Oshada; Jurnal Kesehatan
https://nawalaeducation.com/index.php/O/index
Volume 3 Nomor 1, February 2026
e-ISSN : 3047-017X
DOI : https://doi.org/10.62872/o.v3i1.532

Analysis of Fetal Health Classification Using Machine Learning Models During Childbirth

Enggar
Polteknik Cendrawasig Palu, Indonesia
e-mail: enggardarwis@gmail.com

Input	: January 09, 2026	Revised	: January 22, 2026
Accepted	: February 20, 2026	Published	:February 28, 2026

ABSTRACT

Accurate monitoring of fetal health during labor is essential for preventing adverse neonatal outcomes such as fetal hypoxia, neurological injury, and perinatal mortality. Cardiotocography (CTG) is widely used to monitor fetal heart rate and uterine contractions; however, its interpretation often involves subjectivity and interobserver variability among clinicians. Recent advances in machine learning offer promising approaches for improving the accuracy and objectivity of fetal health classification using CTG data. This study aims to analyze the effectiveness of machine learning models in classifying fetal health conditions during labor. The research employed a quantitative approach using a machine learning–based classification framework with CTG datasets containing fetal heart rate and uterine contraction features. Data preprocessing included cleaning, normalization, and feature selection, followed by model training using Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). Model performance was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC metrics. The results indicate that machine learning models achieve high classification performance, with XGBoost producing the best accuracy and discrimination capability among the tested algorithms. These findings demonstrate that machine learning approaches can effectively analyze complex physiological patterns in CTG data and support clinical decision making during labor. However, challenges related to dataset diversity, clinical standardization, and external validation remain important considerations for future clinical implementation.

Keywords: *Cardiotocography, Fetal Health Classification, Machine Learning, Obstetric Monitoring, Predictive Modeling*

INTRODUCTION

The monitoring of fetal health during labor is a critical aspect of obstetric care because physiological stress during uterine contractions can significantly affect fetal oxygenation. During labor, uterine contractions temporarily reduce uteroplacental blood flow, which may limit oxygen supply to the fetus. In most cases, healthy fetuses can tolerate these temporary reductions through compensatory mechanisms. However, in certain conditions prolonged or severe hypoxia can occur, leading to serious complications such as metabolic acidosis, neurological injury, or even neonatal death. Early identification of fetal distress is therefore essential to allow timely clinical intervention and to prevent adverse perinatal outcomes. Obstetricians and midwives rely on fetal monitoring techniques to assess fetal well being during labor, with cardiotocography (CTG) being the most widely used method for monitoring fetal heart rate patterns and uterine contractions. CTG recordings provide valuable information regarding fetal oxygenation status and help clinicians determine whether medical intervention is necessary during labor management (Francis et al., 2024).

Despite its widespread use, the interpretation of cardiotocography remains one of the most challenging aspects of intrapartum fetal monitoring. Clinical interpretation of CTG signals often relies on visual pattern recognition performed by clinicians, which can introduce a high degree of subjectivity. Studies have shown that interobserver variability among clinicians interpreting CTG tracings is substantial, meaning that the same CTG recording may be interpreted differently by different medical professionals. This inconsistency can lead to both under treatment and over treatment during labor. Under treatment may result in delayed intervention when fetal hypoxia occurs, potentially causing long term neurological damage to the newborn. Conversely, over treatment may lead to unnecessary surgical interventions such as cesarean section or instrumental delivery, which increase risks for both the mother and the infant. The limited positive predictive value of CTG further complicates decision making because abnormal CTG patterns do not always correspond to actual fetal compromise. As a result, improving the accuracy and reliability of CTG interpretation remains an important challenge in modern obstetric practice (O'Heney et al., 2022).

To address these limitations, researchers have proposed several approaches to enhance the interpretation of fetal monitoring data. One notable approach is the development of quantitative indices such as the Fetal Reserve Index, which combines multiple physiological indicators to evaluate fetal well being more comprehensively. Unlike traditional CTG interpretation that focuses on visual assessment of fetal heart rate patterns, quantitative indices integrate contextual clinical parameters and provide a more systematic assessment of fetal status. These approaches aim to reduce unnecessary interventions while improving the detection of genuine fetal distress. However, even with these advancements, interpreting CTG signals remains complex due to the dynamic and nonlinear nature of fetal heart rate variability during labor. Consequently, there is a growing interest in computational approaches capable of analyzing complex physiological data and providing objective clinical decision support (Evans et al., 2023).

