Institut Riset dan Publikasi Indonesia (IRPI)
IJATIS: Indonesian Journal of Applied Technology and Innovation Science
Journal Homepage: https://journal.irpi.or.id/index.php/ijatis
Vol. 3 Iss. 1 February 2026, pp: 19-28
ISSN(P): 3032-7466 | ISSN(E): 3032-7474

Prediction of Fetal Health Using Machine Learning Algorithms

Dinda Mustika[1*], Rindiani Suhadi Putri[2], M. Naufal Dzaky Alhady[3], Kharisma Ummi Khairunnisa[4], Arifah Nur Mahmudah[5]

[1,2,3]Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[4,5]Department of Dirasat Islamiyah, Faculty of Syari'ah Islamiyah,
University of Al Azhar, Egypt

E-Mail: [1]12350323047@students.uin-suska.ac.id,
[2]12350323577@students.uin-suska.ac.id, [3]12350310999@students.uin-suska.ac.id,
[4]ummikharisma07@gmail.com, [5]arifahnurgontor@gmail.com

Received Dec 30th 2025; Revised Feb 14th 2026; Accepted Mar 17th 2026; Available Online Mar 18th 2025
Corresponding Author: Dinda Mustika

Abstract

This study evaluates several machine learning algorithms for predicting fetal health conditions using cardiotocography (CTG) data. The dataset contains 2,126 records with 22 numerical features obtained from Kaggle and is classified into three categories: normal, suspect, and pathological. Four classification models Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression were implemented and evaluated using two data split scenarios (80:20 and 70:30). Model performance was assessed using precision, recall, and F1-score. The results show that Random Forest achieves the best performance with an F1-score of 91% in both split scenarios, indicating stable and accurate classification compared with other models. The contribution of this study is to provide a comparative evaluation of classical machine learning algorithms for CTG-based fetal health prediction. The findings can support the development of decision-support tools to help medical personnel detect and monitor fetal health risks early.

Keywords: Classification, Fetal Health, Machine Learning, Prediction

1. INTRODUCTION

Fetal health risk prediction is an effort to estimate whether the fetus is in a healthy, suspicious, or at risk of disorders based on patterns and changes in the physiological condition of the fetus as well as important predictions so that potential disorders can be known early and medical personnel can take quick action before the fetal condition deteriorates [1]. One form of fetal health is fetal growth conditions, which are seen from fetal body size, heart rate, estimated fetal weight, and amniotic fluid, if all the results are according to the gestational age, it means that the fetus is healthy and if disorders such as stunted growth or abnormal fetal position are found, it indicates a risk to the fetus [2].

In this study, fetal health prediction is performed using Cardiotocography (CTG) data, which captures fetal heart rate patterns and related physiological indicators. By applying supervised machine learning to CTG numerical features, the model can support early screening by categorizing cases into normal, suspect, and pathological [3]. These classifications help medical personnel identify potential abnormalities in fetal conditions more efficiently. As a result, healthcare providers can make faster and more informed decisions to ensure appropriate monitoring and medical intervention during pregnancy.

Based on previous research, a comparison was made among Decision Tree, Random Forest, and SVM to determine the most accurate, stable, and appropriate algorithm for analyzing complex fetal health data. Each algorithm has a different advantage: Decision Tree helps you understand the classification rules. Obviously, Random Forest provides more stable results because it uses multiple decision trees, whereas SVM is known to be effective for non-linear data and achieves high accuracy. Because each method offers different strengths, research needs to measure them all to determine the best approach to predicting fetal condition [4].

Several recent studies have utilized machine learning methods to predict fetal health conditions using CTG data. These studies show that algorithms such as Support Vector Machine (SVM), Decision Trees, and Random Forests can achieve fairly good classification results in detecting fetal conditions [25]. Nevertheless, each algorithm has a different performance level depending on the data characteristics and the processing method used. Therefore, further research is needed to comprehensively compare the performance of several machine learning algorithms for classifying fetal health conditions [26].

Based on previous research, there is still a need to conduct a comparative analysis of machine learning algorithm performance to identify the most effective method for classifying CTG data. This study aims to compare the performance of several machine learning algorithms in predicting fetal health conditions based on CTG data. The main contribution of this study is to provide a comparative analysis of the algorithms' performance, which can serve as a reference for the development of decision support systems in the health sector, especially for monitoring fetal health conditions [25].

Early detection of fetal health disorders is essential to reduce the risk of complications during pregnancy and childbirth. Accurate prediction models can assist medical personnel in identifying potential fetal abnormalities earlier, enabling timely intervention and improving clinical decision-making to support better maternal and fetal health outcomes. The contribution of this research lies in the comprehensive comparison of several supervised machine learning algorithms using CTG data to evaluate their effectiveness in fetal health classification. This study provides clearer insights into algorithm performance and contributes to the development of reliable decision support systems for fetal health monitoring.

