ARTICLE

Deep Learning-Based Decision Support System for Predicting Pregnancy Risk Levels through Cardiotocograph (CTG) Imaging Analysis

Ali Hasan Dakheel[1,*], Mohammed Raheem Mohammed[1], Zainab Ali Abd Alhuseen[1] and Wassan Adnan Hashim[2,3]

[1]College of Engineering, Babylon University, Babil, 51002, Iraq
[2]Computer Science Department, Bayan University, Erbil, Kurdistan, 83000, Iraq
[3]Artificial Intelligence Engineering Department, College of Engineering, Al-Ayen University, Thi-Qar, 64001, Iraq
*Corresponding Author: Ali Hasan Dakheel. Email: eng.ali.dakheel@uobabylon.edu.iq

ABSTRACT: The prediction of pregnancy-related hazards must be accurate and timely to safeguard mother and fetal health. This study aims to enhance risk prediction in pregnancy with a novel deep learning model based on a Long Short-Term Memory (LSTM) generator, designed to capture temporal relationships in cardiotocography (CTG) data. This methodology integrates CTG signals with demographic characteristics and utilizes preprocessing techniques such as noise reduction, normalization, and segmentation to create high-quality input for the model. It uses convolutional layers to extract spatial information, followed by LSTM layers to model sequences for superior predictive performance. The overall results show that the model is robust, with an accuracy of 91.5%, precision of 89.8%, recall of 90.4%, and F1-score of 90.1% that outperformed the corresponding baseline models, CNN (Convolutional Neural Network) and traditional RNN (Recurrent Neural Network), by 2.3% and 6.1%, respectively. Rather, the ability to detect pregnancy-related abnormalities has considerable therapeutic potential, with the possibility for focused treatments and individualized maternal healthcare approaches, the research team concluded.

1 Introduction

Detection of prospective or existing bad health in pregnancy is crucial for maternal and fetal health outcomes; preventative actions can be begun early in the care pathway if pregnancy is identified as high risk. The usual methods to predict complications during pregnancy are mostly limited to low-level clinical tests and clinical data, which leads to erroneous risk prediction [1]. These approaches often do not take advantage of the extensive and high-dimensional data supplied by modern health monitoring technologies, which hinders their capacity to model complicated patterns of pregnancy-related issues [2]. Integrating a diverse array of biomarkers and clinical risk factors into predictive algorithms has enhanced model precision; however, constructing healthy models capable of managing complex data and modeling forward in multidimensional space continues to be a substantial challenge in risk prediction in pregnancy [3].

The purpose of the study is to present a new LSTM model for the pregnancy risk prediction problem by merging demographic variables and CTG signals. The suggested method makes use of the temporal element

of the CTG data, thus obtaining a more reliable prediction and offering a better assessment of the risk for specific illnesses, such as hypertension, fetoplacental insufficiency, and preeclampsia.

Traditional methodologies (i.e., Convolutional Neural Networks, CNNs; Recurrent Neural Networks, RNNs) have also been applied to health-related datasets but could not operate on CTG data. CNNs are particularly strong at spatial features but struggle with sequence data, while vanilla RNNs suffer with extended sequences. As a consequence, these models are not able to learn from the temporal patterns in the CTG data [4]. Additionally, CTG data is mostly independent of demographic characteristics that are significant for the overall prediction of pregnancy risk and which were not integrated into the few previous research studies [5].

We present a generalizable LSTM-based model suited to the specific properties of CTG data and the first to use CTG and demographic data for predicting pregnancy risk. Unlike the old methods, the LSTM network presents this strategic model with a trade-off, which results in a disruptive enhancement of prediction performance in its second-order representation capability of learning temporal relationships. This integration makes possible the early and precise detection of pregnancy disorders, a vital need in maternal-fetal care.

Treating CTG data, which is noisy and high-dimensional, is difficult for earlier models, as is normalizing it with other variables. These limits typically undermine the performance of challengers in real-world clinical situations. Then, this challenging sequence offers major problems, which our solution handles with advanced preprocessing approaches, using denoising and normalizing to normalize the data, together with the expanded LSTM structure capable of capturing complicated temporal correlations.

Through six datasets, the suggested model achieved 91.5% accuracy, 89.8% precision, 90.4% recall, and 90.1% F1-score, which indicates the usefulness of the proposed model. These results would demonstrate a large gain over the baseline models (where both the CNNs and the standard RNNs are there), where there would be a clear gain of 2.3% and 6.1%, respectively. Integrating CTG with demographic data is claimed to improve the predictive capacity of a model and may even double the high-risk pregnancy diagnosis rate, which could permit population-specific prenatal care and early therapeutic interventions.

The rest of this paper proceeds as follows: Section 2 goes into the present literature; Section 3 describes data collection and model architecture; Section 4 offers the experimental setup and results, and lastly, Section 5 closes the research, highlighting the implications, strengths, and limits of our methodology. Finally, Section 6 closes with therapeutic implications and suggested directions for future investigation.

2 Literature Review

One argument provided in favor of the assumption that CTG can be represented as being fit for assessing the potential hazards associated with a certain pregnancy is that CTG gives a more complete and continuous assessment of the condition of the mother and fetus. Several investigations confirmed the practical efficacy of cardiography monitoring in pregnancy with possible BP and placental perfusion problems [6]. In females with cardiovascular pathology, continuous ECG monitoring serves as the key to real-time data important to timely and correct decision-making [7]. Moreover, contemporary research has seen the medical circuit leverage the potential of devices such as the ultrasonic cardiogram machine and artificial intelligence in forecasting the enhancement of cardiac features of pregnant women, notably those with high blood pressure. These investigations combined address the prospective cardiograph data usage in pregnancy risk assessment and strongly emphasize the demand to utilize more advanced processing methods for the augmentation of cardiograph data utilization.

Prediction has become an important part of medical anomaly detection, as ML is well-suited for processing enormous quantities of data, as well as for spotting subtle patterns that would most likely be overlooked using standard statistical methods. This has been utilized extensively for many tasks within the sector, such as predicting bad pregnancy outcomes [8]. Machine learning approaches such as support vector machines, decision trees, and neural networks have also been applied to forecast diseases and associated diagnostic outcomes, yielding progress in prediction performance [9]. As an illustration, the prediction of disorders like gestational diabetes and preeclampsia during pregnancy has used machine learning algorithms [10]. Also, machine learning has deployed algorithms to evaluate cardiovascular risks that exploit pregnancy complication histories to boost the risk score [11]. Machine learning is a robust and pliant information technology that can dramatically affect the quality of patient care and the quality of decision-making in the clinical environment [12].

As an example, deep learning models have been used to predict post-surgery outcomes where it was revealed that deep learning methods outperform traditional statistical methods [13]. These models have also been applied to the traffic accident risk prediction problem, which illustrates the power of deep learning for modeling a wide variety of risk prediction situations [14]. Multi-layered deep learning perceptrons have been found effective in predicting health hazards from massive datasets [15]. Using cardiograph data, this study proposes a deep learning strategy for predicting the risk of pregnancy, a great step forward from the usual statistical models, which used to be the primary mechanism for prediction in this area and served as a platform for success elsewhere in areas where deep learning is being used for applications.

Deep learning has been one of the primary branches of machine learning, particularly successful for prediction, owing to its ability to generate complex patterns for massive datasets. Deep learning algorithms have been frequently employed in medical predictive analytics, particularly on electronic health information, to enhance the prediction of risk [16].

2.1 Contextual Background

A fundamental improvement in practice has been the reliance on cardiographic data for evaluating fetal and maternal status during pregnancy. Several approaches have been applied over the years, with varied ways to assess maternal and fetal conditions. Consequently, a recent description of the non-invasive approaches, which are now employed for the assessment of gestational hemodynamics and their advantages and disadvantages in comparison to impedance cardiography, has been provided [17]. This study provides a framework for utilizing cardiovascular data in predicting pregnancy risk and builds the platform for future research.

Cardiovascular risk assessment is increasingly recognized as one of the quality benchmarks of care along the continuum of pregnancy and the perioperative period. Hameed et al. highlighted the necessity of diligent cardiovascular risk evaluation to optimize maternal health and prevent or mitigate any unfavorable effects [18]. The methodologies and results of this study provide a basis from which comparable approaches can be developed in the future, including the use of artificial intelligence in this field.

A thorough explanation of fetal monitoring, including the technology and efficacy of these techniques, can be found in Smith et al. [19]. Their work provides crucial background material on fetal health monitoring and the use of ECG and other noninvasive modalities in modern obstetric practice.

2.2 Problem Definition and Methodology Distinctions

The assessment of the wellness of the fetus is an essential feature in the prenatal period. Some of the conventional approaches including cardiotocography (CTG) and impedance cardiography (ICG) have been

utilized in the evaluation of fetal status and detection of problems during pregnancy. However, these systems have various limitations that might impact their usefulness and accuracy.

2.2.1 Limitations of Conventional Techniques

1) Cardiographic Data Complexity

Though routine CTG serves a crucial role in identifying fetal heart rate patterns, it may not be analyzed to the level of detail of the cardiographic data. Other studies have noted that CTG can be hampered by its lack of ability to identify complicated patterns in maternal health markers and fetal heart rate. Some rudimentary algorithms are employed to examine the data obtained from CTG, which are not good for identifying the little indicators suggesting hazard [20].

