Published online on the journal's webpage: **http://journal.ittelkom-pwt.ac.id/index.php/dinda**

Journal of Dinda

Data Science, Information Technology, and Data Analytics

Vol. 5 No. 2 (2025) 139 - 146 E-ISSN: 2809-8064

Implementation of Random Forest Algorithm with RFE and SMOTE on Cardiotocography Dataset

Muhammad Ahsani Nur Taqwimi[1*], Buang Budi Wahono[2], Harminto Mulyo[3]

[1*,2,3]Information Engineering, Faculty of Science and Technology, Nahdlatul Ulama Islamic University Jepara

[1*]211240001200@unisnu.ac.id, [2]budihono78@gmail.com, [3]minto@unisnu.ac.id

Abstract

This study examines the implementation of the Random Forest algorithm combined with the Synthetic Minority Oversampling Technique (SMOTE) and Recursive Feature Elimination (RFE) methods for fetal health classification based on Cardiotocography (CTG) data. The dataset used is "Fetal Health Classification" from Kaggle consisting of 2,126 samples with three classes: Normal, Suspect, and Pathological. The main challenges faced are class imbalance and high number of features, which can affect model accuracy. SMOTE is used to increase the proportion of minority data, while RFE is applied to filter the most relevant features and reduce model complexity. The evaluation results show that the developed model is able to achieve 95% accuracy, 93% precision, 89% recall, and 91% F1-score. The ROC-AUC values are 0.9881 (Normal), 0.9789 (Suspect), and 0.9985 (Pathological), respectively. Model validation was performed using 10-fold cross validation which produced an average accuracy of 97.59% with a deviation of ± 0.0097, indicating that the model has stable performance and can be generalized well. The Wilcoxon non-parametric statistical test produced a p-value = 0.0020 (α = 0.05), which proves that the increase in model performance after the application of SMOTE and RFE is statistically significant. Although the model showed high performance in the Normal and Pathological classes, the performance in the Suspect class still needs to be improved. Overall, the combination of Random Forest, SMOTE, and RFE proved effective in improving the performance of fetal health classification.

Keywords: *Cardiotocography, Random Forest, SMOTE, RFE, Fetal Health*

1. Introduction

Pregnancy is a moment that is highly anticipated by most women. During pregnancy, pregnant women need to maintain their health well, because an unhealthy fetus can endanger the condition of the pregnant woman herself [1]. Pregnant women are highly susceptible to health problems, especially infections and serious diseases that have the potential to be fatal. This condition is influenced by physiological changes in the body and different immune system mechanisms. This condition can increase the risk of miscarriage or death during childbirth[2].

According to data from the Indonesian Ministry of Health, maternal and fetal mortality rates in Indonesia have increased from 4,627 cases in 2020 to 7,389 cases in 2021 [3]. This condition emphasizes the importance of accurate fetal monitoring to prevent complications that can endanger the lives of both the mother and the baby. One of the devices used to monitor fetal health is Cardiotocography (CTG). This tool monitors various fetal parameters, such as Fetal Heart Rate (FHR), Uterine Contractions (UC), and heart rate variability, so it can help detect potential health problems early [4].

The Cardiotocography dataset available on the Kaggle platform, namely "Fetal Health Classification", is the source of data in this study. This dataset consists of 2,126 fetal data with 22 attributes, classified by specialist doctors into three categories: Normal, Suspect, and Pathological. However, this dataset faces challenges in the form of data imbalance, where the majority of data is in the Normal category at 77.8%, while the Suspect and Pathological categories are much less at 13.9% and 8.3% respectively [5].

This kind of data imbalance can cause the machine learning model to be biased towards the majority class, so that the prediction performance for the minority class tends to be low [6]. To overcome this problem, the Synthetic Minority Oversampling Technique (SMOTE) is used, which effectively increases the amount of data in the minority class so that the data distribution

Received: 14-04-2025 | Accepted: 23-06-2025 | Published: 05-08-2025

becomes more balanced and the model's sensitivity to the minority class increases [7], [8].