2. MATERIAL AND METHOD

The elements in the research methodology diagram are designed to organize the fetal health dataset, improve the performance of the prediction model, and reduce bias in the analysis. The overall research methodology for fetal health risk prediction can be seen in Figure 1

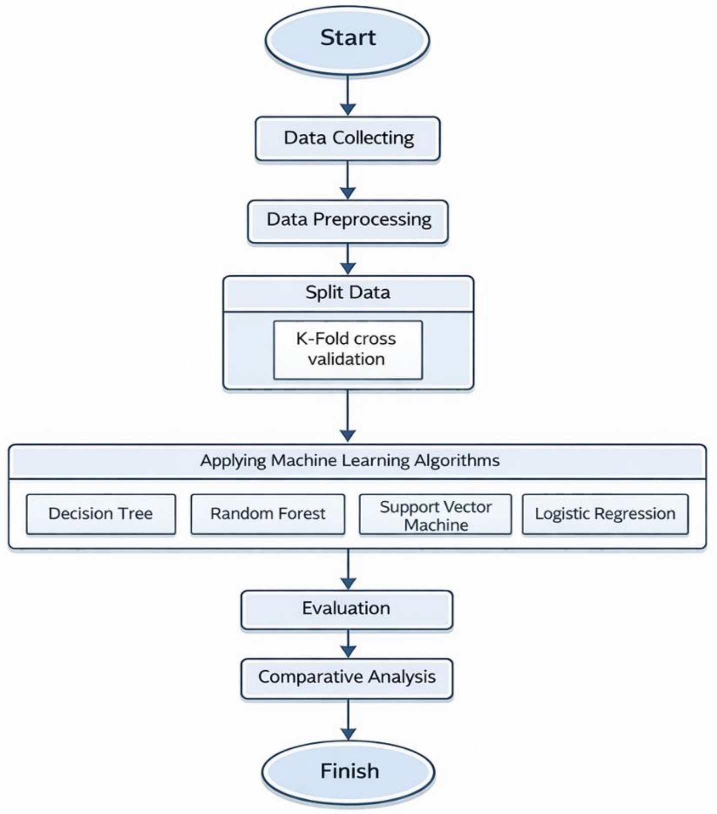

Figure 1. Research Methodology

This study is carried out through several sequential stages. The process begins with a literature review to obtain an overview of previous work on fetal health risk prediction and the application of machine learning to medical data. In the data collection stage, a fetal health dataset is obtained from the chosen source, containing fetal physiological attributes and labels for three risk categories (normal, suspect, and pathological). The dataset is then processed through a preprocessing stage, which includes cleaning incomplete or inconsistent records and normalizing numerical features so that they are ready for analysis. Next, four classification algorithms, Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression, are implemented to construct predictive models of fetal health risk. Finally, in the analysis and results stage, the performance of each model is evaluated using accuracy and other related metrics, and the study concludes by identifying which algorithm yields the best results for the fetal health risk prediction dataset.

2.1. Data Collecting

The data in this study use a fetal health dataset obtained from the Kaggle website (Fetal Health Sets). This dataset consists of 2,126 records with 22 numerical attributes, including baseline fetal heart rate, accelerations, uterine contractions, fetal movements, and several histogram-based features, as well as a target attribute indicating the fetal condition (normal, suspect, or pathological).

Table 1. Data Collecting

Atribut Name	Data Type	Decription
Baseline Value	Numerical	Baseline fetal heart rate during a stable condition.
Abnormal short	Numerical	Number of short-term heart rate decelerations.
Histogram Min	Numerical	Minimum value of the fetal heart rate histogram.
Histogram Max	Numerical	Maximum value of the fetal heart rate histogram.
Histogram Median	Numerical	Median value of the fetal heart rate histogram.

2.2. Data Preprocessing

Data preprocessing is an initial stage that prepares raw data for cleaner, more structured analysis. This process includes removing missing values and duplicate entries, normalizing feature scales, and transforming data, such as converting categorical variables into numerical formats. For image-based data, preprocessing also involves enhancing image quality, reducing noise, adjusting image size, and applying augmentation techniques to increase data variation and prevent overfitting. Additionally, dimensionality reduction is performed to retain only the most relevant features, followed by splitting the data into training and testing sets. These preprocessing steps are essential for improving the accuracy and efficiency of machine learning models, particularly in health-related analyses that require consistent and high-quality data [5].

