

# Intelligent SMOTE Based Machine learning Classification for Fetal State on Cardiotocography Dataset

Pijush Dutta (✉ [pijushdutta009@gmail.com](mailto:pijushdutta009@gmail.com))

Global Institute of Management and Technology <https://orcid.org/0000-0001-5144-315X>

**Shobhandeb Paul**

Gurunanak Institute of Technology

**Madhurima Majumder**

Mirmadan Mohanlal Government Polytechnic

---

## Research Article

**Keywords:** Fetal cardiotocograph, machine learning, SMOTE, Feature selection, Confusion matrix, performance metrics

**Posted Date:** November 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1040799/v1>

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

[Read Full License](#)

---

# Abstract

A major contributor to under-five mortality is the death of children in the 1st month of life. Intrapartum complications are one of the major causes of perinatal mortality. Fetal cardiocotocograph (CTGs) can be used as a monitoring tool to identify high-risk women during labor. The objective of this study was to study the precision of machine learning algorithm techniques on CTG data in identifying high-risk fetuses. CTG data of 2126 pregnant women were obtained from the University of California Irvine Machine Learning Repository. Out of 2126 CTG dataset 78% of them were normal, 14% were suspect, and 8 % had a pathological fetal state. To improve data imbalance SMOTE is applied followed by five different machine learning classification models were trained using CTG data. Sensitivity, precision, and F1 score for each class and overall accuracy of each model were obtained to predict normal, suspect, and pathological fetal states. For the model validity two statistical parameters MCC & kappa (k) are used. SMOTE based all the classification algorithm provides the higher degree of accuracy with minimum value is 96% and RF algorithm had the highest prediction accuracy about 98.01% which is quite satisfactory. Model validation statistical parameters MCC & kappa is maximum achieved by RF about 0.968 & 1 and for SVC is 0.977 & 1 respectively. Finally proposed work also compared with previous state of art techniques.

## 1. Introduction

Globally 2.4 million children died in the first month of life in 2019. There are approximately 7 000 newborn deaths every day, amounting to 47% of all child deaths under the age of 5-years, up from 40% in 1990. In a pregnancy cycle, the fetal heart rate (FHR) is one of the most important evidence about the fetus[1]. The obstetricians are using cardiocotocograph (CTG) to get information that includes FHR and uterine contractions (UC) related to the fetus. The CTG intended for not only to get FHR, but also to observe the mother's contractions and other kinds of fetal monitoring [2]. CTG is a medical test utilized throughout pregnancy which records UC and FHR. This test can be employed by either external or internal techniques. With internal test, a catheter is located in the uterus after a precise quantity of expansion has taken place. In external tests, a pair of sensor nodes is attached to the mother's stomach. The CTG data usually represents two lines. The upper line records the FHR in beats per minute. The lower line records uterine contractions[3]. In order to find fetal risks based on CTG, machine learning techniques turn out to be an increasing trend to produce decision support systems in medicine. Different studies have carried out for the classification of the CTG data[4].

The information taken from CTG is utilized for early identification of a pathological state and can help the obstetrician to predict future problems and hinder before occurring a permanent impairment to the fetus. Throughout the delivery of the baby who is showing to hypoxia can cause a temporary impairment or death. Because of the wrong diagnosis of the FHR pattern recordings and inappropriate treatments employed to the fetus can achieve more than half of these deaths[5]–[7]. While its practicality, there might be some inconsistency in the success of CTG monitoring, predominantly in low-risk pregnancies. If

there is an inaccurately evaluated fetal pain then, it might be results in useless treatments or if there is an inappropriate investigation of fetal wellbeing then it might be excluded essential treatments [8].

CTG data using three different machine learning techniques to predict fetal distress[9]. An employment of statistical features extracted from Empirical Mode Decomposition (EMD)[10].The extracted features from the sub-band decomposition classified as normal or risky. They achieved 86% accuracy for the test data. Another study presented a two-steps examination of fetal heart rate data which permits for effective prediction of the acidemia risk. The FHR signals are classified by Support Vector Machines (SVM), fuzzy, Multilayer perceptron. A new model which utilize the artificial neural network (ANN) to classify the CTG data[11]. The Recall and F-score were employed to assess the performance. Moreover, they proposed the k-means clustering for the CTG classification. Adaptive neuro-fuzzy inference systems (ANFIS) is utilized [12] for the CTG classification. Moreover, SVM and Genetic Algorithm (GA) based classification method was implemented [13].

There several research done for the prediction of CTG on the basis of dataset & classify algorithms are reviewed here. Eight different ML techniques are proposed to classify the normal & pathological fetal state from 1831 dataset with 21 attributes Cardiotocograph dataset. From the result analysis is seen that KNN achieved the maximum accuracy about 98.4% [6]. A hybrid feature selection PSO based ML techniques proposed to predict the classification of fetal heart rate. Among all the ML it is seen that KNN achieved the maximum accuracy about 83.3% [14]. A K Means ML classification algorithm proposed to find out the accuracy from Cardiotocograph dataset contains 21 attributes. In first phase K means algorithm used to eliminate the 7 attributes from the dataset & later on ML techniques are used to find out the accuracy. From the result analysis was conclude that SVM achieved maximum accuracy about 90.64% [15]. A comparative study was performed between SVM & DT on the basis of same dataset used in previous reference. Both the algorithm offers the quite satisfactory accuracy for the classification of dataset [16]. RF classifier model has been used to classify the three different states of fetal state from CTG dataset which contain 2126 dataset with 21 attributes. Maximum accuracy obtained by RF is 93.6 % for seven potential attributes [17]. A correlation feature selection algorithm with four ML technique proposed to classify the fetal state either in normal or pathological. Overall research performed for the same dataset used in previous reference. From the result analysis it is also concludes that moreover all the algorithms have same classification accuracy is about 94.7% [18]. Five different ML algorithms used to classify the states of Fetal from CTG dataset which contain 2126 dataset with 21 attributes. Naïve Bayes acquired the maximum accuracy about 82.32%[19]. A comparative study performed between RBFN, DT, and NB & MLP to predict the states of fetal heat rate .Maximum accuracy obtained by NB about 83.9 % when number of potential attributes is 15 [20]. Prediction of FHR has been performed by hybrid ADB with SVM applied in a CTG datasets which contain 2126 dataset with 21 attributes. Overall research performed into two stage, in first stage PCA used to sorted the potential attributes & later on hybrid classification algorithm used [21]..Maximum accuracy obtained by the proposed model is 98.6% for selected attributes.