In addition, datasets with many features (high dimensions) can also trigger overfitting problems and increase computational complexity [9], [10]. Therefore, this study applies the Recursive Feature Elimination (RFE) feature selection method to simplify the dataset and improve model performance. RFE has been shown to improve model accuracy and AUC by removing less relevant features [11], [12].

For the classification process, the Random Forest algorithm was chosen because of its ability to handle complex data and produce accurate predictions. Random Forest utilizes an ensemble learning approach by building several decision trees and combining the results through majority voting, thereby reducing the risk of overfitting and increasing classification accuracy [13], [14].

Based on this background, there are still gaps in previous studies that mostly only focus on handling data imbalance or feature selection separately, without systematically integrating both in one classification approach. Therefore, the purpose of this study is to develop a fetal health classification model by combining the Random Forest algorithm, SMOTE oversampling technique, and RFE feature selection on the Cardiotocography dataset. This study is expected to contribute to significantly improving classification performance, especially for minority classes, as well as simplifying feature complexity to obtain more optimal and efficient results in medical data processing.

2. Research Methods

This research involves four main stages, namely data collection, data preprocessing, modeling, and model evaluation. The following is a research flow diagram shown in Figure 1.

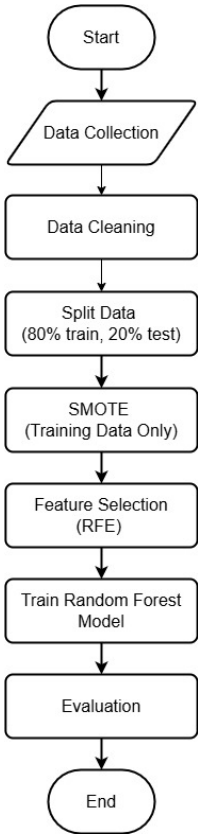

Figure 1. Flowchart

2.1. *Data Collection*

This study uses the Cardiotocography dataset available on the Kaggle platform. The dataset used is titled "Fetal Health Classification", which can be accessed at https://www.kaggle.com/andrewmvd/fetal-health-classification. This dataset is extracted from cardiotocograms and classified by obstetricians [5].

2.2. *Data Preprocessing*

The data preprocessing stage aims to prepare the dataset to suit the model's needs and ensure optimal training. The dataset is divided into 80% for training data and 20% for testing data. This data is processed in three stages, namely data cleaning, SMOTE, and RFE.

2.2.1. *Data Cleaning*

Data cleaning is done to ensure that the dataset used in the analysis meets adequate quality standards. This stage is designed to identify and correct various anomalies in the data, such as the presence of missing values and data duplication. The removal step is carried out to prevent bias that can affect the accuracy of the classification model.

2.2.2. *SMOTE*

SMOTE is applied to overcome the problem of class imbalance in the dataset, namely between the majority class, namely the Normal class, and the minority classes, namely Suspect and Pathological. SMOTE works by generating synthetic samples from the minority class to increase the amount of data in that class, so that the data distribution becomes more balanced [15]. By addressing class imbalance, SMOTE helps reduce bias that may occur in the prediction model, ensuring that the model can predict more accurately for all three classes without the dominance of the majority class. This technique helps the model to be more sensitive to the minority class and increases the classification accuracy of that class. In this study, SMOTE is only applied to the training data after the data splitting process (train_test_split), not to the entire dataset. This aims to prevent data leakage in the test data. The parameters used are k_neighbors = 5 and random_state = 42.

2.2.3. *RFE*

Recursive Feature Elimination (RFE) is a wrapper-based feature selection method used to find the optimal feature subset by iteratively eliminating irrelevant features. This technique is often used in various Machine Learning applications to improve model accuracy and reduce computational complexity [16]. In its application, RFE utilizes a machine learning model to assess the importance of each feature, removes the feature with the lowest contribution at each iteration, and rebuilds the model until the best feature subset is found [17]. In this study, the number of features used in RFE was 18 and the machine learning algorithm used as an estimator in the RFE feature selection process was the Random Forest Classifier, because of its ability to handle high-dimensional data and its stability in feature selection.

