*Nanotechnology Perceptions*
*ISSN 1660-6795*
*www.nano-ntp.com*

# Autonomous Fetal Distress Detection with Cardiotocogram Classification using Auto-Encoder Model Combined with Recurrent Neural Networks

**A. Venkata Sriram[1,2], P. Rajesh Kumar[3], L. Alekhya[4]**

[1]*Lecturer, Electronics and Communication Department, Government Polytechnic, Parvathi Puram, Andhra Pradesh, India*
[2]*Electronics and Communication Department, Andhra University College of Engineering, Visakhapatnam, India*
[3]*Professor, Electronics and Communication Department, Andhra University College of Engineering, Visakhapatnam, India*
[4]*Assistant Professor, Electronics and Communication Department, Lendi Institute of Engineering and Technology, Vizianagaram, Andhra Pradesh, India.*

The scarcity and imbalance of high-quality annotated Cardiotocography (CTG) datasets pose significant challenges in developing most accurate fetal distress detection systems. To address this, a Generative AI based data generation technique and features selection method with classification framework was proposed which combines Auto-Encoders (AEs) and Recurrent Neural networks mainly LSTM and GRU models. The Auto-Encoder is employed to generate synthetic CTG data by capturing and reconstructing complex temporal patterns, effectively augmenting the dataset while preserving its physiological characteristics. These RNN models, designed for sequential data processing and feature selection, is then trained on the enriched dataset to classify CTG signals and detect fetal distress autonomously. Experimental results demonstrate that the proposed framework improves performance metrics for LSTM model with RmsProp optimiser when compared to GRU model interms of classification Accuracy of 98.71%, Sensitivity of 98.76%, Specificity of 99.37%, Precision of 98.65%, Weighted $F_1$ score of 98.71%, Mean MCC of 98.06%, Mean kappa score of 98.05% and averaged AUC of 99.77% within elapsed time of 2.54 minutes. Furthermore, this methodology mitigates the effects of dataset imbalance, and enhances the generalizability of fetal distress detection systems.

## 1. Introduction

Fetal distress is a critical condition during pregnancy or labor that signals potential compromise to the fetus's health, often requiring immediate medical intervention. Timely and accurate detection is paramount for reducing adverse outcomes such as neonatal morbidity or mortality. A Cardiotocogram (CTG) is a widely used diagnostic tool [1] providing valuable insights into fetal well-being when compared to other techniques namely fECG, fMCG and fPCG recordings. Despite its clinical significance, the manual interpretation of CTG traces can be subjective.
Signal processing [2] offers a promising avenue for automating fetal distress detection from CTG data. This study focuses on developing an autonomous fetal distress detection system

using a hybrid approach that combines Auto-Encoder (AE) models and Recurrent Neural Networks (RNNs).

An autonomous Fetal distress detection system with the help of the Deep Neural Networks along with Generative AI models is proposed as shown the Fig 1 below that identifies of the fetal distress by using CTG dataset available in UCI repository. The dataset holds three conditions of the fetus Normal, Pathologic and Suspect thus procuring multi class classification problem. To balance the samples of each class using Generative AI model especially Auto encoders are applied on the dataset to generate synthetic CTG samples of high quality with minimum root mean square error.

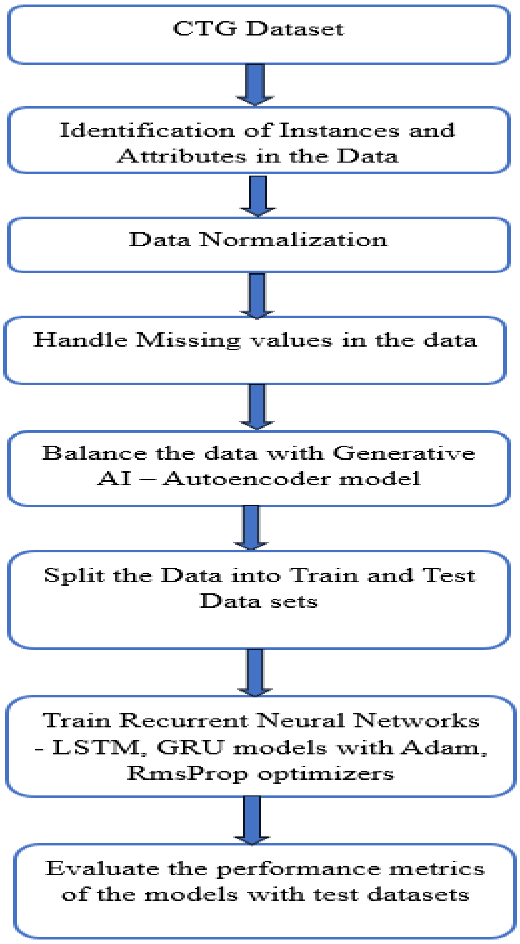



Fig. 1. Steps involved in the Proposed Methodology

## 2. Related Works

Cardiotocograms are critical tools used in obstetrics to monitor fetal well-being and uterine contractions, and advanced classification methods aim to enhance diagnostic accuracy for detecting fetal distress, hypoxia, or any other complications. Automatic detection of any abnormality in the fetus helps in the early diagnosis of the child birth and mother health. A fetal state can be classified by either CTG recordings or by analysing fECG, fMCG or fPCG recordings. Of all these techniques, CTG recordings gave better results for ML-based algorithms.

