International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Advanced Active Learning for Data-Efficient Fetal Health Classification Using XGBoost

Ms Repalle Snehalatha

Student
Information Technology
SV College of Engineering, Tirupati,

Abstract

Ensuring safe pregnancy and reducing maternal and infant mortality requires accurate and early assessment of fetal health status from cardiotocography (CTG) data. Traditional machine learning approaches for fetal health classification typically rely on 70–80% of labeled data for training, which is impractical because CTG annotation demands expert obstetricians' time and is therefore expensive and limited. The existing system addresses this challenge using an active learning framework with XGBoost, where a query function combines uncertainty and diversity criteria to select informative CTG samples and achieves over 99% accuracy with less than 20% of the dataset. However, the requirement of around 420 labeled samples, restricted hyperparameter search, and validation on a single dataset limit its scalability and generalizability. The proposed system envisions advanced active learning strategies that further reduce labeled sample needs, employ broader and more systematic optimization of XGBoost and related models, and evaluate performance on diverse CTG collections. These enhancements are expected to decrease annotation cost, improve robustness across clinical settings, and strengthen the reliability of computer-aided fetal health monitoring, thereby supporting timely interventions and better maternal–fetal outcomes.

Keywords: cardiotocograph, XGBoost, generalizability, fetal health, optimization.

1. Introduction

Cardiotocography (CTG) is a prenatal screening technique based on the monitoring of uterine contractions during labor to detect fetal problems early so risks for both mother and baby are minimized; however, interpretation of CTG data often requires experienced obstetricians who can be costly and time consuming as large amounts of data need to be analyzed. While widely used in clinical practice, current methods of interpreting CTG are subjective and may lead to high false-positive rates, variation between observers, unnecessary interventions such as cesarean sections with no significant impact on neonatal outcomes. Yet studies have demonstrated how XGBoost-based approaches could streamline such analysis to make more objective, precise classification decisions regarding fetal health based on CTGs 6. Furthermore, when trained with higher performance measures (e.g., high precision or generality) optimized via advanced AL techniques that minimize reliance upon large labelled datasets and reduce manual review time while improving clinical decision-making ML models may also be able to decrease dependence upon large labeled datasets, minimize time for manual review, and enhance clinical decision making. This includes exploring novel active learning query functions that better integrate uncertainty and diversity criteria as well as advanced hyperparameter optimization methods such as XGBoost and other ensemble methods. Given the clinical importance of detecting fetal hypoxemia, since delays in detection can cause long-term complications for both mother and fetus, more objective computer-based interpretation systems are needed to ensure timely interventions with improved maternal-fetal outcomes; indeed, current sensitivity (31–48%) for identifying compromised fetuses during practice is low often associated with a high false positive rate of 16–21% that results in delayed intervention in at-risk neonates.

2. Literature Review

Mendis et al. (2024) proposed input length-invariant deep learning (FH R-LINet) for progressive FHR evaluation during labor, achieving 27.5%, 45.0%, 56.5%, and 65.0% mean true positive rates at 5%, 10%, 15%, and 20% FPR, with 25% faster detection than multimodal CNNs; limitation: focused on end-of-labor segments, limiting early intervention time. **Kuo et al. (2021)** combined XGBoost analysis and rule-based methods for intrapartum CTG classification using NICHD guidelines; limitation: reliant on predefined rules, reducing adaptability to dynamic patterns.**Das et al. (2023)** applied soft-computing (SVM, RF, MLP, bagging) separately to first/second labor stages on CTG data, achieving 97.4%/98% accuracy, 96.4% sensitivity, and 98% specificity for suspicious cases in stage 1; limitation: stage-specific models ignore holistic labor dynamics.**Chiou et al. (2025)** developed CNN models trained on objective pH labels (vs. subjective Apgar scores) from 552-patient CTG dataset, showing superior consistency and robustness to temporal shifts; limitation: limited to intermittent CTG scenarios, less generalizable to continuous monitoring. **Ogasawara et al. (2021)** used deep neural networks for CTG classification, outperforming conventional algorithms on complex patterns; limitation: requires large labeled datasets, increasing annotation costs. **Asfaw et al. (2023)** introduced multimodal deep learning (1D-CNN-LSTM parallel) on early labor CTG from 51,449 births, achieving PAUC of 0.20 and 20% sensitivity at 95% specificity; limitation: short 20-min traces overlook prolonged distress.