2.3. *Modeling*

After going through the data pre-processing stage, the next step is to build a model using Random Forest. Random Forest is built using a decision tree as its basic component to carry out the selection process in data classification. Random Forest is very effective in handling classification on large amounts of data. This algorithm consists of three main steps, namely: bootstrap sampling to build a decision tree, making random predictions by each tree, and combining the prediction results from each tree through a majority vote mechanism to select a class [13]. An illustration of Random Forest can be seen in Figure 2.

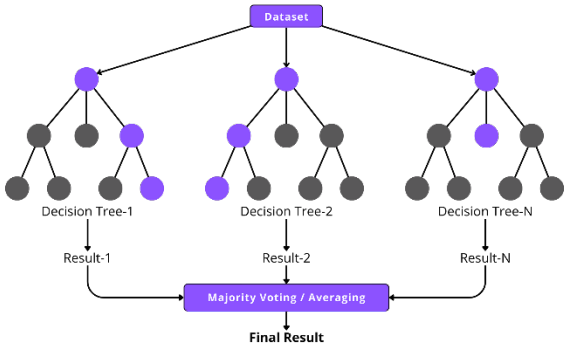

Figure 2. Random Forest Illustration

Random Forest utilizes entropy values to determine the level of attribute uncertainty in the dataset, and information gain is used to select attributes that can separate the dataset into homogeneous groups. This process ensures high accuracy in classification and reduces the risk of overfitting [14]. To calculate the entropy value, the formula listed in formula 1 is used, while to calculate the information gain value, formula 2 is used.

$$Entropy(S) = \sum_{i=1}^{n} - pi * log_2 pi \qquad (1)$$

Where S is the set of cases, n is the number of partitions of S, and pi is the ratio of Si to S.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy(S_i) \qquad (2)$$

Where S is the set of cases, A is an attribute, n is the number of attribute partitions of A, |Si| is the number of cases in the i-th partition, |S| is the number of cases in S.

The Random Forest model in this study uses the following parameters: n_estimators = 300, max_depth = None, class_weight = 'balanced', random_state = 42, max_features = 'sqrt', n_jobs = -1, and oob_score = True.

2.4. *Model Evaluation*

This study uses Accuracy, Precision, Recall, F1-score and ROC-AUC matrices to evaluate the performance of the model in classification. The evaluation is carried out based on the results of the confusion matrix, which provides information on the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP describes the number of positive data that is predicted correctly, while TN describes the number of negative data that is predicted correctly. FP refers to the number of negative data that is incorrectly

predicted as positive, while FN is the number of positive data that is incorrectly predicted as negative [18]. The Confusion Matrix can be seen in Figure 3.

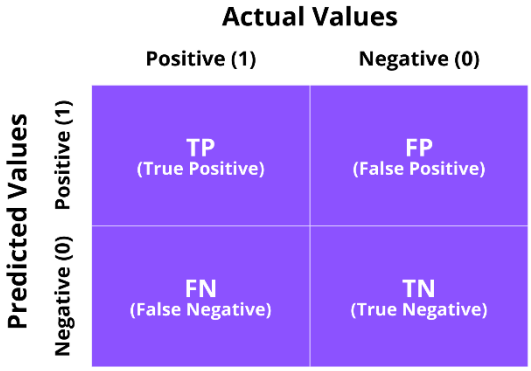

Figure 3. Confusion Matrix

2.4.1. *Accuracy*

Accuracy is defined as the ratio of the number of correct predictions to the total number of prediction data. The formula for accuracy is given in formula 3.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

