An Optimized Single Layer Perceptron-based Approach for Cardiotocography Data Classification

Bader Fahad Alkhamees
Department of Information Systems, College of Computer and Information Sciences
King Saud University, Riyadh 11543, Saudi Arabia

***Abstract*—Uterine Contractions (UC) and Fetal Heart Rate (FHR) are the most common techniques for evaluating fetal and maternal assessment during pregnancy and detecting the changes in fetal oxygenation occurred throughout labor. By monitoring the Cardiotocography (CTG) patterns, doctors can measure fetus state, accelerations, heart rate, and uterine contractions. Several computational and machine learning (ML) methods have been done on CTG recordings to improve the effectiveness of fetus analysis and aid the doctors to understand the variations in their interpretation. However, getting an optimal solution and best accuracy remains an important concern. Among the various ML approaches, artificial neural network (ANN)-based approach has achieved a high performance in several applications. In this paper, an optimized Single Layer Perceptron (SLP)-based approach is proposed to classify the CTG data accurately and predict the fetal state. The approach is able to exploit the advantages of SLP model and optimize the learning rate using a grid search method in which we can arrive at the best accuracy and converge to a local minima. The approach is evaluated on CTG dataset of University of California, Irvine (UCI). The optimized SLP model is trained and tested on the dataset using a 10-fold cross-validation technique to classify the CTG patterns as normal, suspect or pathologic. The experimental results show that the proposed approach achieved 99.20% accuracy compared with the state-of-the-art models.**

Keywords—Cardiotocography; machine learning; artificial neural network (ANN); learning rate; grid search; 10-fold cross-validation

I. INTRODUCTION

Nowadays, development and research have been taking place in the healthcare area, and they will continue to expand the analysis by gathering all relevant data from different devices or methods in medical domains, such as medicine, biotechnology, and biomedical [1, 2]. Electronic Fetal Monitoring (EFM) is one of the common methods used to assess fetal well-being during labor [3, 4]. During labor, the fetus is relatively inaccessible as a result, the clinician evaluations depend just on the available and indirect fetal condition measures [5]. The EFM is the most effective method in assessing the fetal status. Also, it is safety through recording maternal Uterine Pressure (UP) and FHR during labor and delivery, that is, the Cardiotocography (CTG) procedure [6]. In addition, nowadays (over 90%) of the labors are electronically monitored by using sensors in order to measure and record UC and FHR [4]. The EFM is commonly used to assess fetal well-being through labor [4]. In the past, before using the EFM the fetal heart rate was measured through the fetal stethoscope [4]. There was a disadvantage of using the stethoscope because it was not able to detect the FHR subtle changes [6]. On the other hand, the EFM is able to overcome this problem by giving the continuous monitoring of the fetus during pregnancy and labor. The EFM has other effective roles. It is able to give minute-by-minute information on the fetus status. Also, it can note the historical information on fetal status accurately and give insight into the stresses on the fetus. The EFM consists of the Cardiotocography which is continuous recording of the fetal heart rate and the (ST) which is used to analyze the fetal electrocardiogram (FECG) [5, 7].

The CTG is the check-up that is generally done during the last three months of pregnancy. This test is for checking the heartbeats of the baby if it has a regular rate and changeability [8]. In general, the baby's heart rate is 110 to 160 beats for each minute [9]. The Cardiotocography expression composed of: (cardio-) which means the fetal heartbeat, (-toco-) which means Uterine contractions and (-graphy) which means recording [9]. EFM enables us to detect the "at risk fetus" early for medical intervention where the delay or failure in detection of abnormal fetus in the CTG recordings can lead to fetal death or injury; which may be a brain injury [10]. In addition, some reviews on birth-related brain injury cases noted that nearly (50%) of brain-injuries refer to preventable medical errors especially the wrong FHR signal analysis [10]. For example, the recognition of the acceleration and deceleration samples in FHR signal is a fateful matter. Because it is important for the detecting of fetal intrauterine distress where there is any delay or failure in their recognition may cause a fetal injury [10].

More than half of these deaths are the result of improper diagnoses made based on FHR pattern recordings and treatments given to the fetus [11, 12]. Despite its practicality, CTG monitoring may not always be effective, especially in low-risk pregnancies. If fetal pain is incorrectly assessed, it could lead to ineffective treatments, and if fetal well-being is improperly investigated, it could leave out necessary treatments [11]. Difficulties in the interpretation of CTG records require methods for computer-assisted analysis where the computer analysis of classifying the FHR and Uterine Contraction (UC) is generally more accurate than human analysis. Classifying the fetal heart rate and fetal state from CTG patterns using computational methods will provide more accurate and less confused medical intervention decisions which provided in the appropriate time is a very important decision in the fetal life where any fault or delay could lead to fetal death or injury [10].