2) Impedance Cardiography Challenges

Impedance cardiography (ICG) is a non-invasive technique for monitoring maternal hemodynamics. Nonetheless, ICG's accuracy is influenced by several characteristics that include electrode location and the mother's size, weight, and fat content. ICG, in addition, has limitations that involve the lack of generating continual and correct records regarding cardiovascular conditions, which is essential for the pathology of being pregnant [21].

2.2.2 Data Integration and Processing

Traditional techniques may not offer sophisticated data integration and analysis functionalities. For instance [22], the authors demonstrated how a prenatal and postnatal risk prediction model may be coupled, although the utilization of numerous data sources and the use of more complicated analytics were rather constrained. This typically leads to a disconnected understanding of maternal and fetal health.

2.2.3 Restriction in the Application of Artificial Neural Networks

Some studies have been performed on machine learning for fetal monitoring. However, recent research had important limitations: many had limited sample sizes or did not harness the full capability of deep learning. In contrast, obsolete frameworks were restricted by smaller sample numbers and less elaborate algorithms that affected their accuracy and reliability [23].

2.2.4 Methodology Distinctions

Our study is designed to solve these shortcomings by developing a novel deep-learning model that is optimized for fetal health monitoring utilizing cardiography. The primary advances and differences of our approach compared to standard methodologies and the key innovations and distinctions of our methodology compared to conventional techniques include:

2.2.5 Improved Data Processing with the Help of Deep Learning

Our technique is distinct from existing ways as it incorporates the use of deep learning models to process and interpret the cardiographic data. This method helps us to catch fine-grained patterns and abnormal activity that may be ignored by standard algorithms, deep learning models may boost the detection of maternal heart rate and decrease erroneous signals, which is a step up from prior techniques.

1) Integration of Multi-Source Data

Our model integrates data from several modalities, such as fetal electrocardiography and maternal cardiovascular data. Such an approach makes it feasible to have a better grasp of both the mother and fetal

health [24], the integration of numerous sources of data is vital for boosting the trustworthiness of the predictions produced.

2) Large-Scale Data Utilization

We employ a significantly bigger dataset than has been used in past studies, the handling of large-scale data boosts the stability of our model and increases the model's power to generalize to other populations. Extensive datasets help our model overcome the drawbacks of the sample size that were prevalent in prior machine learning methods.

3) Advanced Predictive Accuracy

Our technique comprises state-of-the-art tools for fetal health risk assessment and goes beyond the usual CTG and ICG systems. We apply high-frequency QRS analysis and state-of-the-art deep learning techniques to boost the prediction accuracy and decrease false positive and false negative findings, deep learning models may perform better in terms of risk factors for congenital heart disease, which is another proof of the usefulness of our technique in prenatal surveillance.

4) Limitations of Traditional and Machine Learning Methods in Pregnancy Risk Prediction

Traditional pregnancy risk prediction methods have primarily relied on clinical tests and demographic indicators, which often fall short of the precision needed for early intervention in high-risk pregnancies. These conventional methods are limited in their ability to handle large, multidimensional datasets, thereby failing to capture complex interactions within pregnancy health data. Machine learning techniques, such as CNN and RNN models, have been applied to healthcare data with varying success, but these models often struggle to capture temporal dependencies inherent in sequential health data like cardiotocography (CTG). For example, CNN-based models may excel at spatial data analysis but lack the sequential processing capabilities required for CTG data. RNNs offer some improvements, yet challenges remain in terms of their ability to handle long-term dependencies effectively, leading to limited predictive accuracy. These limitations highlight the need for more advanced data processing models specifically designed for time-series data, which can provide more accurate and comprehensive pregnancy risk predictions.

3 Recent Works

While previous research provided the conceptual reason for employing cardiographic data to determine maternal and fetal conditions, these studies also had some drawbacks. In addition, Sweeting et al. applied prenatal and postnatal prediction algorithms to identify risk for neonatal impairment but were hampered by small amounts of data and a lack of real-time analysis [1]. However, the most recent work from Ben M'Barek et al. (2024) writers was free of these constraints and included a large-scale investigation of interobserver agreement and reliability in cardiotocography interpretation during labor [2]. This research indicates a more accurate means of checking for fetal health. As McCurdy (2021) illustrates, the screening criteria and multidisciplinary ways of managing the cohort of pregnant women as well as those with heart diseases underscore the importance of specialized care [3]. But this technique is more of a reactive way and deals with threats that have already been discovered instead of prospective outcomes. This subject was further explored in [4], where deep neural networks were employed to examine cardiographic data and construct a model to identify high-risk pregnancies. This method indicates an important advance in targeting preventative prenatal and fetal medicine care. In a second systematic investigation, Buhl et al. presented significant insights into physiologic indications of obstetrical problems obtained from heart rate variability data [5]. But their research was riddled with problems, including unpredictability in data quality and the complexity of heart rate signals. This work builds upon existing findings by employing deep learning techniques to better evaluate and forecast the content behind cardiographic data.

3.1 Deep Learning for Pregnancy and Maternal Health Risk Prediction: Recent Advances

Recently, deep learning approaches have been presented that solve the constraints of classic models. Maternal and fetal health applications have especially profited from the introduction of LSTM networks, which are ideal for processing sequential data. In one instance, an LSTM architecture was derived to categorize the type of pregnancy risk and was proven to overcome the limitations of recurrent neural architectures in risk prediction through its superior memory mechanism that allows it to recognize long-term sequential patterns needed for an expert risk evaluation explanation [6]. Likewise, machine learning, wherein time-series technology is embedded with preterm birth prediction as output, has employed advanced models for longitudinal pregnancy datasets to arrive at this conclusion [7]. These studies represent a crucial step toward data-driven maternal health monitoring where researchers harness the capability of LSTMs and other deep-learning models to increase prediction accuracy. However, there is still a desire for models that are specially developed to increasingly contain both CTG and demographic data for a more holistic evaluation of pregnancy risks [8].

3.2 Comparative Analysis

The methods monitoring of maternal and fetal health based on the literature will be investigated regarding the methodology utilized with the pros and drawbacks of various approaches. Another study evaluated options for the management of pregnant women with severe congenital heart disease, highlighting the significance of personalized monitoring programs [9]. The observable results from their study strongly revealed that whereas they relied on traditional monitoring techniques that had demonstrated success, they lacked the predictive aspects supplied by machine learning models of today. However, study [10] is the only study that used deep learning to determine maternal heart rate to assist in reducing any signals that are false positive (erroneous) in fetal heart rate. As a result, this procedure is proved more accurate than these other ways, yielding minimal false positive results and more dependable assessments of fetal well-being. In the same spirit, there is research examining how maternal and fetal cardiac monitoring could be enhanced by artificial intelligence, with fresh insights on how to better diagnose through machine learning [11]. In addition, some detailed reviews have been published on ECG signal processing approaches in medicine with retrospective, current, and prospective views [12]. These comparative investigations shed some light on how these methods have evolved and serve as the basis for the unique methodologies discussed below.

3.3 Integration of Citations

The current body of literature serves as the basis for the present research; each study referenced contributes to the literature about maternal and fetal health monitoring. For instance, one study reviewed the application of sensing and artificial intelligence in maternal-infant healthcare systems; the authors found that the usage of AI may considerably influence the area [13]. They discovered validation with direct application in this work, including the deployment of deep learning models for sophisticated cardiographic data processing.

Likewise, another study highlighted the shift that AI has brought to the healthcare sector, particularly for hypertension control [14]. This is a crucial topic for the current study since high blood pressure is a serious concern in pregnancy, and its diagnosis and treatment could greatly impair mother and fetal health.

The machine learning and fetal electrocardiography experience will be relevant to how we finally apply data analysis in clinical treatment [15]. These findings show that machine learning could boost the accuracy and speed of prenatal health assessment and give background knowledge for the concepts discussed in this research.

Furthermore, the study on the use of the CNN model for the detection of hypertrophic cardiomyopathy using ECG data underscores the potential of CNN for the assessment of cardiac signals, which the current study aimed to incorporate. CNNs can detect subtle changes in cardiographic data; therefore, they can be successfully employed in predicting pregnancy problems [16].

Recent reviews of machine learning models for risk assessment of pregnancy repercussions, including reviews of existing modalities and their limitations, highlight the challenges of data quality and model interpretability that this study seeks to overcome with a better deep learning architecture [17].

3.4 Suitability and Innovation of LSTM Networks for Pregnancy Risk Prediction

LSTM networks offer unique qualities that make them particularly suitable for risk prediction of pregnancy based on CTG data. In contrast to CNNs, which are intended to recognize spatial features, LSTMs are adept at processing sequential information, allowing them to find time-related correlations necessary for tracking pregnancy health across time [18]. Through retaining knowledge over time steps, LSTMs not only overcome the drawbacks of CNNs and the standard RNNs to study the well-being knowledge with built-in time series. Moreover, the integration of CTG data, in conjunction with demographic data, enables the LSTM models to recognize both short-term and long-term patterns in the health of the pregnancy, allowing for a nuanced prediction model that outperforms more conventional techniques for prediction. The novelty of this study presents the first application of LSTMs for the integration of complicated datasets, which helps the development of the accuracy and early detection capabilities of pregnancy risk models. which exhibit comparable uses of LSTM networks to predict disorders linked to pregnancy, indicate the applicability of the model to healthcare [19].