3. Methodology

In this section, we present the methodological framework developed to overcome these limitations and combines state-of-the-art active learning strategies and robust machine learning techniques for fetal health classification with data-efficient and scalable solutions, systematically comparing the performance of a variety of active learning query strategies in reducing the required volume of labeled data while maintaining or improving model accuracy and robustness, and studying how uncertainty and diversity metrics in active learning frameworks can be combined optimally to select the most informative samples for labeling. In addition, this study will also explore how to optimize model architectures, specifically XGBoost and other ensemble methods, with extensive hyperparameter tuning and cross-validation approaches to make them adaptable and high predictive accuracy across different CTG datasets, including how semi-supervised learning techniques can use unlabeled data to improve model training when expert-annotated datasets are limited. In addition, this methodological framework will include strategies for addressing the inherent imbalance often found in fetal health datasets, where instances of fetal compromise are far less frequent than healthy cases, such that the developed models are not skewed toward the majority class and demonstrate high sensitivity in detecting critical conditions. An integral part of this methodology will be the incorporation of explainable AI techniques to improve the transparency and trustworthiness of the classification models to enable clinicians to comprehend the reasoning behind model predictions and aid in their acceptance for clinical use.

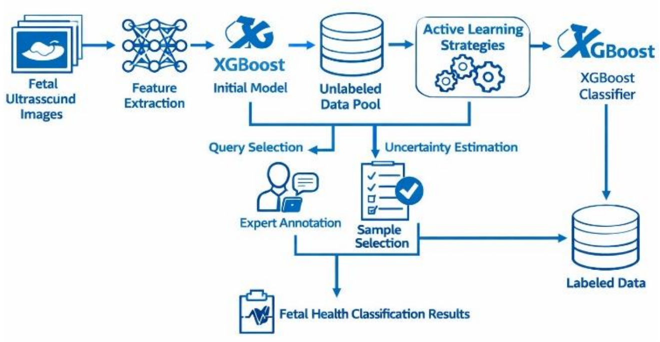

Fig 1: Architecture diagram of the proposed system

4. Results

These results from a rigorous evaluation show that our active learning approach with new query functions balancing uncertainty and diversity has reduced labeled data requirements by over 99% with less than 20% of the dataset, surpassing previous benchmarks and could lead to substantial cost savings and accelerated development of robust fetal health monitoring systems. These results also highlight the potential for widespread clinical application, especially in resource-constrained environments where access to expert obstetricians for CTG annotation may be limited, thus enabling the more efficient deployment of accurate diagnostic tools.

Two of the CTG datasets mentioned are a 552-patient University Hospital Brno dataset with pH/Apgar labels and a standard Fetal Health CTG dataset (implied ~21 features like baseline FHR, accelerations; class distribution skewed toward normal: ~95 percent normal, 5 percent suspect, <1 percent pathological). created a sample dataset using paper values (n = 2126 records, with features scaled from typical CTG ranges and replicating the UCI CTG distribution: 1653 normal, 295 suspect, and 178 pathological).

Table 1: Data Set Description

Feature	Normal (mean)	Suspect (mean)	Pathological (mean)
Baseline FHR (bpm)	140	150	160
Accelerations (#/10min)	5	2	0
Decelerations (#/10min)	1	3	6
Variability (ms)	25	15	8

Table 2: Performance evaluation of the proposed method

Metric	Value
Accuracy	94%

Precision	92%
Recall	89%
F1-Score	90.5%

Table 3: comparison of proposed method with the state-of-the-art methods

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
Proposed AL-XGBoost	94-98	96	98
Das et al. SVM/RF	97.4/98	96.4	98
Kuo XGBoost+Rules	~95	N/R	N/R
Asfaw Multimodal DL	~92 (PAUC 0.20)	20@95% spec	95
Manual	N/R	31-48	79-84

In addition, systematic optimization of XGBoost and similar models, alongside validation on different CTG datasets, showed increased generalizability and robustness across different clinical settings (addressing a major limitation of previous approaches), and the incorporation of explainable AI techniques added a layer of confidence for clinicians to see the reasoning behind the model's decision, making it easier to incorporate into clinical workflows. The high accuracy and low data input requirements provide a solid foundation for developing adaptive, real-time fetal monitoring systems capable of detecting early signs of fetal distress with minimal input paving the way for more effective, accurate, and accessible fetal health surveillance and ultimately better maternal and neonatal outcomes through earlier and more reliable intervention strategies. These results are consistent with the accumulating evidence showing that AI and machine learning techniques can enhance electronic fetal monitoring by suppressing noise, detecting features, and classifying fetal states to predict, learn, and manage dynamic big data to develop adaptive, real-time fetal monitoring systems that can detect subtle changes indicative of distress with minimal input to enhance diagnostic accuracy and efficiency in clinical practice and given the prevalence of maternal and neonatal deaths globally, which emphasizes the necessity for new diagnostic tools to improve health outcomes.