Table 2. Data Preprocessing

No	Baseline Value	Abnormal short	Histogram Min	Histogram Max	Histogram Median
1	212	212	212	212	212
2	132	46	93	164	138
3	98	17	29	17	14
4	106	12	50	122	77
5	126	32	67	152	129

The preprocessing steps in this study are shown in Table 2. First, the dataset was checked for missing values and duplicate entries; none were found, so no data were removed. Second, to ensure all features contributed equally to the model training and to prevent features with larger scales from dominating, feature scaling was applied using StandardScaler. This method standardizes features by subtracting the mean and scaling to unit variance, which is particularly important for distance-based algorithms such as SVMs. Third, the dataset was split into training and test sets at 80:20 and 70:30 ratios to evaluate the stability and generalization ability of the models under different data availability conditions.

2.3. Fetal Health

Fetal health is a condition during pregnancy that is monitored to prevent disorders leading to fetal death in the womb, and one of its important indicators is fetal movement. If fetal movement decreases, it could indicate problems such as placental blood flow disorders or fetal distress [6]. Fetal health is often characterized by efforts to protect embryonic life as a form of life that has moral value, the application of the principle of non-maleficence in every medical procedure, the provision of the best care for the mother and fetus, fair treatment without discrimination, and respect for parental decisions that are limited by moral responsibility for the safety and future of the fetus [7].

2.4. Decision Tree

The Decision Tree method is used to classify fetal health based on Cardiotocography (CTG) data, with three main attributes, namely Baseline Fetal Heart Rate (LB), Accelerations (AC), and Fetal Movements (FM) as inputs to determine fetal condition into three classes: normal, suspect, and pathological. The model was constructed using 2,126 CTG data from the UCI Machine Learning Repository, resulting in 19 nodes of classification and accuracy of up to 98.7%, so it proved to be effective as a tool to predict fetal health conditions, this approach was developed by applying the Decision Tree method, where several models are trained and then combined using the majority voting technique to determine the final result. This makes the classification more stable and accurate than a single model, thereby improving early detection of fetal health anomalies to support medical decisions [8]. The algorithm works by recursively splitting the data based on the feature that yields the purest subsets, often measured by metrics like Entropy (see equation 1).

$$\text{Entropy}(S) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{1}$$

In this study, the Decision Tree model was configured with max_depth=5 to limit the tree's complexity and prevent overfitting, and random_state=42 to ensure reproducibility of results.

2.5. Random Forest (RF)

RF is used to predict fetal health from Cardiotocography (CTG) data by building many random decision trees, then determining normal, suspect, or pathological classification results through majority voting, resulting in more stable and accurate predictions, this approach works very well because each tree learns different patterns, and then all the results are combined to make more powerful decisions and in this way, RF is able to capture complex variations in CTG signals and reduce the risk of prediction errors that may occur if using only one decision tree [9]. Its decision-making process can be conceptually linked to the Bayesian idea of combining evidence, in equation 2.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \tag{2}$$

The Random Forest model was built using 100 decision trees (n_estimators=100). Other parameters were left at their default settings, which have been shown to work well for many classification tasks.

2.6. Logistic Regression

Logistic regression in the context of fetal health in this study was used to predict the severity of anemia in pregnant women, because anemia conditions in mothers have been proven to have a direct impact on fetal health such as the risk of BBLR, prematurity, and neonatal death. The model used is ordinal logistic regression, because the response variables have sequential levels (no anemia, mild anemia, moderate anemia). The analysis showed that upper arm circumference (LILA), body mass index (BMI), and gravida status significantly affected anemia severity, with LILA as the most dominant factor. This model explains about 29.1% of the data variation and predicts conditions with 80% accuracy, helping health workers detect maternal and fetal health risks early and determine appropriate nutritional interventions [10]. The model uses the sigmoid function to estimate the probability that an instance belongs to a particular class (see equation 3).

$$f(x) = \sigma(w^T x + b) = \frac{1}{1+e^{-(w^T x + b)}} \tag{3}$$

For this multiclass problem, the One-vs-Rest (OvR) scheme was employed. The solver used was Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS), which is efficient for small datasets, with max_iter=1000 to ensure convergence.

2.7. Support Vector Machine (SVM)

Support Vector Machine (SVM) is used in this study to classify fetal health conditions using CTG data that represent fetal physiological signals, such as heart rate accelerations, uterine contractions, and variability-related features. SVM is selected because it can handle complex, high-dimensional, and potentially non-linear patterns that are common in medical datasets. A multi-class strategy, such as One-vs-Rest (One-vs-All), can be applied to separate the three classes: normal, suspect, and pathological [11]. SVMs work by finding the optimal hyperplane that maximizes the margin between classes. For non-linear data, it uses a kernel function to map the data into a higher-dimensional space (equation 4).

$$f(x) = w^T \phi(x) + b \tag{4}$$

Given the non-linear nature of medical data, the SVM model was configured with a Radial Basis Function (RBF) kernel. The regularization parameter C was set to 1.0, and gamma was set to 'scale', which automatically adjusts the kernel coefficient based on the data.

