Performance Comparison of Three Classifiers for Fetal Health Classification Based on Cardiotocographic Data

Vijay Khare[*][ID], Sakshi Kumari[ID]

Department of Electronics and Communication Engineering, Jaypee Institute of Information Technology, 201309 Noida, India

[*] Correspondence: Vijay Khare (vijay.khare@mail.jiit.ac.in)

Abstract: The global child mortality rate, which is steadily declining, will be around 26 fatalities per 1000 live births in 2022. Numerous Sustainable Development Goals of the United Nations take into account the declining child mortality rate, which illustrates how far humanity has come. Cardiotocograms (CTGs) are a simple and affordable tool that most professionals choose to reduce infant and mother mortality. Three of the most cutting-edge methodologies are utilized in this research to classify the data, and their results are compared. All three classifiers outperformed the random forest, whose accuracy was 94.3%.

Keywords: Fetal health; Cardiotocograms (CTGs); Machine learning; Support Vector Machine (SVM); Random Forest, Multilayer perceptron

1. Introduction

All mothers want a benign pregnancy, a regular delivery, and a healthy child. The mother and fetus are both negatively impacted by delivery problems. Thus, choosing the right delivery mechanism is of the utmost significance. The most used technique for detecting fetal distress during the antepartum and early postpartum period is cardiotocography (CTG). Four essential and important factors were included in the relevant datasets: baseline fetal heart rate (BL), accelerations (ACC), decelerations (DCL), and variability. Based on these variables, doctors can determine whether the fetal condition is normal, suspicious, or pathological.

The field of knowledge known as machine learning (ML) enables computers to learn without explicit programming. It is one of the most exciting technological developments ever. Unsupervised, supervised, and reinforced machine learning are the three primary types of ML [1]. In this paper, three classification techniques are utilized, namely, support vector machine, random forest, and multilayer perceptron [2].

This paper employs CTG data to monitor the health of the fetus, for CTG data enables the detection of fetal defects and the choice of medical intervention prior to the infant suffering permanent injury. Our investigation was carried out using a number of well-known ML techniques. The most accurate algorithms were found to be random forest, support vector machine, and multilayer perceptron. To the best of our knowledge, these methods have not been compared in previous studies. The accuracy rate of the models utilized in this study is significantly higher than that of earlier studies, indicating that these models are more reliable. Their robustness was demonstrated by numerous model comparisons.

By 2030, the UN wants all countries to cease preventable infant and child deaths, with a goal of reducing under-five mortality to at least 25 per 1,000 live births. Maternal mortality, which includes fatalities during pregnancy and after delivery, accounts for 295 000 deaths in addition to infant mortality (as of 2017). The majority of these deaths (94%) occurred in areas with little resources, and most were preventable [3].

Pregnancy typically lasts nine months. A trimester is a three-month phase of pregnancy. Each trimester marks a new stage in fetus development. Prenatal screenings and routine medical exams are essential. Fetal issues develop when the unborn child develops in the womb. These conditions are categorized as congenital, which means that they exist from birth. Some fetal illnesses are genetic, i.e., they are inherited from one's parents. Most prenatal

illnesses are not known to have a cause. Modern testing techniques are used by the professionals at the Prenatal Care Center to identify fetal anomalies. Early detection is crucial to ensure that a mother and her unborn child receive the best medical treatment possible.

Certain birth defects may be improved in babies after fetal surgery. These highly challenging operations are carried out by our experts while the child is still in the womb. Treatment for some fetal conditions could begin as soon as the baby is born. Unfortunately, not all fetal disorders can be cured. Chest and lung diseases, chromosomal disorders, extremity and skeletal abnormalities, gastrointestinal abnormalities, heart illness, neurological conditions, tumors, and growths are some of the most complicated prenatal conditions.

Cardiotocograms (CTGs) are a quick and affordable approach for medical professionals to assess fetal health and take action to lower infant and mother mortality rates. CTGs emits ultrasound pulses and analyzes their reactions, shedding light on fetal heart rate (FHR), fetal movements, uterine contractions, and other parameters. Here, the authors made an effort to use these parameters to create a model that can categorize the fetus as normal, suspect, or pathological [4, 5].