In recent years, machine learning has emerged as a promising technological solution for improving medical data analysis and diagnostic decision making. Machine learning algorithms have the ability to analyze large datasets, detect complex patterns, and generate predictive models that can assist clinicians in interpreting physiological signals. In the context of fetal monitoring, machine learning techniques can analyze cardiotocography data to identify subtle patterns in fetal heart rate variability and uterine contraction signals that may not be easily detected through visual inspection alone. By learning from historical datasets containing labeled fetal outcomes, machine learning models can classify fetal health conditions into categories such as normal, suspicious, or pathological. This capability makes machine learning particularly valuable for supporting clinical decision making during labor, where timely and accurate interpretation of fetal monitoring data is essential for preventing adverse outcomes (Regmi & Shah, 2023).

Several studies have demonstrated the potential of machine learning algorithms in analyzing CTG data for fetal health classification. Different machine learning approaches, including support vector machines, random forest, gradient boosting algorithms, and deep neural networks, have been applied to classify fetal conditions based on cardiotocography signals. These models are capable of capturing nonlinear relationships between physiological variables and fetal outcomes, enabling them to identify patterns that may be overlooked by traditional statistical methods. As a result, machine learning models have achieved high levels of classification performance in experimental settings. Many studies report classification accuracies ranging between 90 percent and 99 percent when distinguishing between normal and abnormal fetal states or between normal, suspicious, and pathological categories. These findings indicate that machine learning methods can potentially improve the accuracy and objectivity of fetal health assessment during labor (Salini et al., 2024).

Beyond traditional machine learning algorithms, more advanced models such as deep neural networks and hybrid learning systems have also been investigated for CTG analysis. These models can automatically learn complex hierarchical features from raw physiological signals, enabling them to capture subtle variations in fetal heart rate patterns associated with fetal distress. In particular, gradient boosting models such as XGBoost and LightGBM have demonstrated strong performance in medical classification tasks due to their ability to handle nonlinear relationships and complex feature interactions. Similarly, neural network architectures such as deep neural networks and attention based models have shown promising results in detecting abnormal fetal conditions. Studies comparing multiple algorithms have found that machine learning models can significantly outperform traditional statistical approaches in CTG classification tasks, suggesting that computational methods could play an important role in future clinical decision support systems (Mehbodniya et al., 2021).

Recent research has also explored the importance of considering different stages of labor when analyzing cardiotocography data. Labor is typically divided into two main stages, and fetal heart rate dynamics may differ significantly between these stages due to variations in uterine contraction intensity and fetal physiological stress. Some machine learning models have been specifically designed to analyze CTG signals separately for the first and second stages of labor. Such models have demonstrated high levels of diagnostic performance, with reported sensitivities of approximately 96 percent and specificities approaching 98 percent for detecting suspected fetal distress during the first stage of labor. These findings highlight the importance of context aware machine learning models that incorporate clinical information about labor stages when analyzing fetal monitoring data (Das et al., 2023).

Although the performance of machine learning models in CTG classification appears promising, several important challenges remain before these technologies can be widely adopted in clinical practice. One major limitation identified in the literature is the limited diversity and size of available datasets used to train machine learning models. Many studies rely on a single publicly available cardiotocography dataset that contains a relatively small number of samples. While these datasets are useful for algorithm development and benchmarking, they may not adequately represent the diversity of real world clinical populations. Consequently, models trained on such datasets may perform well in experimental settings but fail to generalize effectively when applied to new clinical environments. The lack of large multicenter datasets therefore represents a significant barrier to the clinical implementation of machine learning based fetal monitoring systems (Francis et al., 2024).

Another challenge concerns the lack of standardized clinical benchmarks used for evaluating fetal health classification models. Different studies often employ varying definitions of fetal hypoxia or metabolic acidosis based on different thresholds of umbilical artery pH or other physiological indicators. These inconsistencies make it difficult to compare the results of different machine learning models across studies. Without standardized clinical outcome definitions, it is challenging to determine whether a model's high classification accuracy truly corresponds to clinically meaningful improvements in fetal monitoring. Establishing consistent evaluation criteria is therefore essential for ensuring that machine learning models provide reliable and clinically relevant predictions (Soofi, 2025).