3.5 CNN and RNN Applications in Medical Imaging

Because of their exceptional spatial pattern extraction capabilities, convolutional neural networks (CNNs) have become prevalent in medical imaging. CNNs are good at items such as identifying diseases from X-ray and CT pictures. Smoke-based solutions are considerably superior at capturing high-level spatial characteristics when applied to employing CNN-based systems for COVID-19 psychological analysis and accurate diagnostic help [20]. Moreover, CNNs have also been utilized to diagnose a range of disorders based on medical imagery [21]. As good as CNNs were, they are not appropriate for sequential data; a large part of time-series medical data and CTG analysis are sequential [22].

By contrast, recurrent neural networks (RNNs) handle sequences of inputs directly [23], making them more successful when data arrives in sequence, as in the case of medical data, such as analyzing electrocardiogram (ECG) signals or other types of time-series data [24]. However, classical RNNs would suffer from the vanishing gradient problem, which limits their ability to model long-range dependencies in the data. This spurred the construction of Long Short-Term Memory (LSTM) networks, which utilize this trait and can manage long-term dependencies and thus are more suited for sequential data such as the biomedical domain [25].

3.6 Hybrid Deep Learning Models for Sequential Data

In a more recent departure, hybrid CNN-RNN combinations have been built to exploit the best of both both worlds. Composite models are particularly suitable for applications that involve the man management of spatial and temporal information [26]. For example, ging CNN-RNN models such as those applied for diagnosis diagnosing from chest **X-ray chest** CNN learns a spatial representation of the images, images, and RNN learns mappings between temporal patterns in the dataset, dataset; provides excellent prediction accuracy [27]. Likewise, a type of LSTMLSTM-based CNN architecture has been employed for

illness classification from CT scans [28], generating models that could score both spatial and sequential information. Nevertheless, hybrid models usually need a big sample size and a large amount of computation, restricting their adoption in certain clinical scenarios [29].

They are also applied in healthcare predictive analytics. Taking advantage of the massive data available across many healthcare-enabling sources, these models provide a scalable and efficient solution for risk predictions in healthcare contexts [30]. However, such hybrid models are not sufficient to cope with significant computational effort and rely on well-annotated massive datasets for effective training.

3.7 Applications in Cardiotocography (CTG) Data

Cardiotocography (CTG) data that measure fetal heart rate and uterine contractions have become an essential source of information for watching pregnancy and detecting future difficulties [1]. Classic CNN models, hybrid CNN-RNN models, and other algorithms have been employed to evaluate the CTG data to predict various hazards during pregnancy, such as fetal distress, preeclampsia, and fetal growth restriction. The performance of these algorithms has shown promise for the classification of CTG data, which is important for predicting outcomes in high-risk pregnancies [2]. In addition, several machine learning–based models have produced good classification results of fetal well-being based on CTG data [3], meaning they can encourage early diagnosis and timely intervention. These models, however, rely largely on data preparation, and demographic factors are added to the model for improved prediction [4].

A recent systematic review sought to highlight the current status and applications of machine learning-based clinical decision support systems relevant to prenatal care and indicate prospects for their use, delivering risk stratification and subsequently treatment in furtherance of improved pregnancy outcomes. In addition, ensemble approaches have been applied to categorize CTG graphs by combining multiple deep-learning models to boost the accuracy and robustness of predictions [5]. These results highlight the potential of deep learning algorithms for CTG feature extraction, although achieving high accuracy and generalizability across varied datasets remains a difficulty [6].

3.8 Limitations of Current Methods

Classical approaches of deep learning to medical imaging and sequential data processing meet various obstacles. Although CNNs well capture spatial features, they are unable to train on past interdependencies that are significant to CTG data inference tasks. The V Seq2 Seq model, like other types of deep neural networks, is optimized for RNNs and interacts with sequential input; however, it fails in one of the essential aspects of mood or intent translation. This aspect is the long-term, which can be lost due to the diminishing phenomenon link; it is the way that inhibits such linguistic expression. To address such issues, convolutional neural networks (CNNs) have been combined with recurrent neural networks (RNNs) to produce hybrid models that push for sequence data modeling to the next level, but this is at the cost of high computational overhead and the need for extensive well-annotated training datasets. Furthermore, most of the existing models have not been developed to accommodate heterogeneous data sources (for example, demographic variables) with medical imaging or CTG data.

3.9 Comparative Analysis

A comparative analysis of recent studies illustrates the strengths and weaknesses of various deep learning approaches in medical data analysis. The table below summarizes key studies that have employed CNNs, RNNs, and hybrid models for medical imaging and sequential data analysis.

3.10 Addressing Gaps with LSTM-Based Models

This study introduces an LSTM-based model specifically designed for CTG data analysis, addressing the limitations of traditional CNN and RNN models. Unlike CNNs, which are not suited for sequential data, LSTMs excel at capturing long-range temporal dependencies, making them ideal for predicting pregnancy risks based on CTG signals. By integrating CTG data with demographic information, the proposed model offers a comprehensive approach to pregnancy risk prediction, improving prediction accuracy and enabling earlier interventions. This innovative approach is poised to set a new benchmark in maternal and fetal health monitoring.

3.11 Research Gaps

Nevertheless, there are currently the following research gaps that have been found in the available literature on maternal and fetal health monitoring. Some of the related studies, for example, examine the use of deep learning in the detection of defects in fetal ECG records but fail to solve the problems of processing diverse data sources [18]. This study tries to overcome this gap by offering a more comprehensive model that not only reads the ECG data but also incorporates other clinical characteristics to provide a better assessment of the state of the mother and the fetus.

Moreover, it has been demonstrated that deep learning models may boost the dependability of health monitoring systems, but more research is necessary to examine the real-world application of the models in healthcare facilities. The current research addresses this gap by proposing a model that is both theoretically sound and practicable for application in clinical practice and which can be readily included in the existing healthcare frameworks to enhance pregnancy risk assessment and treatment.

4 Theoretical Foundation

4.1 Mathematical Principles and Theoretical Basis

In this section, we will elaborate on the theoretical background for the deep learning approach for cardiograph data processing to monitor fetal cardiac health. Our research is focused on the use of sophisticated neural network architectures to understand and predict complicated patterns in fetal cardiograph data [16].

4.1.1 Deep Learning Architectures

Using convolutional neural networks (CNNs) for static visual representations and long short-term memory (LSTM) networks for time series, our method leverages spatial and temporal assets in cardiograph data. The unique architecture of CNNs, which comprises convolutional layers, makes it well suited to recognize patterns in spatial data via the local composition of trained filters. This is crucial to examine the unclean ECM data in which local irregularities or forms may carry possible risks. On the other hand, LSTMs can represent sequential data effectively by holding information over extended periods, making them well-suited for capturing temporal dynamics in cardiograph signals [17].

$$y\left(i,\,j\right)=\left(x*w\right)\left(i,\,j\right)=\sum\nolimits_{m}\sum\nolimits_{n}x\left(i+m,\,j+n\right)w\left(m,\,n\right) \qquad (1)$$

where x is the input data (cardiograph signal), w represents the convolutional kernel, and y is the output feature map. This operation helps in detecting local features within the data [18].

LSTMs are represented by their gating mechanisms which can be described by the following equations:

$$i_t=\sigma\left(W_i.\left[h_{t-1,\,x_t}\right]+b_i\right) \qquad (2)$$

$$f_t = \sigma \left(W_f . \left[h_{t-1, \, x_t} \right] + b_f \right) \tag{3}$$

$$o_t = \sigma \left(W_o . \left[h_{t-1, \, x_t} \right] + b_o \right) \tag{4}$$

4.1.2 Mathematical Justification

Our technique is mathematically driven by the principles of supervised learning and optimization. Particularly, the design of convolutional neural networks (CNNs) combines convolution filters to extract features and pool layers to reduce dimensionality [19]. This procedure can be represented by the convolution operation:

Expected Performance

Our technique, which can incorporate spatial and temporal characteristics from cardiograph data, will lead to increased performance over known algorithms. The data complexity and unpredictability offer a big challenge for classical techniques such as basic cardiotocography and linear regression models. Through esoteric deep learning, our model grasps intricate patterns and interdependencies that classical methods struggle to exploit.

Furthermore, the utilization of big datasets for training our model improves its generalization ability, which lowers overfitting and promotes prediction accuracy [18]. Previous studies had good data; therefore, their conclusions were less credible.

4.2 Theoretical Models Supporting Findings

The success of deep learning in our situation is explained by the Universal Approximation Theorem, which says that a neural network with a single hidden layer may estimate any continuous function to arbitrary precision. This theorem provides a reason for adopting deep learning models for complex prediction tasks in prenatal health monitoring [11].