Using effective machine learning (ML) classification methods for classifying the CTG patterns may increase significantly the performance of predicting the fetal state [13]. Moreover, selecting the appropriate values for the parameters of ML methods is considered as one of the factors that effect on the success of ML-based applications on the various medical data samples [13, 14]. Where, if these parameters have inappropriate values then the classification process will be more difficult during training to get optimal or near optimal solutions. The optimization of ML parameters is the process of choosing the most appropriate values on the data in order to increase the accuracy and performance of ML methods [15].

Among the different ML methods, artificial neural network (ANN) has a high capability to model the relationships between output and input data points, which are complex like data of medical field. However, optimizing Single Layer Perceptron (SLP) parameters is also critical to get a high accuracy output of the medical applications. This paper tries to reduce the risk of misclassification and its negative impact by exploiting the advantages of SLP model and learning rate optimization to arrive at the best accuracy and converge to a local minima.

This study intends to design an optimized SLP-based approach for classifying the CTG data accurately and predicting the fetal state. The developed approach selects the best value for learning rate hyper-parameter and converge to a local minima. The main contributions of this research work are itemised in the following lines.

- An accurate approach using optimized SLP model and a grid search algorithm is proposed to classify the CTG patterns as normal, suspect or pathologic.
- By using the grid search algorithm, an optimal or near-optimal value of learning rate is selected to improve learning process of SLP model, arriving at the best accuracy and converge to a local minima.
- The optimized SLP model is trained and tested on the dataset using a 10-fold cross-validation technique to achieve the diversity in learning the SLP model.
- The performance of the optimized SLP model is evaluated and compared with the recent work on the same CTG dataset.

The rest of the paper is structured as follows: Section II provides the related work of the proposed approach. Section III gives an explanation about the materials and methods of the research study. Section IV describes the experimental results with discussion and outcomes. Finally, Section V summarizes the conclusions and future work.

II. Related Work

There are a number of studies and methods that have been presented by many researchers and professionals on medical data and healthcare [16]. Data mining and machine learning techniques are one of these methods, which are applied in different stages of medical applications, including data collection, storage allocation, analytics, pre-processing, and classification or predication [17]. However, the focus of this section aims to evaluate the various classification approaches and methods used on the Cardiotocography (CTG) data patterns. According to Grivell et al. [11], the evaluation of CTG patterns is an important diagnostic procedure, used to measure the maternity during pregnancy and detect the fetal heart rate.

Using ML, e-health software, and traditional pregnancy tests, Akbulut et al. [18] created a framework for prediction in e-health applications that can assist doctors and pregnant women in identifying congenital abnormalities. In order to process the clinical data and maternal predicting for fetal anomalies, nine binary-classification models that have been trained on 96 pregnant women in the clinical datasets were compared for performance. Decision forest models have the highest forecasting accuracy, at 89.5%, during development tests.

Huang [19] used three different machine learning techniques to analyze the CTG data in order to predict fetal distress. Employing statistical features taken from empirical mode decomposition was suggested by Krupa et al. [20]. (EMD). The sub-band decomposition's extracted features were categorized as either normal or risky. For the test data, they achieved an accuracy of 86%. Another study described a two-step analysis of fetal heart rate data that enables accurate risk prediction. Support Vector Machines (SVM), fuzzy, and multilayer perceptron are used to classify the FHR signals. Sundar et al. [21] implemented a new model that classifies the CTG data using ANN. The F-score and Recall were used to evaluate the performance. They also suggested the use of k-means clustering for CTG classification. Moreover, Ocak and Ertunc [22] used adaptive neuro-fuzzy inference systems (ANFIS) for classifying CTG data patterns. Genetic Algorithm (GA) and SVM methods have been implemented for CTG classification by Ocak [23].

In 2019, Potharaju et al. [24] examined the CTG classification using J48, Ridor, Jrip, NBStar, Kstar, and IBk. They used the SMOTE technique to balance the dataset after realizing the CTG data was inherently unbalanced. Their experimental findings show that a balanced dataset performs classification more effectively than one that is unbalanced. After choosing the top six features, they also experimented with three feature selection techniques to assess the effectiveness.

An associative classification (CBA) model has been put forth by Piri and Mohapatra [25] for the analysis of Cardiotocographic fetal evaluation. When choosing a smaller data set with the most crucial features, they also considered the significance of each feature. Zhang and Zhao's [26] developed a decision-support system for clinical fetal risk diagnosis using a cutting-edge machine learning method through pertinent features, which are extracted from the CTG recordings. Therefore, this research showed that hybrid AdaBoost and PCA were effective for classifying the CTG results and determining the fetal status. AUC, sensitivity, and specificity are just a few of the performance classification criteria that have been adopted.