2.4.2. *Precision*

Precision is the ratio of the number of correct predictions for a label to the total number of predictions, both correct and incorrect, for that label. Precision can be calculated by dividing the number of True Positives by the total Predicted Positives, as shown in formula 4.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

2.4.3. *Recall*

Recall is used to evaluate the extent to which the model can avoid False Negatives, which is calculated as the ratio of True Positive to Actual Positive. The recall formula can be found in formula 5.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

2.4.4. *F1-score*

F1-score is the combined average of precision and recall, which provides a balance between the two metrics. F1-score is formulated as shown in formula 6.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

2.4.5. *ROC-AUC*

ROC-AUC is used to measure the ability of the model to distinguish between positive and negative classes. ROC describes the relationship between True Positive Rate (TPR) and False Positive Rate (FPR). TPR can also be called Recall. The TPR and FPR formulas are shown in formulas 7 and 8.

$$TPR = \frac{TP}{TP + FN} \quad (7)$$

$$FPR = \frac{FP}{FP + TN} \quad (8)$$

While AUC shows the value of the area under the ROC curve. The higher the AUC value, the better the model is in classifying. The AUC formula is shown in formula 9.

$$AUC = \int_0^1 TPR(FPR)d(FPR) \quad (9)$$

2.4.6. *K-fold Cross Validation*

K-fold Cross Validation is a model validation method that divides the dataset into K parts or subsets of the same size. Each subset is used in turn as test data, while the rest are used for training. This process is repeated K times, so that each subset becomes test data once. The results of the model evaluation are then calculated based on the average of all iterations [19]. In this study, the fold value used was 10 folds. An illustration of K-fold Cross Validation can be seen in Figure 4.

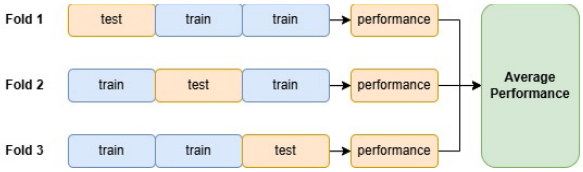

Figure 4. K-fold Cross Validation

2.4.7. *Wilcoxon*

Wilcoxon signed-rank test is a non-parametric statistical test method used to compare two paired samples. This test is useful for determining whether there is a statistically significant difference between two conditions or models, especially when the data distribution is not assumed to be normal [20]. In this study, Wilcoxon was used to compare the performance of the baseline model (without applying SMOTE and RFE) with the final model (using SMOTE and RFE) based on the accuracy results of the 10-fold cross-validation process. The test was carried out at a significance level of 0.05 to ensure that the increase in model accuracy was not due to chance factors, but had a statistically significant basis.

3. Results and Discussion

3.1. *Data Collection*

This data contains 2,126 fetal data with 22 attributes and 3 classes representing fetal health status. These data are used as training and testing data in this study.

3.2. *Data Preprocessing*

The data obtained were then cleaned from duplicate data and missing values to prevent bias that could affect the accuracy of the classification model. From the results of this preparation, 2,113 data can be entered into the modeling stage with the Random Forest algorithm. Of the 2,113 data obtained from the research object, 80% of the data was divided for training, and 20% of the data was used for testing. The following is the data splitting shown in Table 1.

Table 1. Data Splitting

Class	Training Data	Testing Data	Total
Normal (1)	1.316	330	1.646
Suspect (2)	234	58	292
Pathological (3)	140	35	175
Total	1.690	423	2.113

3.2.1. *SMOTE Results*

The training dataset is then balanced using SMOTE. The following is the data distribution before applying SMOTE shown in Figure 5, and the data distribution after applying SMOTE shown in Figure 6.