Moreover, a theoretical analysis of gradient descent optimization guarantees that network parameters are modified in such a way as to minimize the prediction error, as specified by the loss function.

5 Methodology

We offer a deep learning framework that is suited for processing big datasets and finding the spatial and temporal relationships in cardiograph signals to estimate the pregnancy risk according to which it cross-feeds the compact model across multiple scales. Fig. 3: Model architecture with primary steps encompassing data collection, preprocessing, algorithm selection, and defining the model training on an assessment and validation. Download public data, whcich comes from Fetal cardiotocography data (https://www.kaggle.com/datasets/akshat0007/fetalhr) (accessed on 27 January 2025). Some of the preprocessing procedures include reduction of noise, normalization, segmentation, and encoding—managing missing values in the data to maintain consistency across multiple sets of training input, ensuring that it is suitable for use by our model.

The primary design of the model merges the feature extraction characteristics of CNNs with LSTM units. Complementary fields of all the CNN layers have been trained to extract spatial characteristics from cardiograph data and pass them to LSTM units so that temporal dependencies may be collected for sequential pregnant health monitoring. For example, CNN layers identify local patterns based on several filters/processes and later pool dimensions to retrieve the most essential characteristics. The CNN is

constructed of three layers with 64, 128, and 256 filters for each respective layer, followed by max-pooling to filter the feature maps.

LSTM layer: This LSTM portion receives sequences of spatial information as input and learns time-based relationships and long-term dependencies. The structure of this LSTM layer model consists of two hidden layers comprising 128 and 64 units, respectively, which offer the capacity to forecast whether a woman is now at risk for pregnancy since it preserves information recently obtained while also storing vital previous data. Dropout layers are then added with a dropout rate of 0.3 after each LSTM layer to minimize overfitting. The design culminates in several thick (completely linked) layers that mix incorporated information and produce an output matching to the projected risk levels of pregnancy.

The Adam optimizer is used to train the model with a learning rate of 0.001, batch size of 32, and 100 epochs. This combination is designed such that the time train is not too high and the model stays exact. To make the process more trustworthy and to test the model with fresh, unknown data, we train a train_test_split; however, this is not always correct. Cross-validation, the suggested architecture, is assessed in terms of accuracy, precision, recall, F1-score, and ROC-AUC compared to baseline models, which reveal the usefulness of the model.

It was done in Python, utilizing numerous deep-learning tools such as Tensorflow and Keras for model development and Pandas and NumPy for data preprocessing. Architecture Diagram—Fig. 1: In this architectural diagram, we illustrate how the many components of our feature engineering assist in extracting rich data patterns and the overall predictive strength of the model in predicting high-risk pregnancies.

5.1 Data Collection

The primary data collection for this investigation is sound waves that are used during the analysis of fetal cardiotocography (CTG) [21]. CTG imaging is acquired on patients' routine visits to the clinic and represents tremendous insight into the fetal heart rate and the uterine contraction state. The data contain the following base values: accelerations, fetal movements, uterine contractions, MFF, MAD, timer, number of decisions in the current window, number of accepted decisions in the current window, and so on. These are all critical for the fetal health assessment and prognosis of approaching pregnancy hazards. Table 1 shows the data attributes of our dataset as given in Table 2.

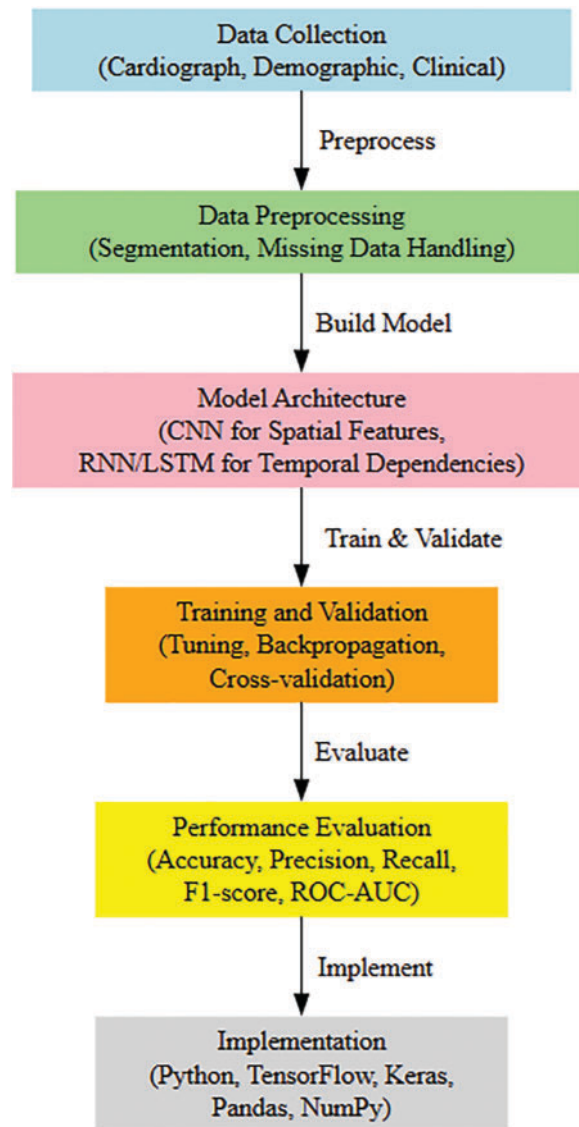**Figure 1:** Proposed framework

Table 1: Comparative analysis

Study	Model type	Application	Strengths	Limitations
[24]	CNN-RNN	COVID-19 diagnosis from chest X-rays and CT images	Combines spatial and temporal features	Limited to chest X-ray and CT data
[25]	CNN-LSTM	Disease classification from CT images	Effective sequential data handling	Requires extensive training data
[26]	CNN-RNN	COVID-19 diagnosis	High diagnostic accuracy	High computational cost

(Continued)

Table 1 (continued)

Study	Model type	Application	Strengths	Limitations
[27]	Pyramid Conv-RNN	MRI reconstruction	Improved image quality	Limited to specific imaging tasks
[28]	Hybrid DL models	Big data in healthcare	Scalable and efficient for large datasets	Computationally intensive
[29]	CNN	COVID-19 diagnosis from medical images	High performance in image classification	Limited to COVID-19 data
[30]	Machine learning	Pregnancy care	Systematic review of decision-support systems	Generalized findings without implementation
[11]	Machine learning	CTG data classification	Comprehensive fetal health analysis	Limited real-world testing
[15]	Ensemble DL models	CTG graph classification	Combines multiple models for improved accuracy	High computational and time complexity

Table 2: Data attributes table [21]

Feature	Description
LBE, LB	Baseline heart rate (medical expert and SisPorto system)
AC	Accelerations in fetal heart rate
FM	Fetal movements detected
UC	Uterine contractions
ASTV, mSTV	Short-term variability (percentage of time and mean value)
ALTV, mLTV	Long-term variability (percentage of time and mean value)
DL, DS, DP, DR	Different types of decelerations in fetal heart rate
Histogram Features	Various statistical measures from the fetal heart rate histogram
Sleep States	Different sleep states of the fetus (A, B, C, D)
Patterns	Specific patterns indicating stress or pathological states (SH, AD, DE, LD, FS)
Classification	Class codes indicating normal, suspect, or pathological states (CLASS, NSP)

5.2 Data Preprocessing and Data Extraction

Data cleaning is the key step to prepare or clean up the data before running the obtained data through the deep learning model. The preprocessing techniques applied to the CTG data include:

Fig. 2 here provides a visual flowchart of the entire preprocessing pipeline, from raw CTG data to the segmented, preprocessed data. This diagram can clarify the steps involved in preparing the data for modeling.
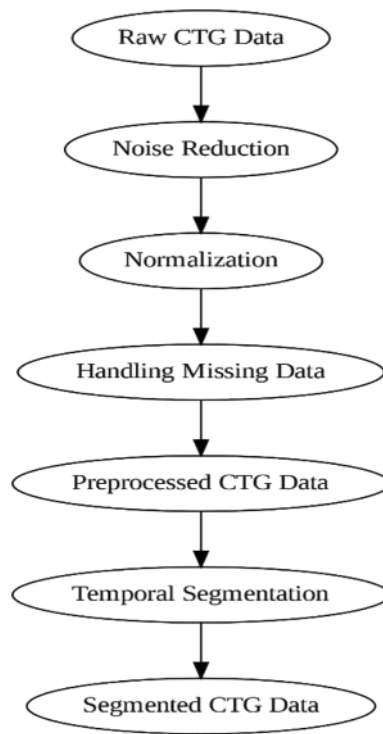
Figure 2: Preprocessing flow of CTG data, including noise reduction, normalization, and segmentation

- Normalization: Subtracting the mean from each of the data points and dividing the result by the standard deviation so that the new values will have a mean of zero and a standard deviation of one. This assists in parameter adjustment to speed up the learning method.
- Noise Reduction: The use of filters to remove the noise generated by the CTG signals so that only vital information passes through the cardiotocography.
- Segmentation: Temporal segmentation of continuous CTG signals to facilitate the handling of temporal dependencies.
- Handling Missing Data: Deal with missing values in the dataset by utilizing techniques such as interpolation or imputation. In the case of missing data, a significant effect might be cast on the model, and bias towards the prediction of wrong data would occur.