Rough Neural Network has been proposed by the authors in [27] as one of the most popular data mining methods for categorizing medical data in the classification of Cardiotocography data. They timed the classification process and measured the accuracy outcomes. The WEKA tool uses a variety of algorithms, including neural networks, decision tables, bagging, and the nearest neighbor, among others, to analyze Cardiotocographic data. In a study, Hoodbhoy et al. [28] evaluated the efficiency of using CTG data to identify high-risk fetuses using ML techniques. They developed the prediction task using XGBoost to obtain the fetal outcome specified in the classification model.

Recently, Yan and Han [29] proposed a Cost-Sensitive Stacked Generalization (CSSG) approach for re-sampling based on two layers analytics model to sample the distribution of imbalanced class in the dataset. With 17 imbalanced public data sets, the results indicated that their method performed better in classification than other ensembles and single algorithms. Brown and Mues [30] have set themselves the task of analyzing various approaches used to analyze credit score data that is unbalanced. As there are typically many more defaulted loans than there are non-defaulting observations in a portfolio, imbalanced data sets frequently occur in a loan scoring environment. The appropriateness of predictions for loan default is examined by random forests, gradient boosting, and SVM, in addition to the use of conventional classification methods such as neural networks, logistic regression, and decision making trees.

Some under-sampling methods have been studied, for example, Shang [31] in his paper introduced two kinds of sampling methods for missing data in the networks. These two methods are uniform and non-uniform random sampling. In uniform random sampling, a fraction of data sample (q) is arranged uniformly at a random manner in the dataset. This kind of sampling is commonly used in some other work of different applications. In non-uniform random sampling, the data sample is generated by selecting the nearest neighbors data of a seed point, then, its second nearest neighbors data and so on until a fraction of data sample $(1 - q)$ in the entire dataset are selected.

Additionally, a popular SMOTE algorithm for cardiovascular data are examined by Rahman and Davis [32]. They also recommended a modified under-sampling technique using clusters that can produce high-quality training samples for classifier design while also balancing the data. Yen and Lee [33] utilized clusters-based under-sampling approach to choice representative data for examining the under-sampling impact on the distribution of unbalanced classes, with the goal of improving the classification accuracy for minority classes. The experimental results of the approach showed that the cluster-based approach outperformed the other under-sampling methods used in earlier studies.

To address imbalanced and limited CTG data, Piri and Mohapatra [34] have proposed a different number of re-sampling methods and two ensemble models to balance the training samples and improve the classification task. They applied several ML models such as Support vector machine (SVM), Random Forest (RF), Decision Tree (DT) K-Nearest Neighbors (KNN), Logistic Regression (LR), and Gaussian Naive Bayes (GNB). They achieved a 95% classification accuracy using RF model on the balanced dataset. From the previous studies and related work, still there is a need to optimize the learning process and improve the accuracy of CTG data classification. In this work, an optimized SLP-based approach is proposed to classify accurately the CTG data as normal, suspect or pathologic for predicting the fetal state.

III. Materials and Methods

This section explains the materials and methods of the research study. It describes the dataset used for obtaining the experimental result and gives an explanation about the methods of the proposed approach. In the following subsections, the dataset, methods, and the approach will be given in more detail.

A. CTG Dataset

In this proposed work, the CTG dataset obtained from the UCI repository [35] is chosen for conducting the experiment of the study. It is prepared for CTG trace analysis and the discovery of fundamental attributes. The dataset contains two attributes (FHR pattern and fetal state) represent the target labels and 21 distinguished features (LB-FHR; AC; FM; UC; etc.). Because the target attribute of the study includes the normal, suspect, and pathological classes, the Fetal state attribute is used as class label to solve a three-class problem and determine the fetal health-state using the other attributes; and the FHR pattern class label is removed from the dataset. According to the fetal state target class, the number of instances in the dataset is distributed into three groups: 1655 instances for normal class, 295 instances for suspect class, and 176 instances for pathological class. The distribution of instances for the three target classes is presented in Fig. 1.

