Annals of Agri-Bio Research **27** (1) : 97-101, 2022

Impact Score : 0.28 (Scopus)

Strategic Analysis of Different Classification Algorithms on CTG Data

MEGHA CHATURVEDI*, SHIKHA AGRAWAL AND SANJAY SILAKARI

Department of Computer Science and Engineering, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal-462 033 (Madhya Pradesh), India
**(e-mail : chaturvedimegha09@gmail.com; Mobile : 91 83494 08181)*

(Received : November 5, 2021; Accepted : January 17, 2022)

ABSTRACT

Early detection of any abnormalities can give further insights into the pregnancy and will provide more time to parents and doctors to prepare for these unnatural circumstances. Cardiotocography (CTG) is a technique used for monitoring fetal heart rate. It is widely used to ensure fetal well-being during pregnancies at high risk. Use of machine-learning techniques automated this task and reduced the chances of diagnostic errors. Deep learning also has powerful algorithms for learning complicated characteristics and higher-level semantics. The principal objective of this paper was to dissect the boundaries of different classification algorithms and contrast their prescient exactnesses to discover the best classifier for ordering fetal well-being.

Key words : Cardiotocography (CTG), machine learning, classification algorithms, perinatal mortality

INTRODUCTION

The perinatal period starts from the 28th week of incubation to 7 days of life after birth. The term "Perinatal Mortality" includes both still births and early neonatal passings. There are a number of factors that are known to increase the risk of prenatal mortality some of them are high parity, malnutrition and severe anemia. Cardiotocography (CTG) is a method of measuring the fetus's heart rate in proportion to the pressure inside the uterus. Electronic Fetal Monitoring is another name for it. The ultimate intention of fetal monitoring is to avert death/morbidity of the fetus due to hypoxia. Fig. 1 shows the major tracings on a CTG, the upper tracing is the fetal heart rate and lower tracing is the uterine contractions.

Sahin and Subasi (2015) deployed eight machine learning algorithms over UCI dataset to classify the ECG data into normal or pathological. A 10-fold cross-verification was done on handout of SVM, ANN, CART, K-NN, logistic regression, C4.5, RF and RBFN. Random forest came out to be a comparatively better algorithm for classifying the available data. IEEE Electron Devices Society (2018) proposed an approach to expect excessive chances of threatening pregnancies primarily based on CART algorithm by using 5 fold cross validation and concluded that CART incorporated testing and cross validation to check the accuracy of fit. Moreover, this work can be extended using bagging and boosting procedures. Subasi *et al.* (2020) came up with a Bagging ensemble classifier technique to classify CTG information as healthful or pathological by the use of bootstrap aggregation along with ensemble classifiers. The base prototypes were created by the use of bootstrap samples of training set and balloting for prediction. To compare the overall performance of different classifiers 10-fold go-validation was done and Random Forest Classifier resulted in accuracy of 99.02%.

Attallah *et al.* (2019) used a pipeline process consisting of segmentation, enhancement, feature extraction and classification as four phases. The main motive of this paper was to devise a method for the early identification of abnormalities in the fetus. They used unique classification algorithms such as LDA, k-nearest algorithm, SVM, machine assemble subspace discriminate analysis to achieve the results. The technique used was able to classify abnormalities from different gestational ages as well as was able to classify different forms of fetal brain abnormality. The method was simple and had a low computational cost. Neocleous *et al.* (2016) presented a method to detect chromosomal abnormalities, potential risk of euploidy, T21 by making use of machine learning techniques such as ANN, SVM and K-NN. Out of these ANN showed exceptionally great results.

