Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Explainable Temporal Risk Trajectory Learning for Multi- Horizon Prediction of Fetal Compromise Using Intrapartum Cardiotocography and Clinical Data

Kiruthiga Devi M
Vel Tech Dr Rangarajan Dr Sagunthala R&D Institute of Science and Technology
Joseph Emerson Raja
emerson.raja@mmu.edu.my
Multimedia University
Rajeswari P
Jaipur National University

Research Article

Keywords: Fetal Distress Prediction, Temporal Risk Trajectory Learning, Explainable Artificial Intelligence (XAI), Cardiotocography (CTG), Multi-Horizon Prediction

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11216423/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Explainable Temporal Risk Trajectory Learning for Multi-Horizon Prediction of Fetal Compromise Using Intrapartum Cardiotocography and Clinical Data

Kiruthiga Devi M[1], Joseph Emerson Raja[2, *], Rajeswari P[3]

[1]Assistant Professor (Sr.G), Department of CSE, School of Computing, Vel Tech Dr Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India.
[1] Postdoctoral Research Scholar, Multimedia University (MMU), Melaka 75450, Malaysia.
kiruthigaprofessor89@gmail.com

[2] Assistant Professor, Faculty of Engineering and Technology, Multimedia University, Melaka 75450,Malaysia.
emerson.raja@mmu.edu.my

[3]Professor, School of Computer and System Sciences, Jaipur National University, Jagatpura, Jaipur. Rajasthan, India.
rajeshwari.purushothaman@jnujaipur.ac.in

Corresponding author: Joseph Emerson Raja

Abstract:

Intrapartum fetal monitoring using cardiotocography (CTG) provides continuous information about fetal heart rate and uterine contraction patterns, but conventional computational approaches often formulate fetal assessment as classification at an individual temporal interval. Such formulations provide limited information about the temporal evolution of estimated fetal risk and its behaviour at future prediction horizons. This study presents an Explainable Temporal Risk Trajectory Learning (E-TRTL) framework for investigating the temporal evolution of model-estimated fetal risk and future fetal-state prediction from intrapartum CTG recordings. The framework was evaluated using the CTU-UHB Intrapartum Cardiotocography Database containing 552 CTG recordings. Quality-controlled fetal heart rate and uterine contraction signals were represented using 5-, 10-, and 20-minute temporal contexts. The framework integrates temporal feature extraction and fusion, temporal risk-trajectory modelling, fetal-state transition analysis, multi-horizon prediction, and model-level explainability. Future-state prediction was evaluated at 5-, 10-, 15-, and 30-minute horizons. Performance was assessed using multiclass classification measures, one-vs-rest ROC-AUC, and continuous risk-trajectory measures. The E-TRTL framework achieved an overall multiclass accuracy of 91.48%, with 89.00% macro precision, 87.00% macro recall, and 88.00% macro F1-score. The one-vs-rest AUC values were 0.884, 0.823, and 0.978 for the Normal, Suspicious, and Pathological categories, respectively. Future-state prediction performance was 95.72%, 94.62%, 93.72%, and 91.48% at the 5-, 10-, 15-, and 30-minute horizons, respectively. Continuous risk-trajectory evaluation produced an RMSE of 0.0442, MAE of 0.0308, $R^2$of 0.9536, EVS of 0.9536, and correlation coefficient of 0.9766. The results indicate that E-TRTL provides a temporal modelling formulation for examining changes in model-estimated fetal risk and evaluating future-state prediction across multiple horizons. The framework extends conventional classification-oriented analysis by jointly considering temporal risk trajectories, fetal-state transitions, forecasting, and model-level interpretation. The estimated risk trajectory represents a model-derived quantity rather than a directly measured clinical severity score. Further evaluation using independent datasets, clinically defined outcomes, calibration and uncertainty analysis, and prospective validation is required to assess generalizability and potential clinical utility.

Keywords: Fetal Distress Prediction; Temporal Risk Trajectory Learning; Explainable Artificial Intelligence (XAI); Cardiotocography (CTG); Multi-Horizon Prediction.

Graphical abstract

The E-TRTL framework fills this void by combining these four types of modelling (continuous risk trajectory modelling; state transition analysis; multi-horizon forecasting; and explainability) into a single functioning pipeline that has been tested using 552 actual CTG recordings in Figure1.

All measures were consistent across all metrics evaluated. The total accuracy of the system was 91.48%; the area under the pathological curve was .978; it predicted 95.72% accurately as to whether or not the fetus would be

pathological within 5 minutes from when the last measurement was taken. It decreased to 91.48% within 30 minutes from when the last reading was taken. Lastly, it had an estimated correlation of .9766 for its estimate of what would happen to the baby's status in terms of where the baby's status really did change over time.

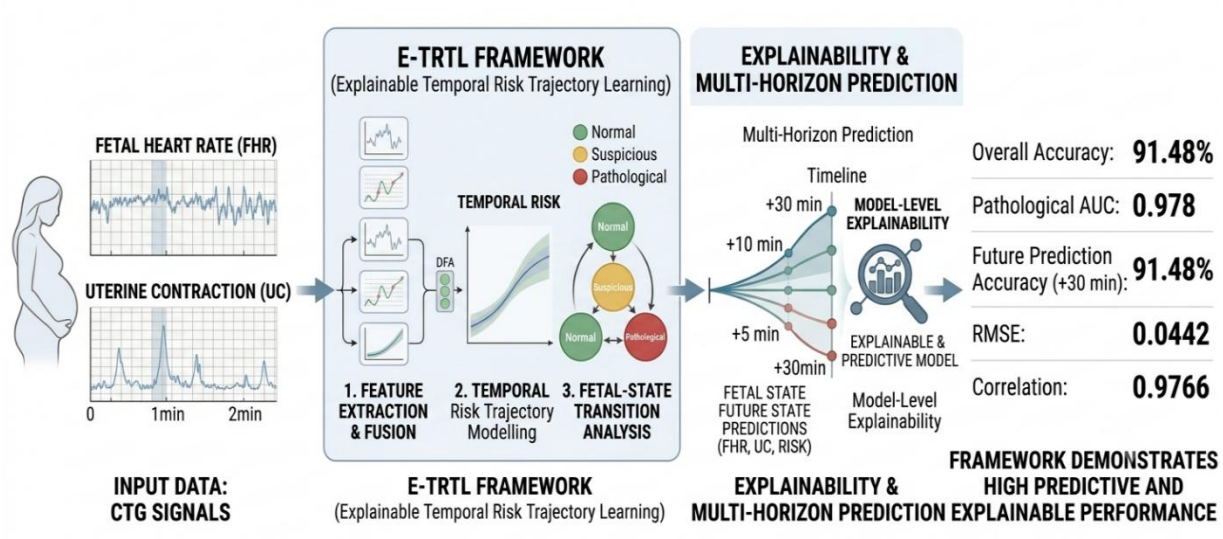


Figure 1. Graphical abstract

1. Introduction

1.1 Background

Fetal distress is still considered as one of the major problems of obstetric care because it is a presenting factor with fetal hypoxia, metabolic acidosis, neurological injury and poor neonatal outcome. Timely recognition of complications in foetus during gestation and labour is essential to the prevention of perinatal morbidity and mortality. A non-invasive method of assessing fetal well-being is the simultaneous recording of both fetal heart rate (FHR) and uterine contraction (UC) signals called cardiotocography (CTG). Although widely used, the assessment of CTG recordings is frequently subjective and very much reliant on clinical skills of the recording clinician, and highly dependent on clinician expertise, leading to substantial inter-observer variability and inconsistent diagnostic decisions [1].

Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) have revolutionized Biomedical signal analysis, allowing the automated extraction of complex temporal patterns from physiological signals. Recent studies have shown that deep learning models can be used to enhance the ability of electronic fetal monitoring (EFM) to detect fetal distress from CTG signals [2]. In the field of fetal compromise detection and minimizing the false-positive rate of conventional CTG interpretation, promising results have been obtained through the use of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and hybrid architectures [3]. Additionally, computer assisted systems for the analysis of CTG have become valuable tools for assisting clinicians in monitoring and decisions during labour [4].

Although considerable progress has been made in automated fetal distress detection, the current methods are mainly used to classify the current fetal condition as normal, suspicious and pathological. These methods give only a small amount of information about any further deterioration in the foetus and do not really assist with a proactive clinical approach. This has led to a growing demand for monitoring systems with such intelligent capabilities as being able to predict the progression of fetal distress and provide early warning signals before sever fetal compromise arises [5].

Automated CTG analysis has increasingly benefited from machine-learning and deep-learning methods for recognizing fetal heart rate abnormalities and classifying fetal condition. However, intrapartum fetal compromise is a temporally evolving process, and assessment of an isolated recording segment does not explicitly describe how estimated fetal risk changes over successive monitoring intervals. From a clinical decision-support perspective, modelling the evolution of risk may therefore provide information complementary to conventional fetal-state classification. This motivates investigation of temporal models that characterize changes in fetal risk and evaluate whether the current monitoring history contains information relevant to fetal condition at subsequent prediction horizons.

Recent research has also investigated progressive analysis of fetal heart rate signals during labour. Mendis et al. developed an input-length-invariant deep-learning approach for progressively detecting fetal compromise using the CTU-UHB dataset and evaluated the temporal availability of FHR information through time-to-predict analysis [1]. This work demonstrates that progressive fetal-compromise detection can incorporate information available at different stages of labour. Table 1 shows the Statement of Significance of the research. Building on this line of research, the present study focuses on a complementary formulation in which model-estimated fetal risk is represented as a temporal trajectory and examined together with fetal-state transitions and predefined future prediction horizons.

Table 1. Statement of Significance

Problem / Issue	Fetal distress during delivery is an ongoing development process but many computerized systems to analyze a woman's labor progress use a static approach analyzing one segment at a time and do not track how a woman's risk of having a distressed baby changes over time or vary by predicting her future condition.
Existing Evidence	There has been research showing that machine learning and deep learning techniques are capable of automatically assessing a woman's state based on cardiotocograph (CTG) traces. Additionally, there has been some exploration into detecting progressive fetal compromise through both the CTG trace and temporal prediction. However, researchers have formulated the inclusion of temporal information differently. There have also been differences with respect to modeling how the risk evolves over time and evaluating future states.
Novel Contribution	The purpose of this paper was to introduce a novel system called Explainable Temporal Risk Trajectory Learning (E-TRTL). The E-TRTL system will include multiple resolutions of CTG traces to model temporal risk trajectories, evaluate the transitions in fetal-states, incorporate obstetric-clinical context data, predict future states of the fetus across multiple horizons, and provide explanations of why predictions were made. Using 552 recordings collected at CTU-UHB, the E-TRTL system obtained a multiclass accuracy of 91.48%. The E-TRTL system provided evaluations of its ability to make accurate predictions of future states of the fetus for five different time horizons including 5-minutes, 10-minutes, 15-minutes and 30-minutes.
Beneficiaries	Researchers and developers working in areas related to fetal monitoring via computers, biomedical signal processing, medical AI, explainable AI, and clinical-decision support predictive analytics may benefit from developing new temporal models to predict the risks associated with fetal compromise.