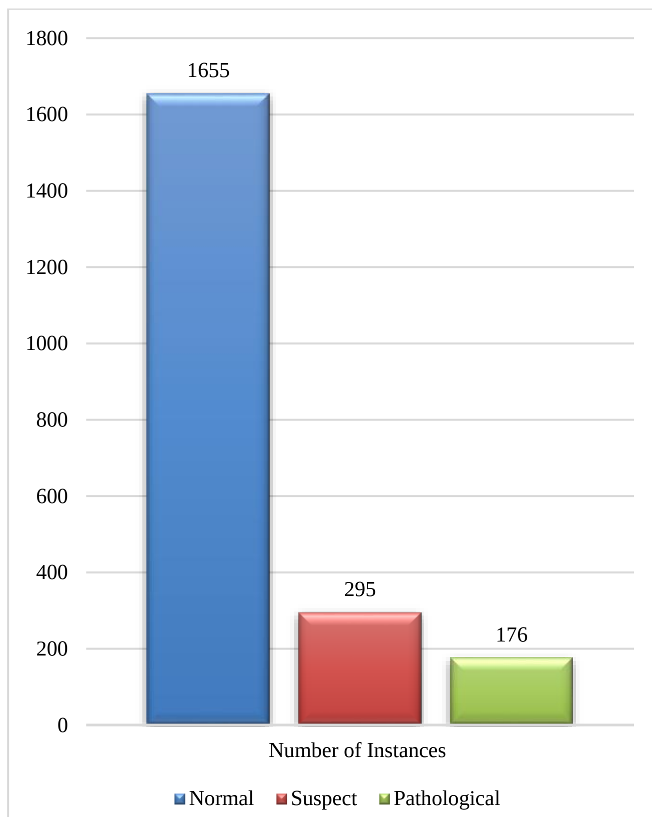

Fig. 1. The Distribution of Instances for the Three Target Classes.

TABLE I. THE ATTRIBUTES INFORMATION OF THE CTG DATASET

No.	Attribute	Description
1	LB–FHR	The baseline (beats per minute)
3	FM	The number of fetal movements for each second
2	AC	The number of accelerations for each second
5	DL	The number of light decelerations for each second
4	UC	The number of uterine contractions for each second
6	DS	The number of severe decelerations for each second
8	ASTV	The percentage of time with abnormal short term variability
7	DP	The number of prolonged decelerations for each second
9	MSTV	The mean value of short term variability
11	MLTV	The mean value of long term variability
10	ALTV	The percentage of time with abnormal long term variability
12	Width	The width of FHR histogram
14	Max	The maximum of FHR histogram
13	Min	The minimum of FHR histogram
15	Nmax	The number of histogram peaks
17	Mode	The histogram mode
16	Nzeros	The number of histogram zeros
18	Mean	The histogram mean
20	Variance	The histogram variance
19	Median	The histogram median
21	Tendency	The histogram tendency
22	CLASS	The FHR pattern class code (1 to 10)
23	NSP	The fetal state class code (N = Normal; S = Suspect; P = Pathologic)

The information of the attributes is given in Table I. As shown in Fig. 1, the number of instance for each class shows that the CTG dataset is imbalanced. Therefore, it is important to build a robust ML method to improve the performance in present of the imbalanced class problem.

B. Single-Layer Perceptron (SLP)

Single-layer perceptron (SLP) is the first and most fundamental model of ANN. The feed-forward neural network is another name for it [36]. The threshold transfer between the nodes serves as the foundation for the SLP’s operation. This is the most basic type of ANN used to solve machine learning problems with linear cases. It is a straightforward neural network with only one layer is known as a single layer perceptron [37]. There are only two layers for input and output in the single layer perceptron (See Fig. 2). The name single layer perceptron refers to the fact that it only has one layer. As opposed to a multilayer perceptron, it lacks hidden layers.

The single layer perceptron computation is the sum of the input vectors multiplied by the corresponding vector weight. The input to an activation function will be the value displayed as the output. Four components make up the perceptron: the input values or one input layer, the weights and biases, the network sum, and the activation function as seen in Fig. 3.

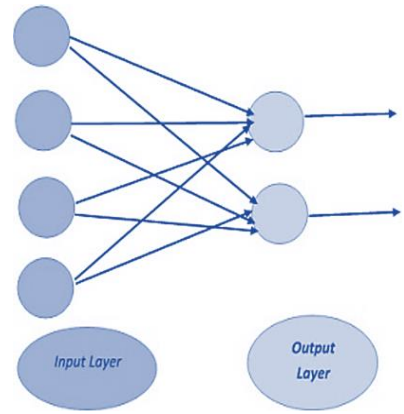

Fig. 2. Two Layers for Input and Output in the Single Layer Perceptron Classes.

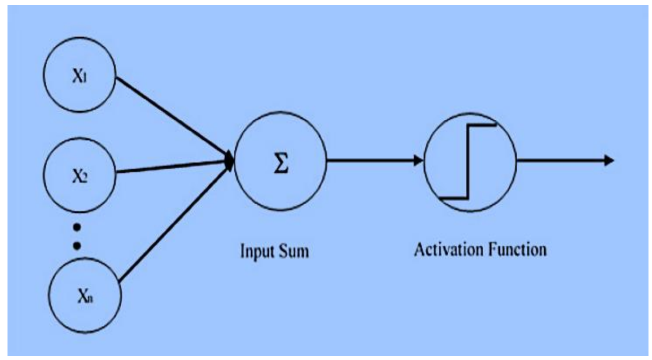

Fig. 3. The Components of Single Layer Perceptron.