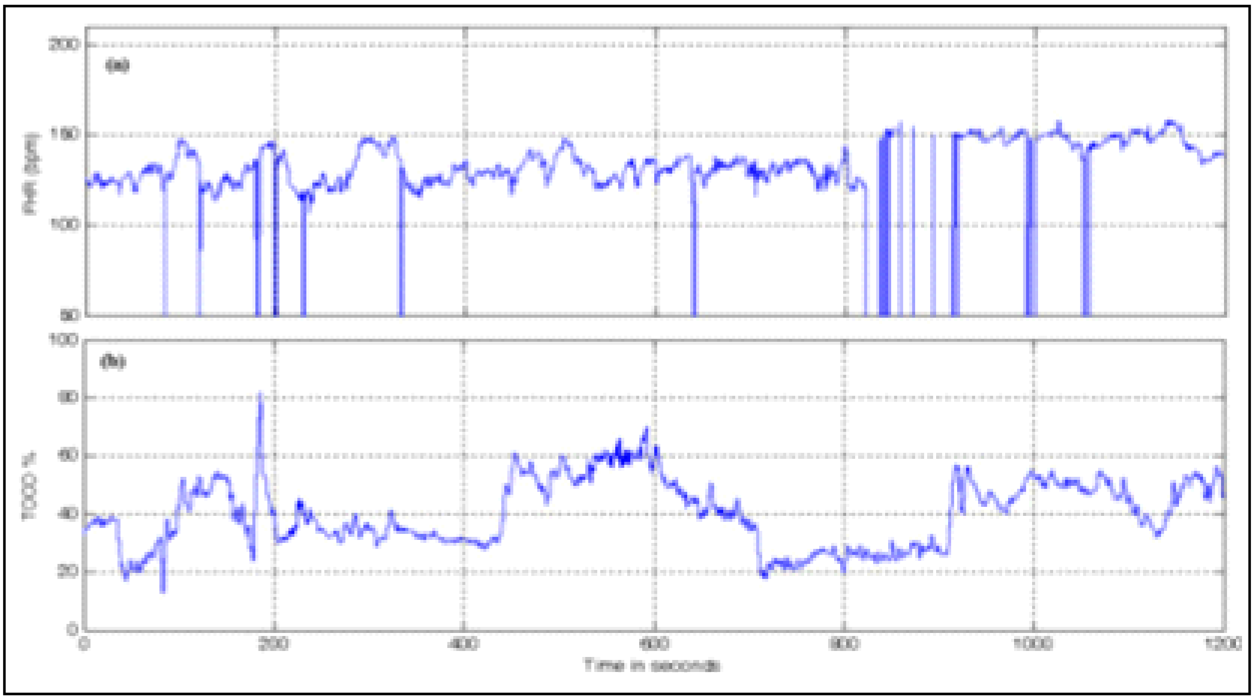

Fig. 1. A typical CTG signal : (a) The FHR signal and (b) The UA signal.
Source : https://openi.nlm.nih.gov/detailedresult.php?img=PMC3033 856_1475-925X-10-6-1&req=4.

Agrawal and Mohan (2019) used R-studio to analyze the proficiency of various algorithms such as Naive Bayes, decision tree, SVM on CTG data. The data were classified into three categories normal, suspect and pathological and the decision tree gave the highest accuracy. Akbulut *et al.* (2018) collected their own maternal clinical dataset from 96 pregnant women and compared nine binary classification algorithms to predict fetal health. They found out those abortions, delivery numbers, age of mother, any disease were the major factors that affected the fetal health. Shah *et al.* (2016) proposed a bagging approach with combination of three decision tree algorithms-J48, REPTree (Reduced Error Pruning Tree) and Random Forest. When performance was evaluated bagging approach with RF showed good results. The only limitation was that the approach was used on publicly available secondary data.

It is an endeavour to dissect, and analyze the effect of data mining strategies by recognizing outliners, executing classification algorithms and bunching the records of the Cardiotocography dataset. Petrozziello *et al.* (2019) used multimodal convolutional neural network on data of 35429 births derived from oxford archives. Qu *et al.* (2020) offered two approaches for classifying US pictures into six typical planes of fontal brains, both based on CNN. They also established that the main reason for the lack of performance for most of the models was the lack of training data.

MATERIALS AND METHODS

The Cardiotocography dataset contained CTG data classified by expert obstetricians. The dataset contained recordings of fetal heart rate (FHR) and uterine contractions. It consisted of 2126 samples out of which 1655 samples were normal, 295 suspicious and 176 samples were pathologic. The dataset was ideal for 10-class or 3-class experiments. The steps completed to enforce the classification algorithms on CTG records the usage of python were stated underneath (Fig. 2) :

- Load the dataset.
- Analyze the data using describe or info function.
- Preprocess the data : Treatment of missing and incorrect data points.
- Data visualization : Using heat maps and histograms and pie charts.
- Applying smote analysis and PCA.
- 5-fold cross validation of the algorithms (hypertuned).