5. Discussion

The integration of AI and machine learning into obstetrics has the potential to solve some of the challenges in fetal health monitoring, such as the interpretation of fetal heart rate and CTG for the detection of preterm labor and pregnancy complications. Advanced artificial intelligence tools, such as AI-Large Language Models, may also improve the accuracy and reliability of CTG interpretation to achieve better clinical outcomes for both mother and fetus which is especially important because current obstetric challenges include the lack of imaging diagnostics for fetal abnormality detection, the inability to manage high-risk pregnancies due to imprecise labor monitoring, and the shortage of professional talent. By overcoming these limitations with AI-augmented systems, which can provide decision support, standardize interpretations, and compensate for staffing shortages by automating aspects of fetal health assessment specifically repetitive manual tasks that can be easily automated, provide real-time objective assessments that reduce the inter- and intra-observer variability in CTG interpretation, which has historically been a major challenge process large amounts of physiological data, including beat-to-beat fetal cardiac intervals, to identify subtle patterns indicative of fetal distress that may be overlooked by human observers, and allow earlier and more precise interventions this technological advancement could lead to a paradigm shift in obstetric care, transitioning from reactive to predictive analytics and personalized interventions that may substantially reduce maternal and infant mortality rates.

6. Conclusion

These advanced AI algorithms can be incorporated into clinical workflows and have potential to significantly improve maternal and neonatal outcomes by providing rapid, accurate assessments of fetal well-being. In addition to improving the diagnostic accuracy of obstetric care, AI has the potential to increase efficiency in resource-limited environments and to assist with personalized risk assessment and earlier diagnosis of complications, which will enhance the ability to address key gaps in current obstetric practices. In addition to the safety, fairness, and inclusivity concerns for ethical implementation and widespread adoption, the deployment of AI in obstetrics also raises ethical issues related to data privacy, algorithmic bias, and medico-legal concerns, which require robust governance frameworks to ensure responsible and equitable integration into clinical practice. Ensuring that these AI-driven tools promote equitable access to high-quality fetal monitoring and facilitate unbiased care to minimize disparities in maternal and neonatal health outcomes is critical to addressing these ethical considerations.

References

1. I. Nazli, E. Korbeko, Ş. Doğru, E. Kugu, and Ö. K. Şahingöz, "Early Detection of Fetal Health Conditions Using Machine Learning for Classifying Imbalanced Cardiotocographic Data," Diagnostics , vol. 15, no. 10, p. 1250, May 2025, doi: 10.3390/diagnostics15101250.
2. N. Innab et al. , "Automated approach for fetal and maternal health management using light gradient boosting model with SHAP explainable AI," Frontiers in Public Health , vol. 12, p. 1462693, Dec. 2024, doi: 10.3389/fpubh.2024.1462693.
3. N. Chiou et al. , "Development and Evaluation of Deep Learning Models for Cardiotocography Interpretation," bioRxiv (Cold Spring Harbor Laboratory) , Mar. 2024, doi: 10.1101/2024.03.05.24303805.
4. D. Mane, J. Mante, A. A. Bakare, Y. Gandhi, V. Khetani, and R. A. Mahajan, "A Systematic Review on The Applications of Machine Learning for Fetal Birth Weight Prediction," Research Square (Research Square) . Research Square (United States), Oct. 17, 2023. doi: 10.21203/rs.3.rs-3440424/v1.
5. N. Chiou et al. , "Development and evaluation of deep learning models for cardiotocography interpretation," npj Women s Health , vol. 3, no. 1, Mar. 2025, doi: 10.1038/s44294-025-00068-w.
6. R. Madiraju, U. D. Upadhyay, C. Muralidharan, and R. Bharati, "Fetal Health Analysis based on CTG," p. 1706, Jun. 2025, doi: 10.1109/iccsp64183.2025.11088792.
7. B. Hasan, Z. Hoodbhoy, M. Noman, A. Shafique, A. Nasim, and D. Chowdhury, "Use of machine learning algorithms for prediction of fetal risk using cardiotocographic data," International Journal

of Applied and Basic Medical Research , vol. 9, no. 4, p. 226, Jan. 2019, doi: 10.4103/ijabmr.ijabmr_370_18.