2.8. Confusion Matrix

Model performance is evaluated using a confusion matrix constructed from the predictions on the test data. For the three fetal health categories (normal, suspect, pathological), the confusion matrix summarizes how many instances from each true class are assigned to each predicted class, yielding the basic counts of true positives, true negatives, false positives, and false negatives. These values are then used to compute several evaluation metrics, such as overall accuracy, precision, recall (sensitivity), specificity, and F1-score. The confusion matrix is shown in Figure 2.

n=165	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

Figure 2. Confusion Matrix

Confusion-matrix-based metrics are widely recommended in biomedical and healthcare classification problems because they provide more detailed information about model errors than accuracy alone. Furthermore, recent work on hierarchical confusion matrices shows that both standard and advanced performance measures for binary and multi-class classification still fundamentally rely on the same confusion-matrix structure [10].

2.9. Literature Review

Fetal cardiotocography has become an important source of data for computational approaches to intrapartum risk prediction, and many recent studies have examined how machine learning models can classify fetal status into normal, suspect, and pathological categories. A recent scoping review found that a wide range of supervised models, including support vector machines, random forests, gradient boosting methods, and neural networks, have been applied to cardiotocography data to predict adverse fetal outcomes and fetal hypoxia [11]. Using a well known public cardiotocography dataset, another study compared several algorithms such as artificial neural networks, long short term memory models, extreme gradient boosting, support vector machines, k nearest neighbour, light gradient boosting, and random forests, and reported that multiple models achieved high accuracy for fetal health classification with light gradient boosting performing particularly well across different experimental settings [12]. A related work on fetal health classification from cardiotocographic data also evaluated several machine learning models and found that random forests offered a favourable balance between accuracy, precision, recall, and F1 score, supporting the use of tree-based ensembles for structured fetal monitoring data [13].

A second group of studies places stronger emphasis on ensemble strategies and on handling class imbalance, which is common because pathological and suspect cases are relatively rare compared with normal cases in cardiotocography datasets. One ensemble learning framework applied to a fetal cardiotocography dataset demonstrated that combining several base learners improved diagnostic performance and led to very high accuracy for three-class fetal health classification[14]. Building on this idea, another investigation of imbalanced cardiotocography data evaluated a broad set of models, including decision trees, extra trees, gradient boosting models, k nearest neighbour, light gradient boosting, random forests, support vector machines, artificial neural networks, and deep neural networks, and showed that the use of synthetic minority oversampling and appropriate rescaling substantially increased balanced accuracy and recall for minority classes [22]. These findings highlight that sampling strategies and ensemble methods are effective for improving recognition of high-risk fetal conditions, and that evaluation should rely on confusion-matrix-based metrics, such as class-wise sensitivity, specificity, and F1 score, rather than overall accuracy alone [23].

Overall, prior studies indicate that classical machine learning methods, including tree-based ensembles and margin-based classifiers, perform strongly on cardiotocography datasets for three-class fetal health prediction [5], [17], [20]. Given that CTG data are structured and numerical, this study compares Decision Tree, Random Forest, Support Vector Machine, and Logistic Regression using confusion-matrix-based metrics to provide detailed evaluation across all classes [15], [16].

3. RESULTS AND DISCUSSION

In this section, the performance of four classification methods, Decision Tree, Random Forest, Support Vector Machine (SVM), and Logistic Regression, is compared to predict fetal health conditions into three categories: normal, suspect, and pathological. The evaluation uses precision, recall, F1-score, and accuracy to measure how well each model performs in making correct predictions. The experiments apply two data split scenarios: 80% training and 20% testing as well as 70% training and 30% testing. The 80:20 split allows the models to learn from more data, thereby improving their pattern recognition. Meanwhile, the 70:30 split provides a larger test set to better assess how well the models can generalize to new, unseen data.

Precision shows how accurate the positive predictions are, recall reflects the model’s ability to capture all true positive cases, F1-score combines both precision and recall, while accuracy represents the overall prediction correctness. These results are then used to compare the effectiveness of each method.

3.1. Data Collecting

Data collection in this study is a process that is carried out by taking the Fetal Health Classification dataset from Kaggle, then the data is adjusted to Indonesian conditions and checked by an experienced obstetrician to be really relevant; after the data is collected, everything is cleaned of missing values, abnormal data, and format dissimilarities, so that the final result becomes neat data and ready to be used to train and test the model Predictions of fetal health.