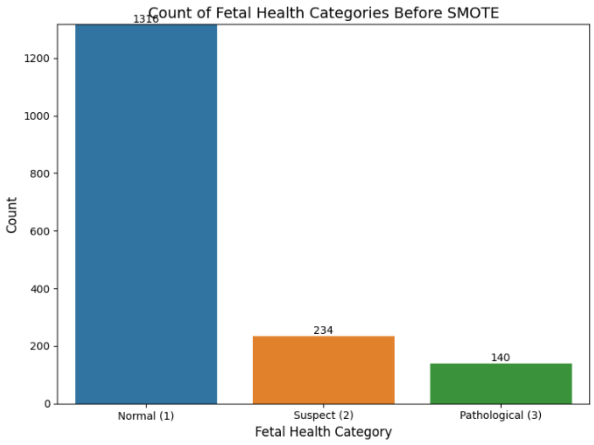

Figure 5. Before SMOTE

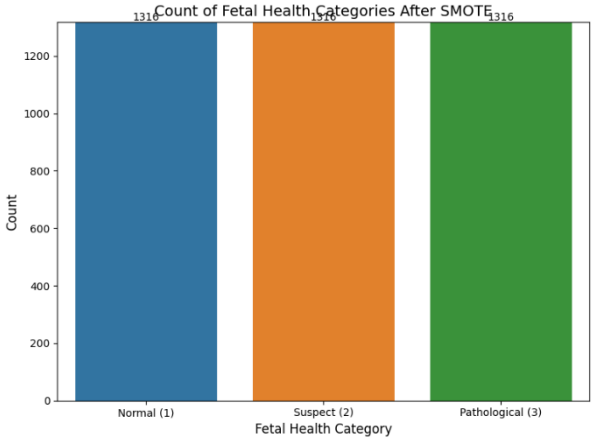

Figure 6. After SMOTE

3.2.2. *RFE Results*

The initial number of features in this dataset is 22 features, with the implementation of feature selection resulting in 18 best features according to the results listed in Table 2. These features have the highest level of importance coefficient compared to the features that were eliminated. A high level of importance indicates that these features have a significant influence on their class. Conversely, a low level of importance indicates that these features do not have much influence on their class. The selected features can be seen in Table 2.

Table 2. RFE Selected Features

No	Feature	Importance
1	abnormal_short_term_variability	0.130856
2	percentage_of_time_with_abnormal_long_term_variability	0.123919
3	histogram_mean	0.092154
4	accelerations	0.088285
5	histogram_median	0.082183
6	mean_value_of_short_term_variability	0.081248
7	prolongued_decelerations	0.078527
8	mean_value_of_long_term_variability	0.057967
9	histogram_mode	0.056562
10	baseline value	0.036006
11	histogram_variance	0.028399
12	histogram_min	0.027136
13	histogram_width	0.026775
14	histogram_max	0.024905
15	uterine_contractions	0.021402
16	fetal_movement	0.017674
17	histogram_number_of_peaks	0.015837
18	histogram_tendency	0.010165

3.5 *Classification Results*

After the data is ready, modeling is done with the Random Forest algorithm. The model is evaluated using test data to measure performance in classifying fetal health. From the results of the implementation of the Random Forest algorithm, the confusion matrix results are obtained as shown in Figure 7.

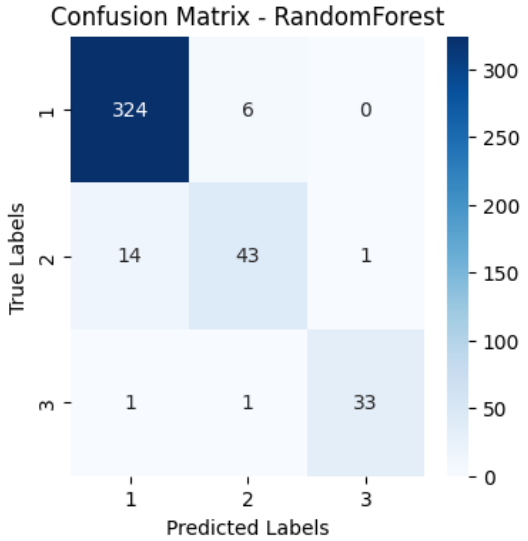

Figure 7. Confusion Matrix Results

Confusion Matrix in Figure 7 shows that most of the values are on the main diagonal, indicating that the model successfully predicted the label correctly. For example, the Normal(1) and Pathological(3) classes were classified well. However, in the Suspect(2) class, there were 14 samples that were misclassified as the Normal class. The following are the results of accuracy, precision, recall, and f1-score shown in Figure 8.