C. Grid Search Technique

A grid search or a parameter sweep, it is just a search process through a manually specified subset of the hyper-parameter space of a learning algorithm [38]. It is considered a traditional method of performing hyper-parameter optimization. A performance metric, typically determined by cross-validation on the training set or evaluation on a hold-out validation set that can guide the grid search technique. Before using grid search, it may be necessary to manually set bounds and discretize certain parameters because the parameter space of a machine learning may include real-valued or unbounded value spaces [39].

For instance, a typical classifier with a kernel function, which has a regularization constant *C* and a kernel hyper-parameter γ. They can be tuned for good performance on unseen data. Since both parameters are continuous, one chooses a limited number of "reasonable" values for each in order to perform a grid search, such as:

$C \in \{10, 100, 1000\}$

$\gamma \in \{0.1, 0.2, 0.5, 1.0\}$

Then, using each pair (*C*, γ) in the Cartesian product of these two sets, grid search trains the classifier. The performance of the classifier is then assessed using a held-out validation set (or by internal cross-validation on the training set, in which case multiple classifiers are trained per pair). The settings that received the highest score during the validation process are finally returned by the grid search algorithm. The hyper-parameter settings are typically evaluated independently of one another, allowing the grid search method to operate in parallel [39].

D. Proposed Approach

The proposed approach aims to build an effective and robust SLP classification model that archives a high accuracy for classifying the CTG patterns and predicting the fetal state. It contains two main phases presented in Fig. 4. The approach exploits the advantages of SLP model and optimizes the learning rate using a grid search method in which we can arrive at the best accuracy and converge to a local minima.

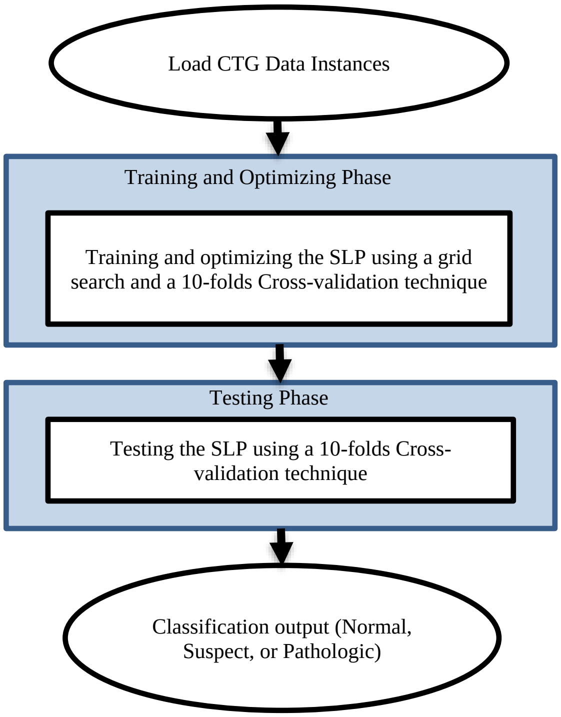

Fig. 4. Flowchart of Proposed Approach.

1) Training and optimizing phase: This subsection explains the training and optimizing phase of the SLP model. The training process is conducted on the CTG dataset using a 10-folds cross-validation technique. In this technique, the dataset is divided into ten subsets and the SLP model is trained using nine out of ten subsets for ten times. The remaining set each time is used in the validation process. The output of this phase is the best model from the ten trained SLP models. Then, the classification result is evaluated through the best model on the whole dataset in the testing phase. The optimizing process of the SLP model is performed to determine the optimal value of the learning rate during the training step by the grid method algorithm. The grid search method selects a set of possible values for the learning rate and evaluates the result for each value.

2) Testing phase: After dividing the dataset to ten subsets and training the SLP model on nine subsets, the remaining subset from the ten subsets is used for testing the trained SLP model. This procedure is repeated for ten times until all ten subsets are involved in the validation process of the trained models. The classification result is obtained by testing the dataset on the best model of trained in the training and optimizing phase for getting the final result of the approach. The classification test result is measured using a number of evaluation metrics such as accuracy, F1-score, precision, and recall, which will be described in the experimental results section.

IV. Experimental Results and Discussions

In this section, the performance of proposed approach is analyzed using a 10-fold cross-validation technique to classify the fetal state from CTG data instances. The experimental results are performed on a laptop with the Intel Core i7, 2.50 GHz CPU, 16GB RAM and Windows 10. The proposed approach is implemented using a WEKA data mining tool. A number of evaluation metrics such as precision, recall, F-score, and accuracy are computed from the classification outputs. These metrics are computed using the following equations:

$$Precision = \frac{TP}{(TP+FP)} \quad (1)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

$$F1\text{-}score = 2 * \left(\frac{Precision*Recall}{Precision+Recall}\right) \quad (3)$$

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (4)$$

Where FN, FP, TN, and TP are the false negative, false positive, true negative, and true positive instances, respectively. The confusion matrix for the classification task is also used. To measure the confusion matrix, FN, FP, TN, and TP are required.