In this paper, several ensemble machine-learning models examined to classify the CTG data as unhealthy or healthy based on the three obstetricians' decisions. The contribution of this paper is to implement Bagging ensemble method to classify the CTG data. To the best of the authors' knowledge, the Bagging ensemble classifiers have not been employed previously for the CTG classification. Hence, this paper compares the performances of the single and ensemble learners in terms of F- measure, accuracy, and ROC area. Hence, in section 2, materials and methods are presented. In section 3, results and discussion presented. Section 4 is conclusion.

## 2. Methods

The dataset was obtained from the University of California Irvine Machine Learning Repository. It comprised of 2126 pregnant women who were in the third trimester of pregnancy. The dataset consisted of 35 attributes used in the measurements of FHR and uterine contractions (UCs) on CTG (Figure 1).

According to the Child Health and Human Development, the core risk variable used to derive the state of fetus includes qualitative and quantitative descriptions of FHR and UCs [22]. The machine learning algorithms used in this study were Decision Tree, Random Forest, KNN, SVC & Linear SVC. The current dataset was split into training and testing folds using K-Fold Cross Validation technique to test the performance of each machine learning model in the training phase.

### 2.1. Feature Selection Approaches:

The increase in diagnosis cost and the huge volume of data produced by different sources consist of the number of attributes. All attributes may not be useful, thus it is necessary to remove them during data preprocessing or feature selection. The feature selected attributes would, in turn, improve the performance to build a better classification. The various feature selection methods such as embedded, ensemble and hybrid methods, filter methods and wrapper methods have been applied to study the fetal heart rate or CTG analysis[23]. In this research we used Symmetrical Uncertainty based feature selection methods [24] [24] and five classification algorithms such as Decision Tree, Random Forest, KNN, SVC & Linear SVC to study the CTG data analysis.

### 2.2. Data Sources:

The publicly UCI machine learning repository has been used to retrieve the Cardiotocograph (CTG) dataset available at <https://archive.ics.uci.edu/ml/datasets/Cardiotocography>. The multivariate data type consists of 2126 instances with 35 attributes, which are numeric. The class attribute consists of 3 distinct values, which are Normal, Suspect, and Pathologic. The frequencies of 2126 instances are as follows: 1655 normal, 295 suspicious and 176 pathologic, indicating the uneven distribution of the observations across the classes, which refers to class imbalance dataset. The imbalanced datasets require special attention because the regular classifiers accuracies are inappropriate to use for class imbalance [5], [25], [26], since these classifiers generally favor the majority class i.e., the class with a large number of instances. The performance of the classifier can be improved by the ensemble of classifiers.

However, the majority of ensembles is static and cannot be applied to imbalanced datasets [27]. Apart from this, based on experimental results, it is known that the performance on the balanced dataset is better than the imbalance dataset [28]. In the view of aforementioned sentences, the dataset used in this study consists of 248 normal fetal state class randomly derived from 1655 instances from UCI repository CTG dataset), keeping other class codes as the same (i.e., 295 suspicious and 176 pathologic) with 23 attributes is shown in Table 1

Table 1  
Characteristics of dataset after SMOTE

	UCI machine repository	Dataset after SMOTE
Attributes	35	23
Normal	1655	248
Suspicious	295	295
Pathologic	176	176
Total instances	2126	719

## 2.3 Attributes Description:

The dataset consists of 23 attributes. The predictable attribute is referred to “NSP: Fetal state class code (N = normal; S = suspect; P = pathologic)”. The description of the attributes is shown in Table 2.

Table 2  
Class description of fetal heart rate

Class values	Description	No of dataset
(N)	Normal	248
(S)	Suspicious	295
(P)	Pathologic	176

The derived dataset consists of 719 instances with 23 attributes (Table 1) has been taken into consideration to build a classification model after normalization of the data. Python-based Scikit learn was used as an analytical tool. A total of seven machine learning (ML) techniques, each (refer to literature review) was used to evaluate the performance of the classifiers and tools. Later, feature selection was also implemented on the aforementioned dataset.

## 3. Result Analysis

To the best of author knowledge, most of the classification model studies have been carried out on the UCI machine learning repository CTG dataset [29], [30]. Thus, there were no studies addressing the derived dataset with the five machine learning techniques. To measure the performance of each classification

algorithm, the accuracy has been taken into accordance. The key outcome of this study was to compare major machine learning algorithms (listed above) with regard to their precision accuracy and sensitivity to predict normal, suspect, or pathologic fetal state based on CTG attributes. Various statistical techniques were used to compare the performance of the algorithms. These included precision, sensitivity or recall, F1 score, and overall accuracy ( $[(\text{true positive} + \text{true negative}) / (\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative})]$ ).

On the provided dataset, the experiment is run, and the results are produced. Each experiment is evaluated to stratified K-fold validation to ensure that the results are free of bias. The major goal is to remove any bias in the outcomes, as feature engineering sometimes leads to the omission of specific characteristics, which might affect overall prediction results. Furthermore, the process of feature engineering is typically highly costly. Machine learning algorithms are provided raw data after some preparation. The findings are then obtained and compared to current state-of-the-art systems. The dataset was examined, and methods were used when needed, and the model was trained to improve the precision.