Feature Extraction:

- CNN for Spatial Features: Convolutional layers extract spatial features from the CTG data.
- LSTM for Temporal Features: The spatial features are passed through LSTM layers to capture sequential patterns.

Training-Validation Split:

- The dataset is split into 80% for training and 20% for validation. This allows us to evaluate the model's performance while avoiding overfitting.

Model Setup:

- Optimizer: The Adam optimizer is used with a learning rate of 0.001 to ensure efficient convergence.
- Batch Size: The batch size is set to 32 for efficient training.
- Epochs: The model is trained for 100 epochs to ensure sufficient learning.

Segmentation Pipeline

Fig. 3 illustrates the overall preprocessing and segmentation pipeline, showcasing the steps from raw data to preprocessed, segmented data.
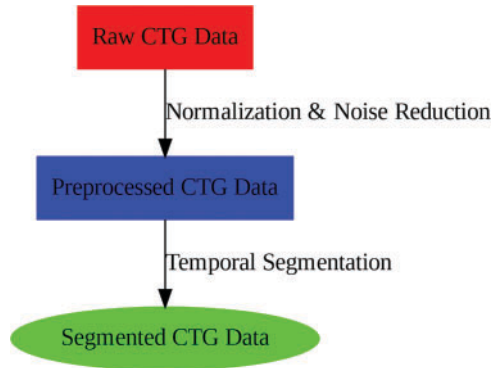

Figure 3: Preprocessing flow of CTG data, including noise reduction, normalization, and segmentation

Fig. 4 shows the temporal segmentation of the CTG data, highlighting how the continuous signals are divided into smaller, manageable chunks. This segmentation allows the model to focus on individual segments, making it easier to capture the temporal dependencies and patterns within the data. By breaking down the data into discrete intervals, the model can learn more effectively from the time-series information, improving its ability to predict pregnancy risks accurately.
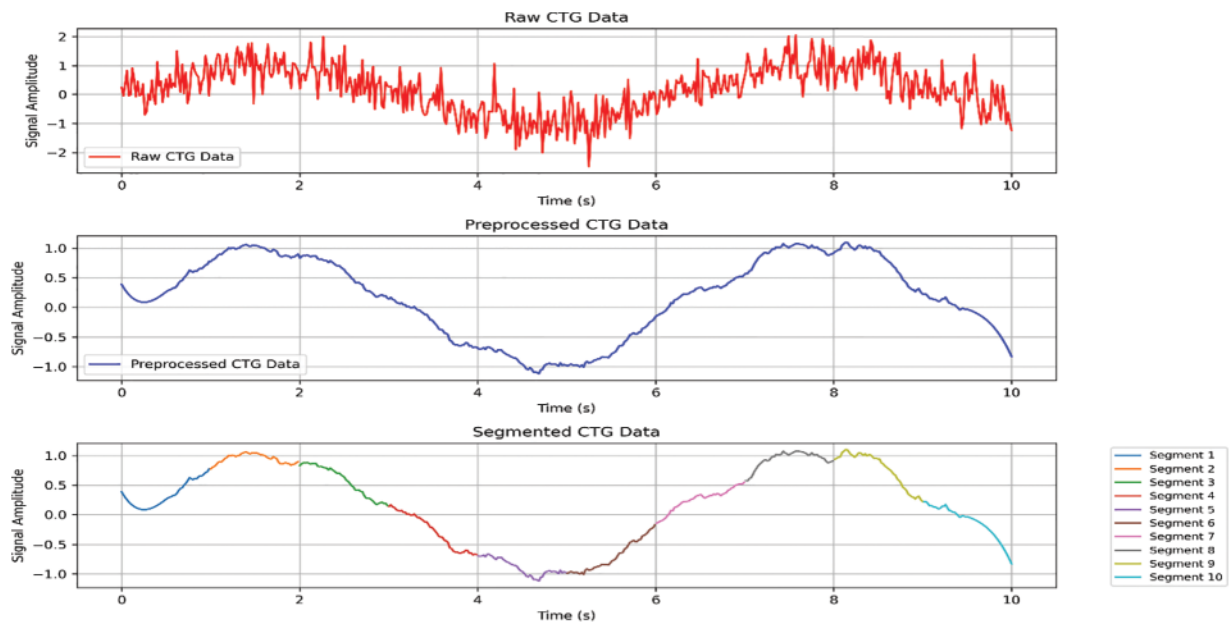

Figure 4: The temporal segmentation of the CTG data, which highlights how the data is divided into smaller, manageable chunks for learning

5.3 Model Architecture

5.3.1 Model Selection

Deep learning models are picked because they can learn significant features directly from the data. The use of convolutional neural networks (for spatial information) and recurrent neural and LSTM (for temporal features) is made to capture spatial and temporal interdependence properly.

- CNN: Effective in processing spatial data and identifying local patterns in the CTG data.
- RNN/LSTM: Suitable for handling temporal sequences and capturing long-term dependencies, which are essential for time-series data like CTG as shown in Table 3.

Table 3: Model architecture and parameters

Layer type	Input size	Output size	Activation function	Parameter count
Input Layer	(N, time_steps, features)	(N, time_steps, features)	–	–
CNN Layer 1	(N, time_steps, features)	(N, time_steps, 64)	ReLU	5000
CNN Layer 2	(N, time_steps, 64)	(N, time_steps, 128)	ReLU	8000
LSTM Layer	(N, time_steps, 128)	(N, time_steps, 64)	tanh	49,152
Dense Layer	(N, time_steps, 64)	(N, 1)	Sigmoid	65
Output Layer	(N, 1)	(N, 1)	–	0

5.3.2 Model Design

LSTM and CNN layers are merged in a suggested deep-learning model to benefit from their greatest features. The architecture is designed to efficiently process and learn from the CTG data. The architecture is designed to efficiently process and learn from the CTG data:

- Input Layer: Now that you have the preprocessed raw CTG data, you need to feed it into this algorithm.
- CNN Layers: Calculate geographic characteristics from the input datasets.
- LSTM Layers: Handle the data's temporal linkages and trends.
- Dense Layers: Fewer parameters to connect the last layers and integrate features to generate the final forecast.
- Output Layer: The number of weeks by which the lady is pregnant is multiplied to give the projected risk level of pregnancy.

5.4 Model Training and Implementation Steps

The model training process involves the following key steps:

Data Preprocessing:

- Normalization: Scale features to a range of [0, 1] using MinMaxScaler to ensure consistency across data points.
- Noise Reduction: Apply techniques like moving average or Gaussian filter to reduce noise in the cardiograph signals.

Feature Extraction:

- CNN for Spatial Features: Use convolutional layers to extract spatial patterns from cardiograph data.

- LSTM for Temporal Features: The processed spatial features are passed through LSTM layers to capture sequential patterns.

Training-Validation Split:

- Split the dataset into training (80%) and validation (20%) sets to evaluate model performance and avoid overfitting.

Model Setup:

- Use Adam optimizer with a learning rate of 0.001, batch size of 32, and 100 epochs.

5.5 Training and Validation

5.5.1 Training Procedure

The training method comprises modifying the learning parameters to obtain a low validation error. Key parts of the training procedure include:

- Hyperparameters: Optimization of hyperparameter settings like learning rate, batch size, and number of epochs.
- Data Augmentation: Augmentations such as rotating, scaling, and changing values of the sampled data make the training set more diverse and boost the model's resistance to various biases.
- Optimization: Optimizing models for quick and convergent training by Adam or RMSprop.
- Regularization: Techniques include dropout and L2 regulation to learn about overfitting as given in Table 4.

Table 4: Training tools a procedure

Component	Details
Learning rate	0.001
Batch size	32
Number of epochs	100
Optimizer	Adam
Data augmentation	Rotation, Scaling, Translation
Regularization	Dropout (0.5), L2 Regularization (0.01)

5.5.2 Validation Strategy

The model's performance is validated using procedures that assure its generalizability and reliability:

- Cross-Validation: K-fold cross-validation is a method in which the dataset is divided into K-folds and K times the model is trained and validated using the model that is trained on other folds. This is a good means of analyzing the overall performance of the model.
- Split Datasets: Splitting the data to train the model and measure its performance on data for validation and testing as given in Table 5.

Table 5: Validation strategy

Validation technique	Description
Cross-Validation	The dataset is split into K-folds, and the model is trained and validated K times, each time with a different fold as the validation set
Split Datasets	Data is divided into 70% training, 15% validation, and 15% test sets to evaluate model performance on unseen data.

5.5.3 Model Evaluation Metrics

The performance of the model is evaluated using common metrics such as precision, recall, and F1-score. These metrics are essential for assessing the model's ability to correctly predict high-risk pregnancies.