2. Methods and Materials

2.1 Data Acquisition

This dataset consist 2,126 records of features extracted from Cardiotocogram exams, which were then classified by three expert obstetricians into 3 classes:
- Normal
- Suspect
- Pathological

The Fetal heart rate (FHR) baseline different ranges of 110 bpm to 150 bpm or 110 bpm to 160 bpm as shown in Figure 1. So, we have 2127 FHR values which are giving different values of acceleration fetal, fetal movement and so on to generate a multiclass model to classify CTG features into the three fetal health states.
Some of the features are:
  i) Fetal accelerations
 ii) Uterine contraction
iii) Short term variability
 iv) Histograms

baseline value	accelerations	fetal movement	uterine contractions	light decelerations	histogram width	histogram min	histogram max	fetal health
120	0	0	0	0	64	62	126	2
132	0.006	0	0.006	0.003	130	68	198	1
133	0.003	0	0.008	0.003	130	68	198	1
134	0.003	0	0.008	0.003	117	53	170	1
132	0.007	0	0.008	0	117	53	170	1
134	0.001	0	0.01	0.009	150	50	200	3
134	0.001	0	0.013	0.008	150	50	200	3
122	0	0	0	0	68	62	130	3
122	0	0	0.002	0	68	62	130	3
122	0	0	0.003	0	68	62	130	3
151	0	0	0.001	0.001	130	56	186	2
150	0	0	0.001	0.001	130	56	186	2
131	0.005	0.072	0.008	0.003	66	88	154	1
131	0.009	0.222	0.006	0.002	87	71	158	1
130	0.006	0.408	0.004	0.005	107	67	174	1
130	0.006	0.38	0.004	0.004	107	67	174	1
130	0.006	0.441	0.005	0.005	125	53	178	1
131	0.002	0.383	0.003	0.005	107	67	174	2
130	0.003	0.451	0.006	0.004	99	59	158	1
130	0.005	0.469	0.005	0.004	112	65	177	1
129	0	0.34	0.004	0.002	128	54	182	3
128	0.005	0.425	0.003	0.003	141	57	198	1

Figure 1. Fetal health data stored in csv file

2.2 Pre-Processing

There were no null values and all target data besides fetal_health are floats. Therefore, we had quickly assessed if we have any replicates then moved into brief EDA. There are a ton of variables so we just make sure our data is relatively balanced. First, we set a plotting function that makes publication ready figures then we plotted a count plot as shown in Figure 2.
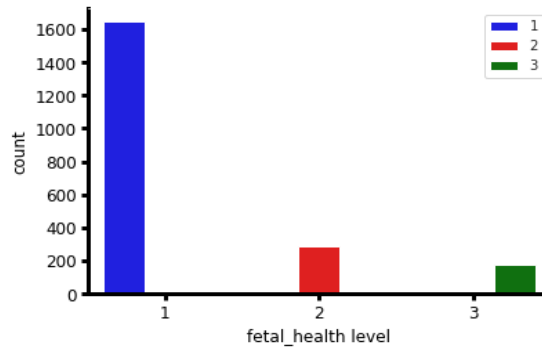
Figure 2. Fetal health count plot