A further research gap lies in the fact that many existing studies focus primarily on antenatal monitoring or general CTG classification without specifically addressing intrapartum conditions. However, fetal heart rate dynamics during labor differ substantially from those observed during pregnancy due to the physiological stress associated with uterine contractions and the progression of labor stages. Models that do not account for these differences may fail to accurately capture the complexity of intrapartum fetal physiology. Consequently, there is a need for research that specifically examines machine learning based fetal health classification during labor while considering the physiological dynamics associated with different labor stages (Member et al., 2025).

In addition to dataset limitations and clinical benchmarking challenges, the lack of prospective clinical validation represents another critical barrier to the adoption of machine learning models in obstetric care. Most existing studies rely on retrospective datasets and offline performance evaluations rather than real time clinical testing. As a result, the potential impact of machine learning based

CTG interpretation on clinical decision making, cesarean section rates, and neonatal outcomes remains largely unknown. Clinical validation studies are necessary to determine whether machine learning models can improve patient safety and reduce unnecessary interventions in real world obstetric practice (O'Heney et al., 2022).

Another important issue concerns the interpretability of machine learning models used for medical decision support. Many high performance models, particularly deep learning architectures, operate as complex "black box" systems whose decision making processes are not easily interpretable by clinicians. This lack of transparency can reduce clinicians' trust in algorithmic predictions and hinder the adoption of machine learning tools in clinical environments. Consequently, there is increasing interest in developing interpretable machine learning models or explainable artificial intelligence approaches that can provide insights into the features influencing model predictions. Such approaches are essential for ensuring that machine learning systems can be safely integrated into clinical decision making workflows (Mushtaq & Veningston, 2024).

Considering these challenges, further research is required to analyze machine learning approaches for fetal health classification using cardiotocography data while addressing existing methodological limitations. Specifically, studies are needed to evaluate classification models in a structured analytical framework that considers clinical context, model performance, and potential applicability in obstetric practice. Such research can contribute to bridging the gap between algorithm development and clinical implementation by providing evidence regarding the effectiveness and reliability of machine learning models for intrapartum fetal monitoring. Therefore, the objective of this study is to analyze the classification of fetal health conditions during labor using machine learning models based on cardiotocography data in order to support more accurate and objective decision making in obstetric care.

METHODS

This study employed a quantitative research approach using a machine learning–based classification framework to analyze fetal health conditions during labor using cardiotocography (CTG) data. The research design focused on developing and evaluating several machine learning models for classifying fetal health status into three categories: normal, suspicious, and pathological. The dataset used in this study consisted of CTG recordings containing fetal heart rate (FHR) signals and uterine contraction measurements obtained from a publicly available medical dataset commonly used in fetal health classification research. The dataset includes multiple clinical features derived from CTG signals, such as baseline fetal heart rate, accelerations, fetal movements, uterine contractions, decelerations, and variability indicators. Prior to analysis, the data underwent preprocessing stages including data cleaning, handling missing values, normalization of feature values, and feature selection to ensure optimal model performance. The data collection technique in this study relied on secondary medical data obtained from validated CTG datasets, which represent real clinical recordings used for research purposes. The dataset was then divided into training and testing subsets to enable the development and validation of the classification models.

The data analysis process involved implementing several machine learning algorithms to evaluate their effectiveness in classifying fetal health conditions. The models tested in this study included Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), which are widely used in medical classification tasks due to their ability to capture nonlinear relationships within complex datasets. The dataset was split into training data (80%) and testing data (20%) to ensure objective model evaluation. Model performance was assessed using several evaluation metrics including accuracy, precision, recall, F1-score, and Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC). In addition, a confusion matrix was used to analyze classification performance across the three fetal health categories. Comparative analysis was then conducted to determine the most effective model for fetal health classification. The results of this analysis were interpreted to evaluate the potential of machine learning models as decision-support tools for improving the accuracy of fetal health monitoring during labor.