3.2. Decision Tree

The table presents the evaluation results of applying the Decision Tree algorithm, see Table 3.

Table 3. Decision Tree Performance Evaluation

Split Data	Precision	Recall	F1-Score
80:20	88%	85%	86%
70:30	88%	86%	87%

The confusion matrix results indicate that the Decision Tree model consistently classifies fetal health data across both split scenarios. The stable precision value (88% for both splits) indicates that the model maintains similar reliability in its positive predictions, while recall and F1-score slightly improve with the 70:30 split. This performance reflects Decision Tree's ability to form interpretable decision rules from CTG features and to differentiate fetal health conditions (normal, suspect, and pathological) with relatively consistent accuracy.

An analysis of the confusion matrix for the Decision Tree (80:20 split) reveals that most misclassifications occur between the 'suspect' and 'pathological' classes. This is expected, as the physiological distinctions between these two categories can be subtle, making them harder for a single tree to separate perfectly compared to the clearly distinct 'normal' cases.

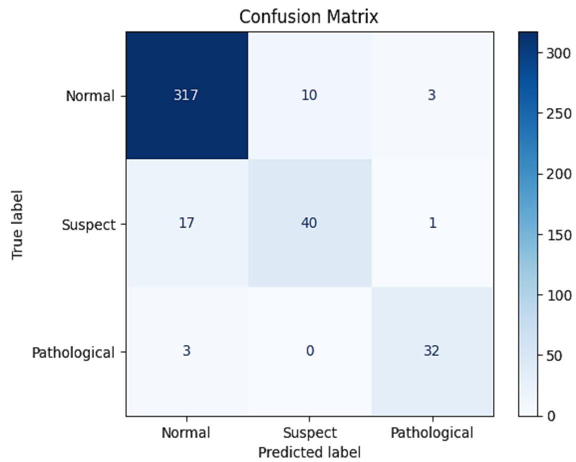

Figure 4. Accuracy Decission Tree

3.3. Random Forest (RF)

Presents the evaluation results of the application of the RF algorithm (Table 4).

Table 4. RF Performance Evaluation

Split Data	Precision	Recall	F1-Score
80:20	95%	88%	91%
70:30	94%	88%	91%

The results of the confusion matrix showed that Random Forest was able to predict 638 normal class data without errors with an accuracy rate of 98.7%. This high performance is in line with the Random Forest concept which combines multiple decision trees through majority voting, so that it is able to recognize complex patterns in CTG (LB, AC, and FM) data stably. This method has been proven to provide more accurate and consistent predictions of fetal health in classifying normal, suspect, and pathological conditions.

The confusion matrix for Random Forest shows a significant reduction in errors compared to the single Decision Tree. Its ensemble nature allows it to correct individual tree biases, leading to higher precision and a more balanced performance across all classes. This is why it achieves the highest F1-score, demonstrating a superior balance between precision and recall (see Figure 5).

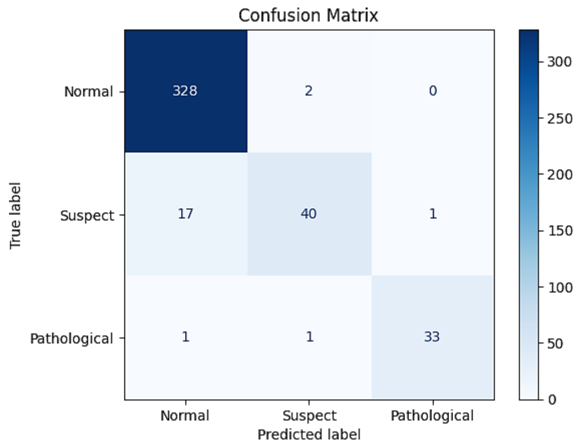

Figure 5. Accuracy Random Forest

3.4. Support Vector Machine (SVM)

Presents the evaluation results of the application of the SVM algorithm (Table 5).

Table 5. SVM Performance Evaluation

Split Data	Precision	Recall	F1-Score
80:20:00	81%	90%	84%
70:30:00	79%	88%	82%

The evaluation results show that SVM achieves the highest recall among the tested models, indicating that it is effective at capturing true cases across fetal health classes. However, the lower precision compared to Random Forest suggests that SVM may generate more false-positive predictions. The confusion matrix helps confirm where these misclassifications occur, especially between classes that share similar CTG characteristics. Overall, SVM is effective at detecting fetal health conditions but may require careful parameter tuning to balance precision and recall.