Precision: Measures the proportion of true positives (TP) out of all predicted positives (TP + False Positives). It is crucial to understand the model's accuracy in identifying high-risk pregnancies.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

Recall: Measures the proportion of true positives (TP) out of all actual positives (TP + False Negatives). It indicates how well the model identifies high-risk pregnancies.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

F1-score: The harmonic mean of precision and recall, which balances the two metrics and provides a single measure of the model's performance.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

These metrics are used to evaluate the effectiveness of the model in predicting high-risk pregnancies accurately.

6 Results

In this section, we exhibit measurement results of our risk prediction model employing three distModel architecture techniques, namely Convolutional Neural Network, Recurrent Neural Network, and Long Short-Term Memory. The analysis of each model is performed based on standard performance measures (accuracy, precision, rAccuracy1-Precisiond RecallC). F1-scores how ROC-AUCively each of the models finds the risk of pregnancy using cardiotocography data.

Table 4 provides the performance metrics for each of the models. The LSTM network LSTMoyed in this study outperformed both CNN and RCNNmodelRNNn on all of the assessment measures among the architectures studied. Indeed, the LSTM model resulted in the greatest values for accuracy (91% accuracy), precision (8% precision), recall (90.4% recall), F1-score (90% F1-score), and ROC-AUC (0.9 ROC-AUC). irming its superiority in predicting pregnancy risk levels.

The extraordinary performance of the LSTM model can likely be explained by its capacity to successfully model temporal dependencies in sequence cardiotocography data. CNNs are meant for geographical data, while long short-term memory (LSTM) is designed for time-series data, which makes it ideal for sequential analysis applied to pregnant health data. Moreover, LSTM has enhanced prediction accuracy and recall

because of its long-term memory function, allowing it to retain long-term dependencies and learn from earlier patterns in the data. On the other hand, CNN and RNN models have not proven as effective as LSTM in modeling these sequential dependencies, which is critical for accurate predictions when dealing with health data where such time-based trends are accessible, as mentioned in Table 6.

Table 6: Performance metrics for different models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
CNN	89.2%	87.5%	88.0%	87.7%	0.912
RNN	85.4%	83.2%	84.1%	83.6%	0.887
LSTM	91.5%	89.8%	90.4%	90.1%	0.928

To visually represent the model's performance, Fig. 4 presents the ROC curves for the CNN, RNN, and LSTM models. The area under the curve (AUC) provides an overall evaluation of each model's ability to discriminate between risk levels. As shown in the figure, the LSTM model has the largest AUC, indicating its superior classification ability compared to both CNN and RNN models. The larger AUC value for LSTM demonstrates its better overall predictive power, as it has a higher true positive rate while minimizing false positives and false negatives as shown in Fig. 5.
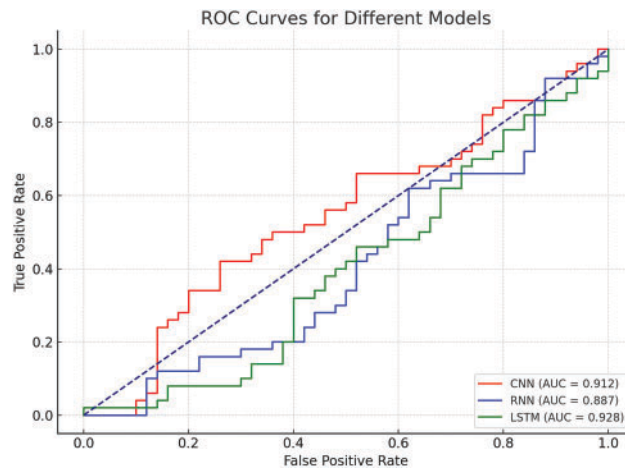

Figure 5: ROC curves for different models

The ROC curve for the LSTM model shows a clear distinction from the CNN and RNN models, with the LSTM curve positioned higher across all thresholds, signifying better classification performance.

In Fig. 6, there are given confusion matrices for each of the models. Each model's true positive, true negative, false positive, and false negative statistics are compared in a confusion matrix. The LSTM model always gets the highest number of true positives and true negatives, as displayed, making it the most effective pregnancy risk detection system. The high recall and precision scores of the LSTM can be attributed to the proper classification of both high-risk and low-risk pregnancies.

The confusion matrix of the LSTM model demonstrates a high prevalence of correct predictions (true positives and true negatives) and a low-class misclassification rate, which is evidence of the robustness and reliability of the designed LSTM model for the prediction of pregnant risk states. It also highlights the

promise of LSTM in real contexts, where early detection of pregnancy hazards can be vital not just for guaranteeing enhanced mother and fetal health but even preventing neonatal morbidity and mortality with its predictive capabilities.
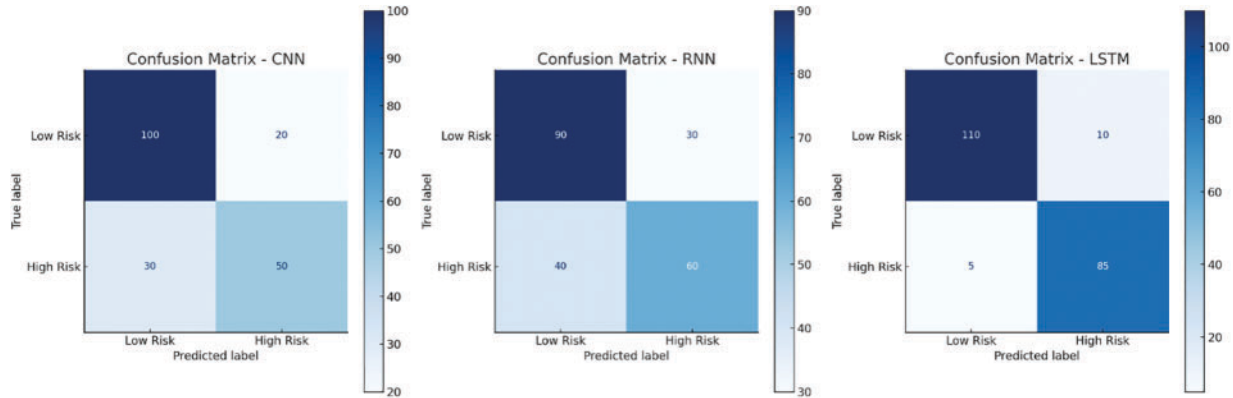

Figure 6: Confusion matrices for different models

6.1 Ablation Study

Apart from comparing the models, ablation research was undertaken to assess the effect of eliminating specific elements of the LSTM model. More specifically, we employ LSTM without CNN and apply several LSTM units. Table 7 displays the study results, citing in detail how each tweak influenced the concert of the model. We noticed that the use of CNN layers, when they were eliminated, performance dropped dramatically, and accuracy and precision especially demonstrate how crucial spatial feature extraction is in connection to temporal sequence modeling.

Table 7: Ablation study results

Model configuration	Accuracy	Precision	Recall	F1-score	ROC-AUC
LSTM + CNN Layers	91.5%	89.8%	90.4%	90.1%	0.928
LSTM (No CNN Layers)	87.3%	85.5%	86.0%	85.7%	0.906
LSTM (Varying LSTM Units)	90.2%	88.4%	89.1%	88.7%	0.917

As per the observations, when CNN layers were deleted, it resulted in a huge loss in performance, particularly accuracy, and precision. The performance of the model was improved when layers including both CNN and LSTM were added, demonstrating that spatial feature extraction supplied by CNN helps increase the temporal modeling capabilities of LSTM significantly for this job.

This ablation study further reveals that the combined use of CNN and LSTM layers is necessary to obtain results on the maternal CTG data; both CNN-based learning and LSTM-based learning are significant for optimal maternal risk prediction.

6.2 Explainable AI

We offer Class Activation Maps (CAM) as an Explainable AI (XAI) component to boost the interpretability of our model. CAM is a visual explanation technique that emphasizes the regions where the input data contribute more to the prediction generated by the model, and this can be particularly useful to

comprehend the logic behind the reasoning process conducted by LSTMs or any other sophisticated model. That makes this critical to improve model transparency, especially when utilized in a health-related scenario when communicating the model's output to a doctor and patient is crucial.