Castillo et al. in [3] used fECG technique recordings with Wavelet Transform as preprocessing and detection of fQRS on ADFECGDB challenge 2013 dataset with k-medoids method for classification.

Chourasia et al. [5] implemented the WT-HT MLP method on an fPCG real-time dataset and achieved an accuracy of 95.7% for multi-class classification with 10 extracted features from the data. Subha V et al. in [6] has used Firefly optimisation based SVM classification on UCI CTG dataset and achieved an overall accuracy of 91.95%. Yılmaz E et al. in [7] integrated particle swarm optimisation and Least square SVM classification models on the same dataset and achieved an accuracy of 91.62%.
Sharma P et al. in [8] implemented with 11 reduced features and obtained 96.21% for random forest 94.36% for Decision tree and 94.67% for k-NN classifiers. Kocamaz A et al. in [9] used Artificial neural networks (ANN) for classification of CTG dataset into Normal, Suspect and Pathological.
Yılmaz E et al. in [10] used Multi-layer perceptron and probabilistic ANN models on CTG dataset and achieved classification accuracy of 90.35% and 92.15% respectively. Das S et al. in [11] implemented Furia-ANN model on UCI CTG dataset.
The proposed work used Recurrent neural network, LSTM and GRU models on UCI CTG dataset by balancing with Autoencoders which outperforms other techniques interms of classification accuracy, $F_1$ score, MCC and Kappa score and other performance metrics as well.

## 3. Preliminary

### 3.1 Dataset Description

The dataset used is obtained from UCI machine learning repository available in free access online [12]. This dataset is attained from Sisporto 2.0 digital acquisition system and made available for free access. It majorly contains Fetal Cardiotocogram signal extracted feature values with measurements from fetal heart rate (FHR) and Uterine contractions (UC) labelled into three classes namely Normal (N), Pathological (P) and Suspect (S) by obstetricians. The data contains around 2126 records with 1655 Normal, 295 Suspect and 176 pathological which is an imbalance dataset. The entire dataset is of size 2126 x 22 which is associated with 21 features for 2126 records and 1 column contains the three labels of the data.

### 3.2 Data Preprocessing and Visualization

Data preprocessing and visualisation is an important step to apply them to machine learning models. Depending upon the visualized and pre-processed data appropriate model can be used. The process involved in the data preprocessing stage is shown in the Fig. 2. To apply these values to the proposed model, initially the missing values must be handled by either removing them or replacing them with mean of the feature column where that value is present.
Once the missing values are handled, the feature scaling is performed using z-score normalisation to improve convergence speed of gradient descent algorithms and enhances the potential of the machine learning models. The correlation between the scaled features is visualized using the heatmap given in the Fig. 3. Below. After proper preprocessing, the data set can now be applied to the proposed methodology as mentioned in the Fig. 1.