The confusion matrix for SVM shows that its high recall often comes at the cost of classifying too many instances as 'suspect' or 'pathological', reflected in its lower precision. The drop in performance from the 80:20 to the 70:30 split suggests that SVM's complex decision boundary is more sensitive to the amount of training data available (see Figure 6).

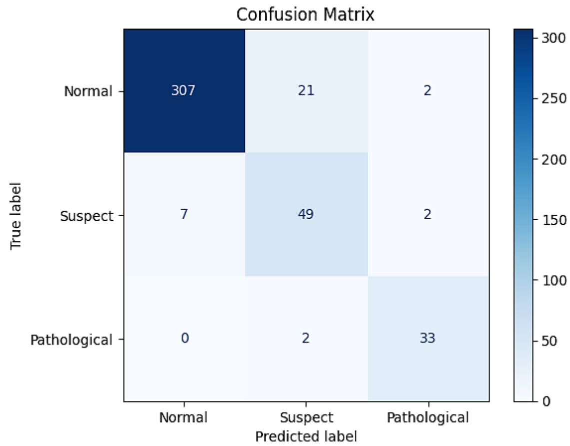

Figure 6. Accuracy SVM

3.5. Logistic Regression

Presents the evaluation results of the application of the Logistic Regression algorithm (Table 6).

Table 6. Logistic Regression Performance Evaluation

Split Data	Precision	Recall	F1-Score
80:20:00	75%	84%	79%
70:30:00	76%	83%	78%

Based on the evaluation results, Logistic Regression produces the lowest overall F1-score among the models, although its performance remains relatively consistent across both split settings. The confusion matrix typically shows more misclassifications for the minority classes (suspect and pathological), indicating that a linear classifier may not fully capture the complex relationships among CTG features. Therefore, Logistic Regression is useful as a baseline model, but it is less optimal than ensemble methods such as Random Forest for fetal health classification.

An examination of the Logistic Regression confusion matrix shows a clear bias towards the majority class ('normal'), with many instances from the 'suspect' and 'pathological' classes being misclassified as 'normal'. This is a common limitation of linear models when faced with non-linearly separable data and imbalanced class distributions (see Figure 7).

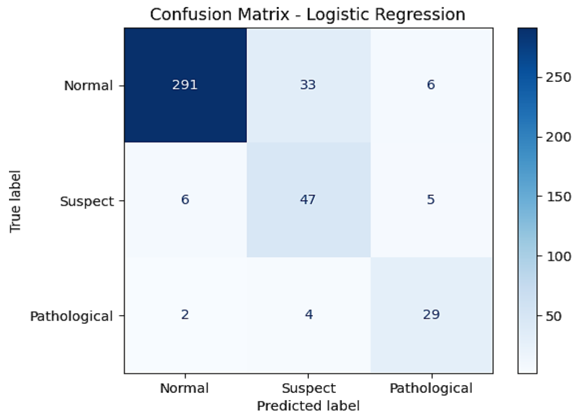

Figure 7. Accuracy Logistic Regression

3.6. Comparative Analysis

A comparative analysis was conducted to compare the performance of the four classification algorithms Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Logistic Regression (LR) under two data-splitting scenarios (80:20 and 70:30). This type of comparison is commonly used to observe how different training–testing proportions affect model generalization and overall performance, can view Figure 8.