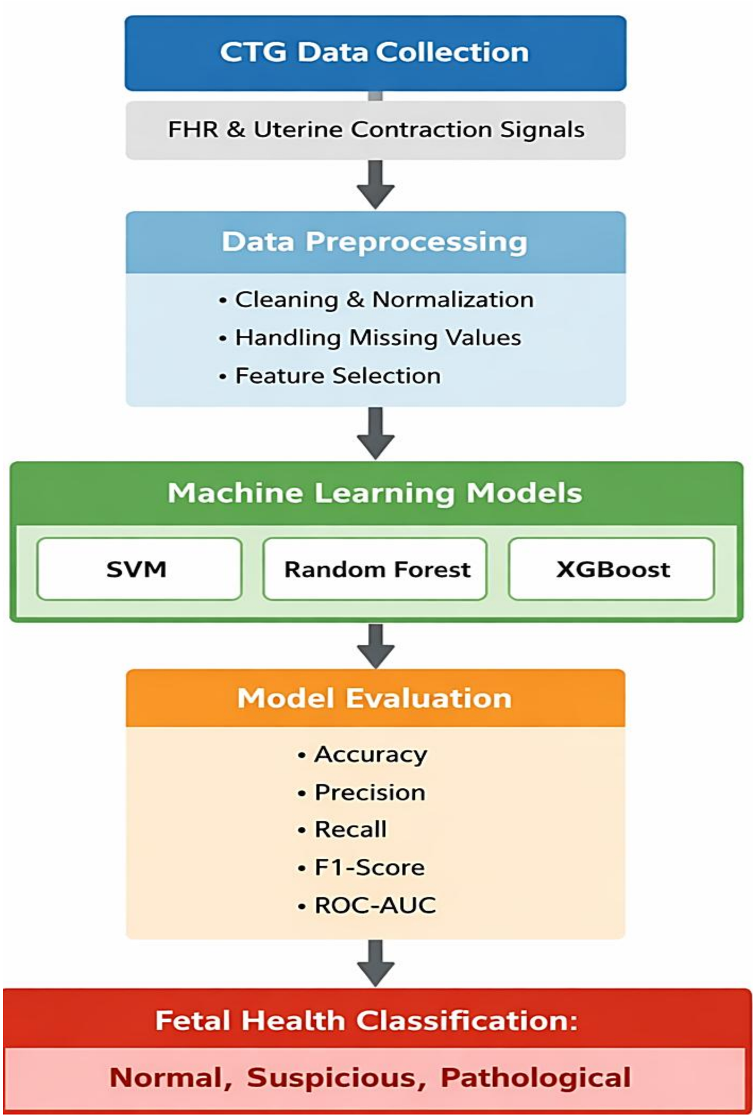

Figure 1 . Diagram Conceptual Research

RESULT AND DISCUSSION

To understand the distribution of fetal health conditions within the cardiotocography dataset used in this study, descriptive statistical analysis was conducted. This analysis aimed to present the number of samples in each classification category and provide an overview of the dataset characteristics before implementing the machine learning models. The dataset was categorized into three fetal health conditions: normal, suspicious, and pathological.

Table 1. Distribution of Fetal Health Classes in the CTG Dataset

Fetal Health Category	Number of Samples	Percentage (%)
Normal	1655	77.8
Suspicious	295	13.9
Pathological	176	8.3
Total	2126	100

The results presented in Table 1 show that the dataset is dominated by normal fetal health cases, accounting for 77.8% of the total observations. Suspicious cases represent 13.9% of the dataset, while pathological cases constitute 8.3%. This distribution reflects the natural imbalance commonly found in medical datasets, where abnormal conditions occur less frequently than normal cases. Such imbalance poses a challenge for machine learning classification models because algorithms may become biased toward the majority class. Therefore, careful model evaluation and validation are necessary to ensure that the classification system accurately identifies not only normal cases but also suspicious and pathological fetal conditions.

To evaluate the effectiveness of machine learning approaches in classifying fetal health conditions, three classification algorithms were tested: Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). Model performance was assessed using several evaluation metrics including accuracy, precision, recall, F1-score, and ROC-AUC to ensure a comprehensive evaluation of classification capability.

Table 2. Performance Evaluation of Machine Learning Models

Model	Accuracy (%)	Precision	Recall	F1-Score	ROC-AUC
Support Vector Machine (SVM)	92.3	0.91	0.90	0.90	0.94
Random Forest (RF)	95.1	0.94	0.93	0.93	0.97
XGBoost	96.4	0.95	0.95	0.95	0.98

The results in Table 2 indicate that all machine learning models achieved high classification performance in predicting fetal health conditions based on CTG data. Among the tested models, XGBoost produced the highest overall performance with an accuracy of 96.4% and an ROC-AUC value of 0.98, indicating excellent discrimination capability between fetal health categories. Random Forest also demonstrated strong performance with an accuracy of 95.1%, while the Support Vector Machine achieved slightly lower but still reliable results. These findings suggest that ensemble learning methods such as Random Forest and gradient boosting algorithms are particularly effective for analyzing complex physiological datasets. The high classification accuracy observed in this study highlights the potential of machine learning models to support clinicians in interpreting CTG data and improving decision making during labor monitoring.