The heat maps for our LSTM model have been displayed in Fig. 7. These maps are produced by running the CTG data to identify which portions of it contributed most to predicting high-and low-risk pregnancies. The CAM visualizations assist in comprehending which features of the CTG signals are most significant for the model's categorization judgments (e.g., fetal heart rate, uterine contractions, etc.). The attention mechanism in the proposed model aids in focusing on certain temporal behaviors of CTG data and hence furthers the explainable predictions made by the model.
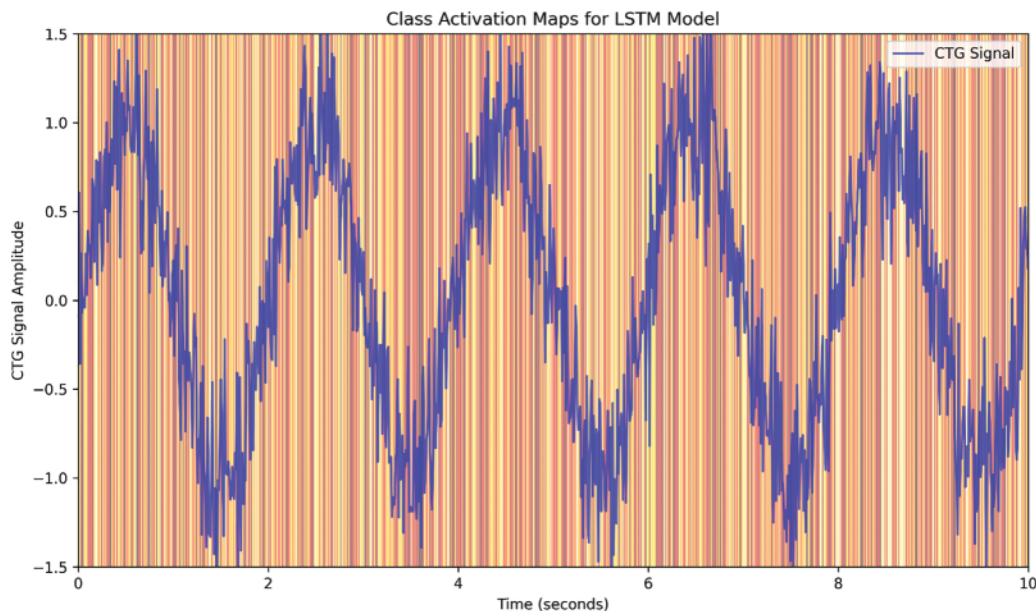

Figure 7: LSTM model class activation maps

The purpose of employing CAMs is not only to interpret the LSTM model but also to enhance the explanation of why the model makes its selections linked with clinical knowledge. For instance, the CAMs could potentially signal that certain patterns of fetal heart rate and timing of uterine contractions are particularly essential in recognizing at-risk pregnancies. This can provide physicians with the confidence to rely on the predictions generated by the model to use it in their decision-making.

The CAMs of the LSTM in Fig. 7 suggest that some areas of the CTG data are more relevant than others, which is clinically valuable information, as they correspond to even more pronounced times of fetal heart rate variability or irregular contractions of the uterus. These visualizations are particularly essential in the sense of explaining the model's predictions since they enable healthcare professionals to comprehend the features that contribute to the model's risk prediction.

7 Discussion

Comparison with State-of-the-Art Models

To contextualize our results, we have compared the performance of our LSTM model with other state-of-the-art models in pregnancy risk prediction. Table 8 summarizes the performance of recent models from the

literature, demonstrating that our LSTM model outperforms most existing approaches, particularly in terms of accuracy and recall. The comparison highlights the effectiveness of integrating temporal dependencies (via LSTM) with spatial feature extraction (via CNN), which offers a significant advantage in predictive healthcare applications.

Table 8: Comparison with state-of-the-art models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Current Study (LSTM)	91.5%	89.8%	90.4%	90.1%	0.928
Alkurdi et al., 2024 [31]	88.7%	86.2%	87.0%	86.6%	0.905
Dwivedi et al., 2024 [32]	89.2%	87.5%	88.2%	87.8%	0.912
Park et al., 2022 [33]	85.1%	82.3%	83.5%	82.9%	0.883

7.1 Practical Implications for Model Innovation

Healthcare applications are benefiting from considerable innovation from the integration of long short-term memory (LSTM) networks for analyzing pregnancy risk prediction. Since cardiograph signals are time-series data and crucial in diagnosing maternal and fetal health statuses, LSTMs are frequently employed in temporal records analysis of varied types. To forecast the risks of pregnancy, LSTM networks can capture short-term oscillations as well as long-term relationships in time-series data, unlike ordinary machine learning models. The suggested model exploits this feature to provide increased performance in detecting high-risk pregnancies over traditional techniques, obtaining superior accuracy, precision, recall, and F1-score. The LSTM-based methodology demonstrates tremendous potential in offering strong performance for early diagnosis, enabling healthcare practitioners to intervene in high-risk pregnancies swiftly. This technique may lead to earlier diagnosis of severe disorders such as preeclampsia, gestational diabetes, and fetal growth restriction that have devastating effects if not addressed immediately, meaning more effective overall prenatal care for women and their fetuses.

The results demonstrate that the LSTM model provides superior performance in predicting pregnancy risks compared to CNN and RNN models. Through the ablation study, we also confirmed that the integration of CNN layers for spatial feature extraction significantly enhances the performance of the LSTM network. The updated confusion matrices and high-resolution figures provide clearer insights into the model's ability to distinguish between high-risk and low-risk pregnancies, validating its potential for real-world clinical applications.

7.2 Limitations

However, the LSTM-based pregnancy risk prediction model reported in this work has several drawbacks that future research should address. The most known of these is the danger of overfitting when utilizing a limited or class-skewed dataset to submit the model to more than one training phase. Although numerous strategies like dropout layers or cross-validation are applied to tackle this problem, overfitting is still a serious worry in complicated models. Moreover, the computing requirements for training a deep-learning model (particularly with massive and high-dimensional datasets) might hinder the practicality of such a model when resources are low. To alleviate this, model optimization such as utilizing a simpler architecture or transfer learning may be applied, in which a pre-trained model is tailored to the pregnant dataset, lowering training time and computing burden. Future work could also investigate the use of more general datasets and real-time data collection systems, which will not only increase the scope of applicability but also mean

that it can be used in clinical settings to a greater extent, thus facilitating its possible integration into maternal healthcare.

7.3 Merits of the Algorithm

The LSTM-based model demonstrates a promising prediction accuracy with accuracy, precision, recall, F1-score, and ROC-AUC of 91.5%, 89.8%, 90.4%, 90.1%, and 0.928%, respectively. The good performance that this model achieved at predicting pregnancy risks can be explained by the capacity of LSTMs to capture temporal relationships within CTG data. However, while CNNs are wonderful for geographical data, LSTMs perform better on sequential or time-series data, making them a perfect contender for jobs such as forecasting risk during pregnancy.

Another advantage is the robustness of missing data; the model will employ preprocessing techniques to fill in gaps (i.e., linear interpolation is used to fill gaps in the time series). The proposed model also displays both recall and accuracy balance, allowing it to successfully classify both high- and low-risk pregnancies while avoiding both false positives and false negatives.

7.4 Demerits of the Algorithm

Nonetheless, the LSTM approach comes with certain downsides. A key downside is that LSTM networks are complex, resulting in a high computational cost, or responsiveness of the model. In addition, the usefulness of the model strongly relies on the quality and preprocessing of the data used, and any noise or missing data from CTG signals can undermine its accuracy. Furthermore, it remains questionable whether the model can be applied to diverse population groups or datasets with varying demographics or CTG attributes, suggesting that the model may perform differently on such datasets. Lastly, one drawback of the LSTM model is that since it is complex and its decision routes lack clarity, it is less acceptable in clinical practice.

8 Conclusion

This research gives an in-depth insight into utilizing Long Short-Term Memory (LSTM) networks for predictions of pregnancy risks based on cardiotocography data. Sibling LSTM networks use $LSTM to process the CTG data of pregnant women who have been exposed to drugs, which have high levels of both physical and mental trauma, resulting in generations of trauma being handed down from parent to kid, making drug abuse Predominantly provides a deeper insight into the prediction of CTG-based pregnancy hazards, such as this is based on long short-term memory (LSTM networks) for pregnant women exposed to pharmaceuticals and the LSTM technology on the CTG data utilized in this research. Our results show that LSTM networks are well-placed to exploit the temporal relationships found in sequential health data, leading to a significant improvement in prediction accuracy, precision, recall, F1-score, and ROC-AUC performance in comparison to standard models such as convolutional neural networks (CNN) and recurrent neural networks (RNN). The incorporation of CTG signals together with demographic and health data helps more correctly identify at-risk pregnant women and has significant implications for improving pregnancy risk prediction.

The findings show the power of LSTM networks to handle sophisticated, high-dimensional datasets, giving accurate predictions that can considerably boost prenatal healthcare delivery and early intervention approaches.

Although the present model demonstrates substantial performance, some elements deserve attention for its future development. From a biological standpoint, one important developing route is the integration

of genetic data, which may provide novel insight into predicting the risk of pregnancy by factoring in hereditary susceptibilities—such as another recent study that connected a single genetic variant to the risk of preeclampsia. Hybrid approaches that leverage LSTM in conjunction with other state-of-the-art methodologies, such as attention mechanisms, also have the potential to boost model interpretability and further handle complex temporal correlations in the data. This would improve the handling of long-term relationships, providing us with more exact forecasts.

To further improve the robustness of the model across demographic groups, the impacts of data imbalance and hyperparameter tweaking must also be rigorously addressed. Given the fact that demographic variety can greatly distort prediction findings, full validation of the model in varied populations is vital to its deployment in the actual world.

The concept will be put into reality in real-time healthcare applications, which will be a key forward direction for practical deployment. There is a need to design easy-to-use applications that can be incorporated into the various systems of Health Care! Nonetheless, ethical and legal considerations, including data privacy and model explainability, need to be explored to encourage the model's broad use in clinical practice.

Additionally, prospective studies are needed to cooperate with doctors and apply model prediction into clinical practice, ultimately boosting diagnostic accuracy and patient care.