Classification Report:

	precision	recall	f1-score	support
Normal(1)	0.96	0.98	0.97	330
Suspect(2)	0.86	0.74	0.80	58
Pathological(3)	0.97	0.94	0.96	35
accuracy			0.95	423
macro avg	0.93	0.89	0.91	423
weighted avg	0.94	0.95	0.94	423

Figure 8. Classification Report

The results of using the Random Forest Model on the dataset used in this study showed 95% accuracy, 93% precision, 89% recall, and 91% f1-score. These results are explained in detail in Figure 8. The Suspect(2) class has lower precision, recall, and f1-score values, indicating the need for optimization in the pre-processing or classification model to improve model performance in the Suspect(2) class. The following is the ROC graph shown in Figure 9 below.

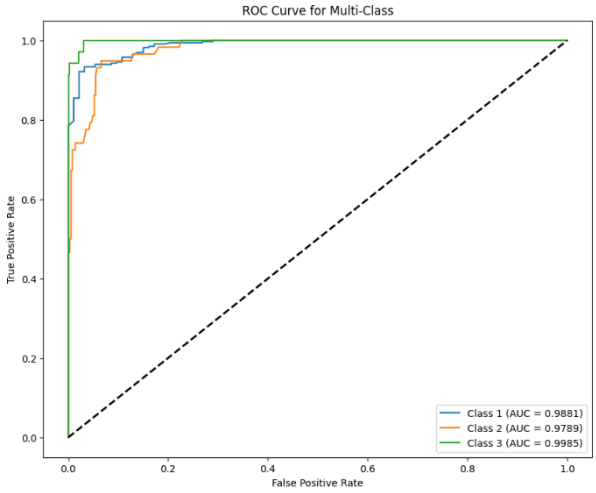

Figure 9. ROC Chart

From the ROC graph, the AUC value for class 1 (Normal) is 0.9881, class 2 (Suspect) 0.9789, and class 3 (Pathological) 0.9985. Overall, these results indicate that the model has good performance in classifying fetal health.

```
10-Fold - Cross Validation
0.9873417721518988
0.9822784810126582
0.9772151898734177
0.959493670886076
0.9696202531645569
0.9620253164556962
0.9848101265822785
0.979746835443038
0.9873096446700508
0.9695431472081218

Cross Validation Accuracy Mean: 0.9759 ± 0.0097
```

Figure 10. Cross Validation

Figure 10 shows that the results of the accuracy evaluation of each fold produce fairly high and consistent values, ranging from 0.9595 to 0.9873. The average accuracy obtained from all folds is 0.9759 with a standard deviation of ± 0.0097. This shows that the developed model has stable performance and is able to generalize well to different data variations in each training and testing subset.

```
Cross Validation Accuracy (10-Fold)
without SMOTE and RFE   |       with SMOTE + RFE
0.9349112426035503      |       0.9873417721518988
0.9585798816568047      |       0.9822784810126582
0.9526627218934911      |       0.9772151898734177
0.9112426035502958      |       0.959493670886076
0.9349112426035503      |       0.9696202531645569
0.9171597633136095      |       0.9620253164556962
0.9289940828402367      |       0.9848101265822785
0.9644970414201184      |       0.979746835443038
0.9467455621301775      |       0.9873096446700508
0.9349112426035503      |       0.9695431472081218

Wilcoxon Test:
Statistic = 0.0000, p-value = 0.0020
→ Statistically significant difference (p < 0.05)
```

Figure 11. Wilcoxon

In Figure 11, the test results show that the cross-validation accuracy value of the baseline model (without SMOTE and RFE) ranges from 0.9112 to 0.9645, while the model with SMOTE and RFE has a higher and stable accuracy in the range of 0.9595 to 0.9873. Based on the Wilcoxon test, a statistical value of 0.0000 and p-value = 0.0020 were obtained. Since the p-value is smaller than the significance level $\alpha = 0.05$, it can be concluded that there is a statistically significant difference between the two models.