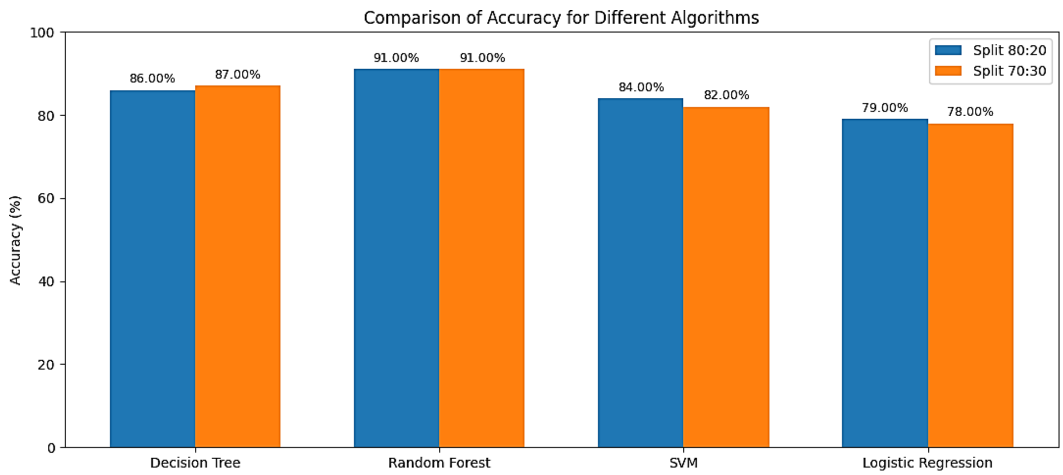

Figure 8. Accuracy Comparative Analysis

Based on the accuracy comparison shown in your figure, Random Forest consistently achieves the highest accuracy in both scenarios, reaching 91% for 80:20 and 91% for 70:30. This indicates that RF is not only strong in predictive performance but also stable across different split settings. In contrast, Decision Tree achieves 86% (80:20) and improves slightly to 87% (70:30), suggesting it performs well but remains below RF. Meanwhile, SVM shows a moderate accuracy of 84% (80:20) and decreases to 82% (70:30), indicating that its performance is more sensitive when the test proportion increases. Finally, Logistic Regression produces the lowest accuracy among all models, at 79% (80:20) and 78% (70:30), suggesting that a linear approach is less capable of capturing the complexity of fetal health patterns in CTG features.

The key difference in performance lies in each model's capacity to handle the complex, non-linear relationships within the CTG data. Random Forest excels by aggregating many decision trees, effectively capturing these complexities while maintaining stability. Logistic Regression, being linear, is too simplistic. SVM, while capable of non-linear classification with its RBF kernel, shows more variance, suggesting it may be harder to tune than the ensemble method. Decision Tree offers a good balance of interpretability and performance but is slightly less accurate than its ensemble counterpart.

4. CONCLUSION

The experimental results show that each machine learning algorithm demonstrates distinct capabilities for classifying fetal health conditions from CTG data. Among the evaluated models, Random Forest achieves the best performance with an F1-score of 91% in both data split scenarios. This indicates that ensemble-based approaches are more effective at capturing complex patterns in CTG features and yield more stable predictions. Decision Tree also produces relatively strong and consistent results, with F1-scores of 86% and 87%. Its structure allows clear classification rules to be generated, making the model easier to interpret. However, its performance is slightly lower than that of Random Forest because it relies on a single tree [25]. Support Vector Machine shows high recall, indicating that the model is effective in detecting fetal health cases. Nevertheless, its precision is lower, suggesting it produces more false positives. Meanwhile, Logistic Regression performs the worst, suggesting that linear models may not fully capture the complex relationships in CTG data.

REFERENCES

[1] V. Mayya, S. K. S, U. Kulkarni, D. K. Surya, and U. R. Acharya, “An empirical study of preprocessing techniques with convolutional neural networks for accurate detection of chronic ocular diseases using fundus images,” Applied Intelligence, vol. 53, no. 2, pp. 1548–1566, Jan. 2023, doi: 10.1007/s10489-022-03490-8.

[2] A. Sivasubramanian, D. Sasidharan, S. V, and V. Ravi, “Efficient Feature Extraction Using Light-Weight CNN Attention-Based Deep Learning Architectures for Ultrasound Fetal Plane Classification,” Oct. 2024, doi: 10.1007/s13246-025-01566-6.

[3] R. Krista and U. Albab, “Pentingnya Pemeriksaan Kehamilan (ANC) di Puskesmas Pasar Rebo: Studi Potong Lintang Deskriptif The Importance of Antenatal Care in Puskesmas Pasar Rebo, a Descriptive Cross Sectional Study.”

[4] Petral, “Labelisasi Otomatis Dan Segmentasi Citra Jantung Janin Menggunakan Deep Learning.”

[5] R. Loa Wanda, “Preprocessing Data Untuk Sistem Peramalan Tingkat Kedisiplinan Mahasiswa.”

[6] M. F. Darkani and N. Khairina, “Klasifikasi Kesehatan Janin Pada Ibu Hamil Menggunakan Metode Support Vector Machine,” Incoding: Journal of Informatics and Computer Science Engineering, vol. 5, no. 2, pp. 160–170, May 2025, doi: 10.34007/incoding.v5i2.830.

[7] S. Elsa Situmeang and N. Putri Savina, “Analisis Perbandingan Metode Decision Tree, Random Forest, dan Support Vector Machine (SVM) dalam Memprediksi Kesehatan Janin”, doi: 10.12962/j27213862..

[8] A. Damayanti and A. Baita, “Comparison of Support Vector Machine (SVM) and Random Forest (RF) Algorithm Performance with Random Undersampling Technique to Predict Gestational Diabetes Mellitus Risk,” 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[9] E. U. Oti, M. O. Olusola, F. C. Eze, and S. U. Enogwe, “Comprehensive Review of K-Means Clustering Algorithms,” International Journal of Advances in Scientific Research and Engineering, vol. 07, no. 08, pp. 64–69, 2021, doi: 10.31695/ijasre.2021.34050.