1.2 Problem Statement

Automated analysis of intrapartum cardiotocography (CTG) has increasingly employed machine-learning and deep-learning methods to support the identification of fetal compromise from fetal heart rate (FHR) and uterine contraction patterns. Most reported approaches formulate the assessment primarily as a classification problem, in which a CTG recording or temporal segment is assigned to a fetal condition or outcome category [3], [5]. Although these approaches have demonstrated the potential of computational CTG analysis, classification at a single recording or segment level provides limited information about how estimated fetal risk evolves over successive monitoring intervals.

Fetal compromise develops within a temporal physiological context, and clinically relevant changes in FHR characteristics may occur progressively during intrapartum monitoring. Recent reviews of automated CTG analysis and intrapartum fetal hypoxia prediction have highlighted continuing challenges related to temporal modelling, generalizability, validation, and translation of predictive models into clinically meaningful decision support [3], [5]. These observations motivate investigation of approaches that represent changes in estimated fetal risk over time rather than relying solely on isolated fetal-state predictions.

Another relevant consideration is the prediction horizon. A model that identifies an abnormal condition at the current monitoring interval addresses a different task from a model that estimates fetal condition at a subsequent time point. For an early-warning formulation, the prediction at time t should therefore be based only on observations available up to that time, while the target corresponds to fetal condition at a predefined future horizon. Evaluating multiple prediction horizons can additionally determine how predictive performance changes as the interval between the available observations and the future target increases.

Interpretability is also important when predictive models are considered for medical decision-support research. Although explainable artificial intelligence methods can provide information about the features contributing to model predictions, such explanations should be interpreted as model-derived associations rather than evidence of causal clinical relationships [6]. Accordingly, an interpretable temporal modelling approach should provide both the estimated evolution of fetal risk and information regarding the available signal-derived and clinical factors associated with its predictions.

Based on these considerations, this study formulates fetal assessment as a temporal risk-trajectory modelling and multi-horizon forecasting problem. The proposed Explainable Temporal Risk Trajectory Learning (E-TRTL) framework uses multi-resolution CTG input windows of 5, 10, and 20 minutes to capture short-, intermediate-, and longer-duration temporal characteristics. These representations are subsequently used to model the evolution of estimated fetal risk and fetal-state transitions, followed by future-state forecasting at predefined horizons of 5, 10, 15, and 30 minutes. Available clinical-obstetric information is incorporated as contextual information alongside CTG-derived representations. The formulation is intended to examine temporal model behaviour and forecasting performance on the evaluated dataset rather than to establish a clinically validated warning system or a causal model of fetal compromise.

1.3 Research Motivation

The motivation for this study arises from the need to characterize fetal condition as an evolving temporal process rather than solely as an isolated classification outcome. During intrapartum monitoring, fetal heart rate (FHR) patterns and uterine activity are observed continuously, and changes in these signals over successive monitoring intervals may provide information about the evolution of fetal condition. Automated CTG analysis has demonstrated the potential of computational approaches for identifying fetal compromise; however, extending such analysis toward temporal modelling may provide a more explicit representation of how estimated fetal risk changes during monitoring [1], [3].

A further motivation is the need to investigate fetal condition at different prediction horizons. Recent literature has emphasized the importance of improving prediction of intrapartum fetal compromise while also identifying limitations related to temporal modelling, validation, and clinical translation

[5]. Rather than estimating only the fetal state associated with the currently observed interval, modelling the temporal evolution of risk enables investigation of whether information available up to a given monitoring point is associated with fetal condition at subsequent time intervals. This provides the basis for evaluating fetal-state prediction over multiple predefined horizons.

Clinical and obstetric information associated with CTG recordings may also provide contextual information complementary to signal-derived temporal representations. Integrating such available information with CTG-derived features provides an opportunity to examine relationships between fetal monitoring patterns and the clinical context represented in the dataset. However, such integration should be restricted to variables actually available in the study dataset and should not be interpreted as continuous maternal physiological monitoring.

Interpretability provides an additional motivation for the proposed framework. In medical artificial intelligence, understanding the factors associated with a model prediction is important for evaluating model behaviour and identifying whether learned relationships are clinically plausible [6]. Therefore, in addition to fetal-state prediction, this study investigates temporal risk visualization and feature-attribution analysis to provide information about the factors associated with estimated fetal risk.

These considerations motivate the development of the Explainable Temporal Risk Trajectory Learning (E-TRTL) framework, which integrates multi-resolution CTG representation, temporal risk-trajectory learning, fetal-state transition analysis, contextual clinical-obstetric information, multi-horizon forecasting, and model-level explanation. The framework uses 5-, 10-, and 20-minute CTG analysis windows to represent temporal characteristics at different scales and evaluates future-state prediction at 5-, 10-, 15-, and 30-minute forecasting horizons. The framework is evaluated as a dataset-level modelling approach, and the reported findings are not interpreted as evidence of clinical effectiveness without independent external validation [5].

1.4 Contribution of the Proposed work

This study investigates temporal modelling of fetal risk from intrapartum cardiotocography by extending fetal assessment beyond isolated segment-level classification. The main contributions of the study are summarized as follows:

1. Temporal risk-trajectory formulation: The study formulates fetal assessment as a temporal risk-modelling problem in which changes in estimated fetal risk are represented across successive intrapartum monitoring intervals. The Temporal Risk Trajectory Learning (TRTL) component is used to characterize the evolution of estimated risk over time rather than relying solely on independent fetal-state predictions.

2.Transition-aware multi-horizon fetal-state prediction: Fetal-state transition modelling is incorporated to characterize changes among fetal states over successive monitoring intervals. The framework further investigates fetal-state prediction at multiple predefined future horizons to examine how predictive performance varies with increasing prediction duration.

3. Integration of CTG-derived representations and available clinical information: Temporal representations derived from fetal heart rate and uterine contraction signals are integrated with the clinical-obstetric variables available in the study dataset. This integration is used to investigate the complementary contribution of signal-derived and clinical information to fetal-risk estimation.

4. Explainable temporal risk assessment: The framework incorporates temporal risk visualization and feature-attribution analysis to examine changes in estimated fetal risk and identify the available features associated with model predictions. These explanations are intended to support interpretation of model behaviour and are not considered evidence of causal clinical relationships.

The individual signal-processing, deep-learning, graph-learning, attention, and explainability techniques incorporated in E-TRTL are established computational methods. The contribution of this study is positioned in their integration around multi-resolution CTG representation, temporal risk-trajectory learning, fetal-state transition modelling, multi-horizon forecasting, contextual information integration, and model-level interpretation. Thus, the proposed framework is not

presented as a collection of individually novel algorithms, but as an integrated analytical formulation for investigating temporal fetal-risk behaviour and future-state prediction.

2. Related Works

Cardiotocography (CTG) is widely used for intrapartum fetal surveillance through simultaneous monitoring of fetal heart rate (FHR) and uterine activity. Clinical interpretation considers characteristics such as baseline FHR, variability, accelerations, decelerations, and uterine contractions according to established monitoring guidelines [4]. However, variability in visual interpretation has motivated research into computerized methods that can provide more consistent assessment of CTG recordings [3], [8].

Early computational approaches generally relied on signal processing and manually derived CTG characteristics followed by conventional machine-learning classifiers. Features describing FHR variability, accelerations, decelerations, and other CTG characteristics have been evaluated using methods such as support vector machines, decision trees, random forests, nearest-neighbour approaches, and ensemble classifiers [9]–[12]. These studies demonstrate that CTG-derived features can support automated fetal-state classification. Nevertheless, differences in feature construction, outcome definitions, datasets, and validation strategies can make direct comparison of reported performance difficult [3], [8].

Deep-learning approaches have subsequently been investigated to reduce dependence on manually engineered representations and to learn discriminative characteristics directly from CTG-derived inputs. Convolutional and temporal deep-learning architectures have been studied for fetal compromise detection, including approaches based on raw FHR recordings and time–frequency representations [1], [2], [13]. More recent studies have continued to investigate deep-learning methods for automated CTG interpretation [14], [16]. Collectively, these studies indicate that learned representations can capture information relevant to fetal-state assessment; however, model performance depends on the dataset, target definition, input representation, and evaluation protocol.

An important distinction in automated fetal monitoring is between current-state assessment and future-state prediction. Many published CTG studies formulate the task as classification of a recording or segment according to a fetal condition or clinical outcome [8]–[14]. Such classification is useful for identifying patterns associated with fetal compromise, but it does not by itself describe how estimated risk changes across successive monitoring intervals. A recent systematic review of prediction models for intrapartum fetal hypoxia also identified methodological and validation challenges that remain relevant to the development and clinical evaluation of predictive approaches [5]. These observations support further investigation of temporal formulations that explicitly evaluate risk evolution and prediction across defined future horizons.

Interpretability represents another consideration in medical artificial intelligence. Explainable artificial intelligence (XAI) methods are intended to provide information about the contribution of model inputs to individual or overall predictions, thereby supporting examination of model behaviour [6]. In CTG analysis, interpretable deep-learning approaches have also been investigated for fetal hypoxia detection [19]. Methods such as feature attribution can therefore complement predictive performance by indicating which model inputs are associated with a particular prediction. Nevertheless, such explanations describe relationships learned by the model and should not automatically be interpreted as causal physiological mechanisms.

Taken together, previous research has established the feasibility of machine-learning and deep-learning approaches for automated CTG assessment. The literature also indicates continuing interest in reliable validation, temporal modelling, and interpretable prediction [3], [5], [8]. These considerations provide the basis for investigating whether modelling the evolution of estimated fetal risk across successive monitoring intervals can complement conventional fetal-state classification.

Recent studies demonstrate the feasibility of automated and progressive CTG analysis using machine-learning and deep-learning approaches. In particular, Mendis et al. investigated progressive

fetal-compromise detection using input-length-invariant deep learning and the CTU-UHB dataset [1]. Other studies have investigated deep-learning-based CTG interpretation, temporal representations, clinical-context integration, and explainability [3], [8] – [19]. These studies provide a methodological basis for the present work. However, the approaches differ in how temporal information is represented and evaluated. The present study focuses on an integrated formulation in which multi-resolution CTG windows are used to construct temporal representations, model-estimated risk is represented as a trajectory, fetal-state transitions are characterized, and future-state prediction is evaluated at predefined forecasting horizons. Thus, the contribution of E-TRTL is positioned primarily in the integration of these components within a common temporal-risk formulation, rather than in claiming that the individual underlying algorithms are new.

The distinction from progressive FHR detection is important for positioning the present study. Mendis et al. [1] investigated progressively available FHR information and time-to-predict for fetal-compromise detection, whereas the present study uses multi-resolution CTG input windows of 5, 10, and 20 minutes to derive temporal representations and subsequently models estimated risk evolution, fetal-state transitions, and future-state prediction at 5-, 10-, 15-, and 30-minute horizons. The two approaches therefore share an interest in temporal fetal assessment but address different analytical formulations.

2.1 Research Gaps

Although automated CTG analysis has progressed from conventional feature-based machine learning to deep-learning approaches, many studies formulate fetal assessment through recording-level or segment-level classification. Progressive FHR analysis has also been investigated using the CTU-UHB dataset, including input-length-invariant deep learning and time-to-predict analysis [1]. However, progressive fetal-compromise detection and explicit representation of a model-estimated risk trajectory constitute different analytical formulations. The present study focuses on representing estimated fetal risk across successive temporal observations and examining this trajectory together with fetal-state transitions and future prediction horizons.