The evaluation results of the approach are also compared with the various recent existing methods and approaches designed for CTG fetal state classification. The experimental parameters of the SLP model are established in Table II.

The values of the model's parameters in Table II are selected experimentally where the learning rate is chosen to be optimized using a grid search method with 10-fold cross-validation technique. The grid search method starts the searching operation among a set of possible values, which are 0.0005, 0.001, 0.005, 0.01, 0.1, 0.15, 0.2, 0.25, and 0.3. Then, the model is trained at each value of these possible learning rate values. The best value of learning rate hyper-parameter is 0.1. Fig. 5 shows the results of accuracy at these different values of learning rate.

TABLE II. Experimental Parameters of Proposed Approach SLP Model

Parameter	**Value**
Hidden nodes	19
Momentum	0.2
Batch Size	100
Learning Rate	0.1
Number of Epochs	500

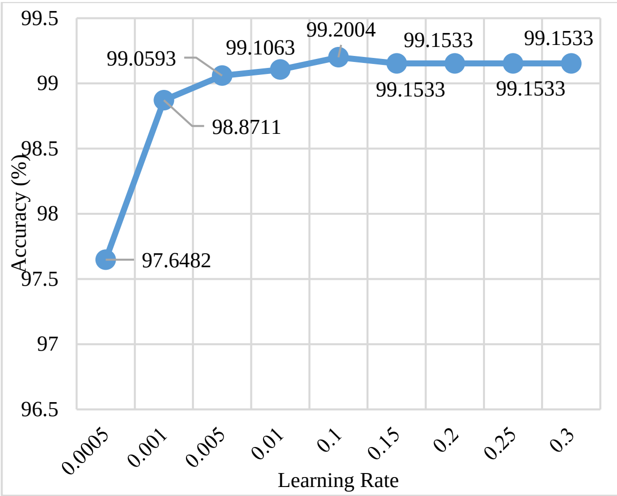

Fig. 5. Accuracies of SLP Model at different Learning Rate.

Also, Fig. 6 demonstrates the classification confusion matrix of optimized SLP model. We can see that 1649 from 1655 normal instances are classified correctly as normal class and 285 of 295 suspect instances are classified correctly as suspect class, as well as 175 of 176 pathologic instances are classified correctly as pathologic class. These classification results confirm the ability of SLP model to classify the CTG patterns in accurate manner.

Predicted Classes \ Actual Classes	Normal	Suspect	Pathologic
Normal	1649	6	0
Suspect	9	285	1
Pathologic	0	1	175

Fig. 6. Confusion Matrix of SLP Model Test Classification.

According the result of confusion matrix in Fig. 6, Table III displays the results of other evaluation metrics. It shows that the model achieves 99.2% of accuracy and 0.992 for weighted average F1-score, recall, and precision metrics.

To highlight the performance of the proposed approach, the accuracy result obtained on the dataset is compared with the other models proposed in the recent current study [34] for classifying CTG data into normal, suspect, and pathologic classes. The visualization of this comparative analysis is presented in Fig. 7.

TABLE III. RESULTS OF EVALUATION METRICS FOR SLP MODEL

Class Name	**Precision**	**Recall**	**F-score**
Normal	0.995	0.996	0.995
Suspect	0.976	0.966	0.971
Pathologic	0.994	0.994	0.994
Weighted Avg.	0.992	0.992	0.992
Accuracy	99.2004%		

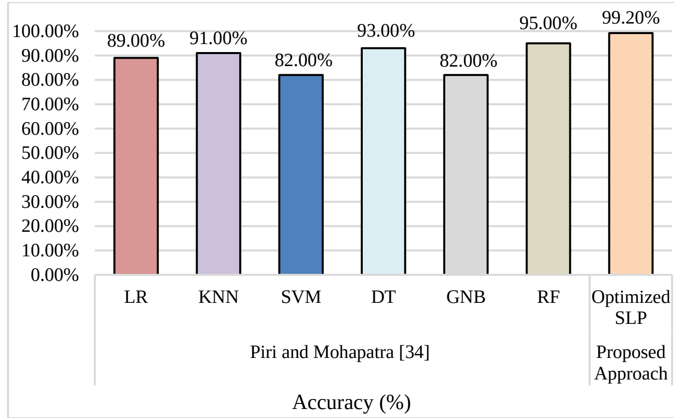

Fig. 7. Comparison Results of Accuracy for Developed SLP Model and other Models in the State-of-the Art.

