parameter spaces more flexibly and efficiently. This is evidenced by the model's ability to achieve high accuracy (up to 94%) even when searching a complex parameter space, consisting of eight XGBoost hyperparameters. Thus, the ABC method has proven effective in addressing optimization challenges in predictive models with wide search spaces.

This predictive model can serve as a valuable tool for medical practitioners by enabling early detection of potential complications in hemodialysis patients. Its integration into clinical decision-support systems may improve patient outcomes, reduce mortality risks, and contribute to better healthcare management for chronic kidney disease patients undergoing hemodialysis. Future research may explore additional machine learning techniques, real-time model deployment, and external validation to further refine predictive capabilities.

REFERENCES

[1] N. Z. Aditama, H. Kusumajaya, and N. Fitri, "Faktor-faktor yang berhubungan dengan kualitas hidup pasien gagal ginjal kronis," vol. 6, no. 1, 2024. https://doi.org/10.33862/jnsr.v1i2.466

[2] Tim Penulis Kementrian Kesehatan, "Laporan Nasional Riskesdas 2018." Badan Penelitian dan Pengembangan Kesehatan. [Online]. Available : https://repository.badankebijakan.kemkes.go.id/id/eprint/3514/

[3] L. ISROIN, Manajemen cairan pada pasien hemodialisis untuk meningkatka n kualitas hidup. Unmuh Ponorogo Press, 2016. [Online]. Available: https://umpopress.umpo.ac.id/media/2024-01-18/manajemen-cairan-pada-pasien-hemodialisis-untuk-meningkatkan-kualitas-hidup/mobile/index.html

[4] Rizki Muliani, L. A. Fauziah, and Sumbara, "Komorbiditas dan Lama Menjalani Hemodialisis dengan Kualitas Hidup pada Klien yang Menjalani Hemodialisis," Window Health J. Kesehat., pp. 533-544, May 2022. https://doi.org/10.33096/woh.v5i02.24

[5] T. Irr, "INDONESIAN RENAL REGISTRY 2020". [Online]. Available : https://www.indonesianrenalregistry.org/

[6] H. J. Murff, V. L. Patel, G. Hripcsak, and D. W. Bates, "Detecting adverse events for patient safety research: a review of current methodologies," J. Biomed. Inform., vol. 36, no. 1-2, pp. 131-143, Feb. 2003, https://doi.org/10.1016/j.jbi.2003.08.003

[7] S. Chaudhuri et al., "Artificial intelligence enabled applications in kidney disease," Semin. Dial., vol. 34, no. 1, pp. 5-16, Jan. 2021, https://doi.org/10.1111/sdi.12915

[8] F. Ratantja Kusumajati, B. Rahmat, and A. Junaidi, "Implementation of Balancing Data Method using Smotetomek in Diabetes Classification using XGBoost," J. Ilm. Kursor, vol. 12, no. 4, pp. 201-212, Dec. 2024, https://doi.org/10.21107/kursor.v12i4.410

[9] F. A. Mohammad, A. M. Rizki, and A. N. Sihananto, "Peramalan Tingkat Inflasi di Indonesia menggunakan Artificial Bee Colony Dan Xgboost," J. Inform. Dan Tek. Elektro Terap., vol. 12, no. 3, Aug. 2024, https://doi.org/10.23960/jitet.v12i3.4827

[10] P. Septiana Rizky, R. Haiban Hirzi, and U. Hidayaturrohman, "Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang," J Stat. J. Ilm. Teori Dan Apl. Stat., vol. 15, no. 2, pp. 228-236, Dec. 2022, https://doi.org/10.36456/jstat.vol15.no2.a5548

[11] W.-H. Hsieh, C. C.-Y. Ku, H. P.-C. Hwang, M.-J. Tsai, and Z.-Z. Chen, "Model for Predicting Complications of Hemodialysis Patients Using Data From the Internet of Medical Things and Electronic Medical Records," IEEE J. Transl. Eng. Health Med., vol. 11, pp. 375-383, 2023, https://doi.org/10.1109/JTEHM.2023.3234207

[12] M. Othman, A. M. Elbasha, Y. S. Naga, and N. D. Moussa, "Early prediction of hemodialysis complications employing ensemble techniques," Biomed. Eng. OnLine, vol. 21, no. 1, p. 74, Oct. 2022, https://doi.org/10.1186/s12938-022-01044-0