### 3.1 Performance Measure

One of the most blatant misrepresentations about machine learning model assessment is that every dataset, regardless of its type, can be quantified using the same evaluation matrices. The majority of machine learning models are judged on their accuracy[31]–[39]. When working with an unbalanced dataset, this deliberate proves to be deceiving. As a result, several appropriate assessment matrices, as well as accuracy, are employed. Precision, recall, the F1 measure, and the ROC curve were used to evaluate the proposed study[40], [41]. The accuracy ratio is the number of correct predictions divided by the total number of inputs. The confusion matrices are obtained by calculating the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values (FN).  $TP / (FN + TP)$  and  $FP / (FP + TN)$  are two considerations that are computed as  $TP / (FN + TP)$  and  $FP / (FP + TN)$ . Another statistic common screening a model's classification accuracy is the receiver operation curve (ROC)[42].

Table 3. Performance of SMOTE based Classifier Algorithms

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC
Decision Tree	96.8%	96.8%	96.5%	96.4%	0.89
Random Forest	98.01%	97.8%	97.7%	97.5%	0.96
KNN	96.2%	96.2%	96%	96%	0.92
SVC	97.7%	97.5%	97%	97%	0.96
Linear SVC	97%	97%	97%	97%	0.64

Table 3 shows the tabular form result analysis of Average Accuracy, Precision, F1 score, Precision Area under ROC & Computational Time for SMOTE based Logistic Regression, Random Forest, Decision tree,

KNN & SVM models when trained and tested on the Tabular data consisting of actual 540 records. Results are obtained after principle component analysis & SMOTE. Average parameters are calculated for both negative & positive classes' cases. We found that SMOTE based Random Forest performed best among the entire SMOTE based algorithm with the Average Accuracy, precision, Recall & F1 Score values of 98.01%, 97.8%, 97.7% & 97.5 % respectively. However the SMOTE based Random Forest and SMOTE based KNN & SVM have least computational time & maximum area under ROC with the values of 0.10 sec & 96% respectively.

Figure 3 shows graphical comparison for the entire SMOTE based machine learning algorithm on the basis of computational time. From the plot we found that SMOTE based Random Forest have least computational time of 0.010 sec while SMOTE based Decision Tree have a maximum time of 0.031 sec.

### 3.2. Classification model evaluation

The reason for the assessment of a classification model is to achieve a solid evaluation of the model that is known as the model's predictive performance. Diverse execution parameters can be utilized.

Provided that the model is dependent on training set and has speculation property which is basis for the quality assessment. For any assessment measure, it is imperative to recognize its incentive for a specific dataset performance, particularly the training set performance, and its true generalization performance. The created model's training performance is determined by assessing the model on the training set. However, the aim of classification models is not to categorize the training data. Suitable evaluation processes are required to dependably evaluate the unfamiliar values of the assumed performance measures on the whole domain [43], [44].

### 3.3 Statistical Measurement

The Mathews correlation coefficient (MCC) is a metric for evaluating binary classification quality[45]–[47]. The Matthews correlation coefficient is a contingency matrix technique of calculating the Pearson product-moment correlation coefficient between actual and predicted values that is unaffected by the unbalanced datasets issue. MCC is the only binary classification rate that awards a high score only if the binary predictor accurately predicts the majority of positive and negative data instances. It has a range of  $[-1, +1]$ , with extreme values of  $-1$  and  $+1$  for perfect misclassification and perfect classification, respectively, and for coin tossing classifier  $MCC=0$ . Equation (1) is demonstrating the MCC.

$$MCC = \frac{(TP*TN-FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}} \quad (1)$$

The kappa (k) statistic is a key parameter for judging the model's consistency[48]–[50]. It compares the outcome of the suggested model to the outcome of the randomly classified technique. The kappa statistic's value ranged from 0 to 1. The model's expected effect is represented by a value near to 1, whereas 0 indicates that the model is flawed. (2), (3), and (4) demonstrate the kappa statistic's equation.

$$P(A) = \frac{(TP+TN)}{N} \quad (2)$$

$$P(E) = \frac{[(TP+TN)*(TP+FP)*(TN+FN)]}{N^2} \quad (3)$$

$$K = \frac{[P(A)-P(E)]}{1-P(E)} \quad (4)$$

In present research range of kappa value 0.702 to 1 indicates proposed model attains great consistency. Both the values of MCC & kappa for all the algorithms are shown in Table 4.

**Table 5.** Statistical Measure after experimentation

Algorithm	MCC	Kappa(k)
Decision Tree	0.741	0.958
Random Forest	0.968	1
KNN	0.789	0.95
SVC	0.957	1
Linear SVC	0.541	0.857

### 3.4 Comparison with existing system

The proposed work's findings are being compared to the results of other state-of-the-art existing system in order to ascertain the proposed work's trustworthiness.

**Table 6.** Comparative study between proposed models with existing model

Reference	Algorithm used	Outcomes from the research
[14]	PSO based KNN & SVM	Overall accuracy for PSO feature selection based KNN achieved the maximum accuracy 88.5%
[15]	SVM & hybrid K means SVM	Maximum accuracy obtained by the K means SVM with 90.64% , where k=10
[17]	Random Forest	Random Forest with seven important feature classify the CTG data with maximum accuracy 93.6%
[18]	Bagging approach with three different decision tree algorithms : Random forest , REP Tree & J48 & correlation feature selection were used	All the proposed algorithm achieved overall accuracy was about to 90%.
[20]	Naive Bayes, Decision Tree, Multi Layer Perceptron and Radial Basis Function	Maximum accuracy obtained by Decision Tree for 15 potential attributes about 93.3%
Present Research	Decision Tree, Random Forest , SVC , KNN, Linear SVC	Maximum accuracy obtained by SMOTE based Random Forest is about 98.01% for 23 attributes

## 4. Conclusion

Classification of accuracy from CTG dataset is a one of major challenges in the medical diagnosis system. Delayed detection of pathologic fetal state based on CTG attributes may caused serious health issue of mother & baby so early diagnosis is important. In modern research for early detection in medical diagnosis ML techniques have been introduced. ML techniques are the subsection of AI which has capability to learn the large amount of unlabeled & unstructured data in few seconds. In this research we proposed existing techniques for diagnosis of early detection of pathologic fetal state on CTG datasets.