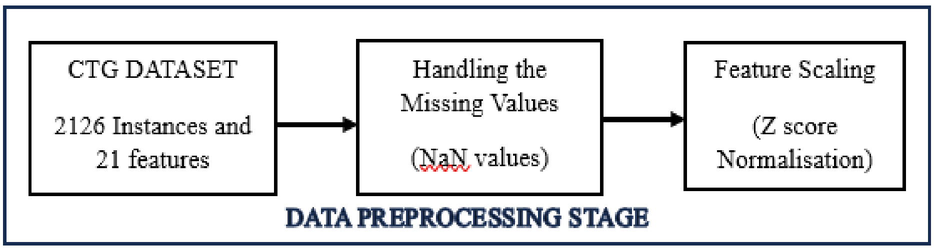



Fig. 2. CTG Data preprocessing

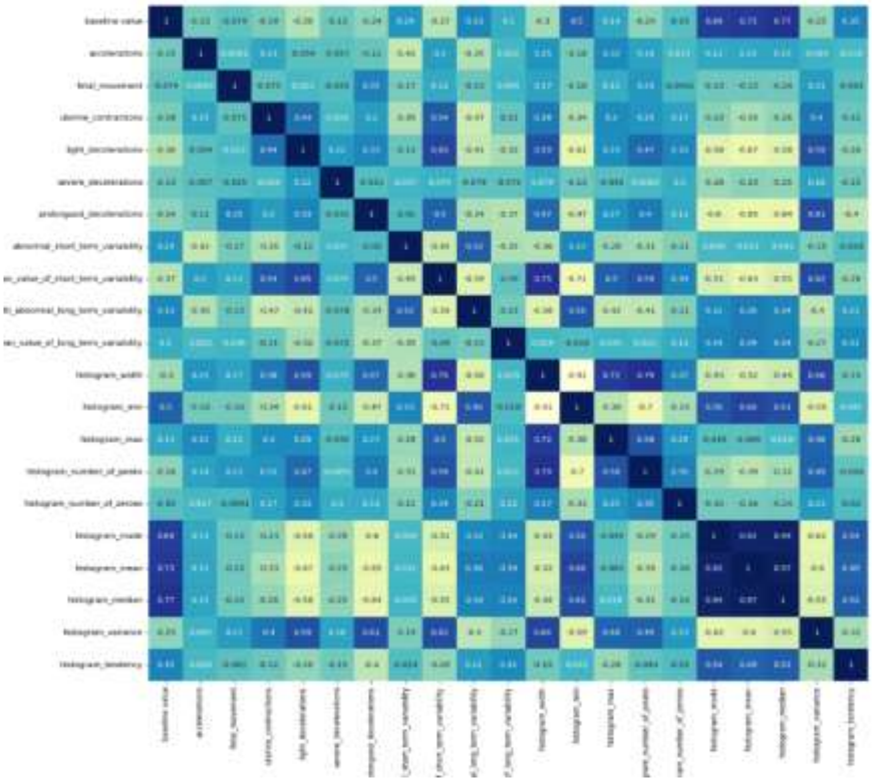


Fig. 3. Heatmap for processed CTG data

### 3.3 Balancing of the Dataset with Auto-Encoders

A Generative AI based approach namely Auto-Encoder model was employed as a technique to generate the samples of the CTG classes that are aptly correlated to the original CTG data. Auto-Encoders are a kind of neural network models comprising of an Encoder, Latent space and a decoder as given in the Fig 4. The encoder component compresses the applied input data into latent space by extracting essential features. The decoder reconstructs the data from the latent space, attempting to reproduce the input as authentically as possible. The quality of the decoder is crucial for generating realistic and high-quality samples.

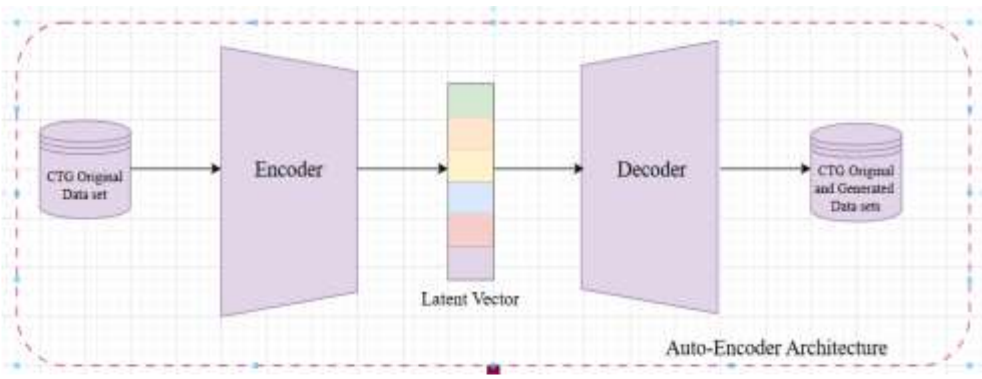



Fig. 4. Internal Architecture of Auto-Encoder

These networks are built in such a way to reduce the error between the original and reconstructed data samples. Out of the many applications of Auto-encoders like dimensionality reduction, Image synthesis Anomaly detection etc., Data Augmentation is one of the most predominantly used application which generate synthetic samples from the learned latent space, a base to Generative AI model.

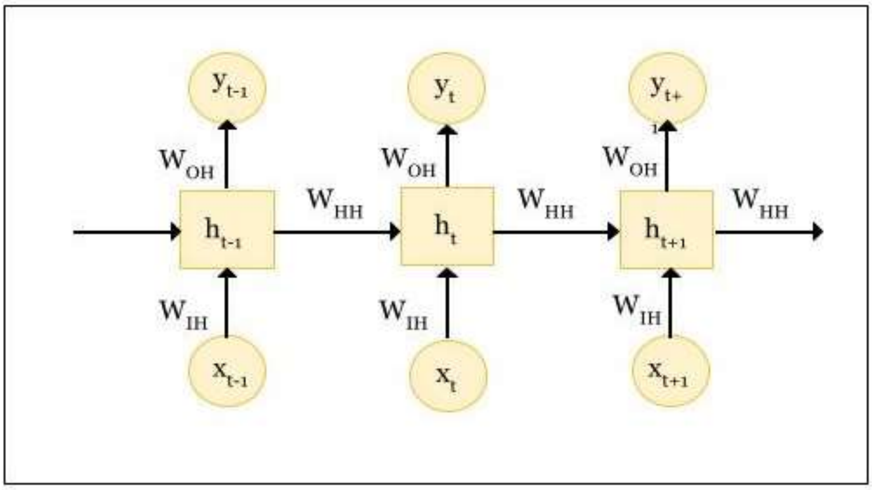



Fig. 5. A basic architecture of RNN

In LSTM-RNN the hidden layer of basic RNN as shown in Fig.5 is replaced by an LSTM cell as given in Fig.6 below.