Second, evaluation at a single assessment point provides limited information about how predictive performance changes when the target state is located at different future intervals. The present study therefore evaluates future-state prediction at 5-, 10-, 15-, and 30-minute forecasting horizons. These horizons are distinct from the 5-, 10-, and 20-minute input windows used for multi-resolution CTG representation. The purpose is to characterize how forecasting performance changes with prediction distance rather than to define a clinically validated intervention window.

Third, Feature attribution can provide information about inputs associated with a prediction, but it does not by itself describe the temporal evolution of estimated risk. Combining feature attribution with trajectory and state-transition representations can therefore provide complementary information about both model inputs and the temporal behaviour of the prediction process.

Based on these gaps, the present study investigates temporal risk-trajectory learning as an extension of conventional fetal-state assessment. Rather than assuming that existing classification or progressive detection approaches are inadequate, the study examines whether explicit modelling of estimated risk evolution, fetal-state transitions, and multiple future prediction horizons can provide complementary information for characterizing fetal condition from intrapartum CTG recordings. The framework integrates established signal-processing, feature-learning, temporal modelling, contextual fusion, forecasting, and explainability techniques around this temporal formulation.

Based on these gaps, the present study investigates temporal risk-trajectory learning as an extension of conventional fetal-state assessment. Rather than assuming that existing classification approaches are inadequate, the study examines whether explicit modelling of risk evolution, fetal-state transitions, and multiple prediction horizons can provide complementary information for characterizing fetal condition from intrapartum CTG recordings. Explainability is incorporated to examine the model-derived factors associated with the estimated risk and its temporal progression.

3. Proposed Methodology

The Explainable Temporal Risk Trajectory Learning (E-TRTL) framework is developed to investigate the temporal evolution of fetal risk from intrapartum cardiotocography (CTG) recordings. The framework uses fetal heart rate (FHR) and uterine contraction (UC) signals together with the clinical-obstetric information available for the corresponding CTG records. Rather than treating each analysed interval only as an independent classification instance, E-TRTL represents changes in estimated fetal risk across successive monitoring intervals and evaluates fetal-state prediction at predefined future horizons.

The framework consists of signal preprocessing, temporal representation learning, feature extraction and fusion, temporal risk-trajectory modelling, fetal-state transition estimation, multi-horizon prediction, and model interpretation. These components are organized to examine three related aspects of fetal assessment: the fetal condition represented by the available observations, the evolution of estimated risk over time, and the relationship between the observation history and fetal condition at subsequent prediction horizons.

The proposed E-TRTL framework is organized as an integrated temporal modelling pipeline rather than as a collection of independent prediction modules. The processing begins with CTG signal quality assessment and preprocessing, followed by multi-resolution representation using 5-, 10-, and 20-minute input windows. The resulting representations are used for temporal and spectral feature extraction and fusion, after which the temporal modelling stage estimates fetal risk across successive observations. The estimated risk sequence is subsequently used to characterize fetal-state transitions and support future-state forecasting at 5-, 10-, 15-, and 30-minute horizons. Available clinical-obstetric information is incorporated as contextual information, while explainability methods are used to examine model-derived factors associated with the predictions. Thus, the methodological contribution lies in connecting these stages through a common temporal-risk formulation.

The following subsections describe the mathematical formulation and individual components of the E-TRTL framework.

3.1 Problem Formulation

Intrapartum fetal monitoring using cardiotocography (CTG) produces continuously varying fetal heart rate (FHR) and uterine contraction (UC) signals, whose characteristics may change throughout the monitoring period. CTG-based computational methods have commonly formulated fetal assessment as the classification of CTG recordings or temporal segments into predefined fetal or outcome categories [3], [5]. Recent studies have also investigated progressive or temporally varying fetal-state prediction using CTG/FHR data, including progressive fetal-compromise detection using the CTU-UHB dataset [1]. Therefore, the present study does not consider temporal prediction alone as the primary novelty; instead, it formulates fetal assessment around the temporal evolution of model-estimated fetal risk, followed by fetal-state transition analysis and future-state prediction.

Let the CTG observation at temporal position t be represented as

$$X_t = [FHR_t, UC_t], \quad (1)$$

where FHR_t and UC_t denote the fetal heart rate and uterine contraction observations, respectively. The CTU-UHB database provides synchronized FHR and UC signals sampled at 4 Hz together with associated clinical and obstetric information [23].

To capture information at different temporal scales, the CTG observations are represented using three temporal contexts:

$$X_t^{(w)}, w \in \{5,10,20\}, \quad (2)$$

where denotes the duration of the temporal context in minutes. At a sampling frequency of 4 Hz, the corresponding representations contain 1200, 2400 and 4800 samples, respectively, for the two CTG channels.

The available clinical-obstetric information is represented by a contextual feature vector

$$C = [c_1, c_2, \dots, c_m], \quad (3)$$

where c_i denotes an available clinical or obstetric variable associated with the corresponding CTG record. Such contextual information is available in the CTU-UHB database in addition to the FHR and UC signals [23].

The multi-resolution CTG representations are transformed into a latent temporal representation using the learned temporal feature function

$$\mathbf{H}_t = f_\theta\left(\mathbf{X}_t^{(5)}, \mathbf{X}_t^{(10)}, \mathbf{X}_t^{(20)}\right) \quad (4)$$

where $f_\theta(.)$ denotes the learned temporal feature-extraction and fusion function and $\mathbf{H}_t$ denotes the resulting CTG-derived temporal representation.

The CTG representation is subsequently combined with the available clinical-obstetric context:

$$\mathbf{Z}_t = g_\phi(\mathbf{H}_t, \mathbf{C}) \quad (5)$$

Where $g_\phi(.)$ denotes the contextual integration function and $\mathbf{Z}_t$ represents the integrated feature representation used for subsequent prediction.

The model then estimates fetal risk from the integrated representation as

$$r_t = h_\psi(\mathbf{Z}_t) \quad (6)$$

where r_t denotes the model-estimated fetal risk at temporal position $\mathbf{t}$, and $h_\psi(.)$denotes the risk-estimation function.

The sequence of risk estimates over successive temporal observations is represented as $\mathcal{R} = \{r_1, r_2, \dots, r_T\}$ (7)

and is referred to as the temporal risk trajectory. This representation allows the model-estimated risk to be examined across successive monitoring intervals rather than considering each prediction independently. In this formulation, the trajectory describes changes in the model-estimated risk within the analysed recording and does not represent a directly measured physiological risk quantity.

The fetal state at temporal position t is represented as

$$y_t \in \{\text{Normal}, \text{Suspicious}, \text{Pathological}\} \quad (8)$$

Accordingly, the temporal sequence of fetal-state predictions can be expressed as

$$\mathcal{Y} = \{y_1, y_2, \dots, y_T\} \quad (9)$$

The sequence enables the framework to characterize persistence and transitions between successive model-estimated fetal states.

For future-state prediction, the temporal information available up to time is used to estimate the fetal state at a predefined future horizon h:

$$\hat{\mathbf{y}}_{t+h} = p_\omega(y_{t+h} \mid \mathbf{Z}_{\leq t}) \quad (10)$$

Where $h \in \{5,10,15,30\}$

represents the prediction horizon in minutes and $\hat{y}_{t+h}$denotes the predicted fetal-state distribution at the corresponding future time.

Thus, the proposed formulation explicitly distinguishes between temporal context and prediction horizon. The 5-, 10- and 20-minute representations are used to characterize the available CTG temporal context, whereas 5-, 10-, 15- and 30-minute horizons are used to evaluate future-state prediction. The overall objective of the E-TRTL framework is therefore to learn a temporal representation that connects multi-resolution CTG observations, model-estimated risk evolution, fetal-state transitions and future-state prediction, while incorporating available clinical-obstetric information. Explainability analysis is subsequently used to examine the model-derived features contributing to the predictions.

The methodological contribution of E-TRTL is not the individual use of convolutional, recurrent, attention or explainability techniques, since these approaches are established in CTG and medical signal analysis [3], [6], [8], [19], [22]. Rather, the proposed framework organizes these components around an explicit temporal-risk trajectory formulation, in which multi-resolution CTG observations are transformed into a sequence of model-estimated risk values, followed by state-transition characterization and multi-horizon future-state prediction.

3.2 Overall E-TRTL Architecture

The proposed Explainable Temporal Risk Trajectory Learning (E-TRTL) framework is designed to model the temporal evolution of **model-estimated fetal risk** from quality-controlled cardiotocography (CTG) signals while incorporating available clinical-obstetric information. The framework consists of sequential processing stages comprising signal-quality assessment, preprocessing, multi-resolution temporal representation, temporal feature learning, attention-based CTG fusion, clinical-context integration, temporal risk estimation, temporal risk trajectory construction, fetal-state transition modelling, multi-horizon prediction, and explainability analysis.

The overall architecture of the proposed framework is illustrated in Figure 2. The input consists of synchronized fetal heart rate (FHR) and uterine contraction (UC) signals obtained from the CTU-UHB database. The signals are first subjected to signal-quality assessment and preprocessing to identify invalid observations and reduce the influence of low-quality temporal segments. The quality-controlled signals are subsequently represented using 5-, 10- and 20-minute temporal contexts.The multi-resolution representations provide different temporal views of the same monitoring process. The 5-minute representation provides a relatively short temporal context, whereas the 10- and 20-minute representations provide progressively longer temporal contexts. At the sampling frequency of 4 Hz, the corresponding CTG inputs contain 1200, 2400 and 4800 samples, respectively, with two signal channels corresponding to FHR and UC.Each temporal representation is processed through a corresponding temporal feature-learning branch. The temporal encoders transform the input sequences into latent feature representations while retaining their temporal structure. In the implemented architecture, the three branches produce representations with a common feature dimension, allowing information from the different temporal resolutions to be subsequently combined.

The resulting temporal representations are processed using an attention-based fusion mechanism. The attention mechanism assigns learned relative weights to the representations obtained from the three temporal resolutions and combines them to produce a compact CTG representation. In the implemented model, the fused CTG representation has a dimensionality of 64 features. The learned weights represent the relative contribution of the temporal representations to the fused model representation and are not interpreted as direct measures of clinical importance.The fused CTG representation is then integrated with the available clinical-obstetric information associated with the corresponding recording. The clinical-obstetric features provide contextual information complementary to the dynamic information

extracted from the CTG signals. The resulting integrated representation is used for subsequent temporal risk estimation and fetal-state prediction.The model-estimated risk values obtained from successive temporal observations are organized into a temporal risk trajectory. Instead of treating individual temporal predictions as completely independent outputs, this representation allows changes in the estimated risk to be examined across successive monitoring observations. The trajectory represents a model-derived quantity and should not be interpreted as a directly measured physiological risk or as a clinical severity score.

The corresponding fetal-state predictions are represented as a sequence of Normal, Suspicious and Pathological states. Successive predictions are examined to characterize the persistence and transition of the model-estimated fetal state over time. This temporal representation provides a basis for evaluating whether the predicted state remains unchanged or changes between successive observations.The resulting temporal representation is subsequently used to estimate fetal state at predefined future horizons of 5, 10, 15 and 30 minutes. These prediction horizons are evaluated separately to examine the behaviour of the framework as the temporal distance between the available observation and the future prediction increases. The framework therefore distinguishes between the temporal contexts used as model inputs and the future horizons used for prediction.