In last decades there are several research performed for the detection of pathologic fetal state in terms of accuracy. All the approaches used same dataset (CTG) for their training & testing model. In ML there are number shortcomings for the prediction of Diabetes like accuracy & identification of potential attributes etc. Hence, a model must be designed in such a manner in future so that it able to overcome these shortcomings. In this research we performed the overall research into three stages: In first stage imbalance CTG data oversampled by SMOTE, in second stage we used hyper parameter tuning of training dataset to reduce the model's complexity & make a trade-off between these components and final stage we applied six Machine learning technique for testing data classification. For the classification of fetal state of CTG dataset we used five different ML techniques, DT, SVC, KNN, RF & Linear SVC.

All the proposed model are evaluated on the basis of confusion matrix parameters like accuracy, recall, precision, F1 score, and AUC, computational time parameters like training & testing time & statistical parameters for model consistency like MCC & k. Here we identify the best model on the basis of six parameters which is explain in previous sections, SMOTE based RF model achieved best accuracy of

98.01% as compared to other models used. Further maximum AUC provided by RF & SVC model indicates it is optimal classifiers for CTG dataset. However both RF & SVC also achieved the maximum value of MCC & k indicating higher degree of model consistency. On the contrary, future aspects of present research can be upgraded by the collecting the large real-time dataset from IoT based device & prediction can be done the performance of ML techniques by means of reduction of signal bandwidth & less computational time.

## References

1. M. F. Hurtado-Sánchez, D. Pérez-Melero, A. Pinto-Ibáñez, E. González-Mesa, J. Mozas-Moreno, and A. Puertas-Prieto, "Characteristics of Heart Rate Tracings in Preterm Fetus," *Medicina (Mex.)*, vol. 57, no. 6, p. 528, 2021.
2. Y. Lu, Y. Qi, and X. Fu, "A framework for intelligent analysis of digital cardiocographic signals from loMT-based foetal monitoring," *Future Gener. Comput. Syst.*, vol. 101, pp. 1130–1141, 2019.
3. S. Chan, M. Arjuna, and N. L. Nik Ahmad Zuky, "Cardiotocography waveform analysis using image extraction technique," 2020.
4. H. Sahin and A. Subasi, "Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques," *Appl. Soft Comput.*, vol. 33, pp. 231–238, 2015.
5. F. Marzbanrad, L. Stroux, and G. D. Clifford, "Cardiotocography and beyond: a review of one-dimensional Doppler ultrasound application in fetal monitoring," *Physiol. Meas.*, vol. 39, no. 8, p. 08TR01, 2018.
6. A. Subasi, B. Kadasa, and E. Kremic, "Classification of the cardiotocogram data for anticipation of fetal risks using bagging ensemble classifier," *Procedia Comput. Sci.*, vol. 168, pp. 34–39, 2020.
7. V. Nagendra, H. Gude, D. Sampath, S. Corns, and S. Long, "Evaluation of support vector machines and random forest classifiers in a real-time fetal monitoring system based on cardiotocography data," in *2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2017, pp. 1–6.
8. M. Belzile, A. Pouliot, A. Cumyn, and A. M. Côté, "Renal physiology and fluid and electrolyte disorders in pregnancy," *Best Pract. Res. Clin. Obstet. Gynaecol.*, vol. 57, pp. 1–14, 2019.
9. A. E. Permanasari and A. Nurlayli, "Decision tree to analyze the cardiotocogram data for fetal distress determination," in *2017 international conference on sustainable information engineering and technology (SIET)*, 2017, pp. 459–463.
10. S. Aziz, M. U. Khan, Z. A. Choudhry, A. Aymin, and A. Usman, "ECG-based biometric authentication using empirical mode decomposition and support vector machines," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019, pp. 0906–0912.
11. H. Ellethy, S. S. Chandra, and F. A. Nasrallah, "The detection of mild traumatic brain injury in paediatrics using artificial neural networks," *Comput. Biol. Med.*, vol. 135, p. 104614, 2021.

12. Y. Fei *et al.*, "Automatic Classification of Antepartum Cardiotocography Using Fuzzy Clustering and Adaptive Neuro-Fuzzy Inference System," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 1938–1942.
13. A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Computer-aided diagnosis of lung nodules in computed tomography by using phylogenetic diversity, genetic algorithm, and SVM," *J. Digit. Imaging*, vol. 30, no. 6, pp. 812–822, 2017.
14. G. Georgoulas, C. Stylios, V. Chudacek, M. Macas, J. Bernardes, and L. Lhotska, "Classification of fetal heart rate signals based on features selected using the binary particle swarm algorithm," in *World Congress on Medical Physics and Biomedical Engineering 2006*, 2007, pp. 1156–1159.
15. N. Chamidah and I. Wasito, "Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine," in *2015 international conference on advanced computer science and information systems (ICACSIS)*, 2015, pp. 37–41.
16. D. Jagannathan and M. Phil, "Cardiotocography-a comparative study between support vector machine and decision tree algorithms," *Int. J. Trend Res. Dev.*, vol. 4, no. 1, 2017.
17. M. Arif, "Classification of cardiotocograms using random forest classifier and selection of important features from cardiotocogram signal," *Biomater. Biomech. Bioeng.*, vol. 2, no. 3, pp. 173–183, 2015.
18. S. A. A. Shah, W. Aziz, M. Arif, and M. S. A. Nadeem, "Decision trees based classification of cardiotocograms using bagging approach," in *2015 13th international conference on frontiers of information technology (FIT)*, 2015, pp. 12–17.
19. D. Bhatnagar and P. Maheshwari, "Classification of cardiotocography data with WEKA," *Int. J. Comput. Sci. Netw.-IJCSN*, vol. 5, no. 2, 2016.
20. V. Subha, D. Murugan, J. Rani, K. Rajalakshmi, and T. Tirunelveli, "Comparative analysis of classification techniques using Cardiotocography dataset," *Int. Jour Res. Inf. Technol.*, vol. 1, no. 12, pp. 274–280, 2013.
21. 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.
22. Z. Hoodbhoy, M. Noman, A. Shafique, A. Nasim, D. Chowdhury, and B. Hasan, "Use of machine learning algorithms for prediction of fetal risk using cardiotocographic data," *Int. J. Appl. Basic Med. Res.*, vol. 9, no. 4, p. 226, 2019.
23. I. Rafique, M. Dilawar, A. Umer, and M. A. Hassan, "Classification of Cardiotocography Data for Fetal Health Using Feature Selection Techniques," in *Computer Science On-line Conference*, 2021, pp. 34–44.
24. M. S. Devi, S. Sridevi, K. K. Bonala, R. H. Dadi, and K. V. K. Reddy, "Oversampling Response Stretch based Fetal Health Prediction using Cardiotocographic Data," *Ann. Romanian Soc. Cell Biol.*, pp. 1448–1464, 2021.
25. J. Xu, Z. Chen, J. Zhang, Y. Lu, X. Yang, and A. Pumir, "Realistic preterm prediction based on optimized synthetic sampling of EHG signal," *Comput. Biol. Med.*, vol. 136, p. 104644, 2021.