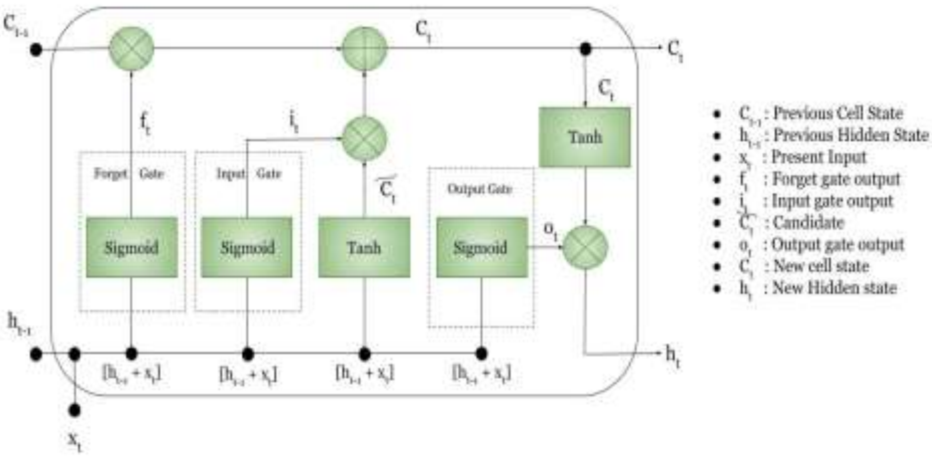



Fig. 6. Long Short-Term Memory Networks basic cell unit

*GRU Architecture:*
Unlike LSTM, it has only two gates in its basic unit cell namely Update gate and Reset gate which play a key role in calculating the present hidden state output vector $h_t$ as shown in the Fig 7 below. Following are the equations are used to analyse the GRU process,

$r_t$= Sigmoid ($W_r\, x_t + U_r\, h_{t-1}$) (3)

$\tilde{h}_t$ = tanh ($W_h\, x_t + U_h(r_t \odot h_{t-1})$) (4)

$z_t$ = Sigmoid ($W_z\, x_t + U_z\, h_{t-1}$ ) (5)

$h_t = (1- zt ) \odot ht\text{-}1 + zt \odot \tilde{h}t$ (6)

Where r and z are commonly the reset and update gates. It can be observed that, GRU is most simple than LSTM, and is more efficient interms of computational cost and adaptability. But LSTMs are more accurate for longer datasets.

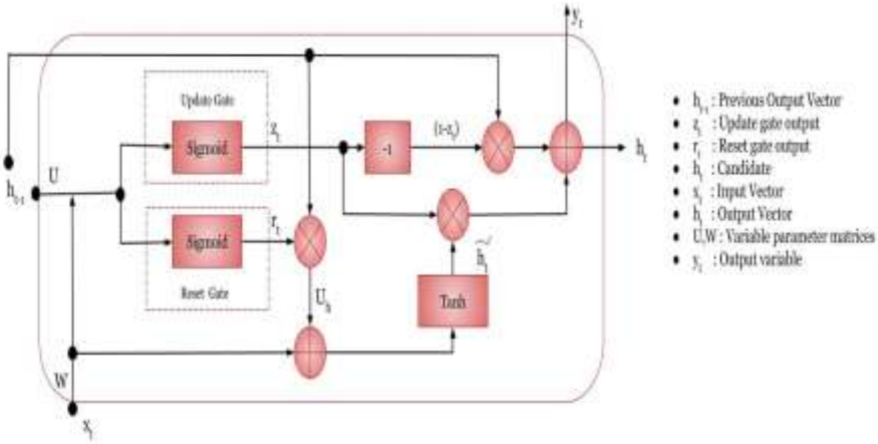



Fig. 7. Gated Recurrent Unit basic cell architecture

Adam and RMSprop optimiser algorithms are used in the work for implementing fetal distress detection system. ADAM is an adaptive moment estimation optimizer that works well for a wide range of deep learning problems. A popular default choice. Whereas RMSprop is an adaptive learning rate optimizer that can handle non-stationary objectives, like those found in recurrent networks.

## 4. Experimental Results

The entire experimental analysis of the proposed work was performed on a computer having an Intel i5 microprocessor and 16 GB Ram and Nvidia GPU of 2 GB, while coding were carried out on MATLAB 2024a version software. The CTG dataset [12] of size 2126 samples with 21 features and 3 classes Normal, Pathological and Suspect is used to analyse the fetal distress and build a detection model. Since the dataset is imbalanced with uneven number of samples for each class, relevant data is generated with Autoencoders with elapsed time of 114.45 seconds making the dataset balanced with 1936 samples for each class. Hence the size of the dataset becomes 5808 x 21 with multi class classification. The balanced CTG dataset along with labels column are divided into train and test dataset in the ratio of 4:1 making 4646 samples for train and 1162 for test data.

The train dataset is now applied to Recurrent neural networks, especially LSTM and GRU models to train the networks with two optimisers ADAM and RMSprop. The test dataset is given to these models with appropriate RNN net training options, like Max Epoch set to 200, MiniBatch size 158, Learning rate to 0.001 for both the optimiser algorithms. The performance metrics for multi-class classification problem as given in [16] were evaluated and given in the Table 1 below. The four combination of confusion matrices are given in the figures below.