4. Conclusion

This study shows that the combination of the Random Forest algorithm with the Synthetic Minority Oversampling Technique (SMOTE) and Recursive Feature Elimination (RFE) techniques can significantly improve the performance of fetal health classification. The model built produces an accuracy of 95%, precision of 93%, recall of 89%, and F1-score of 91%. The high ROC-AUC values for the three classes (Normal 0.9881, Suspect 0.9789, and Pathological 0.9985) indicate the model's ability to effectively distinguish fetal health conditions.

The SMOTE method successfully overcomes class imbalance in the dataset, while RFE simplifies the number of features from 22 to 18 without reducing model accuracy. Validation was performed using 10-fold cross-validation which produced an average accuracy of 97.59% with a deviation of ± 0.0097, indicating that the model has stable performance and can be generalized well. The Wilcoxon non-parametric statistical test produced a p-value = 0.0020 ($\alpha = 0.05$), which proves that the increase in model performance after the application of SMOTE and RFE is statistically significant.

This model has the potential to be applied as a decision support system in the medical world, especially in fetal health monitoring based on Cardiotocography (CTG) data. With a fast and accurate classification process, this model can be integrated into a cardiotocography system or mobile application to assist medical personnel in making early diagnoses.

However, there are several limitations in this study. The model was tested using a clean dataset, so its performance on real-time data containing noise, missing values, or irregularities still needs to be evaluated further. In addition, the performance in the Suspect class is still low compared to the other two classes, which indicates the need for improvements in the pre-processing or feature selection aspects. In the future, model development can be directed at hyperparameter optimization, exploration of alternative algorithms such as LightGBM, and direct testing on clinical data to improve the robustness and accuracy of the model in real-world environments.

References

[1] M. R. Santoso and P. Musa, "REKOMENDASI KESEHATAN JANIN DENGAN PENERAPAN ALGORITMA C5.0 MENGGUNAKAN CLASSIFYING CARDIOTOCOGRAPHY DATASET," *J. Simantec*, vol. 9, no. 2, pp. 65–76, Jun. 2021, doi: 10.21107/simantec.v9i2.10730.

[2] B. Abu-Raya, C. Michalski, M. Sadarangani, and P. M. Lavoie, "Maternal Immunological Adaptation During Normal Pregnancy," *Front. Immunol.*, vol. 11, Oct. 2020, doi: 10.3389/fimmu.2020.575197.

[3] K. K. RI, *Profil Kesehatan Indonesia 2021*. Jakarta: Kementerian Kesehatan RI, 2022.

[4] A. Mehbodniya *et al.*, "Fetal health classification from cardiotocographic data using machine learning," *Expert Syst.*, vol. 39, no. 6, Jul. 2022, doi: 10.1111/exsy.12899.

[5] D. Ayres-de-Campos, J. Bernardes, A. Garrido, J. Marques-de-S, and L. Pereira-Leite, "Sisporto 2.0: A program for automated analysis of cardiotocograms," *J. Matern. Fetal. Med.*, vol. 9, no. 5, pp. 311–318, Sep. 2000, doi: 10.1002/1520-6661(200009/10)9:5<311::AID-MFM12>3.0.CO;2-9.

[6] A. Thumpati and Y. Zhang, "Towards Optimizing Performance of Machine Learning Algorithms on Unbalanced Dataset," in *Artificial Intelligence & Applications*, Academy & Industry Research Collaboration Center, Oct. 2023, pp. 169–183. doi:

10.5121/csit.2023.131914.