Clearly, the data is imbalanced and we can't plan on performing an upsample till initial modeling is complete. Instead of plotting a pair plot, we can plot a correlation matrix to observe the pearson correlation coefficients as shown in Figure 3. Remember though that correlation does not imply causation. This will also guide us to predicting what the feature selection (KBest) will decide are the most important features as well later [6, 7].
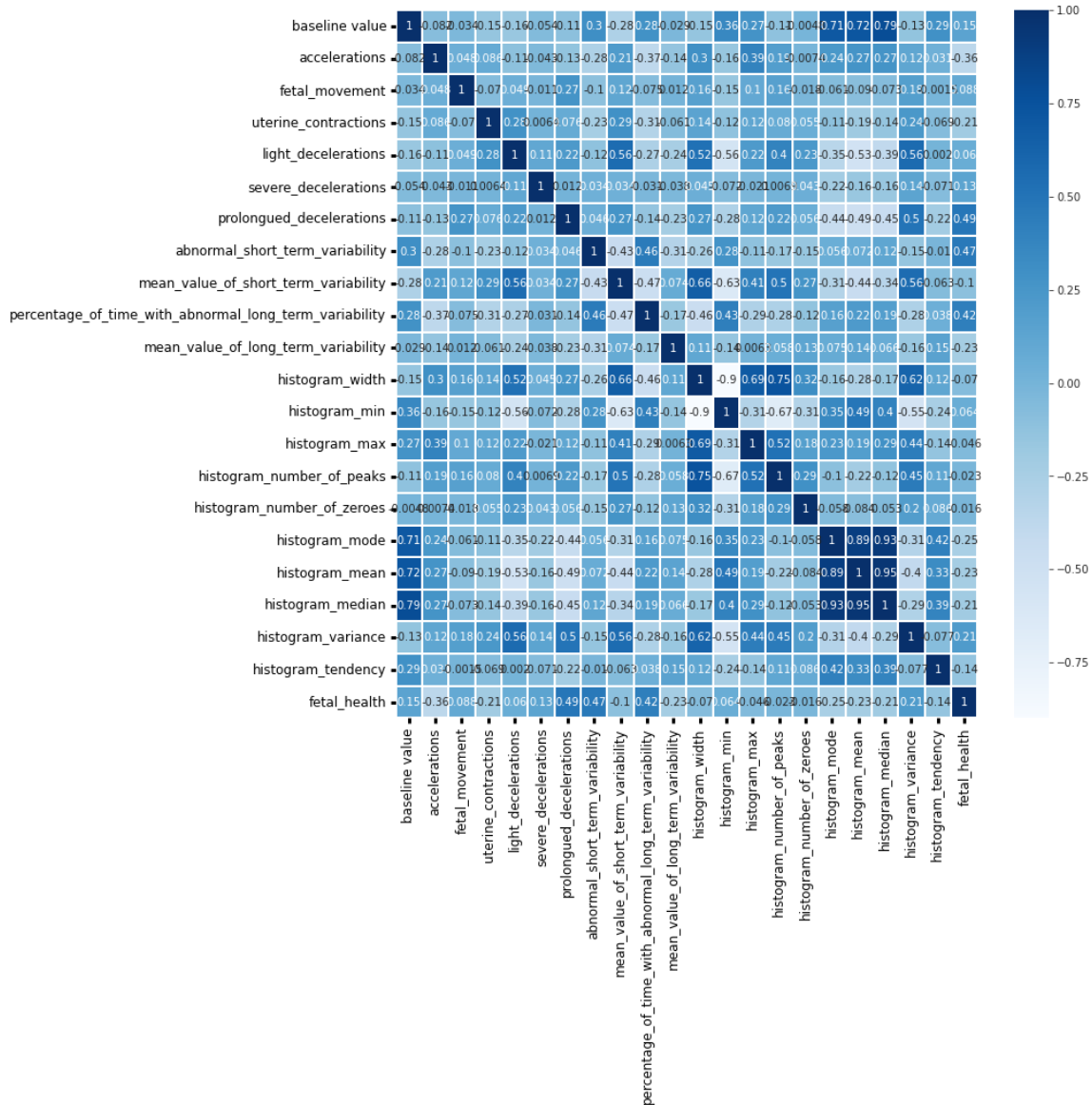