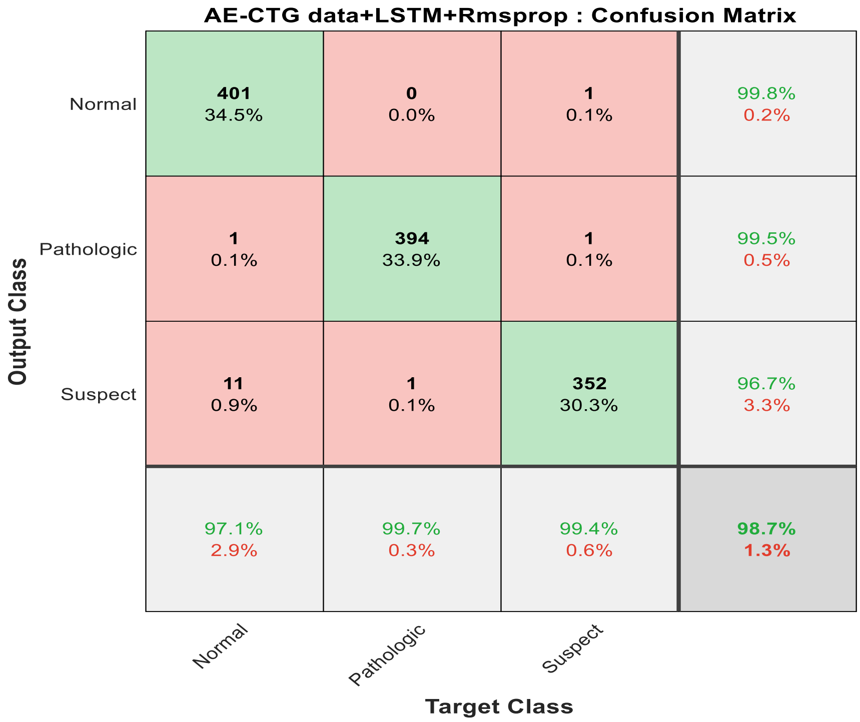



Fig. 8. Confusion matrix of LSTM model with RMSprop optimiser

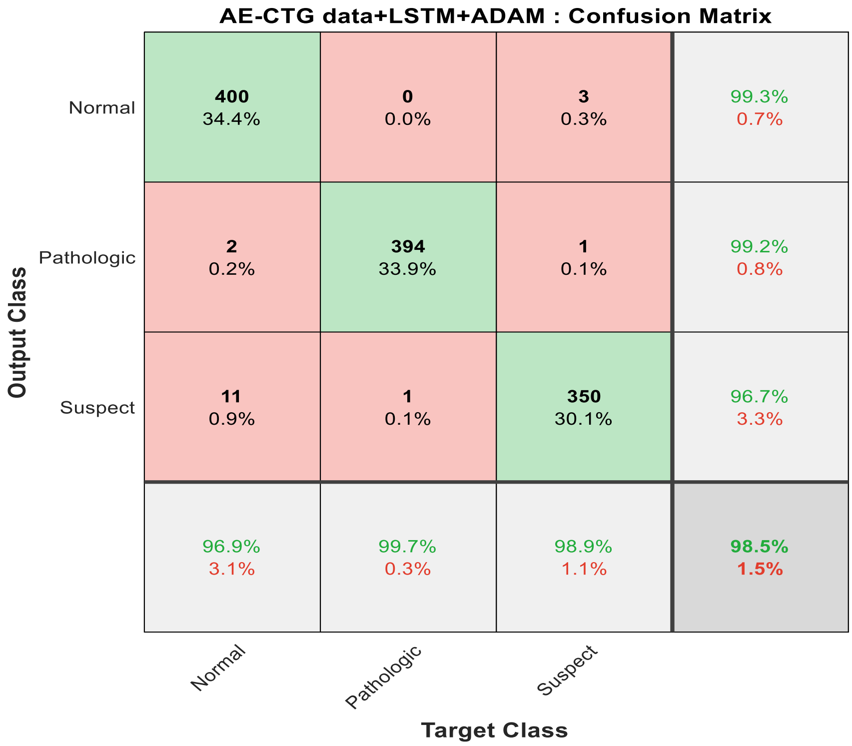



Fig. 9. Confusion matrix of LSTM model with ADAM optimiser

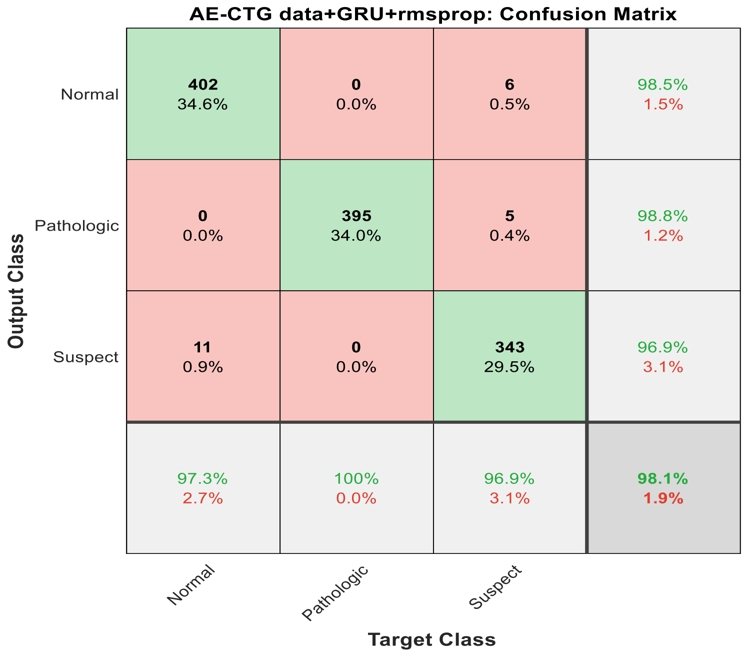



Fig. 10. Confusion matrix of GRU model with RMSprop optimizer

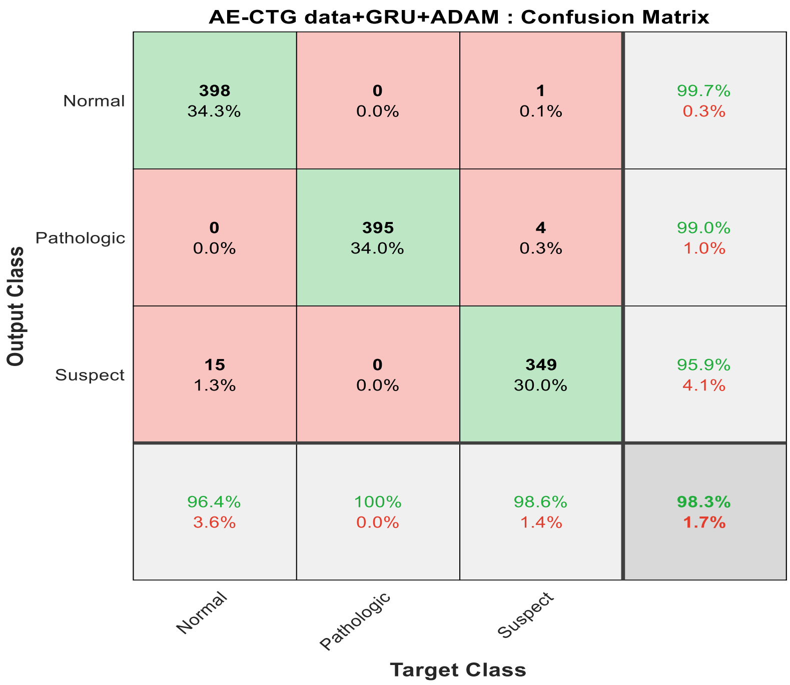


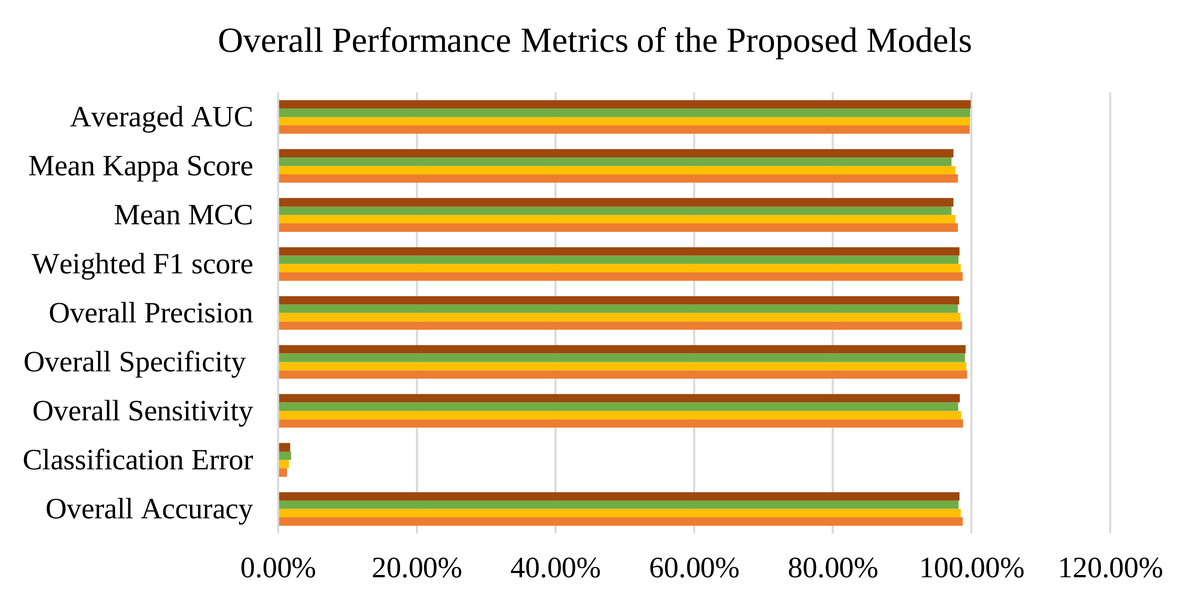

Fig. 11. Confusion matrix of GRU model with ADAM optimiser

From the confusion matrices given in the Fig 8, Fig 9, Fig 10 and Fig 11, the LSTM model with RMSprop optimiser has achieved an accuracy of 98.7% with Autoencoder generated data which outperforms other existing models applied on the CTG dataset. Also, LSTM with ADAM optimiser gave 98.45% which is also next best to other state of the art techniques. Similarly, GRU model with RMSprop gave 98.11% and the same with ADAM gave 98.28% overall classification accuracy. The remaining overall performance metrices are tabulated in the Table 1. given below.