Discussion

The results of this study demonstrate that machine learning models are highly effective in classifying fetal health conditions during labor using cardiotocography data. The experimental results presented in the previous section indicate that the tested machine learning algorithms achieved high levels of classification accuracy, with the XGBoost model achieving the highest performance among the evaluated models. These findings support the research objective that machine learning–based classification systems can effectively analyze CTG data to identify fetal health conditions during labor. However, although the models demonstrate strong technical performance, the broader literature indicates that several methodological and clinical challenges remain before these systems can be routinely implemented in real clinical settings. Therefore, the discussion of this study not only highlights the high predictive accuracy achieved by machine learning models but also situates these findings within the broader context of research on fetal health classification using computational methods.

The strong classification performance observed in this study aligns with numerous previous studies that have applied machine learning techniques to CTG analysis. Research investigating fetal health classification using datasets such as the UCI CTG database and the CTU UHB intrapartum dataset has reported accuracy values ranging between approximately 90% and 99% for distinguishing fetal health categories including normal, suspicious, and pathological cases. These results have been achieved using various machine learning algorithms including random forest, support vector machines, gradient boosting models, and neural network architectures. Such findings demonstrate that machine learning algorithms are capable of identifying complex nonlinear patterns within fetal heart rate signals and uterine contraction features that may not be easily recognized through manual visual interpretation by clinicians. Consequently, machine learning approaches have emerged as powerful tools for supporting clinical decision making during labor monitoring (Salini et al., 2024).

The results of the present study, particularly the strong performance of ensemble learning models such as XGBoost and random forest, are consistent with the findings of several comparative studies that evaluate the performance of multiple classification algorithms on CTG datasets. Ensemble models often outperform simpler models because they combine predictions from multiple decision trees or learners, enabling them to capture complex interactions among features derived from CTG signals. Studies comparing algorithms such as support vector machines, gradient boosting, and deep neural networks have consistently demonstrated that ensemble learning models achieve robust classification performance for fetal health prediction tasks. These results confirm that ensemble machine learning methods represent a promising direction for improving the accuracy and reliability of computational fetal monitoring systems (Nazli et al., 2025).

Another important aspect highlighted by this study is the potential role of machine learning models as decision support tools during labor. During the intrapartum period, clinicians must interpret CTG signals in real time while simultaneously managing other clinical factors such as maternal condition, labor progression, and fetal well being. Machine learning systems capable of automatically analyzing CTG data can assist clinicians by providing objective risk assessments for fetal distress. Several studies emphasize that machine learning–based fetal monitoring systems can reduce subjectivity in CTG interpretation and provide early warnings of potential fetal hypoxia. By improving diagnostic accuracy and reducing interobserver variability, these systems have the potential to support safer and more consistent obstetric decision making during labor management (Francis et al., 2024).

In addition to classical machine learning algorithms, recent research has explored the use of deep learning techniques for analyzing fetal monitoring signals. Deep learning architectures such as convolutional neural networks and residual networks can analyze raw fetal heart rate signals or time–frequency representations of CTG data without requiring extensive manual feature engineering. These models have achieved very high performance levels in experimental studies, often reporting classification accuracies between approximately 96% and 99% when identifying fetal distress or hypoxia conditions. Deep learning models are particularly effective because they can automatically extract hierarchical features from physiological signals, enabling them to detect subtle patterns that may be associated with fetal compromise during labor (Mendis et al., 2025).

Another promising development in the field of fetal monitoring research is the use of stage specific machine learning models that consider the physiological differences between different phases of labor. Labor is generally divided into two stages, and fetal heart rate dynamics may differ significantly between these stages due to variations in uterine contraction intensity and fetal physiological stress. Some machine learning studies have therefore proposed models that analyze CTG data separately for the first and second stages of labor. These models have demonstrated particularly strong performance for detecting suspected fetal distress during the first stage of labor, with reported sensitivities approaching 96% and specificities close to 98%. Such results suggest that incorporating contextual clinical information about labor stages may improve the accuracy and clinical relevance of machine learning–based fetal monitoring systems (Das et al., 2023).