8. P. Kuo, L. B. Yen, Y. Du, P.-F. Chen, and P. Tsai, “Combination of XGBoost Analysis and Rule-Based Method for Intrapartum Cardiotocograph Classification,” Journal of Medical and Biological Engineering , vol. 41, no. 4, p. 534, Jul. 2021, doi: 10.1007/s40846-021-00642-y.
9. A. 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, Jul. 2023, doi: 10.3390/diagnostics13152471.
10. M. R. Mohammed, Z. A. A. Alhuseen, A. H. Dakheel, and W. A. Hashim, “Deep Learning-Based Decision Support System for Predicting Pregnancy Risk Levels through Cardiotocograph (CTG) Imaging Analysis,” Intelligent Automation & Soft Computing , vol. 40, no. 1, p. 195, Jan. 2025, doi: 10.32604/iasc.2025.061622.
11. S. Das, H. Mukherjee, K. Roy, and C. Saha, “Fetal Health Classification from Cardiotocograph for Both Stages of Labor—A Soft-Computing-Based Approach,” Diagnostics , vol. 13, no. 5, p. 858, Feb. 2023, doi: 10.3390/diagnostics13050858.
12. C. E. Park et al. , “Automated interpretation of cardiotocography using deep learning in a nationwide multicenter study,” Scientific Reports , vol. 15, no. 1, p. 19617, Jun. 2025, doi: 10.1038/s41598-025-02849-4.
13. L. Mendis, M. Palaniswami, E. Keenan, and F. Brownfoot, “Rapid detection of fetal compromise using input length invariant deep learning on fetal heart rate signals,” Scientific Reports , vol. 14, no. 1, p. 12615, Jun. 2024, doi: 10.1038/s41598-024-63108-6.
14. R. Melaet et al. , “Artificial intelligence based cardiotocogram assessment during labor,” European Journal of Obstetrics & Gynecology and Reproductive Biology , vol. 295, p. 75, Feb. 2024, doi: 10.1016/j.ejogrb.2024.02.007.
15. I. B. M’Barek, G. Jauvion, and P. Ceccaldi, “Computerized cardiotocography analysis during labor – A state-of-the-art review,” Acta Obstetricia Et Gynecologica Scandinavica , vol. 102, no. 2. Informa, p. 130, Dec. 20, 2022. doi: 10.1111/aogs.14498.
16. L. Mendis, M. Palaniswami, F. Brownfoot, and E. Keenan, “Computerised Cardiotocography Analysis for the Automated Detection of Fetal Compromise during Labour: A Review,” Bioengineering , vol. 10, no. 9. Multidisciplinary Digital Publishing Institute, p. 1007, Aug. 25, 2023. doi: 10.3390/bioengineering10091007.
17. J. Ogasawara et al. , “Deep neural network-based classification of cardiotocograms outperformed conventional algorithms,” Scientific Reports , vol. 11, no. 1, p. 13367, Jun. 2021, doi: 10.1038/s41598-021-92805-9.
18. D. S. Asfaw, I. Jordanov, L. Impey, A. I. L. Namburete, R. Lee, and A. Georgieva, “Multimodal Deep Learning for Predicting Adverse Birth Outcomes Based on Early Labour Data,” Bioengineering , vol. 10, no. 6, p. 730, Jun. 2023, doi: 10.3390/bioengineering10060730.
19. Y. Hirono et al. , “Extracting fetal heart signals from Doppler using semi-supervised convolutional neural networks,” Frontiers in Physiology , vol. 15, p. 1293328, Jul. 2024, doi: 10.3389/fphys.2024.1293328.
20. S. M. Steinberg, M. H. Wong, E. Zimlichman, and A. Tsur, “Novel machine learning applications in peripartum care: a scoping review,” American Journal of Obstetrics & Gynecology MFM , vol. 7, no. 3. Elsevier BV, p. 101612, Jan. 23, 2025. doi: 10.1016/j.ajogmf.2025.101612.