[13] J. Dong et al., "Machine learning-based intradialytic hypotension prediction of patients undergoing hemodialysis: A multicenter retrospective study," Comput. Methods Programs Biomed., vol. 240, p. 107698, Oct. 2023, https://doi.org/10.1016/j.cmpb.2023.107698

[14] K. G. Pebriantari and I. P. Astuti Dewi, "Hubungan Komplikasi Intra Hemodialisis dengan Kualitas Hidup pada Pasien Chronic Kidney Disease (Ckd) Stage V yang menjalani Hemodialisis di Ruang Hemodialisa BRSU Tabanan Tahun 2017," J. Ris. Kesehat. Nas., vol. 2, no. 1, pp. 9-17, Apr. 2018, https://doi.org/10.37294/jrkn.v2i1.95

[15] N. A. Pramudhyta and M. S. Rohman, "Perbandingan Optimasi Metode Grid Search dan Random Search dalam Algoritma XGBoost untuk Klasifikasi Stunting," J. MEDIA Inform. BUDIDARMA, vol. 8, no. 1, p. 19, Jan. 2024, https://doi.org/10.30865/mib.v8i1.6965

[16] R. Basuki, Z. J. H. Tarigan, H. Siagian, L. S. Limanta, D. Setiawan, and J. Mochtar, "The effects of perceived ease of use, usefulness, enjoyment and intention to use online platforms on behavioral intention in online movie watching during the pandemic era," Int. J. Data Netw. Sci., vol. 6, no. 1, pp. 253-262, 2022, https://doi.org/10.5267/j.ijdns.2021.9.003

[17] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani, and M. Idhom, "Exploratory

Data Analysis and Machine Learning Algorithms to Classifying Stroke Disease," IJCONSIST J., vol. 2, no. 02, pp. 77-82, Jun. 2021, https://doi.org/10.33005/ijconsist.v2i02.49

[18] A. T. Damaliana, T. Trimono, and D. A. Prasetya, "Ensemble Tree untuk Memprediksi Level Resiko Maternal Mortality di Bangladesh," Pros. Semin. Nas. SAINS DATA, vol. 2, no. 1, pp. 24-29, Nov. 2022, https://doi.org/10.33005/senada.v2i1.36

[19] R. Luo, L. Guo, X. Li, J. Tuo, C. Lei, and Y. Zhou, "An eXtreme Gradient Boosting Algorithm Combining Artificial Bee Colony Parameters Optimized Technique for Single Sand Body Identification," IEEE Access, vol. 9, pp. 156894-156906, 2021, https://doi.org/10.1109/ACCESS.2021.3129830

[20] A. T. Damaliana, A. Muhaimin, and D. A. Prasetya, "Forecasting the Occupancy Rate of Star Hotels In Bali using the Xgboost and SVR Methods". Vol. 12, No.1, 2024. https://doi.org/10.26714/jsunimus.12.1.2024.24-33

[21] A. Muhaimin, W. Wibowo, and P. A. Riyantoko, "Multi-label Classification Using Vector Generalized Additive Model via Cross-Validation," J. Inf. Commun. Technol., vol. 22, 2023, https://doi.org/10.32890/jict2023.22.4.5

[22] A. Z. Nur, H. Suyono, and M. Aswin, "Application of Hybrid GA-PSO to Improve the Performance of Decision Tree C5.0," J. Ilm. Kursor, vol. 10, no. 4, Dec. 2020, https://doi.org/10.21107/kursor.v10i4.248

[23] M. Nasrudin, "MRI-Based Brain Tumor Instance Segmentation Using Mask R-CNN," Comput. Eng. Appl. J., vol. 13, no. 03, pp. 1-9, Oct. 2024, https://doi.org/10.18495/comengapp.v13i03.490

[24] F. N. Riyana Putri, R. R. Isnanto, and A. Sugiharto, "Skin Rash Classification System using Modified Densenet201 Through Random Search for Hyperparameter Tuning," J. Ilm. Kursor, vol. 12, no. 4, pp. 179-190, Dec. 2024, https://doi.org/10.21107/kursor.v12i4.418