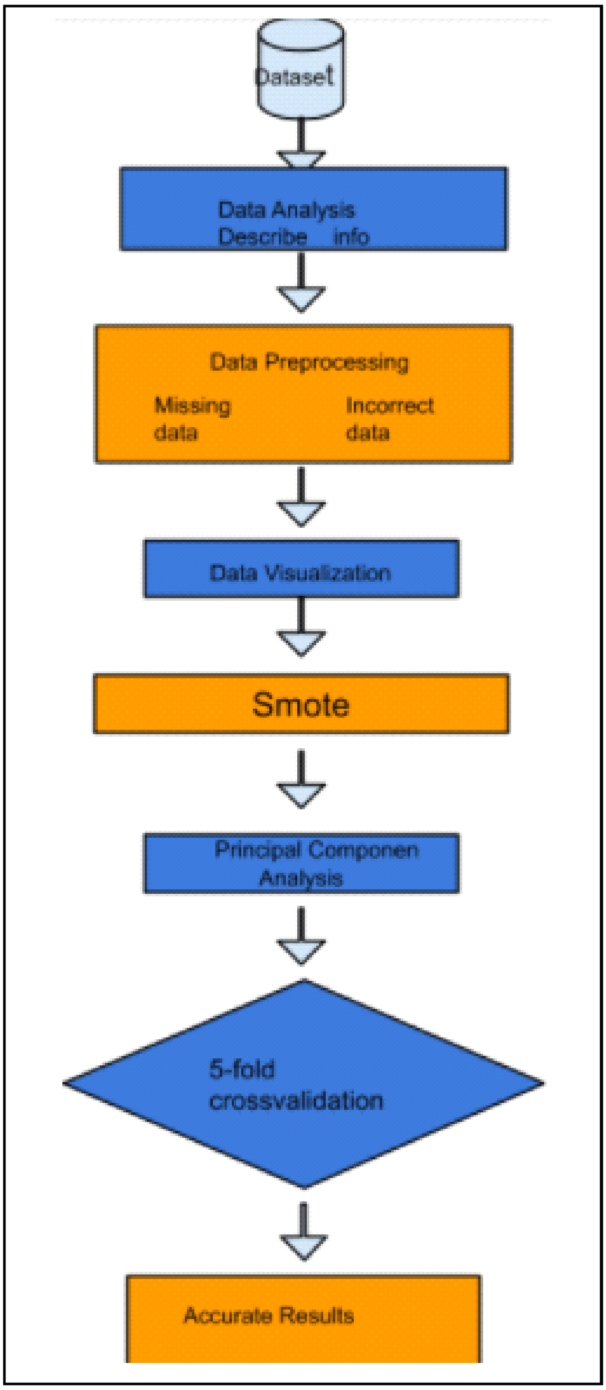

Fig. 2. Process flow.

SMOTE stands for Synthetic Minority Oversampling Technique which is primarily used to increase the samples in minority class to match up the samples in the majority class. This was important because ML models were generally biased and they produced results in the favour of the majority class as the datasets were highly imbalanced. Principal Component Analysis (PCA) was used for reducing the dimensionality of dataset so that the dataset became easier to interpret and the information loss was minimized (Fig. 3). All the models were hyperparameter tuned to get the most accurate results.

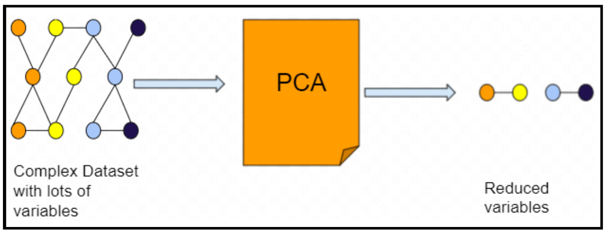

Fig. 3. PCA.

Following were the performances used in this research :

Accuracy was the proportion of correctly anticipated observations to the total number of observations. In a confusion matrix, TP represented true positive, TN referred to true negative, FP to false positive, and FN indicated false negative.

Precision was the parameter used to assess the range of correct predictions made.

Based on the TP and FN values, another parameter called recall was defined as how many true positives were recalled (discovered) i. e. how many right hits were also found.

To find an optimum point where both, the precision and recall values were high another parameter called f1-score was calculated**.** It was a harmonic mean of the precision and recall values.

RESULTS AND DISCUSSION

Random forest classifier (RFC) was a classification model that was used for dividing the dataset into classes or labels. It was an ensemble or collection of decision trees where a decision chart was basically a flow chart that was used for separating the data based on some condition. The final predicted value was the majority class i. e. the class predicted by a maximum number of decision trees in the forest (Table 1).

XGB stands for extreme gradient boosting, it was a state-of-the-art machine learned algorithm that used a technique called boosting to achieve accurate results. In terms of speed and accuracy it surpassed other ML models, the only setback was it was a heavy computational model and required very high RAM, CPU and GPU. It took some time to learn

Table 1. Performance measures of the dataset using random Forest Classifier

Algorithm	Random Forest Classifier		
Class	Normal	Suspect	Pathologic
Precision	0.87	0.80	0.93
Recall	0.84	0.87	0.87
F1-score	0.86	0.83	0.90
Accuracy		0.86%	

Table 2. Performance measures of the dataset using XGB Classifier

Algorithm	XGB Classifier		
Class	Normal	Suspect	Pathologic
Precision	0.98	0.95	0.98
Recall	0.95	0.98	0.98
F1-score	0.96	0.96	0.98
Accuracy		97%	

the feature variables through the training data and then made predictions on the test data (Table 2).