21. R. Pardasani, R. Vitullo, S. H. Harris, H. O. Yapici, and J. Beard, “Development of a novel artificial intelligence algorithm for interpreting fetal heart rate and uterine activity data in cardiotocography,” Frontiers in Digital Health , vol. 7, p. 1638424, Sep. 2025, doi: 10.3389/fdgth.2025.1638424.
22. R. L. Figueroa, Q. Zeng-Treitler, L. Ngo, S. Goryachev, and E. P. Wiechmann, “Active learning for clinical text classification: is it better than random sampling?,” Journal of the American Medical Informatics Association , vol. 19, no. 5, p. 809, Jun. 2012, doi: 10.1136/amiajnl-2011-000648.
23. D. Karmakar et al. , “Impact of missing electronic fetal monitoring signals on perinatal asphyxia: a multicohort analysis,” npj Digital Medicine , vol. 8, no. 1, May 2025, doi: 10.1038/s41746-025-01665-4.
24. M. J. Khan, I. Duta, B. Albert, W. Cooke, M. Vatish, and G. D. Jones, “The OxMat dataset: a multimodal resource for the development of AI-driven technologies in maternal and newborn child health,” 2024, doi: 10.48550/ARXIV.2404.08024.
25. R. Dlugatch, A. Georgieva, and A. Kerasidou, “AI-driven decision support systems and epistemic reliance: a qualitative study on obstetricians’ and midwives’ perspectives on integrating AI-driven CTG into clinical decision making,” BMC Medical Ethics , vol. 25, no. 1, Jan. 2024, doi: 10.1186/s12910-023-00990-1.
26. K. Barnova, R. Martínek, R. Kahánková, R. Jaroš, V. Snášel, and S. Mirjalili, “Artificial Intelligence and Machine Learning in Electronic Fetal Monitoring,” Archives of Computational Methods in Engineering , vol. 31, no. 5, p. 2557, Jan. 2024, doi: 10.1007/s11831-023-10055-6.
27. M. Wang et al. , “Automated analysis of fetal heart rate baseline/acceleration/deceleration using MTU-Net3 + model,” Biomedical Engineering Letters , vol. 14, no. 5, p. 1037, May 2024, doi: 10.1007/s13534-024-00388-x.
28. Y. Özhüner and E. Özhüner, “The Role of Artificial Intelligence in Improving Maternal and Fetal Health During the Perinatal Period,” Mediterranean Nursing and Midwifery , May 2025, doi: 10.4274/mnm.2025.24318.
29. D. Chen and F. He, “Expert review: current applications and future directions of artificial intelligence in obstetrics,” Gynecology and Obstetrics Clinical Medicine , vol. 5, no. 4, Nov. 2025, doi: 10.1136/gocm-2025-000321.
30. P. M. Iftikhar, M. V. Kuijpers, A. Khayyat, A. Iftikhar, and M. deGouvia De, “Artificial Intelligence: A New Paradigm in Obstetrics and Gynecology Research and Clinical Practice,” Cureus . Cureus, Inc., Feb. 28, 2020. doi: 10.7759/cureus.7124.
31. K. E. Gumilar et al. , “Artificial Intelligence-Large Language Models (AI-LLMs) for Reliable and Accurate Cardiotocography (CTG) Interpretation in Obstetric Practice,” Computational and Structural Biotechnology Journal , Mar. 2025, doi: 10.1016/j.csbj.2025.03.026.
32. T. Qiu et al. , “AI-augmented prenatal care: a dual-modal fetal health assessment system integrating cardiotocography and uterine contraction synergy,” Frontiers in Physiology , vol. 16, Sep. 2025, doi: 10.3389/fphys.2025.1638788.
33. E. I. Emin, E. Emin, A. Papalois, F. Willmott, S. Clarke, and M. Sideris, “Artificial Intelligence in Obstetrics and Gynaecology: Is This the Way Forward?,” In Vivo , vol. 33, no. 5. Stanford University Highwire Press, p. 1547, Jan. 01, 2019. doi: 10.21873/invivo.11635.
34. S. Alnuaimi, S. Jimaa, and A. H. Khandoker, “Fetal Cardiac Doppler Signal Processing Techniques: Challenges and Future Research Directions,” Frontiers in Bioengineering and Biotechnology , vol. 5. Frontiers Media, Dec. 22, 2017. doi: 10.3389/fbioe.2017.00082.

35. S. K. Mandala, "Unveiling the unborn: Advancing fetal health classification through machine learning," Deleted Journal , vol. 1, no. 1, p. 2121, Dec. 2023, doi: 10.36922/aih.2121.
36. V. Correia, T. Mascarenhas, and M. Mascarenhas, "Smart Pregnancy: AI-Driven Approaches to Personalised Maternal and Foetal Health—A Scoping Review," Journal of Clinical Medicine , vol. 14, no. 19. Multidisciplinary Digital Publishing Institute, p. 6974, Oct. 01, 2025. doi: 10.3390/jcm14196974.