Figure 3. Correlation matrix to observe the pearson correlation coefficients

2.3 Feature Selection

Using k best selection and f_classif as score function as shown in Figure 4, we visualize the result by seaborn library using bar chart [8].
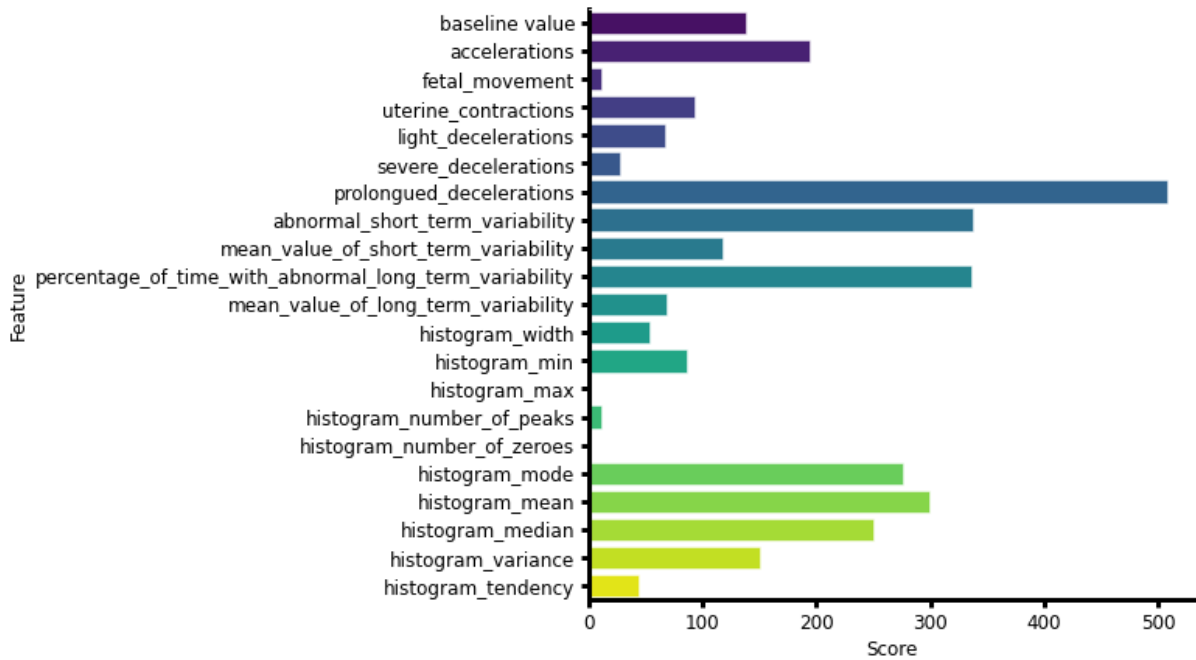

Figure 4. Features score

Next, we selected features that scored more than 200 and generates the features into a list. We add the Level string to be used to make new data frame. We create new data frame with selected features as shown in Figure 5.

	prolongued_decele ration	abnormal_short_t erm_variability	percentage_of_ time_with_abn ormal_long_ter m_variability	histogra m_mode	histogra m_mean	histogra m_medi an	fetal _hea lth
0	0.0	73.0	43.0	120.0	137.0	121.0	2.0
1	0.0	17.0	0.0	141.0	136.0	140.0	1.0
2	0.0	16.0	0.0	141.0	135.0	138.0	1.0
3	0.0	16.0	0.0	137.0	134.0	137.0	1.0
4	0.0	16.0	0.0	137.0	136.0	138.0	1.0

Figure 5. Data frame of selected features

We were left with 6 features that were selected as the most important. Since we have a reduced feature amount lets plot a quick pairplot to spot some differences as shown in Figure 6.