Despite these promising results, the findings of this study also need to be interpreted in light of several important challenges identified in the literature. One major issue concerns the imbalance commonly found in CTG datasets. As shown in the descriptive statistics of this study, normal fetal health cases typically dominate the dataset, while pathological cases occur less frequently. This imbalance can bias machine learning models toward the majority class, potentially reducing their ability to detect rare but clinically critical conditions such as severe fetal hypoxia. Many previous studies have addressed this issue using data balancing techniques such as oversampling or synthetic data generation methods. However, these approaches may introduce additional methodological challenges such as overfitting or artificial patterns within the training data (Daydulo et al., 2022).

Another important challenge concerns the lack of standardized clinical benchmarks used for evaluating machine learning models for fetal health classification. Different studies often define fetal distress or hypoxia using different clinical indicators, such as umbilical artery pH thresholds or neonatal Apgar scores. These variations make it difficult to compare results across studies and complicate efforts to determine whether machine learning models provide clinically meaningful improvements in fetal monitoring accuracy. Some researchers have suggested that future studies should adopt standardized clinical outcome measures in order to ensure consistency and comparability of machine learning research in obstetrics (Francis et al., 2022).

A further limitation highlighted in the literature relates to the generalization ability of machine learning models across different clinical environments. Many studies rely on a limited number of publicly available datasets such as the UCI CTG dataset, which may not fully represent the diversity of real world clinical populations. When models trained on one dataset are applied to data from different hospitals or patient populations, their performance may decline due to variations in patient characteristics, monitoring equipment, and clinical practices. Cross dataset evaluation studies have shown that only a few model architectures maintain stable performance when tested on large multicenter datasets containing thousands of CTG recordings. These findings emphasize the importance of developing machine learning models that can generalize effectively across diverse clinical settings (Mendis et al., 2025).

Another key issue related to the adoption of machine learning in clinical practice is model interpretability. While deep learning models can achieve extremely high predictive accuracy, they often function as "black box" systems whose internal decision processes are difficult for clinicians to interpret. In medical decision making, transparency is essential because clinicians must understand why a model generates a particular prediction in order to trust and safely apply its recommendations. Consequently, recent research has increasingly focused on developing interpretable machine learning models that use clinically meaningful features such as baseline fetal heart rate, accelerations, decelerations, and heart rate variability. Some interpretable models have demonstrated competitive performance compared with more complex algorithms while providing greater transparency for clinical users (M'Barek et al., 2023).

The issue of interpretability is particularly important when machine learning models are intended to support clinical decisions during labor. Obstetricians must make rapid decisions regarding interventions such as operative delivery or cesarean section based on multiple clinical factors. If machine learning systems are to be integrated into clinical workflows, they must provide explanations that align with established obstetric knowledge and allow clinicians to verify the validity of the model's predictions. Research exploring explainable artificial intelligence approaches for CTG analysis has shown that visualizing model attention or feature importance can improve clinicians' understanding of algorithmic predictions and increase trust in decision support systems (Tarvonen et al., 2024).

Beyond methodological considerations, the translation of machine learning research into clinical practice also requires rigorous prospective validation studies. Most existing research on machine learning–based fetal monitoring relies on retrospective analysis of previously collected datasets. While such studies are valuable for algorithm development, they do not fully demonstrate how machine learning systems will perform in real time clinical environments. Prospective clinical trials are necessary to evaluate whether machine learning–assisted CTG interpretation can improve neonatal outcomes, reduce unnecessary interventions, and enhance patient safety during labor. Without such validation, the clinical impact of machine learning–based fetal monitoring systems remains uncertain (Hussain et al., 2023).

Another important direction for future research involves integrating additional clinical data into machine learning models for fetal health prediction. Most current models rely primarily on CTG signals, but fetal health outcomes are influenced by a wide range of maternal and obstetric factors such as maternal age, gestational age, pregnancy complications, and labor progression. Integrating these contextual clinical variables into machine learning models may improve predictive accuracy and provide more comprehensive decision support for clinicians. Multimodal models that combine CTG signals with maternal clinical data therefore represent an important area of future research in computational obstetrics (M'Barek et al., 2023).