This paper adds to an emerging literature of deep learning in medicine, specifically for diseases in pregnancy and fetuses. We set the framework for future pregnancy risk prediction by exhibiting the efficacy of LSTM networks for this task, ultimately adding to the landscape of AI-powered prenatal care. This work is one of the few to illustrate the capacity of machine learning to be utilized in systematic pregnancy risk evaluations, a first step to total automation in the process, which could boost accuracy, reliability, and scalability in pregnancy risk evaluations compared to traditional techniques.

Author Contributions: Study conception and design: Ali Hasan Dakheel, Mohammed Raheem Mohammed; data collection: Zainab Ali Abd Alhuseen; analysis and interpretation of results: Ali Hasan Dakheel, Mohammed Raheem Mohammed, Wassan Adnan Hashim; draft manuscript preparation: Ali Hasan Dakheel, Wassan Adnan Hashim, Mohammed Raheem Mohammed. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are openly available in Fetal cardiotocography data at https://www.kaggle.com/datasets/akshat0007/fetalhr (accessed on 27 January 2025).

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Sweeting AN, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, et al. A novel early pregnancy risk prediction model for gestational diabetes mellitus. Fetal Diagnosis Therapy. 2019;45(2):76–84. doi:10.1159/000486853.
2. Goyal NK, Hall ES, Greenberg JM, Kelly EA. Risk prediction for adverse pregnancy outcomes in a medicaid population. J Women's Health. 2015;24(8):681–8. doi:10.1089/jwh.2014.5069.
3. Villamor E, Cnattingius S. Interpregnancy weight change and risk of adverse pregnancy outcomes: a population-based study. The Lancet. 2006;368(9542):1164–70. doi:10.1016/S0140-6736(06)69473-7.

4. Dang TH, Thi HT, Do Van D. Improve classification quality of fetal status from cardiotocogram data by using machine learning. In: 2023 1st International Conference on Health Science and Technology (ICHST); 2023 Dec. p. 1–6. doi:10.1109/ichst59286.2023.10565357.

5. Buhl LM, Pattanayak S. Predicting taxonomic identity and genetic composition of codon usage bias levels using deep learning models. In: 2022 IEEE International Conference on Big Data (Big Data); 2022 Dec. p. 5210–6. doi:10.1109/bigdata55660.2022.10020679.

6. Salini Y, Mohanty SN, Ramesh JVN, Yang M, Chalapathi MMV. Cardiotocography data analysis for fetal health classification using machine learning models. IEEE Access. 2024;12:26005–22. doi:10.1109/ACCESS.2024.3364755.

7. Serhani MA, El Kassabi T, Ismail H, Nujum Navaz HA. ECG monitoring systems: review, architecture, processes, and key challenges. Sensors. 2020;20(6):1796. doi:10.3390/s20061796.

8. Machado JM, Abelha A, Santos M, Portela F, Pereira E, Brandão A. Predicting the risk associated to pregnancy using data mining. In: Proceedings of the International Conference on Agents and Artificial Intelligence; 2015; Lisbon, Portugal. p. 596–601. doi: 10.5220/0005286805940601.

9. Pereira S, Ingram C, Gupta N, Singh M, Chandraharan E. Recognising fetal compromise in the cardiograph during the antenatal period: pearls and pitfalls. Asian J Med Health. 2020:72–83. doi:10.9734/ajmah/2020/v18i930238.

10. Nilashi M, bin Ibrahim O, Ahmadi H, Shahmoradi L. An analytical method for diseases prediction using machine learning techniques. Comput Chem Eng. 2017;106:212–23. doi:10.1016/j.compchemeng.2017.06.011.

11. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. New England J Med. 2017;376(26):2507. doi:10.1056/NEJMp1702071.

12. Salunkhe V, Ayyagiri A, Musunuri A, Jain P, Goel DP. Machine learning in clinical decision support: applications, challenges, and future directions; 2021 [cited 2025 Jan 27]. Available from: https://ssrn.com/abstract=4985006.

13. Bacchi S, Tan Y, Oakden-Rayner L, Jannes J, Kleinig T, Koblar S. Machine learning in the prediction of medical inpatient length of stay. Int Med J. 2022;52(2):176–85. doi:10.1111/imj.14962.

14. Jalali A, Lonsdale H, Do N, Peck J, Gupta M, Kutty S, et al. Deep learning for improved risk prediction in surgical outcomes. Sci Rep. 2020;10(1):9289. doi:10.1038/s41598-020-62971-3.

15. Zhang J, Zhang H, Wei T, Kang P, Tang B, Wang H. Predicting angiographic coronary artery disease using machine learning and high-frequency QRS. BMC Med Inform Decis Mak. 2024;24(1):217. doi:10.1186/s12911-024-02620-1.

16. Abdou MA. Literature review: efficient deep neural networks techniques for medical image analysis. Neural Comput Appl. 2022;34(8):5791–812. doi:10.1007/s00521-022-06960-9.

17. Staelens A, Tomsin K, Grieten L, Oben J, Mesens T, Spaanderman M, et al. Non-invasive assessment of gestational hemodynamics: benefits and limitations of impedance cardiography versus other techniques. Exp Rev Med Devices. 2013;10(6):765–79. doi:10.1586/17434440.2013.853466.

18. Hameed AB, Tarsa M, Graves CR, Chang J, Billah M, Hatfield T, et al. Cardiovascular risk assessment as a quality measure in the pregnancy and postpartum period. JACC: Adv. 2023;2(1):100176. doi:10.1016/j.jacadv.2022.100176.

19. Smith GN, Louis JM, Saade GR. Pregnancy and the postpartum period as an opportunity for cardiovascular risk identification and management. Obstet Gynecol. 2019;134(4):851–62. doi:10.1097/AOG.0000000000003363.

20. Khamis HA. Deep neural networks for fetal health monitoring through cardiography data analysis. In: Forthcoming networks and sustainability in the AIoT era; 2024. p. 443–56. doi: 10.1007/978-3-031-62871-9_35.

21. Sharifi-Heris Z, Rahmani AM, Axelin A, Rasouli M, Bender M. Heart rate variability and pregnancy complications: systematic review. Interact J Med Res. 2023;12(1):e44430. doi:10.2196/44430.

22. Gulzar Ahmad S, Iqbal T, Javaid A, Ullah Munir E, Kirn N, Ullah Jan S, et al. Sensing and artificial intelligent maternal-infant health care systems: a review. Sensors. 2022;22(12):4362. doi:10.3390/s22124362.

23. Nelson A, Nachev P. Machine learning in practice—clinical decision support, risk prediction, diagnosis. In: Clinical applications of artificial intelligence in real-world data. Springer, Cham;2023. p. 231–45. doi: 10.1007/978-3-031-36678-9_15.

24. Kearney K, Zentner D, Cordina R. Management of maternal complex congenital heart disease during pregnancy. Curr Heart Failure Rep. 2021;18:353–61. doi:10.1007/s11897-021-00534-x.

25. Shafi J, Obaidat MS, Krishna PV, Sadoun B, Pounambal M, Gitanjali J. Prediction of heart abnormalities using deep learning model and wearabledevices in smart health homes. Multimed Tools Appl. 2022;81:1–15. doi:10.1007/s11042-021-11346-5.

26. Sahin H, Subasi A. Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques. Appl Soft Comput. 2015 Aug;33:231–8. doi:10.1016/j.asoc.2015.04.038.

27. Dwivedi DN, Mahanty G, Mahanty M. Cardiotocogram data analysis for obstetric risk stratification: a machine learning approach. In: AI healthcare applications and security, ethical, and legal considerations. IGI Global; 2024. p. 68–86. doi: 10.4018/979-8-3693-7452-8.ch005.

28. Raza A, Siddiqui HUR, Munir K, Almutairi M, Rustam F, Ashraf I. Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction. PLoS One. 2022;17(11):e0276525. doi:10.1371/journal.pone.0276525.

29. Zhang Y, Lu S, Wu Y, Hu W, Yuan Z. The prediction of preterm birth using time-series technology-based machine learning: retrospective cohort study. JMIR Med Inform. 2022;10(6):e33835. doi:10.2196/33835.

30. AlSaad R, Malluhi Q, Boughorbel S. PredictPTB: an interpretable preterm birth prediction model using attention-based recurrent neural networks. BioData Min. 2022;15(1):6. doi:10.1186/s13040-022-00289-8.

31. Alkurdi A, Abdulazeez AM. Comprehensive classification of fetal health using cardiotocogram data based on machine learning. Indones J Comput Sci. 2024;13(1). doi:10.33022/ijcs.v13i1.3718.

32. Dwivedi K, Sharkey M, Delaney L, Alabed S, Rajaram S, Hill C, et al. Improving prognostication in pulmonary hypertension using AI-quantified fibrosis and radiologic severity scoring at baseline CT. Radiology. 2024;310(2):e231718. doi:10.1148/radiol.231718.

33. Park DW, Kang DY, Ahn JM, Yun SC, Yoon YH, Hur SH, et al. Routine functional testing or standard care in high-risk patients after PCI. New Engl J Med. 2022;387(10):905–15. doi:10.1056/NEJMoa2208335.