26. K. Madasamy and M. Ramaswami, "Data imbalance and classifiers: impact and solutions from a big data perspective," *Int. J. Comput. Intell. Res.*, vol. 13, no. 9, pp. 2267–2281, 2017.
27. B. Krawczyk, A. Cano, and M. Woźniak, "Selecting local ensembles for multi-class imbalanced data classification," in 2018 *International joint conference on neural networks (IJCNN)*, 2018, pp. 1–8.
28. D. Ballabio, F. Grisoni, and R. Todeschini, "Multivariate comparison of classification performance measures," *Chemom. Intell. Lab. Syst.*, vol. 174, pp. 33–44, 2018.
29. K. Agrawal and H. Mohan, "Cardiotocography analysis for fetal state classification using machine learning algorithms," in 2019 *International Conference on Computer Communication and Informatics (ICCCI)*, 2019, pp. 1–6.
30. N. Sevani, I. Hermawan, and W. Jatmiko, "Feature Selection based on F-score for Enhancing CTG Data Classification," in 2019 *IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 2019, pp. 18–22.
31. P. Dutta and A. Kumar, "Application of an ANFIS model to optimize the liquid flow rate of a process control system," *Chem. Eng. Trans.*, vol. 71, pp. 991–996, 2018.
32. P. Dutta and A. Kumar, "Design an intelligent calibration technique using optimized GA-ANN for liquid flow control system," *J. Eur. Systèmes Autom.*, vol. 50, no. 4–6, p. 449, 2017.
33. P. Dutta and A. Kumar, "Design an intelligent flow measurement technique by optimized fuzzy logic controller," *J. Eur. Systèmes Autom.*, vol. 51, no. 1–3, p. 89, 2018.
34. P. Dutta and A. Kumar, "Intelligent calibration technique using optimized fuzzy logic controller for ultrasonic flow sensor," *Math. Model. Eng. Probl.*, vol. 4, no. 2, pp. 91–94, 2017.
35. P. Dutta and A. Kumar, "Modeling and optimization of a liquid flow process using an artificial neural network-based flower pollination algorithm," *J. Intell. Syst.*, vol. 29, no. 1, pp. 787–798, 2020.
36. S. Mandal, P. Dutta, and A. Kumar, "Modeling of liquid flow control process using improved versions of elephant swarm water search algorithm," *SN Appl. Sci.*, vol. 1, no. 8, pp. 1–16, 2019.
37. P. Dutta and A. Kumar, "Modelling of Liquid Flow control system Using Optimized Genetic Algorithm," *Stat. Optim. Inf. Comput.*, vol. 8, no. 2, pp. 565–582, 2020.
38. P. Dutta, R. Agarwala, M. Majumder, and A. Kumar, "PARAMETERS EXTRACTION OF A SINGLE DIODE SOLAR CELL MODEL USING BAT ALGORITHM, FIREFLY ALGORITHM & CUCKOO SEARCH OPTIMIZATION," *Ann. Fac. Eng. Hunedoara*, vol. 18, no. 3, pp. 147–156, 2020.
39. P. Dutta, S. K. Biswas, S. Biswas, and M. Majumder, "Parametric optimization of Solar Parabolic Collector using metaheuristic Optimization".
40. P. Dutta, S. Paul, and A. Kumar, "Comparative analysis of various supervised machine learning techniques for diagnosis of COVID-19," in *Electronic Devices, Circuits, and Systems for Biomedical Applications*, Elsevier, 2021, pp. 521–540.
41. P. Dutta, S. Paul, A. J. Obaid, S. Pal, and K. Mukhopadhyay, "Feature Selection based Artificial Intelligence Techniques for the Prediction of COVID like Diseases," in *Journal of Physics: Conference Series*, 2021, vol. 1963, no. 1, p. 012167.

42. P. DUTTA and A. KUMAR, "Flow sensor Analogue: Realtime prediction Analysis using SVM & KNN," presented at the Emerging trends in Engineering and Science (ETES 2018), 2018.
43. F. DIMAIO *et al.*, "Accounting for Safety Barriers Degradation in the Risk Assessment of Oil and Gas Systems by Multistate Bayesian Networks," *Reliab. Eng. Syst. Saf.*, vol. 216, p. 107943, 2021.
44. P. Schneider and K. Böttinger, "High-performance unsupervised anomaly detection for cyber-physical system networks," in *Proceedings of the 2018 workshop on cyber-physical systems security and privacy*, 2018, pp. 1–12.
45. D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, 2020.
46. D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Min.*, vol. 14, no. 1, pp. 1–22, 2021.
47. C. Halimu, A. Kasem, and S. S. Newaz, "Empirical comparison of area under ROC curve (AUC) and Mathew correlation coefficient (MCC) for evaluating machine learning algorithms on imbalanced datasets for binary classification," in *Proceedings of the 3rd international conference on machine learning and soft computing*, 2019, pp. 1–6.
48. H. Wu, S. Yang, Z. Huang, J. He, and X. Wang, "Type 2 diabetes mellitus prediction model based on data mining," *Inform. Med. Unlocked*, vol. 10, pp. 100–107, 2018.
49. Y. Zheng, R. Zhang, M. Huang, and X. Mao, "A pre-training based personalized dialogue generation model with persona-sparse data," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, vol. 34, no. 05, pp. 9693–9700.
50. G. Qian, W.-S. Lei, M. Niffenegger, and V. F. González-Albuixech, "On the temperature independence of statistical model parameters for cleavage fracture in ferritic steels," *Philos. Mag.*, vol. 98, no. 11, pp. 959–1004, 2018.