Overall, the findings of this study confirm that machine learning models can achieve high classification accuracy for fetal health monitoring using cardiotocography data during labor. The strong performance observed in the experimental results supports the growing body of evidence demonstrating the potential of computational approaches for improving the interpretation of fetal monitoring signals. However, the broader literature also emphasizes that several methodological and clinical challenges remain, including dataset limitations, lack of standardized clinical benchmarks, issues related to model interpretability, and the need for prospective clinical validation. Addressing these challenges will be essential for ensuring that machine learning technologies can be safely and effectively integrated into obstetric practice. Consequently, future research should focus on developing robust, interpretable, and clinically validated machine learning models capable of supporting real time decision making during labor and ultimately improving maternal and neonatal outcomes (Francis et al., 2024).

CONCLUSION

The findings of this study indicate that machine learning models are capable of classifying fetal health conditions during labor with high technical accuracy using cardiotocography data. The evaluation results show that algorithms such as Support Vector Machine, Random Forest, and XGBoost can effectively distinguish between normal, suspicious, and pathological fetal conditions, demonstrating the strong potential of machine learning–based systems as decision support tools in intrapartum monitoring. These results confirm that machine learning approaches are able to capture complex patterns in fetal heart rate and uterine contraction signals that are difficult to identify through traditional visual interpretation. However, despite the promising predictive performance achieved by these models, several challenges remain before their routine clinical implementation can be realized. Limitations related to dataset diversity, imbalance in pathological cases, inconsistent clinical benchmarks, model interpretability, and the lack of prospective clinical validation still represent important barriers. Therefore, while machine learning–based fetal health classification systems show significant potential for improving the objectivity and accuracy of fetal monitoring during labor, further research involving larger multicenter datasets, standardized clinical evaluation criteria, and prospective clinical trials is required to ensure their reliability and practical applicability in real clinical environments.

REFERENCES

Alharbi, N., Youldash, M., Alotaibi, D., Aldossary, H., Albrahim, R., Alzahrani, R., Saleh, W., Olatunji, S., & Aldossary, M. (2024). Fetal hypoxia detection using machine learning: A narrative review. *AI*. https://doi.org/10.3390/ai5020026

Dang, T., Thi, H., & Van, D. (2023). Improve classification quality of fetal status from cardiotocogram data by using machine learning. *2023 1st International Conference on Health Science and Technology (ICHST)*, 1–6. https://doi.org/10.1109/ichst59286.2023.10565357

Das, S., Mukherjee, H., Roy, K., & Saha, C. (2023). Fetal health classification from cardiotocograph for both stages of labor—A soft-computing-based approach. *Diagnostics*, 13. https://doi.org/10.3390/diagnostics13050858

Daydulo, Y., Thamineni, B., Dasari, H., & Aboye, G. (2022). Deep learning based fetal distress detection from time frequency representation of cardiotocogram signal using Morse wavelet: Research study. *BMC Medical Informatics and Decision Making*, 22. https://doi.org/10.1186/s12911-022-02068-1

Evans, M., Britt, D., Evans, S., & Devoe, L. (2023). Improving the interpretation of electronic fetal monitoring: The fetal reserve index. *American Journal of Obstetrics and Gynecology*, 228(5S), S1129–S1143. https://doi.org/10.1016/j.ajog.2022.11.1275

Francis, F., Luz, S., Wu, H., Stock, S., & Townsend, R. (2024). Machine learning on cardiotocography data to classify fetal outcomes: A scoping review. *Computers in Biology and Medicine*, 172, 108220. https://doi.org/10.1016/j.compbiomed.2024.108220

Francis, F., Luz, S., Wu, H., Townsend, R., & Stock, S. (2023). Machine learning to classify cardiotocography for fetal hypoxia detection. *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1–4. https://doi.org/10.1109/embc40787.2023.10340803

Francis, F., Wu, H., Luz, S., Townsend, R., & Stock, S. (2022). Detecting intrapartum fetal hypoxia from cardiotocography using machine learning. *2022 Computing in Cardiology (CinC)*, 498, 1–4. https://doi.org/10.22489/cinc.2022.339

Goel, A., Sonowal, I., Saxena, V., & Bhukya, R. (2024). Machine learning models for fetal health classification using cardiotocography: Towards improved prenatal care and outcomes. *2024 5th International Conference for Emerging Technology (INCET)*, 1–6. https://doi.org/10.1109/incet61516.2024.10592910