Finally, explainability analysis is applied to examine the model-derived features associated with the predictions. The resulting explanations are interpreted as model-attribution information and are not considered evidence of causal physiological relationships.The framework integrates these processing stages around the temporal risk trajectory, providing a unified representation that connects multi-resolution CTG observations with model-estimated risk evolution, fetal-state transitions and future-state prediction. The individual learning and explainability mechanisms are established techniques; the contribution of the proposed framework lies in their organization within this temporal-risk modelling formulation.

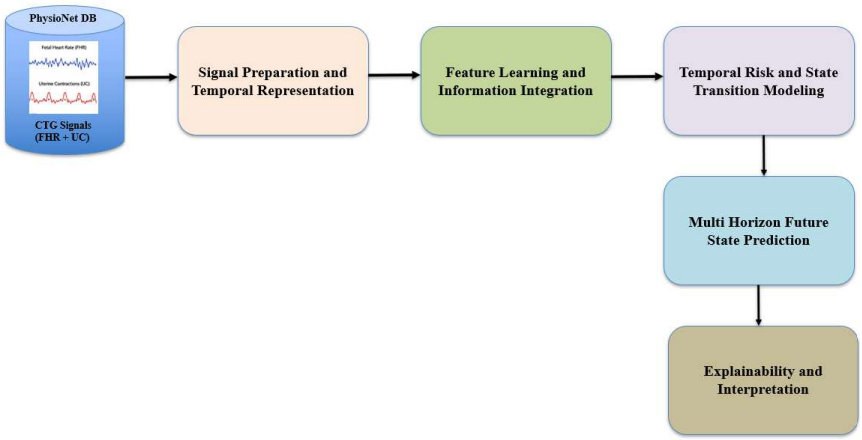


Figure 2. Overall architecture of the proposed E-TRTL framework.

3.3 Data Inputs

The E-TRTL framework uses intrapartum cardiotocography recordings from the CTU-UHB Intrapartum Cardiotocography Database. The database contains 552 CTG recordings and provides simultaneously recorded fetal heart rate (FHR) and uterine contraction (UC) signals sampled at 4 Hz [23]. These two time-series signals constitute the primary physiological inputs to the proposed framework.

For a given CTG record, the signal input is represented as

$$\mathbf{X} = [\mathbf{x}^{FHR}, \mathbf{x}^{UC}] \tag{11}$$

Where $\mathbf{x}^{FHR}$ and $\mathbf{x}^{UC}$ denote the FHR and UC time-series signals, respectively. For a recording containing samples, the FHR signal is represented as

$$\mathbf{x}^{FHR} = [x_1^{FHR}, x_2^{FHR}, \dots, x_T^{FHR}] \quad (12)$$

and the corresponding UC signal is represented as

$$\mathbf{x}^{UC} = \left[x_1^{UC}, x_2^{UC}, \dots, x_T^{UC}\right] \quad (13)$$

In addition to the CTG signals, the database provides clinical and obstetric information associated with individual recordings [23]. The clinical and obstetric variables used in this study are represented by a contextual feature vector:

$$\mathbf{C} = [c_1, c_2, \dots, c_m] \quad (14)$$

where m denotes the number of clinical-obstetric variables included in the analysis. In the implemented dataset, 13 clinical-obstetric features are used as contextual inputs after data preparation and missing-value handling.

The CTG signals and clinical-obstetric information are processed through separate representation pathways before their learned representations are integrated at the feature-fusion stage. This separation preserves the temporal structure of the CTG signals while allowing record-level clinical-obstetric information to provide additional contextual information. The CTG pathway therefore represents dynamic physiological observations, whereas the clinical pathway represents associated record-level contextual variables. For multi-resolution temporal analysis, the CTG signals are subsequently organized into 5-, 10-, and 20-minute input windows. At a sampling frequency of 4 Hz, these correspond to 1200, 2400, and 4800 samples, respectively, for each of the two CTG channels. These windows provide different amounts of preceding temporal context for subsequent feature extraction, fusion, and temporal-risk modelling

$$\mathbf{X}^{(5)} \in \mathbb{R}^{1200\times 2},$$

$$\mathbf{X}^{(10)} \in \mathbb{R}^{2400\times 2},$$

$$\mathbf{X}^{(20)} \in \mathbb{R}^{4800\times 2}\} \quad (15)$$

Here, the two columns correspond to FHR and UC, respectively. The three representations correspond to aligned temporal contexts used to provide short-, intermediate- and longer-duration information to the temporal feature-learning stages.

3.4 Signal Preprocessing and Quality Assessment

The recorded FHR and UC signals were subjected to preprocessing and signal-quality assessment before temporal window generation. The preprocessing procedure consisted of FHR validity checking, short-gap interpolation, wavelet-based denoising, and window-level signal-quality assessment.For FHR, observations were considered valid when they were finite and within the range of 50–220 beats per minute (bpm). Invalid observations were identified before subsequent processing. The validity of a signal observation x_t is defined as

$$q_t = \begin{cases} 1, & 50 \le x_t \le 220 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where $q_t = 1$ indicates a valid FHR observation and $q_t = 0$ indicates an invalid observation.

Consecutive invalid FHR observations were identified as signal gaps. Short internal gaps of up to 10 s were filled using linear interpolation when valid observations were available on both sides of the gap. The interpolated value is given by

$$\hat{x}_t = x_{t_1} + \frac{t-t_1}{t_2-t_1}\left(x_{t_2} - x_{t_1}\right) \quad (17)$$

where t_1 and t_2 denote the positions of the valid observations immediately before and after the gap, respectively. Gaps exceeding the specified duration were not interpolated.

The interpolated FHR signal was subsequently denoised using a Daubechies-4 (db4) wavelet with a decomposition level of 4. Soft-thresholding was applied to the detail coefficients, followed by inverse wavelet reconstruction to obtain the denoised FHR signal. The UC signal was retained as the synchronized second channel corresponding to the FHR segment.

Signal quality was assessed using the Signal Quality Index (SQI). For a temporal FHR window, the SQI was calculated as

$$SQI = \frac{N_{\text{valid}}}{N_{\text{total}}} \times 100 \quad (18)$$

A minimum SQI threshold of 70% was applied during window generation. The SQI was evaluated for the 5-, 10-, and 20-minute temporal windows, and a multi-resolution sample was retained only when all three corresponding windows satisfied the quality criterion. Residual missing FHR observations within an accepted window were subsequently handled before construction of the final model input.

3.5 Temporal Window Generation and Multi-Resolution Representation

The quality-controlled FHR and UC signals were segmented into temporal windows to represent fetal monitoring information at different temporal scales. Since the CTU-UHB database provides synchronized FHR and UC signals sampled at 4 Hz [23], three temporal contexts of 5, 10, and 20 minutes were considered in the E-TRTL framework.

For a temporal window of duration L minutes, the corresponding number of samples is calculated as

$$N_L = F_S \times 60 \times L \quad (19)$$

where N_L denotes the number of samples in the temporal window, F_S represents the sampling frequency, and L denotes the window duration in minutes. With Hz, the 5-, 10-, and 20-minute windows correspond to 1200, 2400, and 4800 samples, respectively.

For each temporal position, the FHR and UC signals were jointly represented as a two-channel temporal input:

$$X^{(L)} = \left[x_{1:N_L}^{FHR}, x_{1:N_L}^{UC}\right] \quad (20)$$

where $X^{(L)}$ denotes the CTG representation for a window of duration, while $x_{1:N_L}^{FHR}$ and $x_{1:N_L}^{UC}$ represent the corresponding FHR and UC sequences.

The three temporal resolutions provide complementary temporal contexts for feature learning. The 5-minute representation captures relatively short-term variations in the CTG signals, whereas the 10- and 20-minute representations provide progressively longer temporal contexts. The three representations are generated from corresponding temporal locations so that they describe the same underlying monitoring episode at different temporal scales.

Only temporal windows satisfying the signal-quality criterion defined in Section 3.4 were retained for subsequent processing. Window generation was performed separately within the training, validation, and testing partitions to maintain the separation between the respective data subsets. The record identifier associated with each window was also preserved, allowing the temporal window to remain linked to its corresponding CTG recording and clinical-obstetric information. The retained CTG

windows were normalized before being provided to the learning model. The normalization parameters were estimated exclusively from the training partition and subsequently applied to the validation and testing partitions. Thus, information from the validation and testing data was not used in determining the normalization parameters.

The resulting multi-resolution CTG representations are organized into three temporal pathways:5-minute representation,10-minute representation,20-minute representation.where the second dimension corresponds to the FHR and UC channels. These multi-resolution representations are subsequently provided to the corresponding temporal feature-learning branches of the E-TRTL framework. The learned representations from the different temporal resolutions are then combined through the attention-based CTG fusion stage described in the following section.

3.6 Temporal Feature Learning and Attention-Based CTG Fusion

The 5-, 10-, and 20-minute CTG representations generated in Section 3.5 are processed through separate temporal feature-learning branches. For each temporal resolution $r \in 5,10,20$, the corresponding CTG input $X(r)$is transformed into a latent temporal representation:

$$H^{(r)} = E_r\left(X^{(r)}\right) \tag{21}$$

where $E_r(.)$denotes the temporal feature encoder and $H^{(r)}$represents the learned features at temporal resolution.

The resolution-specific representations are projected into a common feature space to obtain $Z^{(r)}$ and these representations are subsequently combined using an attention-based fusion mechanism. The normalized attention weight for each temporal resolution is defined as

$$\alpha_r = \frac{\exp(e_r)}{\sum_{k\in\{5,10,20\}} \exp(e_k)} \tag{22}$$

where e_r denotes the learned attention score for the corresponding temporal representation. The attention weights satisfy

$$\sum_{r\in\{5,10,20\}} \alpha_r = 1 \tag{23}$$

The fused CTG representation is obtained by the weighted combination of the three resolution-specific representations:

$$Z_{CTG} = \sum_{r\in\{5,10,20\}} \alpha_r Z^{(r)} \tag{24}$$

where $Z^{(r)}$ denotes the projected representation at temporal resolution r , and Z_{CTG} denotes the resulting fused CTG representation. In the implemented architecture, the fused CTG representation is represented in a 64-dimensional feature space. The resulting representation is subsequently passed to the clinical-context integration stage.

3.7 Clinical-Context Integration

The fused CTG representation obtained from the multi-resolution temporal feature-learning stage is integrated with the clinical-obstetric information associated with the corresponding CTG recording. The clinical information provides record-level contextual variables that complement the temporal information learned from the FHR and UC signals.

Let the clinical-obstetric feature vector for a recording be represented as

$$c = [c_1, c_2, \dots, c_m] \tag{25}$$

where c_j denotes the j-th clinical-obstetric variable and m represents the number of clinical variables included in the model.

The clinical variables are transformed into a numerical representation before integration with the CTG pathway. Let the resulting clinical representation be denoted by

$$Z_C = E_C(c) \tag{26}$$

where $E_C(.)$denotes the clinical-context encoding function and Z_C represents the encoded clinical-obstetric information.

The encoded clinical representation is subsequently combined with the fused CTG representation Z_{CTG} obtained in Section 3.6. The integrated multimodal representation is defined as

$$Z_M = Concat(Z_{CTG}, Z_C) \tag{27}$$

where $Concat(.)$denotes feature concatenation and Z_M represents the integrated CTG–clinical representation.

The resulting representation contains information from both the dynamic CTG signals and the corresponding record-level clinical-obstetric variables. This integrated representation is subsequently provided to the temporal risk modelling stage for estimation of the model-derived fetal risk trajectory and future fetal-state predictions.