## Declarations

## Funding:

For this research authors does not get any fund.

## Conflict of Interest:

All the author declares that there has no conflict of interest.

## Ethical approval:

This article does not contain any studies with human participants or animals performed by any of the authors.

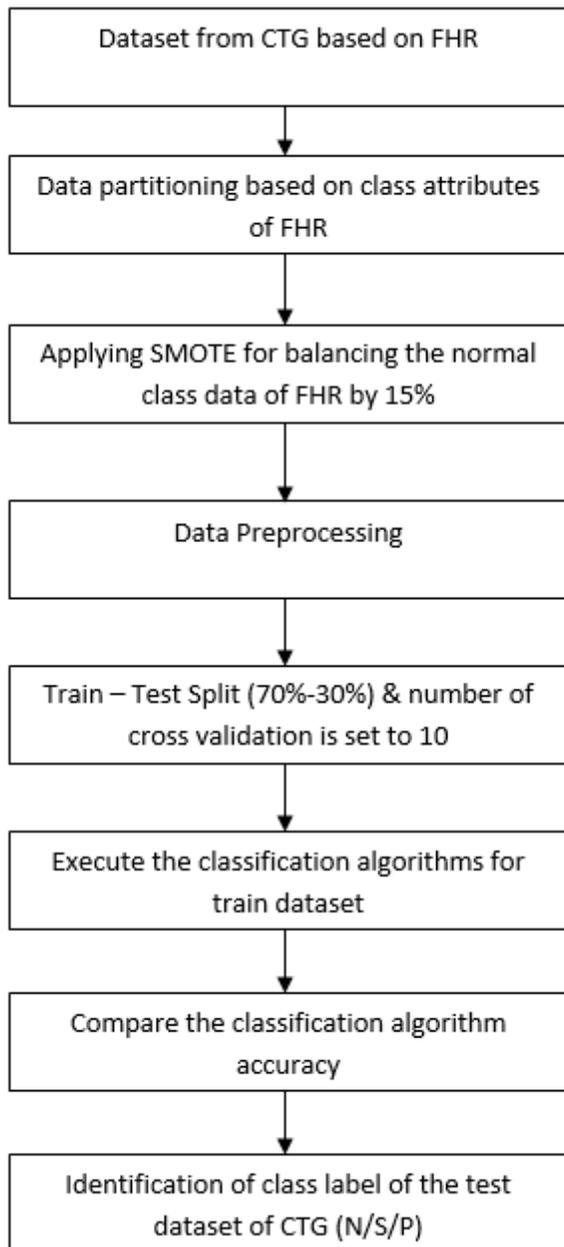
# Figures

	LBE	LB	AC	FM	UC	ASTV	MSTV	ALTV	MLTV	DL	...	C	D	E	AD	DE	LD	FS	SUSP	CLASS	NSP
0	120.0	120.0	0.0	0.0	0.0	73.0	0.5	43.0	2.4	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	9.0	2.0
1	132.0	132.0	4.0	0.0	4.0	17.0	2.1	0.0	10.4	2.0	...	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	6.0	1.0
2	133.0	133.0	2.0	0.0	5.0	16.0	2.1	0.0	13.4	2.0	...	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	6.0	1.0
3	134.0	134.0	2.0	0.0	6.0	16.0	2.4	0.0	23.0	2.0	...	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	6.0	1.0
4	132.0	132.0	4.0	0.0	5.0	16.0	2.4	0.0	19.9	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	1.0