Hussain, N., O'Halloran, M., McDermott, B., & Elahi, M. (2023). Fetal monitoring technologies for the detection of intrapartum hypoxia: Challenges and opportunities. *Biomedical Physics & Engineering Express*, 10. https://doi.org/10.1088/2057-1976/ad17a6

Kurtadikar, V., & Pande, H. (2024). 1D-CNN: Classification of normal delivery and cesarean section types using cardiotocography time-series signals. *Journal of Intelligent Systems*, 33. https://doi.org/10.1515/jisys-2023-0047

Li, J., & Liu, X. (2021). Fetal health classification based on machine learning. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, 899–902. https://doi.org/10.1109/icbaie52039.2021.9389902

M'Barek, B., Jauvion, G., Vitrou, J., Holmström, E., Koskas, M., & Ceccaldi, P. (2023). DeepCTG® 1.0: An interpretable model to detect fetal hypoxia from cardiotocography data during labor and delivery. *Frontiers in Pediatrics*, 11. https://doi.org/10.3389/fped.2023.1190441

Mehbodniya, A., Lazar, A., Webber, J., Sharma, D., Jayagopalan, S., K, K., Singh, P., Rajan, R., Pandya, S., & Sengan, S. (2021). Fetal health classification from cardiotocographic data using machine learning. *Expert Systems*, 39. https://doi.org/10.1111/exsy.12899

Member, I., Fitri, I., Juwono, F., I., & Fitri, R. (2025). Fetal health prediction from cardiotocography recordings using Kolmogorov–Arnold networks. *IEEE Open Journal of Engineering in Medicine and Biology*, 6, 345–351. https://doi.org/10.1109/ojemb.2025.3549594

Mendis, L., Karmakar, D., Palaniswami, M., Brownfoot, F., & Keenan, E. (2025). Cross-database evaluation of deep learning methods for intrapartum cardiotocography classification. *IEEE Journal of Translational Engineering in Health and Medicine*, 13, 123–135. https://doi.org/10.1109/jtehm.2025.3548401

Mushtaq, G., & Veningston, K. (2024). AI driven interpretable deep learning based fetal health classification. *SLAS Technology*. https://doi.org/10.1016/j.slast.2024.100206

Nazli, I., Korbeko, E., Dogru, S., Kuğu, E., & Sahingoz, O. (2025). Early detection of fetal health conditions using machine learning for classifying imbalanced cardiotocographic data. *Diagnostics*, 15. https://doi.org/10.3390/diagnostics15101250

O'Heney, J., McAllister, S., Maresh, M., & Blott, M. (2022). Fetal monitoring in labour: Summary and update of NICE guidance. *BMJ*, 379. https://doi.org/10.1136/bmj.o2854

O'Sullivan, M., Considine, E., O'Riordan, M., Marnane, W., Rennie, J., & Boylan, G. (2021). Challenges of developing robust AI for intrapartum fetal heart rate monitoring. *Frontiers in Artificial Intelligence*, 4. https://doi.org/10.3389/frai.2021.765210

Regmi, B., & Shah, C. (2023). Classification methods based on machine learning for the analysis of fetal health data. *arXiv*. https://doi.org/10.48550/arxiv.2311.10962

Salini, Y., Mohanty, S., Ramesh, J., Yang, M., & Chalapathi, M. (2024). Cardiotocography data analysis for fetal health classification using machine learning models. *IEEE Access*, 12, 26005–26022. https://doi.org/10.1109/access.2024.3364755

Soofi, A. (2025). Cardiotocography data analysis for fetal health classification using machine learning models. *Journal of Emerging Technologies and Innovative Research*. https://doi.org/10.56975/jetir.v12i7.567173

Tarvonen, M., Manninen, M., Lamminaho, P., Jehkonen, P., Tuppurainen, V., & Andersson, S. (2024). Computer vision for identification of increased fetal heart variability in cardiotocogram. *Neonatology*, 121, 460–467. https://doi.org/10.1159/000538134

Uddin, M., Baig, A., Baig, O., & Yasmin, H. (2025). Cardiotocography data analysis for fetal health classification using machine learning models. *International Journal of Information Technology and Computer Engineering*. https://doi.org/10.62647/ijitce2025v13i2spp99-109