3.8 Temporal Risk Trajectory Learning

The integrated CTG–clinical representation obtained in Section 3.7 is used for temporal risk modelling across successive monitoring observations. The objective of this stage is to represent the evolution of model-estimated fetal risk over time rather than treating each temporal observation as an independent prediction.

Let $Z_{M,t}$ denote the integrated CTG–clinical representation at temporal observation t. The sequence of integrated representations is processed using a bidirectional long short-term memory (BiLSTM) network to capture temporal dependencies between successive observations. The resulting temporal representation is expressed as

$$H_t^{TR} = BiLSTM\left(Z_{M,t}\right) \tag{28}$$

where H_t^{TR} denotes the temporal representation learned from the integrated CTG–clinical information at observation t.

The temporal representation is subsequently used to estimate the probabilities of the three fetal states. The state-probability vector is defined as

$$p_t = Softmax(W_r H_t^{TR} + b_r) \tag{29}$$

where W_r and b_r denote the learnable parameters of the risk-estimation layer. The resulting probability vector is represented as

$$p_t = [p_t^N, p_t^S, p_t^P] \tag{30}$$

Where p_t^N, p_t^S, p_t^P denote the estimated probabilities of the Normal, Suspicious, and Pathological fetal states, respectively. The probabilities satisfy

$$p_t^N + p_t^S + p_t^P = 1 \tag{31}$$

The model-derived risk values associated with successive temporal observations are organized according to their chronological order to form a temporal risk trajectory:

$$R = [R_1, R_2, \dots, R_T] \quad (32)$$

where R_t denotes the model-derived risk representation at temporal observation t and T, and denotes the number of successive observations included in the trajectory. The resulting trajectory represents the temporal behaviour of the model-estimated risk and is not interpreted as a directly measured physiological quantity or as a clinically validated risk score.

The continuous risk trajectory is subsequently used to represent the corresponding fetal-state sequence. Two predefined threshold values, τ_1 and τ_2 are used to divide the risk representation into three ordered state regions:

$$s_t = \begin{cases} N, & R_t < \tau_1 \\ S, & \tau_1 \le R_t < \tau_2 \\ P, & R_t \ge \tau_2 \end{cases} \quad (33)$$

where s_t denotes the model-estimated fetal state N, S and , P, and correspond to the Normal, Suspicious, and Pathological states, respectively.

The resulting temporal state sequence is represented as

$$S = [s_1, s_2, \dots, s_T] \quad (34)$$

The ordered state sequence allows changes between successive fetal-state estimates to be represented as temporal transitions:$s_t \rightarrow s_{t+1}$

A transition may represent persistence of the same estimated state or a change between the Normal, Suspicious, and Pathological states. Consequently, the temporal risk trajectory provides a common representation for examining both the magnitude of estimated risk and its evolution across successive monitoring observations.

The trajectory and state sequence are subsequently used in the transition-analysis stage to characterize temporal persistence and changes in fetal-state estimates. The resulting temporal information is then passed to the multi-horizon prediction stage, where future fetal-state distributions are evaluated at the predefined 5-, 10-, 15-, and 30-minute horizons. This explicitly distinguishes the temporal contexts used to construct the model representation from the future intervals used for prediction.

3.9 Fetal-State Transition Modelling

The fetal-state sequence obtained from the temporal risk trajectory in Section 3.8 is used to characterize changes and persistence between successive model-estimated fetal states. The three states considered in the framework are Normal (N), Suspicious (S), and Pathological (P).

For two consecutive temporal observations, the state transition is represented as

$$s_t \rightarrow s_{t+1} \quad (35)$$

where s_t and s_{t+1} denote the estimated fetal states at consecutive temporal observations.

The possible transitions can be represented using a three-state transition matrix:

$$P_{ij} = Pr(s_{t+1} = j | s_t = i) \quad (36)$$

where, $i, j \in \{\mathrm{N}, \mathrm{S}, \mathrm{P}\}$,and P_{ij} denotes the estimated transition probability from state i to state j. The transition matrix is expressed as

$$P = \begin{bmatrix} P_{NN} & P_{NS} & P_{NP} \\ P_{SN} & P_{SS} & P_{SP} \\ P_{PN} & P_{PS} & P_{PP} \end{bmatrix} \quad (37)$$

where the diagonal elements represent state persistence, while the off-diagonal elements represent state changes between successive temporal observations. Each row of the transition matrix satisfies

$$\sum_{j\in\{N,S,P\}} P_{ij} = 1 \quad (38)$$

The transition representation allows the temporal behaviour of the model-estimated fetal state to be characterized in terms of persistence and changes between the three states.

For a complete temporal sequence, the state transitions are represented as

$$\mathcal{T} = \{(s_1, s_2), (s_2, s_3), \dots, (s_{T-1}, s_T)\} \quad (39)$$

where $\mathcal{T}$ denotes the set of consecutive state transitions in the analysed temporal sequence.

The transition representation is subsequently used as temporal information for future-state prediction. In this way, the framework does not rely only on the current estimated fetal state but also retains information about how the estimated state has evolved across preceding observations. This provides the transition-aware component of the E-TRTL formulation and connects the temporal risk trajectory to the multi-horizon prediction stage.

3.10 Multi-Horizon Future-State Prediction

The temporal risk trajectory and fetal-state transition representation obtained in Sections 3.8 and 3.9 are used to predict the future fetal state at predefined horizons of 5, 10, 15, and 30 minutes.

Let the set of prediction horizons be

$$\mathcal{H} = \{5,10,15,30\} \quad (40)$$

For a temporal observation t , the predicted fetal-state probability distribution at horizon $h \in H$ is defined as

$$\hat{p}_{t+h} = f_h(\mathcal{Z}_t) \quad (41)$$

where $\mathcal{Z}_t$ denotes the temporal representation available up to observation t , $f_h(.)$ denotes the prediction function for horizon h , and $\hat{p}_{t+h}$represents the predicted probabilities of the Normal, Suspicious, and Pathological states.

The predicted fetal state is obtained as

$$\hat{s}_{t+h} = \arg\max_{c\in\{N,S,P\}} \hat{p}^{c}_{t+h} \quad (42)$$

where c denotes the three fetal-state classes. The 5-, 10-, and 20-minute windows represent the temporal contexts used as model inputs, whereas the 5-, 10-, 15-, and 30-minute intervals represent the future prediction horizons.

3.11 Explainability Analysis

Explainability analysis is applied to examine the model-derived features associated with the fetal-state predictions. The analysis considers the contribution of the learned representation to the final prediction and provides complementary information on the features associated with the model output.

Feature-attribution analysis is used to identify features that have greater influence on the predicted fetal state. The resulting feature contributions are examined together with the temporal risk trajectory to provide an interpretable representation of the model's prediction behaviour.

The complete processing sequence of the proposed E-TRTL framework integrates the individual stages described in Sections 3.3–3.11. The synchronized FHR and UC signals are first subjected to signal-quality assessment and preprocessing, followed by the generation of 5-, 10-, and 20-minute temporal representations. These representations are processed through temporal feature-learning branches and attention-based CTG fusion. The resulting CTG representation is then integrated with the available clinical-obstetric information. The integrated representation is subsequently used for model-estimated fetal-risk calculation and temporal risk trajectory construction. The trajectory is used to characterize the sequence and transitions of the estimated fetal states, followed by future-state prediction at 5-, 10-, 15-, and 30-minute horizons. Finally, SHAP-based feature attribution is used to examine the model-derived features associated with the predictions.

4. Dataset and Experimental Setup

The E-TRTL framework was evaluated using the CTU-UHB Intrapartum Cardiotocography Database, which contains 552 intrapartum CTG recordings with synchronized fetal heart rate (FHR) and uterine contraction (UC) signals sampled at 4 Hz [23]. The database also provides associated clinical and obstetric information for the corresponding recordings. The recordings were subjected to the signal-quality assessment and preprocessing procedures described in Section 3.4 and represented using 5-, 10-, and 20-minute temporal contexts, corresponding to 1200, 2400, and 4800 samples, respectively, for the two CTG channels. The fetal-state categories considered in the study were Normal, Suspicious, and Pathological.

The 552 recordings were partitioned at the record level into 386 training recordings, 83 validation recordings, and 83 testing recordings. Temporal windows generated from an individual recording were retained within the same partition to prevent windows from the same recording from being distributed across different subsets. Following temporal representation and fusion of the signal-derived and clinical-obstetric features, a 203-dimensional feature representation was provided as input to the E-TRTL multi-horizon prediction model. The model was implemented using TensorFlow/Keras and consists of fully connected layers with 128 and 64 units, respectively, using ReLU activation, with a dropout rate of 0.3 between the layers. Four sigmoid output units were used to generate predictions for the 5-, 10-, 15-, and 30-minute prediction horizons. The model was trained using the Adam optimizer for 20 epochs with a batch size of 64. Mean squared error (MSE) was used as the loss function, and mean absolute error (MAE) was used as the evaluation metric for the prediction outputs. The 5-, 10-, and 20-minute windows represent the temporal contexts used to construct the model input, whereas the 5-, 10-, 15-, and 30-minute intervals represent the future prediction horizons. The prediction outputs were evaluated separately for each predefined horizon using the evaluation measures described in the following section.

5. Performance Metrics

The performance of the proposed **E-TRTL framework** was evaluated using accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). For the multiclass fetal-state classification task, macro-averaged precision, recall, and F1-score were considered to provide equal weighting to the Normal, Suspicious, and Pathological classes. One-vs-rest ROC analysis was used to assess the discriminative performance of each fetal-state category. Confusion matrices were additionally considered to examine class-wise prediction behaviour.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{43}$$

For binary or class-wise evaluation, precision, recall, and F1-score are defined as:

$$Precision = \frac{TP}{TP+FP} \quad (44)$$

$$Recall = \frac{TP}{TP+FN} \quad (45)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (46)$$

$$AUC = \int_0^1 TPR(FPR)\, d(FPR) \quad (47)$$

where,

$TPR = \frac{TP}{TP+FN}$ and $FPR = \frac{FP}{FP+TN}$

where:

- TP= True Positives
- TN= True Negatives
- FP= False Positives
- FN= False Negatives

For multiclass evaluation, the macro-average is calculated by independently evaluating each class and then taking the arithmetic mean across the three fetal-state categories. The ROC-AUC values were calculated using a one-vs-rest strategy for the Normal, Suspicious, and Pathological classes.

6. Results and Discussions

The proposed Explainable Temporal Risk Trajectory Learning (E-TRTL) framework was tested using CTU-UHB Intrapartum Cardiotocography Database. The model was evaluated by the following metrics: Accuracy, Precision, Recall, F1-Score, and Area Under the ROC Curve (AUC). The proposed framework showed good fetal distress prediction ability through the effective modelling of fetal state transition processes as well as temporal risk evolution processes.