From the comparison analysis, we can see that among the different models in [34], the DT and RF have the best accuracy results, which are 93% and 95%, respectively. However, the optimized SLP model has the highest accuracy (99.2%) compared with DT and RF.

V. Conclusions and Future Work

Fetal state detection from the CTG data using classification methods has become an important task for developing various medical applications. In this paper, an optimized SLP-based approach is introduced for improving CTG data classification and analysis. The approach is able to select an optimal learning rate value of SLP model and make the learning process converge to a local minima, achieving a best accuracy result. The experimental results of the research are conducted on CTG dataset and by using a number of evaluation metrics, such as accuracy, F1-score, precision, ad recall. The optimized SLP model is trained and tested on the dataset using a 10-fold cross-validation technique to generalize the training phase of developed model. The obtained results on the test process revealed that the optimized SLP model has the highest accuracy result (99.2%) compared with the models used in the recent study of related work. According the accuracy results of the state-of-the art models, the accuracy of optimized SLP model is increased by 4.2%. Therefore, the SLP has the ability to classify the CTG patterns in an accurate manner. In future work, a combination of different methods and deep learning models will be explored to classify the CTG data and detect fetal heart state.

Acknowledgment

“This research was supported by the Researchers Supporting Project number (RSP2022R493), King Saud University, Riyadh, Saudi Arabia.”

References

[1] P. R. Williamson, D. G. Altman, J. M. Blazeby, M. Clarke, D. Devane et al., "Developing core outcome sets for clinical trials: Issues to consider," Trials, vol. 13, no. 1, pp. 1-8, 2012.

[2] G. Acampora, D. J. Cook, P. Rashidi and A. V. Vasilakos, "A survey on ambient intelligence in healthcare," Proceedings of the IEEE, vol. 101, no. 12, pp. 2470-2494, 2013.

[3] C. Edwin and S. Arulkumarn, "Electronic fetal heart rate monitoring in current and future practice," Journal Obstet Gynecol India, vol. 58, no. 2, pp. 121-130, 2008.

[4] A. Sweha, T. W. Hacker and J. Nuovo, "Interpretation of the electronic fetal heart rate during labor," American family physician, vol. 59, no. 9, p. 2487, 1999.

[5] Z. Alfirevic, D. Devane and G. M. Gyte, "Continuous cardiotocography (ctg) as a form of electronic fetal monitoring (efm) for fetal assessment during labour," Cochrane database of systematic reviews, no. 5, 2013.

[6] Y. Oyelese, M. Chavez and A. M. Vintzileos, "Assessment of fetal well-being: Fetal heart rate monitoring and the fetal biophysical profile," in Clinical maternal-fetal medicine online: CRC Press, 2021, pp. 67.1-67.17.

[7] I. Amer-Wåhlin and K. Maršál, "St analysis of fetal electrocardiography in labor," in Seminars in Fetal and Neonatal Medicine, 2011, vol. 16, no. 1, pp. 29-35: Elsevier.

[8] P. Gupta, A. Saxena, T. Garg, A. Khatri and R. Mittal, "A prototype for realtime monitoring of fetal health using a pressure sensitive material and sensor based belt," in 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2018, pp. 935-940: IEEE.

[9] A. Martin, "Fetal heart rate during labour: Definitions and interpretation," Journal de gynecologie, obstetrique et biologie de la reproduction, vol. 37, pp. S34-45, 2008.

[10] L. A. Cibils, "Clinical significance of fetal heart rate patterns during labor: Ii. Late decelerations," American Journal of Obstetrics Gynecology, vol. 123, no. 5, pp. 473-492, 1975.

[11] R. M. Grivell, Z. Alfirevic, G. M. Gyte and D. Devane, "Antenatal cardiotocography for fetal assessment," Cochrane database of systematic reviews, no. 9, 2015.

[12] M. Cesarelli, M. Romano and P. Bifulco, "Comparison of short term variability indexes in cardiotocographic foetal monitoring," Computers in biology medicine, vol. 39, no. 2, pp. 106-118, 2009.

[13] Y. Lu, Y. Gao, Y. Xie and S. He, "Computerised interpretation systems for cardiotocography for both home and hospital uses," in 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 2018, pp. 422-427: IEEE.

[14] A. Çetin, "Use of three-tier classification system during intrapartum electronic fetal heart rate monitoring," Basic Clinical Sciences, vol. 1, no. 4, pp. 71-76, 2010.

[15] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," Neurocomputing, vol. 415, pp. 295-316, 2020.

[16] J. Benneyan, R. Lloyd and P. Plsek, "Statistical process control as a tool for research and healthcare improvement," BMJ Quality Safety, vol. 12, no. 6, pp. 458-464, 2003.