[7] P.-I. Zhang *et al.*, "Real-time AI prediction for major adverse cardiac events in emergency department patients with chest pain," *Scand. J. Trauma. Resusc. Emerg. Med.*, vol. 28, no. 1, p. 93, Dec. 2020, doi: 10.1186/s13049-020-00786-x.

[8] R. Bhuvaneswari and R. S, "Optimizing Lung Cancer Classification Prediction based on Random Forest and SMOTE Techniques," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, IEEE, Jul. 2024, pp. 1–6. doi: 10.1109/ICEEICT61591.2024.10718616.

[9] H. Jeon and S. Oh, "Hybrid-Recursive Feature Elimination for Efficient Feature Selection," *Appl. Sci.*, vol. 10, no. 9, p. 3211, May 2020, doi: 10.3390/app10093211.

[10] C. T. Tran, "ENSEMBLE LEARNING APPROACHES FOR CLASSIFICATION WITH HIGH-DIMENSIONAL DATA," *J. Sci. Tech.*, vol. 12, no. 01, Jun. 2023, doi: 10.56651/lqdtu.jst.v12.n1.659.ict.

[11] B. Langenberger, T. Schulte, and O. Groene, "The application of machine learning to predict high-cost patients: A performance-comparison of different models using healthcare claims data," *PLoS One*, vol. 18, no. 1, p. e0279540, Jan. 2023, doi: 10.1371/journal.pone.0279540.

[12] R. J. Majed, R. M. Al-Heddi, and A. M. Zeki, "Parkinson's Disease Detection Using Data Mining Models: A Comparative Study," in *2023 4th International Conference on Data Analytics for Business and Industry (ICDABI)*, IEEE, Oct. 2023, pp. 290–294. doi: 10.1109/ICDABI60145.2023.10629302.

[13] L. Budianti and Suliadi, "Metode Weighted Random Forest dalam Klasifikasi Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *Bandung Conf. Ser. Stat.*, vol. 2, no. 2, pp. 103–110, Jul. 2022, doi: 10.29313/bcss.v2i2.3318.

[14] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *Stata J. Promot. Commun. Stat. Stata*, vol. 20, no. 1, pp. 3–29, Mar. 2020, doi: 10.1177/1536867X20909688.

[15] M. I. C. Rachmatullah, "The Application of Repeated SMOTE for Multi Class Classification on Imbalanced Data," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 22, no. 1, pp. 13–24, Nov. 2022, doi: 10.30812/matrik.v22i1.1803.

[16] E. S. Septiany, H. H. Handayani, T. Al Mudzakir, and A. F. N. Masruriyah, "Optimasi Metode Support Vector Machine Menggunakan Seleksi Fitur Recursive Feature Elimination dan Forward Selection untuk Klasifikasi Kanker Payudara," *TIN Terap. Inform. Nusant.*, vol. 5, no. 2, pp. 144–154, Jul. 2024, doi: 10.47065/tin.v5i2.5324.

[17] Y. Yin *et al.*, "IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset," *J. Big Data*, vol. 10, no. 1, p. 15, Feb. 2023, doi: 10.1186/s40537-023-00694-8.

[18] A. V. Doroshenko and D. Y. Savchuk, "Research of data mining methods for classification of imbalanced data sets," *Ukr. J. Inf. Technol.*, vol. 6, no. 1, pp. 48–57, 2024, doi: 10.23939/ujit2024.01.048.

[19] B. G. Marcot and A. M. Hanea, "What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?," *Comput. Stat.*, vol. 36, no. 3, pp. 2009–2031, Sep. 2021, doi: 10.1007/s00180-020-00999-9.

[20] Y. Zhang and L. Feng, "Adaptive rank-based tests for high dimensional mean problems," *Stat. Probab. Lett.*, vol. 214, p. 110226, Nov. 2024, doi: 10.1016/j.spl.2024.110226.