5 rows × 35 columns

Figure 1

Attributes of CTG datasets



**Figure 2**

Flowchart of the proposed model

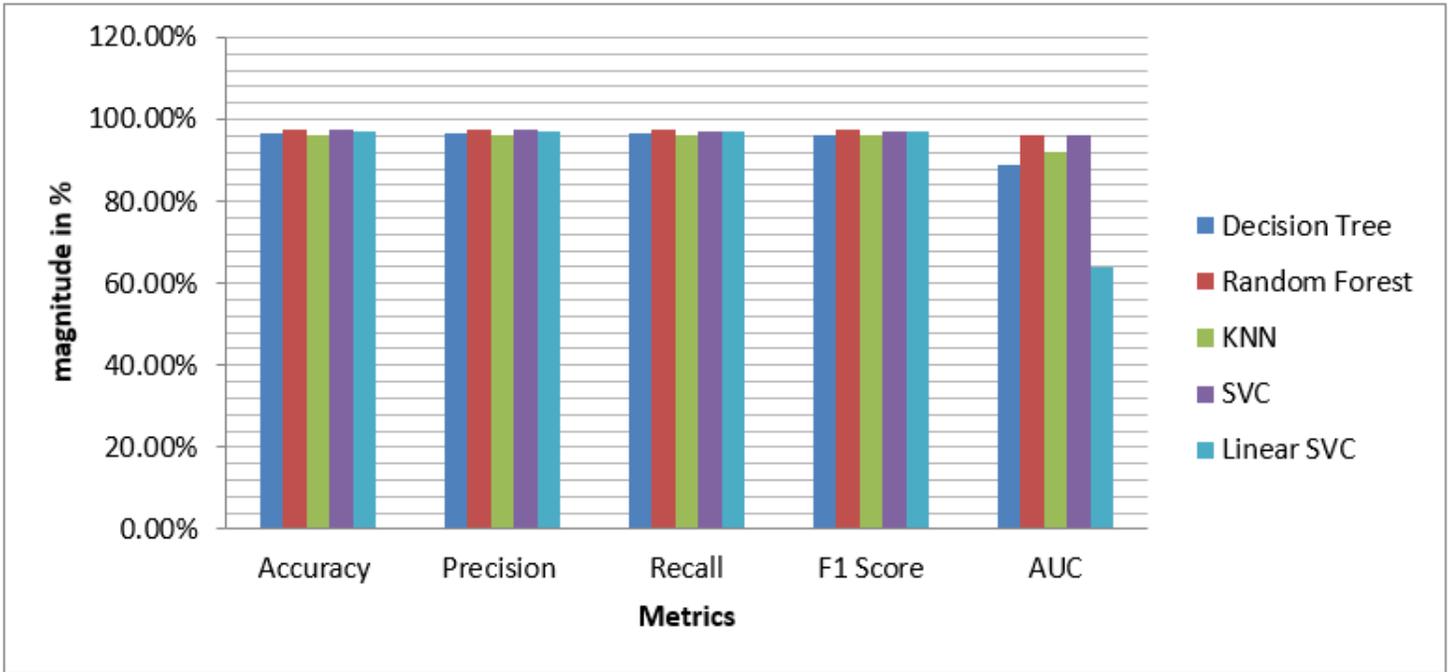


Figure 3

graphical comparison for the entire SMOTE based machine learning algorithm on the basis of computational time