[17] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi et al., "Data mining in healthcare and biomedicine: A survey of the literature," Journal of medical systems, vol. 36, no. 4, pp. 2431-2448, 2012.

[18] A. Akbulut, E. Ertugrul and V. Topcu, "Fetal health status prediction based on maternal clinical history using machine learning techniques," Computer methods programs in biomedicine, vol. 163, pp. 87-100, 2018.

[19] M.-L. Huang and Y.-Y. Hsu, "Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network," 2012.

[20] N. Krupa, M. A. MA, E. Zahedi, S. Ahmed and F. M. Hassan, "Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine," Biomedical engineering online, vol. 10, no. 1, pp. 1-15, 2011.

[21] C. Sundar, M. Chitradevi and G. Geetharamani, "An analysis on the performance of k-means clustering algorithm for cardiotocogram data clustering," International Journal Computer Science Application, vol. 2, no. 5, pp. 11-20, 2012.

[22] H. Ocak and H. M. Ertunc, "Prediction of fetal state from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems," Neural Computing Applications, vol. 23, no. 6, pp. 1583-1589, 2013.

[23] H. Ocak, "A medical decision support system based on support vector machines and the genetic algorithm for the evaluation of fetal well-being," Journal of medical systems, vol. 37, no. 2, pp. 1-9, 2013.

[24] S. P. Potharaju, M. Sreedevi, V. K. Ande and R. K. Tirandasu, "Data mining approach for accelerating the classification accuracy of cardiotocography," Clinical Epidemiology Global Health, vol. 7, no. 2, pp. 160-164, 2019.

[25] J. Piri and P. Mohapatra, "Exploring fetal health status using an association based classification approach," in 2019 International Conference on Information Technology (ICIT), 2019, pp. 166-171: IEEE.

[26] Y. Zhang and Z. Zhao, "Fetal state assessment based on cardiotocography parameters using pca and adaboost," in 2017 10th international congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI), 2017, pp. 1-6: IEEE.

[27] B. Amin, M. Gamal, A. Salama, I. El-Henawy and K. Mahfouz, "Classifying cardiotocography data based on rough neural network," International Journal of Advanced Computer Science Applications, vol. 10, no. 8, 2019.

[28] Z. Hoodbhoy, M. Noman, A. Shafique, A. Nasim, D. Chowdhury et al., "Use of machine learning algorithms for prediction of fetal risk using cardiotocographic data," International Journal of Applied Basic Medical Research, vol. 9, no. 4, p. 226, 2019.

[29] J. Yan and S. Han, "Classifying imbalanced data sets by a novel re-sample and cost-sensitive stacked generalization method," Mathematical Problems in Engineering, vol. 2018, 2018.

[30] I. Brown and C. Mues, "An experimental comparison of classification algorithms for imbalanced credit scoring data sets," Expert Systems with Applications, vol. 39, no. 3, pp. 3446-3453, 2012.

[31] Y. Shang, "Subgraph robustness of complex networks under attacks," IEEE Transactions on Systems, Man, Cybernetics: Systems, vol. 49, no. 4, pp. 821-832, 2017.

[32] M. M. Rahman and D. N. Davis, "Addressing the class imbalance problem in medical datasets," International Journal of Machine Learning Computing, vol. 3, no. 2, p. 224, 2013.

[33] S.-J. Yen and Y.-S. Lee, "Cluster-based under-sampling approaches for imbalanced data distributions," Expert Systems with Applications, vol. 36, no. 3, pp. 5718-5727, 2009.

[34] J. Piri and P. Mohapatra, "Imbalanced cardiotocography data classification using re-sampling techniques," in Proceedings of international conference on machine intelligence and data science applications, 2021, pp. 681-692: Springer.

[35] K. Bache and M. J. U. h. a. i. u. e. m. Lichman, "Uci machine learning repository. University of california, irvine, school of information and computer sciences (2013)," pp. 0162-8828, 2019.

[36] G. P. Zhang, "Neural networks for classification: A survey," IEEE Transactions on Systems, Man, Cybernetics, Part C, vol. 30, no. 4, pp. 451-462, 2000.

[37] X. Zhang, X. Chen, J. Wang, Z. Zhan and J. Li, "Verifiable privacy-preserving single-layer perceptron training scheme in cloud computing," Soft Computing, vol. 22, no. 23, pp. 7719-7732, 2018.

[38] L. Zahedi, F. G. Mohammadi, S. Rezapour, M. W. Ohland and M. H. Amini, "Search algorithms for automated hyper-parameter tuning," arXiv preprint arXiv:.14677, 2021.

[39] J. Bergstra, R. Bardenet, Y. Bengio and B. Kégl, "Algorithms for hyper-parameter optimization," Advances in neural information processing systems, vol. 24, 2011.