6.1 CTG Signal Acquisition and Preprocessing

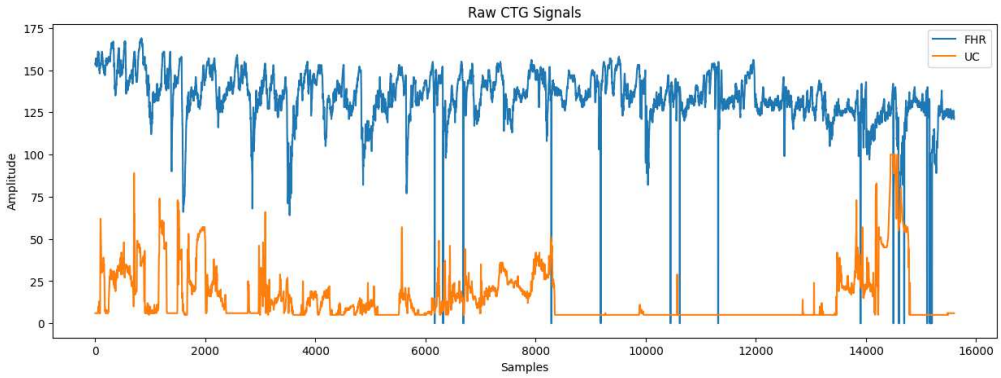


Figure 3: Raw CTG Signals

Figure 3 presents a representative raw CTG recording containing synchronized fetal heart rate (FHR) and uterine contraction (UC) signals. These signals constitute the primary physiological inputs used by the E-TRTL framework. The FHR signal shows temporal variations in heart rate, while the UC signal represents uterine contraction activity. Adaptive Signal Enhancement

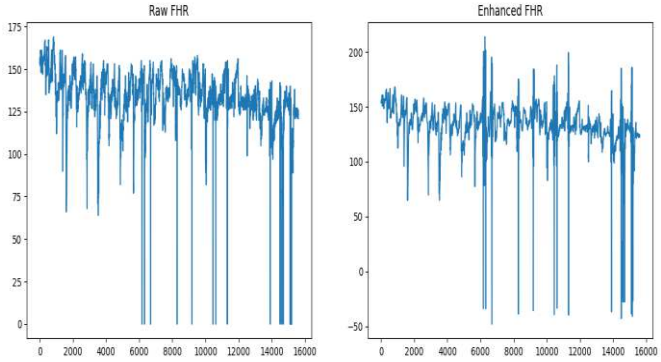


Figure 4: Signal Enhancement

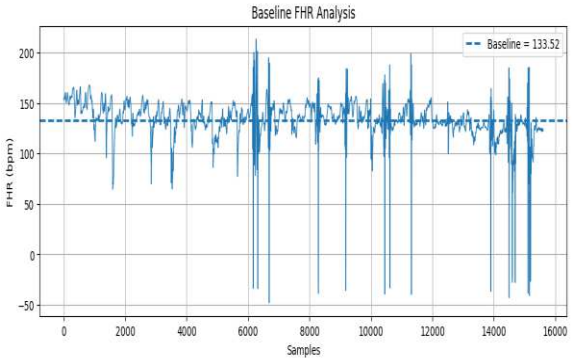


Figure 5: Baseline FHR Analysis

Figure 4 presents the processed CTG signal after the preprocessing procedure described in Section 3.4. The processed signal provides the quality-controlled representation used for the subsequent temporal modelling stages. Figure 5 represents the baseline FHR and signal-quality analysis for the representative recording. The baseline provides a reference for the temporal FHR behaviour within the analysed segment, while the signal-quality analysis is used to assess the suitability of the recording for subsequent processing.

6.2 Multi-Resolution Temporal Representation

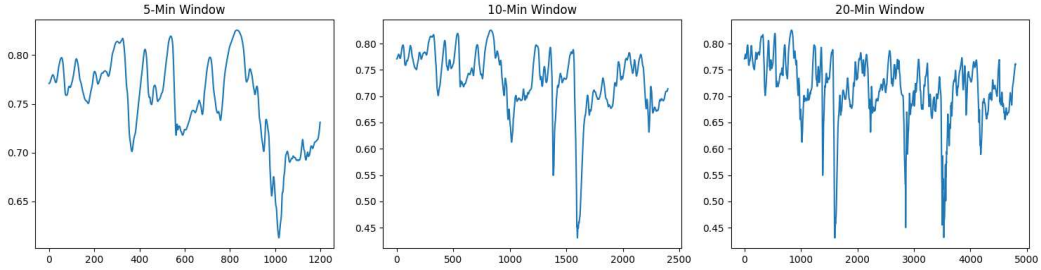


Figure 6: Multi-Resolution Temporal Window Generation

Figure 6 illustrates the multi-resolution temporal representation generated using 5-, 10-, and 20-minute CTG windows. The three windows provide short-, intermediate-, and longer-duration temporal contexts of the synchronized FHR and UC signals.The 5-minute window represents a shorter temporal context, while the 10- and 20-minute windows provide progressively longer temporal information. These representations are subsequently processed through the corresponding temporal feature-learning branches and combined using the attention-based CTG fusion mechanism described in Section 3.6.

6.3 Temporal Risk Trajectory

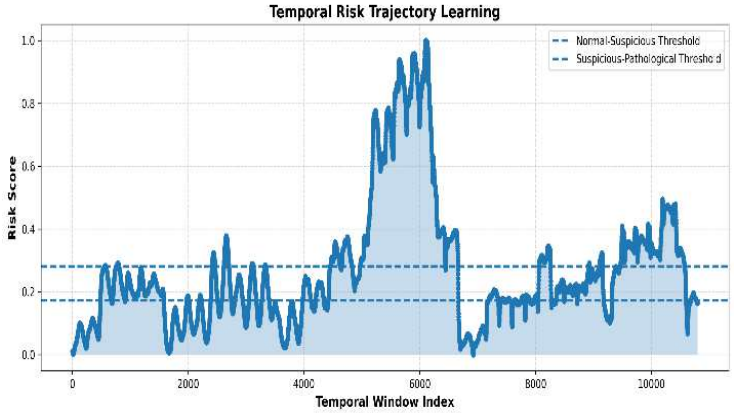


Figure 7. Temporal Risk Trajectory

Figure 7 presents the temporal risk trajectory generated from successive CTG observations. The vertical axis represents the model-estimated risk score, while the horizontal axis represents the temporal window index. The two horizontal dashed lines indicate the thresholds used to separate the Normal, Suspicious, and Pathological risk regions. The estimated risk varies across the temporal sequence, with relatively lower values during several portions of the recording and higher values in the middle portion of the trajectory. A pronounced increase is observed approximately between temporal windows 5000 and 6500, where the estimated risk reaches values close to 1.0. The risk subsequently decreases and remains at comparatively lower levels, with additional moderate variations toward the later windows. These temporal variations indicate that the E-TRTL representation captures changes in model-estimated risk across successive observations. The resulting risk trajectory is subsequently used for fetal-state transition modelling and multi-horizon prediction.

6.4 Temporal Fetal-State Evolution

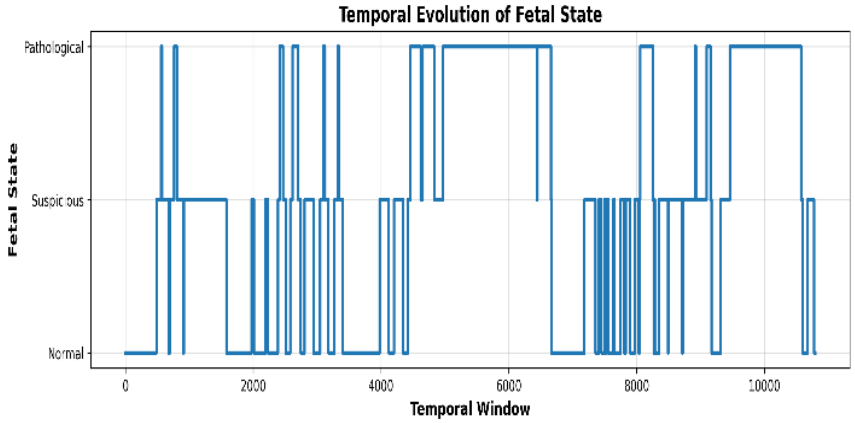


Figure 8. Temporal Evolution of Fetal State

Figure 8 presents the temporal evolution of the model-predicted fetal state across successive temporal windows. The estimated risk values are mapped to the three predefined states—Normal, Suspicious, and Pathological using the threshold levels described in Section 3.8. The figure shows changes between the three predicted states across the monitoring sequence. Several intervals exhibit persistence of the same predicted state, while other intervals show transitions between Normal, Suspicious, and Pathological states. A number of longer Pathological-state intervals are also observed in the latter portions of the sequence.

This temporal state sequence provides the basis for the fetal-state transition modelling described in Section 3.9, where successive predicted states are examined to characterize state persistence and transitions over time.

6.5 Fetal-State Transition Analysis

Figure 9 presents the temporal transition structure among the three model-predicted fetal states, namely Normal (N), Suspicious (S), and Pathological (P). Each node represents a predicted state, while the directed connections represent transitions observed between successive temporal observations. Self-

transitions indicate persistence of the same predicted state across consecutive observations.The transition network shows that the predicted states exhibit both persistence and changes across the temporal sequence. The represented transitions include persistence within the Normal, Suspicious, and Pathological states, together with transitions between selected states.This transition representation provides a compact description of the temporal relationship between successive model-predicted states and forms the basis for the subsequent multi-horizon prediction stage.

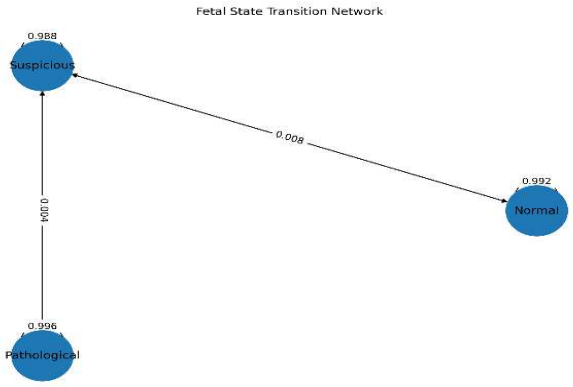


Figure 9. Fetal-State Transition Network

6.6 Multi-Horizon Prediction

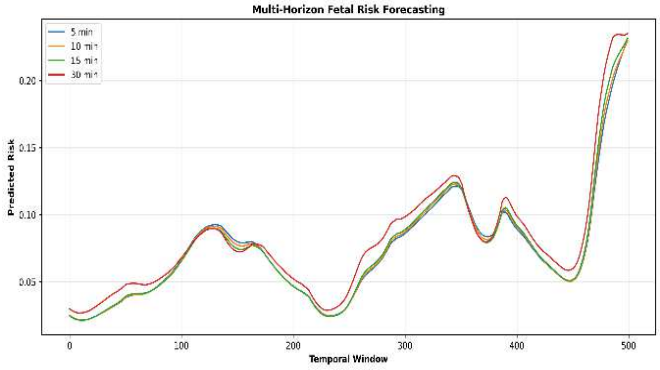
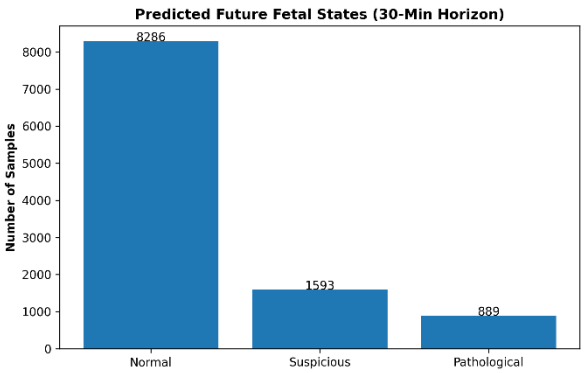


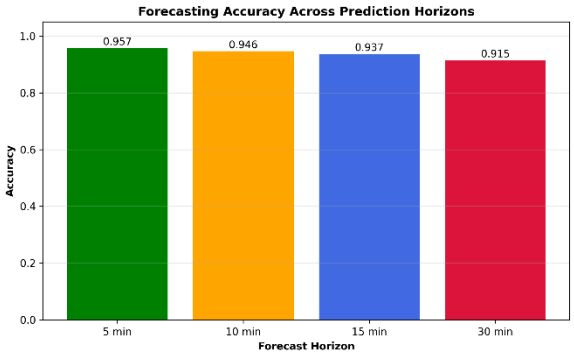
Figure 10. Multi-Horizon Fetal Risk Forecasting

Figure 10 presents the model-estimated risk forecasts for the predefined 5-, 10-, 15-, and 30-minute prediction horizons. The forecasts are generated from the temporal representation described in the preceding sections and provide estimates at different future intervals.