Table 1. Overall Performance Metrics of the AE generated CTG data with RNN Algorithms

| **Performance Metrics** | **LSTM (Rmsprop)** | **LSTM (Adam)** | **GRU (Rmsprop)** | **GRU (Adam)** |
|---|---|---|---|---|
| **Overall Accuracy** | **98.71%** | 98.45% | 98.11% | 98.28% |
| **Classification Error** | 1.29% | 1.55% | 1.89% | 1.72% |
| **Overall Sensitivity** | 98.76% | 98.49% | 98.08% | 98.32% |
| **Overall Specificity** | 99.37% | 99.24% | 99.06% | 99.16% |
| **Overall Precision** | 98.65% | 98.40% | 98.06% | 98.21% |
| **Weighted $F_1$ score** | 98.71% | 98.45% | 98.11% | 98.28% |
| **Mean MCC** | 98.06% | 97.67% | 97.13% | 97.41% |
| **Mean Kappa Score** | 98.05% | 97.66% | 97.12% | 97.39% |
| **Averaged AUC** | 99.77% | 99.79% | 99.80% | 99.87% |
| **Elapsed Time** | 2.54 min | 2.19 min | 4.69 min | 3.44 min |

The comparison of the overall performance metrices is given in the Fig 12, which give a detailed bar graph on metrics of two RNN models, LSTM and GRU with two optimisers ADAM and RMSprop applied on each model, a total of four combination of results are obtained and are compared. Clearly, LSTM with RMSprop optimiser is having best results interms of overall performance metrics. Also the other models LSTM with ADAM, GRU with RMSprop and GRU with ADAM gave better results when compared to other state of the art techniques.




| | Overall Accuracy | Classification Error | Overall Sensitivity | Overall Specificity | Overall Precision | Weighted F1 score | Mean MCC | Mean Kappa Score | Averaged AUC |
|---|---|---|---|---|---|---|---|---|---|
| GRU (Adam) | 98.28% | 1.72% | 98.32% | 99.16% | 98.21% | 98.28% | 97.41% | 97.39% | 99.87% |
| GRU (Rmsprop) | 98.11% | 1.89% | 98.08% | 99.06% | 98.06% | 98.11% | 97.13% | 97.12% | 99.80% |
| LSTM (Adam) | 98.45% | 1.55% | 98.49% | 99.24% | 98.40% | 98.45% | 97.67% | 97.66% | 99.79% |
| LSTM (Rmsprop) | 98.71% | 1.29% | 98.76% | 99.37% | 98.65% | 98.71% | 98.06% | 98.05% | 99.77% |