[10] S. Elsa Situmeang and N. Putri Savina, “Analisis Perbandingan Metode Decision Tree, Random Forest, dan Support Vector Machine (SVM) dalam Memprediksi Kesehatan Janin”, doi: 10.12962/j27213862.

[11] A. Damayanti and A. Baita, “Comparison of Support Vector Machine (SVM) and Random Forest (RF) Algorithm Performance with Random Undersampling Technique to Predict Gestational Diabetes Mellitus Risk,” 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[12] A. Muniasamy and A. Alasiry, “Deep learning: The impact on future eLearning,” International Journal of Emerging Technologies in Learning, vol. 15, no. 1, pp. 188–199, 2020, doi: 10.3991/IJET.V15I01.11435.

[13] E. U. Oti, M. O. Olusola, F. C. Eze, and S. U. Enogwe, “Comprehensive Review of K-Means Clustering Algorithms,” International Journal of Advances in Scientific Research and Engineering, vol. 07, no. 08, pp. 64–69, 2021, doi: 10.31695/ijasre.2021.34050.

[14] A. Sivasubramanian, D. Sasidharan, S. V, and V. Ravi, “Efficient Feature Extraction Using Light-Weight CNN Attention-Based Deep Learning Architectures for Ultrasound Fetal Plane Classification,” Oct. 2024, doi: 10.1007/s13246-025-01566-6.

[15] B. W. Kurniadi, H. Prasetyo, G. L. Ahmad, B. Aditya Wibisono, and D. Sandya Prasvita, “Comparative Analysis of SVM and CNN Algorithms for Fruit Classification.” 2021.

[16] S. Sathyanarayanan and B. R. Tantri, “Confusion Matrix-Based Performance Evaluation Metrics,” African Journal of Biomedical Research, vol. 27, no. 4S, pp. 4023–4031, 2024, doi: 10.53555/AJBR.v27i4S.4345.

[17] K. Riehl, M. Neunteufel, and M. Hemberg, “Hierarchical confusion matrix for classification performance evaluation,” J R Stat Soc Ser C Appl Stat, vol. 72, no. 5, pp. 1394–1412, 2023, doi: 10.1093/jrsssc/qlad057.

[18] F. Francis, S. Luz, H. Wu, S. J. Stock, and R. Townsend, “Machine learning on cardiotocography data to classify fetal outcomes: A scoping review,” Comput Biol Med, vol. 172, p. 108220, 2024, doi: 10.1016/j.compbiomed.2024.108220.

[19] N. Rahmayanti, H. Pradani, M. Pahlawan, and R. Vinarti, “Comparison of machine learning algorithms to classify fetal health using cardiotocogram data,” Procedia Comput Sci, vol. 197, pp. 162–171, 2022, doi: 10.1016/j.procs.2021.12.130.

[20] A. Mehbodniya et al., “Fetal health classification from cardiotocographic data using machine learning,” Expert Syst, vol. 39, no. 6, p. e12899, 2021, doi: 10.1111/exsy.12899.

[21] R. S. Kuzu and Y. Santur, “Early Diagnosis and Classification of Fetal Health Status from a Fetal Cardiotocography Dataset Using Ensemble Learning,” Diagnostics, vol. 13, no. 15, p. 2471, 2023, doi: 10.3390/diagnostics13152471.

[22] I. Nazli, E. Korbeko, S. Dogru, E. Kugu, and O. K. Sahingoz, “Early Detection of Fetal Health Conditions Using Machine Learning for Classifying Imbalanced Cardiotocographic Data,” Diagnostics, vol. 15, no. 10, p. 1250, 2025, doi: 10.3390/diagnostics15101250.

[23] D. Neijzen and G. Lunter, “Unsupervised learning for medical data: A review of probabilistic factorization methods,” Stat Med, vol. 42, no. 30, pp. 5541–5554, 2023, doi: 10.1002/sim.9924.

[24] G. Mushtaq, K. Veningston, and L. Walker, “AI driven interpretable deep learning based fetal health classification,” SLAS Technol, vol. 29, p. 100206, 2024, doi: 10.1016/j.slast.2024.100206.

[25] Hilmi, F., Taqiyassar, K., Romero, N., Pratama, P., & Kusuma, S. C. (2025). Analisis Perbandingan Model Machine Learning Tree-Based Dan Non-Tree-Based U Ntuk Tugas Klasifikasi Comparative Analysis Of Tree-Based And Non-Tre-Based Machine. 12(4).

[26] Komputer, J., No, V., Hal, N., & Alfidyah, M. (2025). Optimasi Algoritma Machine Learning untuk Prediksi Kinerja Sistem Komputer. 1(1), 1–7.