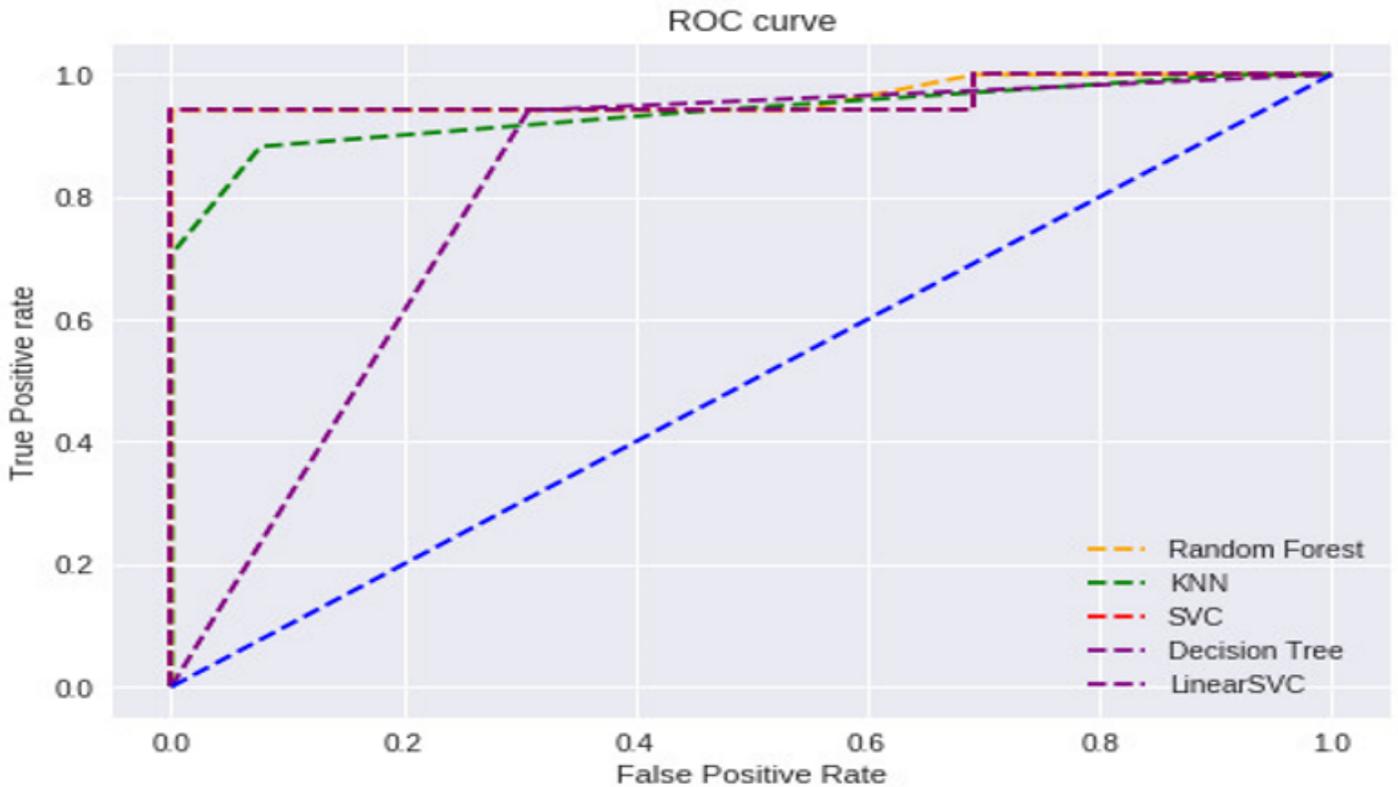


Figure 4

Comparative study of proposed ML techniques on ROC Curve

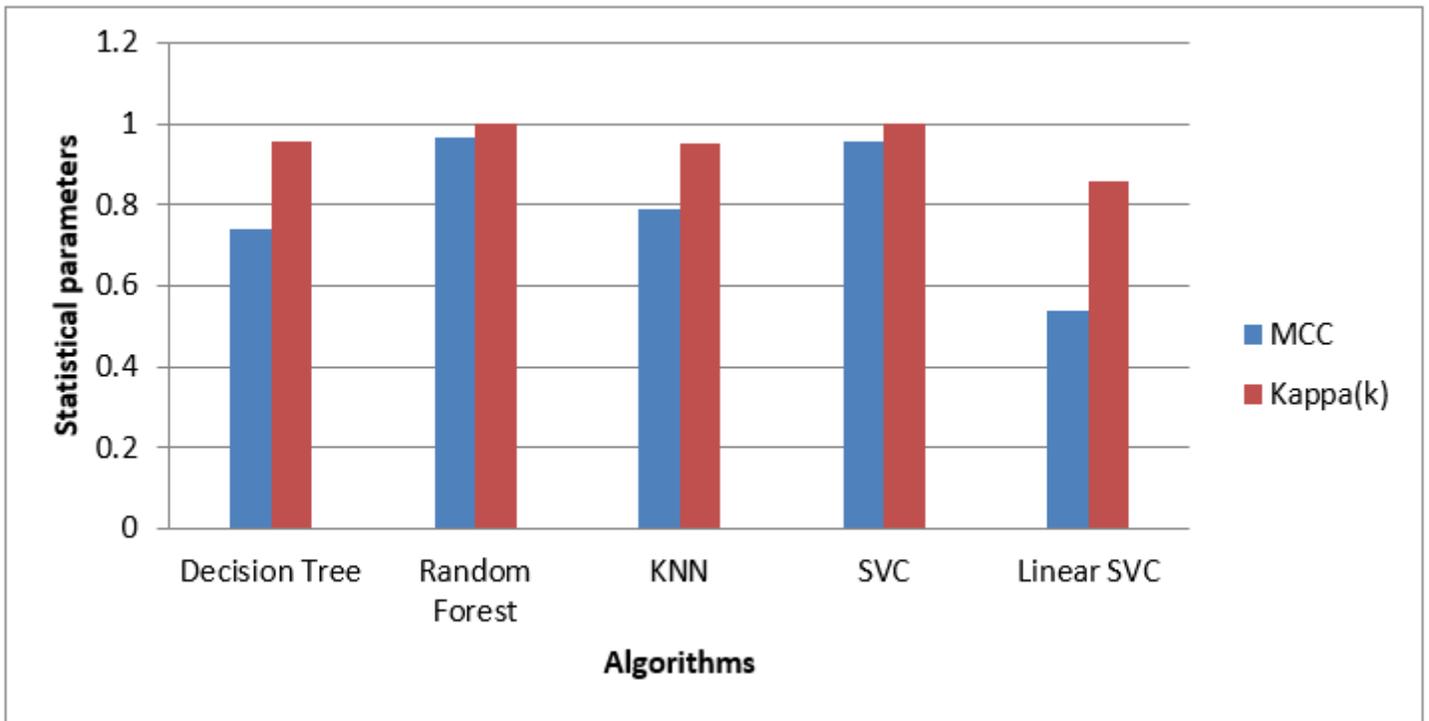


Figure 5

Comparative study based on Statistical parameters

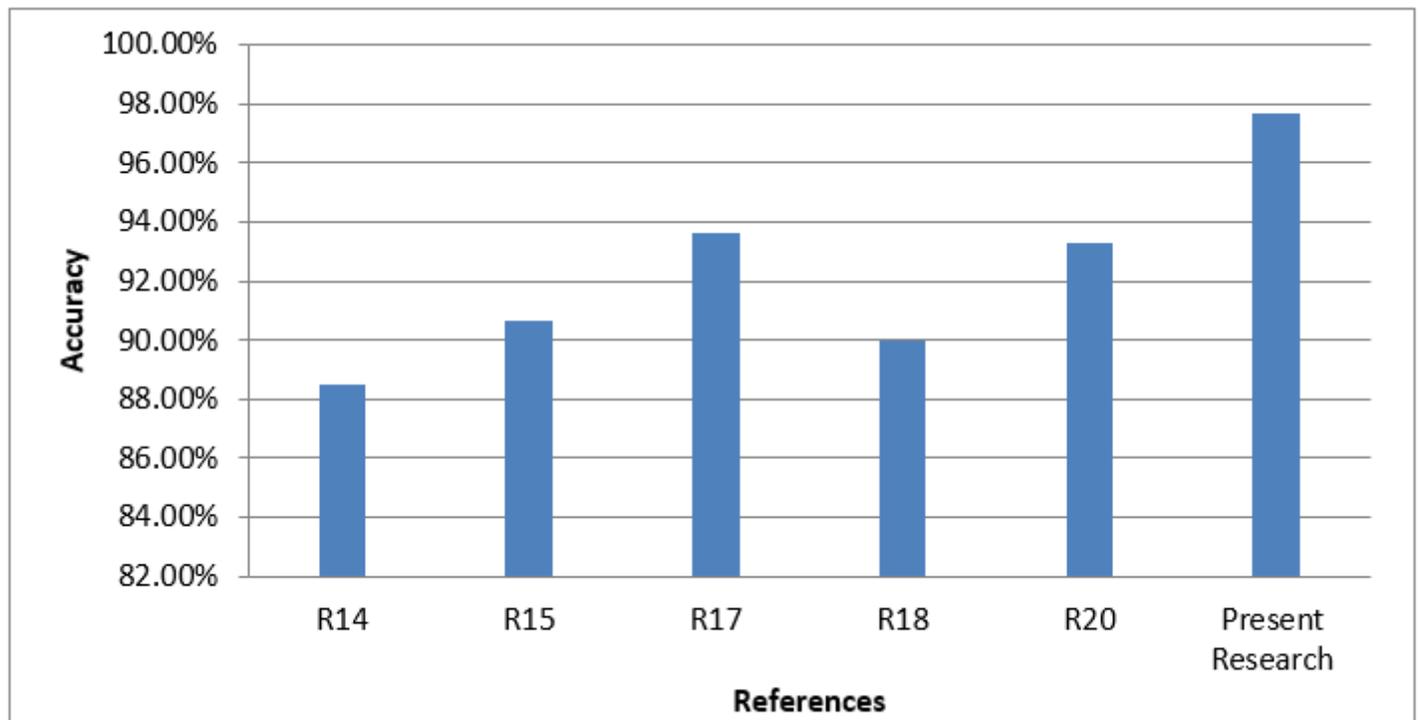


Figure 6

Comparative study between state of art techniques with present research