Fig. 12. Examination of Overall Performance Metrics of the Proposed Models

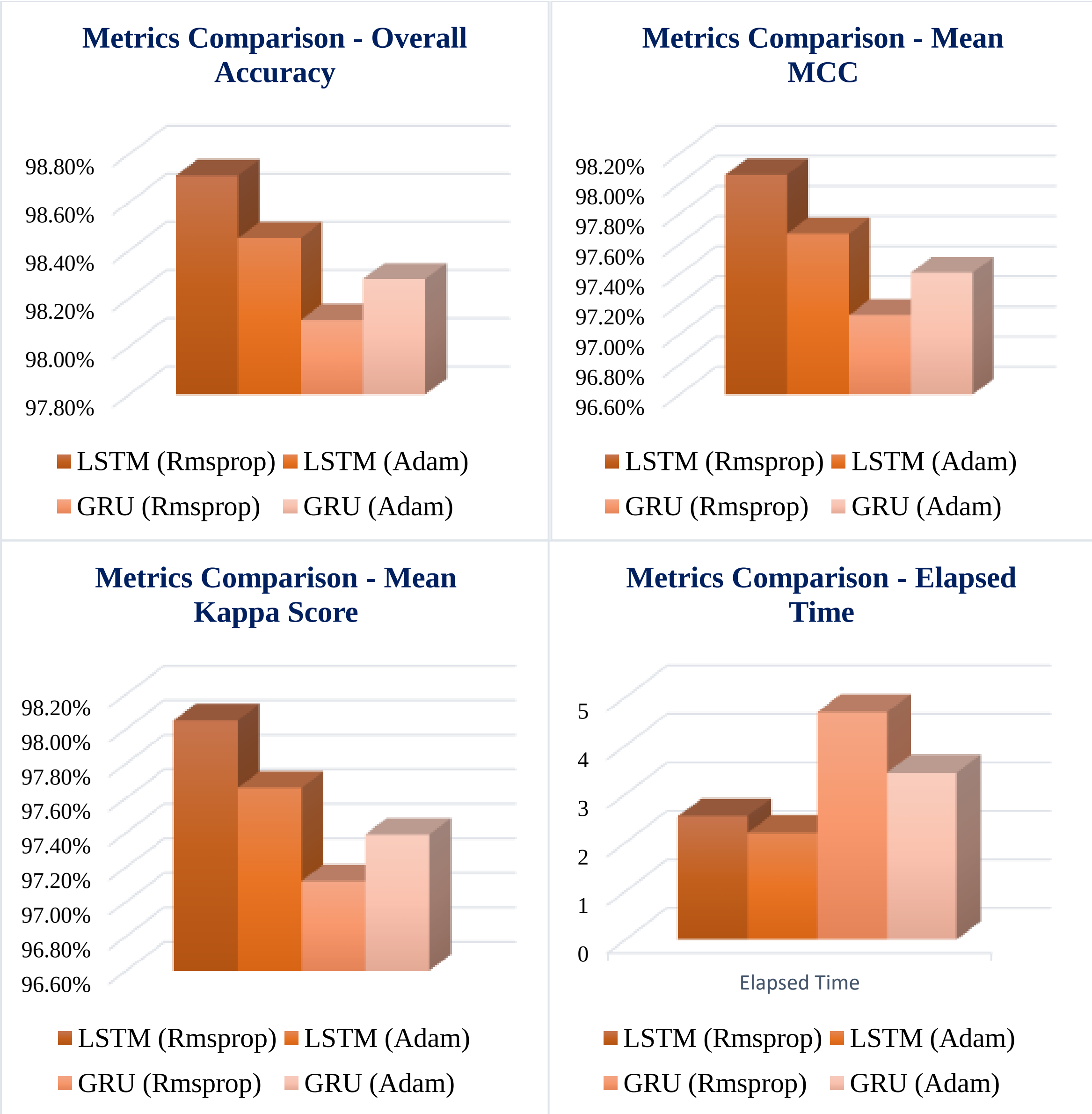



Fig. 13. Metrics Comparison – Data Analysis Graphs

## 5. Conclusion and Future work

The fetal distress detection and analysis of the CTG recordings helps in early diagnosis of fetal health if any abnormalities persist. This helps in reducing the fetal death rate. The CTG dataset which is Imbalanced type is applied to the Auto-Encoder model and make the data balanced by generating samples for each class, Normal, Pathological and Suspect with 1936 samples each. This balanced dataset is split into train and test data, where the train data is applied to Recurrent Neural networks namely LSTM and GRU models with ADAM and RMSprop optimisers which reduce the feature set automatically and build a network that predict the samples of the test dataset and evaluate the various performance metrics of the multi class classification.

Experimental findings of the discussed methodology show that Auto Encoder balanced data when given to RNNs algorithms outperforms other state of the art techniques. This dataset applied to LSTM model with RMSprop optimiser gave best results in terms Overall accuracy as 98.71%, Sensitivity as 98.76%, Specificity as 99.37%, Precision as 98.65%, Weighted F1 score as 98.71%, Mean MCC as 98.06% and Mean Kappa score as 98.05% with elapsed time of 2.54 minutes and average AUC-ROC as 99.77%. This model can be applied to real time system to build IoT based CTG Fetal distress detection model. A quantum-based machine learning algorithms can be further used to reduce the elapsed time and increase the efficiency of the model.

**CRediT authorship contribution statement**

Mr. A Venkata Sriram: Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Dr. P Rajesh Kumar: Writing – review & editing, Supervision Validation, Resources, Project administration, Methodology. Dr. L Alekhya: Writing – review & editing, Data preprocessing, Validation, Software and programming in Python and MATLAB version 2024a.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data Availability**

The Data that is used in the work is available online in UCI machine learning repository with free access.

## References

[1] Ayres-de-Campos D, Spong CY, Chandraharan E; "FIGO Intrapartum Fetal Monitoring Expert Consensus Panel. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography" Int J Gynaecol Obstet. 2015 Oct;131(1):13-24.

[2] Kahankova R, Martinek R, Jaros R, Behbehani K, Matonia A, Jezewski M, Behar JA." A Review of Signal Processing Techniques for Non-Invasive Fetal Electrocardiography". IEEE Rev Biomed Eng. 2020;13:51-73. Epub 2019 Aug 29. PMID: 31478873.

[3] Castillo E, Morales DP, García A et al "A clustering-based method for single-channel fetal heart rate monitoring". PLoS ONE 13(6):e0199308.

[4] Snider D, Xiaowei Xu, "Mining fetal magnetocardiogram data for high-risk fetuses" IEEE international conference on bioinformatics and biomedicine workshops (BIBMW), 2011. IEEE, Atlanta, GA, pp 1066–1068".

[5] Chourasia VS, Tiwari AK, "Fetal heart rate variability analysis from phonocardiographic recordings". J Mech Med Biol 11(05):1315–1331, 2011.

[6] Subha V, Murugan D, "Foetal state determination using support vector machine and firefly optimisation". Int J Knowl Comput Syst 2(2):7–12, 2014.

[7] Yılmaz E, Kılıkçıer Ç, "Determination of fetal state from cardiotocogram using LS-SVM with particle swarm optimization and binary decision tree". Comput Math Methods Med 2013:1–8.

[8] Sharma P, Sharma K,"Fetal state health monitoring using novel Enhanced Binary Bat Algorithm", Computers and Electrical Engineering, Volume 101, 2022, 108035, ISSN 0045-7906.

[9] Kocamaz A et al.," A study of artificial neural network training algorithms for classification of cardiotocography signals". Bitlis Eren Univ J Sci Technol 7(2):93–103, 2017.

[10] Yılmaz E, "Fetal state assessment from cardiotocogram data using artificial neural networks". J Med Biol Eng 36(6):820–832, 2016.

[11] Das S, Roy K, Saha CK, "Fuzzy membership estimation using ANN: a case study in CTG analysis". Proceedings of the 3rd International conference on frontiers of intelligent computing: theory and applications (FICTA) 2014, vol 327. Springer, Cham, p 221–228, 2015.

[12] Campos.D. et al., Cardiotocography. UCI Machine Learning Repository, 2010.

[13] Hassan and A. Mahmood, “Convolutional Recurrent Deep Learning Model for Sentence Classification.” IEEE Access 6:13949-13957, 2018.
[14] Hochreiter, J. Schmidhuber, “Long Short-Term Memory”, Neural Computation, 9(8):1735-1780, 1997.
[15] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evalua-tion of gated recurrent neural networks on sequence modeling", 2014.
[16] A. Venkata Sriram et al. “Autonomous Fetal Distress Detection based on performance of Machine learning algorithms”. *Journal of Computational Analysis and Applications (JoCAAA)*, *33*(06), 1136–1147, 2024.