KNN was a supervised ML algorithm that was used for both classification and regression problems. It worked on the principle that things that were similar stayed close to one another or "birds of feathers flock together". It was easy to understand the only setback that it tended to perform slow when the size of the data was huge (Table 3).

Table 3. Performance measures of the dataset using KNN Classifier

Algorithm	KNN Classifier		
Class	Normal	Suspect	Pathologic
Precision	1.00	0.91	0.98
Recall	0.90	0.99	0.99
F1-score	0.94	0.99	0.99
Accuracy		0.96%	

SVM was a very popular model that solved a wide range of problems. Support vector classifier played a pivotal role in multi-class classification (Table 4).

In this research, cardiotocographic data were used for determining the health of the fetus, the data were obtained from the UCI repository. The data were divided into training and testing data with a probability of 0.75 and 0.25. The results showed that the accuracy of RFC and

Table 4. Performance measures of the dataset using Support Vector Classifier

Algorithm	SVC Classifier		
Class	Normal	Suspect	Pathologic
Precision	1.00	0.97	0.99
Recall	0.97	1.00	1.00
F1-score	0.98	0.98	1.00
Accuracy		99%	

XGB was 86 and 97%, respectively, for KNN it was 96% and SVC resulted in an accuracy of 99%. Hence, taking these accuracies into account SVC and XGB performed fairly better than the other algorithms.

CONCLUSION

Fetal mortality was an issue that needed grave attention. Machine learned techniques were proving to be a great help in medical science by opening up the possibilities of automation of tasks and at the same time high accuracy. In this research accuracy, precision, recall and f1-score were used for performance measurements, and as the results, state SVC and XGB turned out to perform exceptionally better than KNN and RFC. In the future, even better models could be developed with higher accuracies on larger datasets.

REFERENCES

Agrawal, K. and Mohan, H. (2019). Cardiotocography analysis for fetal state classification using machine learning algorithms. In : *International Conference on Computer Communication and Informatics, ICCCI 2019.* Institute of Electrical and Electronics Engineers Inc., U. K.

Akbulut, A., Ertugrul, E. and Topcu, V. (2018). Fetal health status prediction based on maternal clinical history using machine learning techniques. *Computer Methods and Programs in Biomedicine* **163** : 87-100.

Attallah, O., Gadelkarim, H. and Sharkas, M. A. (2019). Detecting and classifying fetal brain abnormalities using machine learning techniques. In : *Proceedings–17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018.* Institute of Electrical and Electronics Engineers Inc. pp. 1371-1376.

IEEE Electron Devices Society (2018). Institute of Electrical and Electronics Engineers, Vaigai College of Engineering. *Proc. 2018 International Conference on Intelligent Computing and Control Systems (ICICCS)*, June 14-15, 2018.

Neocleous, A. C., Nicolaides, K. H. and Schizas, C. N. (2016). First trimester non-invasive prenatal diagnosis : A computational intelligence approach. *IEEE J. Biomedical and Health Informatics* **20** : 1427-1438.

Petrozziello, A., Redman, C. W. G., Papageorghiou, A. T., Jordanov, I. and Georgieva, A. (2019). Multimodal convolutional neural networks to detect fetal compromise during labour and delivery. *IEEE Access* **7** : 112026-112036.

Qu, R., Xu, G., Ding, C., Jia, W. and Sun, M. (2020). Deep learning-based methodology for recognition of fetal brain standard scan planes in 2D ultrasound images. *IEEE Access* **8** : 44443-44451.

Sahin, H. and Subasi, A. (2015). Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques. *App. Soft Com. J.* **33** : 231-238.

Shah, S. A. A., Aziz, W., Arif, M. and Nadeem, M. S. A. (2016). Decision trees based classification of cardiotocograms using bagging approach. In : *Proc. 13th International Conference on Frontiers of Information Technology FIT 2015.* Institute of Electrical and Electronics Engineers Inc. pp. 12-17.

Subasi, A., Kadasa, B. and Kremic, E. (2020). Classification of the cardiotocogram data for anticipation of fetal risks using bagging ensemble classifier. In : *Procedia Computer Science. Elsevier B.V.* pp. 34-39.