The forecasting curves show variation in the estimated risk across the temporal sequence. The predicted trajectories exhibit broadly similar temporal patterns across the four horizons, while differences in the estimated values are observed as the prediction horizon increases. These differences reflect the change in the temporal distance between the available observation and the corresponding future target.



(a) Predicted Future Fetal States at the 30-minute Horizon



(b) Performance Across Prediction Horizons.

Figure 11. Future-State Prediction and Performance Across Prediction Horizons.

Figure 11 presents the future-state prediction results and the model performance across the predefined prediction horizons. Figure 11(a) shows the distribution of the model-predicted fetal states at the 30-minute horizon. Among the evaluated temporal samples, 8,286 were predicted as Normal, 1,593 as Suspicious, and 889 as Pathological. Figure 11(b) presents the reported performance values for the 5-, 10-, 15-, and 30-minute prediction horizons. The corresponding values are 0.957, 0.946, 0.937, and 0.915, respectively. The values show a gradual reduction in the reported performance measure as the prediction horizon increases. These results indicate that the prediction performance varies with the temporal distance of the future prediction. The multi-horizon formulation therefore provides a separate assessment of model behaviour at each predefined prediction interval.

6.7 Explainability Analysis

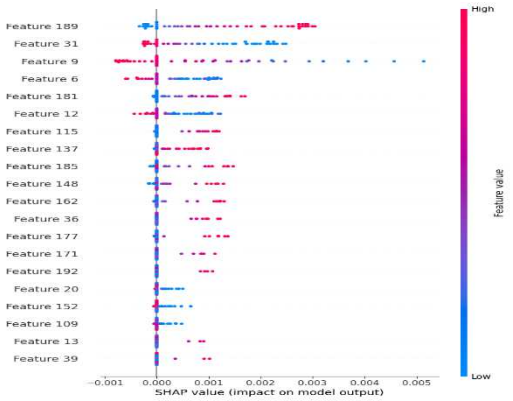


Figure 12: SHAP Summary Plot of Model Feature Contributions

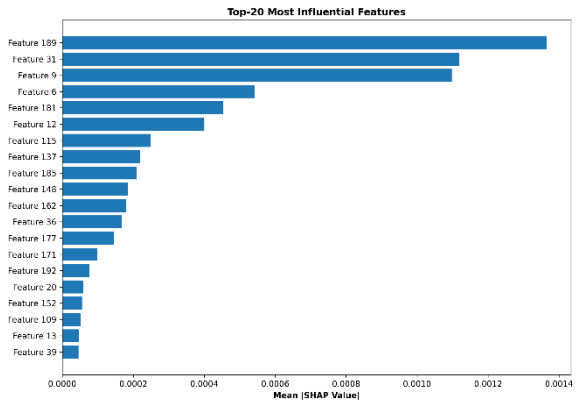


Figure 13: Top-20 Features Ranked by Mean Absolute SHAP Value

Figure 12 presents the SHAP summary plot of the model predictions. The plot shows the distribution of SHAP values for the displayed input features, with the horizontal position indicating the contribution of a feature to the model output and the colour representing the corresponding feature value. Figure 13 presents the top-20 features ranked according to their mean absolute SHAP values. Feature 189 has the highest mean absolute SHAP value among the displayed features, followed by Features 31, 9, 6, and

181. The ranking provides a model-based view of the relative contribution of the input features to the predictions.These SHAP results are interpreted as model-derived feature contributions and do not establish causal physiological relationships or independent clinical importance of the identified features.

6.8 Multi-Class Classification Evaluation

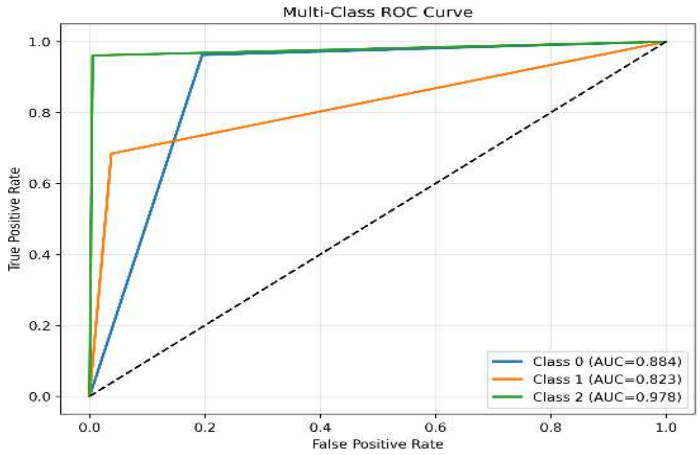


Figure 14. Multi-Class ROC Curve Analysis of E-TRTL for Normal, Suspicious, and Pathological Fetal States

The E-TRTL framework achieved an overall multiclass classification accuracy of 91.48%. The corresponding macro-averaged precision, recall, and F1-score were 89.00%, 87.00%, and 88.00%, respectively. Macro averaging was used to provide equal consideration to the Normal, Suspicious, and Pathological categories. These results provide an overall assessment of the multiclass classification performance, while the subsequent ROC-AUC analysis provides class-wise discrimination measures. The multi-class classification performance of the E-TRTL framework was evaluated using one-vs-rest Receiver Operating Characteristic (ROC) analysis for the Normal, Suspicious, and Pathological fetal-state classes. Figure 14 presents the corresponding ROC curves and class-wise Area Under the Curve (AUC) values. The obtained AUC values are 0.884 for the Normal class, 0.823 for the Suspicious class, and 0.978 for the Pathological class. The results show different levels of class-wise discrimination, with the highest AUC observed for the Pathological class and the lowest for the Suspicious class. The comparatively lower AUC for the Suspicious class indicates that this class is more difficult to distinguish from the Normal and Pathological classes in the evaluated data. This difference in class-wise performance is consistent with the three-state formulation adopted in the E-TRTL framework, where Suspicious represents an intermediate state between Normal and Pathological categories. Overall, the ROC-AUC analysis provides a class-wise assessment of the discriminative behaviour of E-TRTL on the evaluated dataset.

6.9 Multi-Horizon Prediction Performance

The E-TRTL framework was evaluated at four prediction horizons of 5, 10, 15, and 30 minutes to assess its performance at different future temporal intervals in Figure 15. The corresponding prediction accuracies were 95.72%, 94.62%, 93.72%, and 91.48%, respectively. The highest accuracy was obtained at the 5-minute horizon (95.72%), while the accuracy at the 30-minute horizon was 91.48%. A gradual reduction in accuracy is observed as the prediction horizon increases. This indicates that prediction performance decreases when the model is required to estimate the fetal state farther ahead from the available temporal observations in Table 2. These results provide a horizon-wise evaluation of the implemented E-TRTL framework and demonstrate how prediction performance varies with the

forecasting interval. The results are interpreted as model performance on the evaluated dataset and are not considered as evidence of clinical effectiveness.

Table 2. Prediction Accuracy of E-TRTL Across Different Forecasting Horizons

Prediction Horizon	Accuracy (%)
5 min	95.72
10 min	94.62
15 min	93.72
30 min	91.48

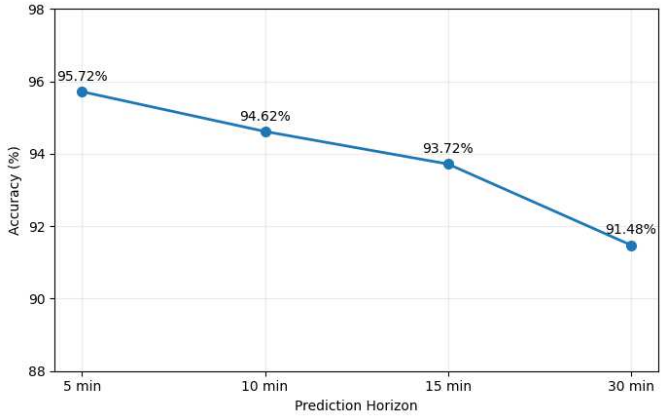


Figure 15. Prediction Accuracy of E-TRTL Across Different Forecasting Horizons

6.10 Continuous Risk-Trajectory Forecasting

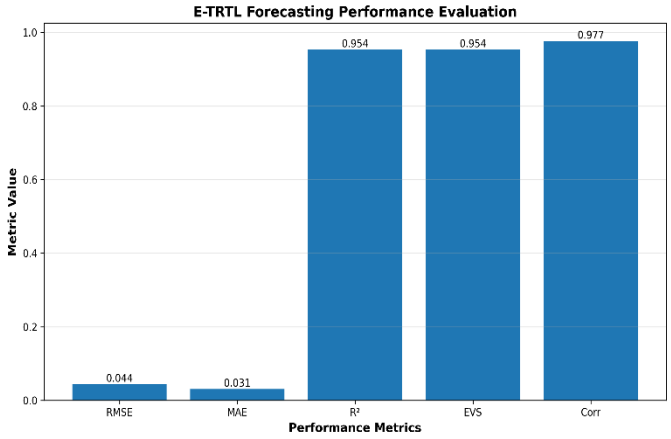


Figure 16: E-TRTL Forecasting Performance Evaluation

The continuous forecasting performance of the E-TRTL framework was evaluated by comparing the predicted and observed fetal-risk trajectories in Figure 16. The obtained Root Mean Square Error (RMSE) was 0.0442, while the Mean Absolute Error (MAE) was 0.0308. The coefficient of determination (R^2) and Explained Variance Score (EVS) were both 0.9536, with a correlation coefficient of 0.9766. The relatively low RMSE and MAE indicate a small numerical difference between the predicted and observed risk values in the evaluated data. The R^2 and EVS values indicate that a substantial proportion of variation in the observed risk trajectory is represented by the predicted

trajectory. The correlation coefficient further indicates a strong association between the predicted and observed risk sequences. These results provide a quantitative assessment of the agreement between the predicted and observed continuous risk trajectories.

6.11 Comparative Performance and Discussion

The proposed E-TRTL framework was compared with previously reported classification-oriented approaches to provide contextual reference for its multiclass risk-state performance. The comparison includes CNN, BiLSTM, CNN-BiLSTM, and the previously developed EM-MHF framework. Accuracy and macro-averaged precision, recall, and F1-score are considered to provide a balanced assessment across the three risk-state categories.

Table 3. Comparison of CNN, BiLSTM, CNN-BiLSTM, and the previously developed EM-MHF framework.

Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1-Score (%)
CNN	94.2	88.5	90.1	89.0
BiLSTM	95.0	91.0	93.0	92.0
CNN-BiLSTM	96.1	94.2	95.4	94.8
EM-MHF	97.0	96.0	93.0	98.0
Proposed E-TRTL	**91.48**	**89.00**	**87.00**	**88.00**

The CNN, BiLSTM, CNN-BiLSTM, and EM-MHF values are previously reported reference results. The E-TRTL values correspond to the present study. Since the approaches may differ in data partitioning, target construction, preprocessing, and evaluation procedures, the values are provided for contextual comparison and are not interpreted as a direct same-protocol performance ranking.In Table 3 the E-TRTL framework achieved an overall accuracy of 91.48%, with macro precision, macro recall, and macro F1-score of 89.00%, 87.00%, and 88.00%, respectively. The use of macro-averaged measures is particularly relevant for multiclass evaluation because it gives equal consideration to the individual risk-state categories rather than allowing the largest category to dominate the aggregate measure. The same evaluation principle was used in the previously reported EM-MHF study. The class-wise ROC analysis produced one-vs-rest AUC values of 0.884 for Normal, 0.823 for Suspicious, and 0.978 for Pathological. These values indicate that the discrimination performance differs across the three evaluated categories, with the Suspicious category showing comparatively lower discrimination than the other two categories.

Beyond multiclass classification, E-TRTL evaluates future-state prediction at multiple temporal horizons. The prediction accuracies were 95.72% at 5 minutes, 94.62% at 10 minutes, 93.72% at 15 minutes, and 91.48% at 30 minutes. The reduction in performance with increasing prediction horizon indicates that forecasting becomes more difficult as the temporal distance from the available observation increases. The continuous risk-trajectory analysis produced an RMSE of 0.0442, MAE of 0.0308, of 0.9536, EVS of 0.9536, and correlation coefficient of 0.9766. These measures provide complementary information regarding the agreement between the predicted and reference risk trajectories. The estimated risk trajectory represents a model-derived quantity and should not be interpreted as a directly measured physiological or clinical severity score. Overall, the comparison is interpreted in terms of modelling scope rather than numerical superiority. The referenced baseline approaches primarily provide classification-oriented predictions, whereas E-TRTL additionally models the temporal evolution of estimated risk and evaluates future-state prediction at multiple predefined horizons. Thus, the principal distinction of E-TRTL is its integration of temporal risk-trajectory modelling, state-transition analysis, multi-horizon forecasting, and explainability within a common framework.

7. Conclusion

This study presented an Explainable Temporal Risk Trajectory Learning (E-TRTL) framework for investigating the temporal evolution of model-estimated fetal risk from intrapartum cardiotocography signals. The framework integrates multi-resolution CTG representation, temporal feature learning, risk-trajectory modelling, fetal-state transition analysis, multi-horizon prediction, and model-level interpretation within a common analytical formulation.The multiclass evaluation achieved an accuracy of 91.48%, with macro precision, macro recall, and macro F1-score of 89.00%, 87.00%, and 88.00%, respectively. The one-vs-rest ROC-AUC values were 0.884, 0.823, and 0.978 for the Normal, Suspicious, and Pathological categories, respectively. In addition to multiclass evaluation, the framework investigated future-state prediction at 5-, 10-, 15-, and 30-minute horizons, with corresponding prediction performance of 95.72%, 94.62%, 93.72%, and 91.48%. The continuous risk-trajectory analysis produced an RMSE of 0.0442, MAE of 0.0308, of 0.9536, EVS of 0.9536, and correlation coefficient of 0.9766.These results indicate that the E-TRTL formulation can be used to examine changes in model-estimated fetal risk across successive monitoring intervals and to investigate future-state prediction at multiple temporal horizons. The comparison with previously reported approaches is considered contextual rather than evidence of direct numerical superiority because the compared approaches may use different data partitions, target definitions, preprocessing procedures, and evaluation protocols.The explainability analysis provides additional information about the features associated with model outputs. These explanations are intended to facilitate interpretation of model behaviour and are not considered evidence of causal physiological relationships. Similarly, the estimated risk trajectory represents a model-derived quantity rather than a directly measured clinical severity score. Future studies may also investigate computationally efficient versions of the framework suitable for real-time deployment and evaluate whether the temporal risk information can provide useful complementary support to established fetal-monitoring practices.

Author Contributions: Conceptualization, Kiruthiga Devi M; methodology, Kiruthiga Devi M; software, Kiruthiga Devi M, Joseph Emerson Raja, and Rajeswari P; validation, Kiruthiga Devi M, Joseph Emerson Raja, and Rajeswari P; formal analysis, Kiruthiga Devi M; investigation, Kiruthiga Devi M; resources, Joseph Emerson Raja and Rajeswari P; data curation, Kiruthiga Devi M; writing—original draft preparation, Kiruthiga Devi M; writing—review and editing, Joseph Emerson Raja and Rajeswari P; visualization, Kiruthiga Devi M; supervision, Joseph Emerson Raja and Rajeswari P; project administration, Joseph Emerson Raja. All authors have read and agreed to the published version of the manuscript.

Compliance with Ethical Standards

Research involving Human Participants and/or Animals: This study used the publicly available CTU-UHB Intrapartum Cardiotocography Database, which is accessible through PhysioNet at https://www.physionet.org/content/ctu-uhb-ctgdb/1.0.0/. The dataset contains de-identified intrapartum cardiotocography recordings and associated clinical information. It is made available for research purposes in accordance with the access and ethical-use requirements specified by the PhysioNet repository. No new data were collected directly from human participants for this study, and the authors did not have access to personally identifiable information.

Disclosure of potential conflicts of interest:

Funding: This research received no external funding. The article processing charge (APC) was funded by Multimedia University (MMU), Malaysia.

Institutional Review Board Statement: Not applicable. This study used the publicly available CTU-UHB Intrapartum Cardiotocography Database from PhysioNet, which contains de-identified data. No new human participants were recruited, and no identifiable personal information was collected by the authors.

Informed Consent Statement: Not applicable. This study used a publicly available, de-identified dataset and did not involve the recruitment of human participants or the collection of identifiable personal information by the authors.

Data Availability Statement: https://www.physionet.org/content/ctu-uhb-ctgdb/1.0.0/ The data used in this study can be freely downloaded from PhysioNet's CTU-UHB Intrapartum Cardiotocography Database and are accessible on basis of the guidelines given by repository.

Clinical trial number: Not Applicable.

Acknowledgements: The authors would like to thank their respective institutions for providing the facilities and support necessary to carry out this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

[1] Mendis, L., Palaniswami, M., Keenan, E., & Brownfoot, F. C. (2024). Rapid detection of fetal compromise using input length invariant deep learning on fetal heart rate signals. Scientific Reports, 14(1), 12615.

[2] Frasch, M. G., Strong, S. B., Nilosek, D., Leaverton, J., & Schifrin, B. S. (2021). Detection of preventable fetal distress during labor from scanned cardiotocogram tracings using deep learning. Frontiers in Pediatrics, 9, 736834.

[3] Mendis, L., Palaniswami, M., Brownfoot, F. C., & Keenan, E. (2023). Computerised cardiotocography analysis for the automated detection of fetal compromise during labour: A review. Bioengineering, 10(9), 1007.

[4] Ayres-de-Campos, D., Spong, C. Y., & Chandraharan, E. (2015). FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. International Journal of Gynecology & Obstetrics, 131(1), 13–24.

[5] Giles-Clark, H. J., Skinner, S. M., Linn, M. M., Rolnik, D. L., & Mol, B. W. (2025). Prediction models for intrapartum fetal hypoxia: A systematic review. European Journal of Obstetrics & Gynecology and Reproductive Biology, 308, 99–111.

[6] Tjoa, E., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793–4813.

[7] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2021). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4–24.

[8] Francis, F., Luz, S., Wu, H., Stock, S. J., & Townsend, R. (2024). Machine learning on cardiotocography data to classify fetal outcomes: A scoping review. Computers in Biology and Medicine, 172, 108220.

[9] Salini, Y., Mohanty, S. N., Ramesh, J. V. N., Yang, M., & Chalapathi, M. M. V. (2024). Cardiotocography data analysis for fetal health classification using machine learning models. Biomedical Signal Processing and Control, 89, 105786.

[10] Mehbodniya, A., Lazar, A. J. P., Webber, J., Sengan, S., & Shankar, K. (2022). Fetal health classification from cardiotocographic data using machine learning. Expert Systems, 39(4), e12899.

[11] Alam, M. T., Khan, M. A. I., Dola, N. N., & Almalki, F. A. (2022). Comparative analysis of machine learning methods for fetal health classification using cardiotocography. Applied Sciences, 12(12), 6055.

[12] O'Sullivan, M., Gabruseva, T., Boylan, G., O'Riordan, M., Lightbody, G., & Marnane, W. (2021). Classification of fetal compromise during labour: Signal processing and feature engineering of the cardiotocograph. Artificial Intelligence in Medicine, 120, 102166.

[13] Daydulo, Y. D., Alim, G. M. N., Islam, M. R., & Moni, M. A. (2023). Deep learning-based fetal distress detection from time–frequency representation of cardiotocography signals. Biomedical Signal Processing and Control, 80, 104286.

[14] Chiou, N., Young-Lin, N., Kelly, C., Cattiau, J., Tiyasirichokchai, T., Diack, A., Koyejo, S., Heller, K., & Asiedu, M. (2025). Development and evaluation of deep learning models for cardiotocography interpretation. **npj Women's Health, 3**, Article 21.

[15] Wang, X., Yu, K., Liang, X., Wang, L., & Han, C. (2026). Artificial intelligence-assisted cardiotocography: Unified model for signal reconstruction, fetal heart rate analysis, and variability assessment. IEEE Journal of Biomedical and Health Informatics.

[16] Park, C. E., Kim, J., Lee, S., & Choi, H. (2025). Automated interpretation of cardiotocography using deep learning. Scientific Reports, 15(1), 1–15.

[17] I. Nazli, E. Korbeko, S. Dogru, E. Kugu, and O. K. Sahingoz, “Early detection of fetal health conditions using machine learning for classifying imbalanced cardiotocographic data,” Diagnostics, vol. 15, no. 10, Art. no. 1250, 2025.

[18] A. H. Dakheel, M. R. Mohammed, Z. A. A. Alhuseen, 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, pp. 195–220, 2025.

[19] I. Ben M'Barek, Q. Jauvion, M. Vitrou, M. Holmström, M. Koskas, and P. F. Ceccaldi, “DeepCTG 1.0: An interpretable model to detect fetal hypoxia during labor and delivery from cardiotocography signals,” Frontiers in Digital Health, vol. 5, Art. no. 1186599, 2023.

[20] Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In Proceedings of the International Conference on Learning Representations (ICLR 2017).

[21] R. Pardasani, R. Vitullo, S. 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, Art. no. 1638424, 2025.

[22] Lundberg, S.M. and Lee, S.I. (2017) A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems (NeurIPS), Long Beach, 4-9 December 2017, 4766-4777

[23] J. Chudáček, V. Spilka, M. Burša, P. Janků, L. Hruban, M. Huptych, and G. Lhotská, “Open access intrapartum CTG database,” BMC Pregnancy and Childbirth, vol. 14, no. 1, pp. 1–12, 2014. DOI: 10.1186/1471-2393-14-16.