2.4 Splitting the Data and Scaling

First, the data will be split so we can train a scaler model to apply to an unknown (test) data set. We will save 25% of the data for testing as shown in Figure 7. The data will then be split by standard scaler using the formula $Z=\frac{(X0-\mu)}{\sigma}$. This can help reduce the effect of outliers when modeling later.
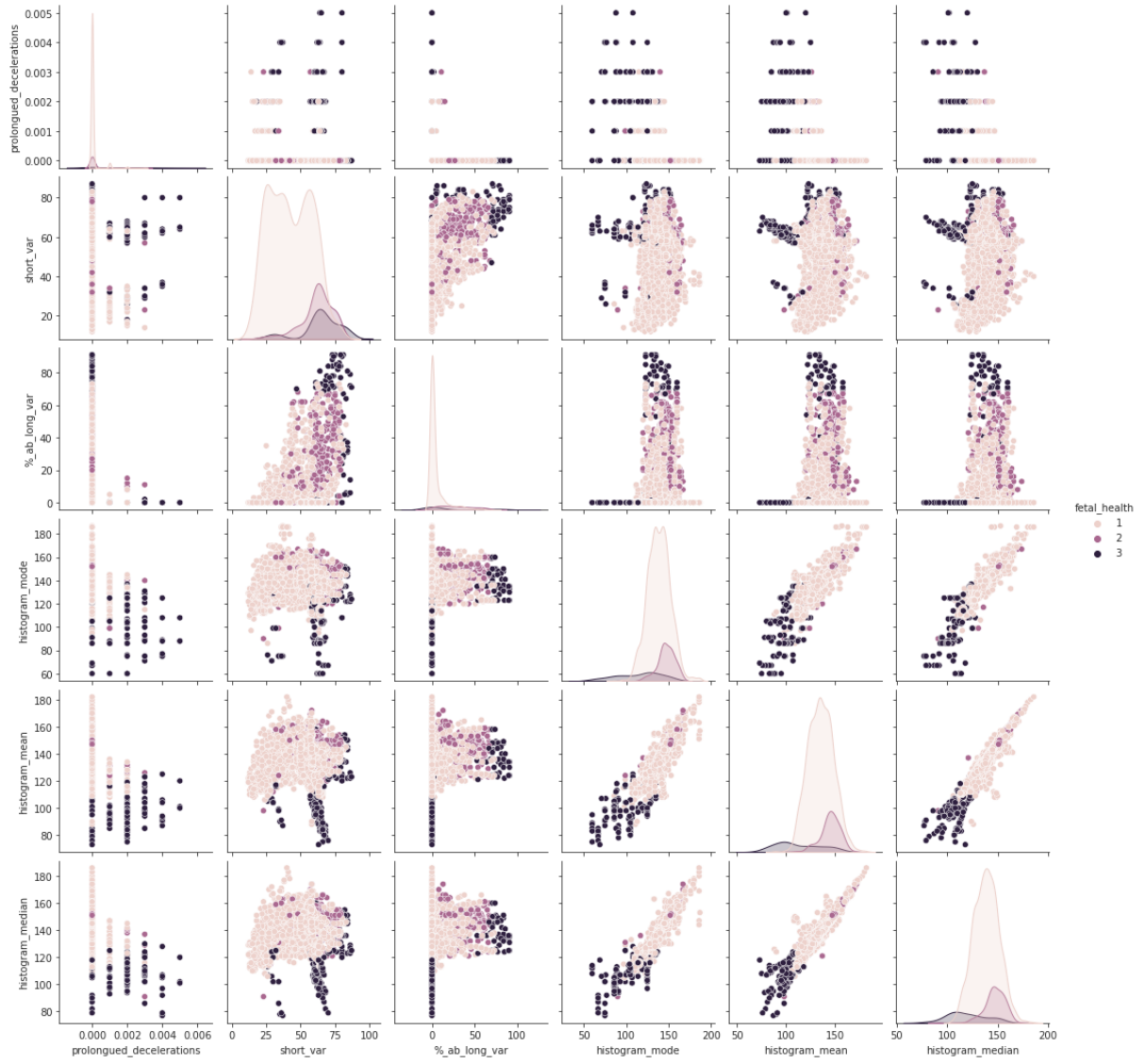
As per the task, stratify will be used.

Figure 6. Pairplot to spot differences

```
X_train shape : (1479, 21)
Y_train shape : (1479,)
X_test shape : (634, 21)
Y_test shape : (634,)
```

Figure 7. Splitting of data and scaling

2.5 Classification

In this paper, three classifiers are utilized to classify the Cardiotocographic data as follow [9-11].

2.5.1 Support vector machine

It is used to generate the optimal line or decision boundary that can divide n-dimensional space into the classes so that we can simply place fresh data points in the proper category in the future. The (soft-margin) SVM classifier is computed by minimizing an expression of the form.

$$[\frac{1}{n} \sum_{i=1}^{n} \max(0, 1 - y_i(w^T x_i - b))] + \lambda \|w\|^2$$

We focus on the soft-margin classifier since, as noted above, choosing a sufficiently small value for λ yields the hard-margin classifier for linearly classifiable input data.

In this paper, using the support vector machines classifiers (SVC) generate hyperplanes for separation and score on a yes (1) no (1) basis as shown in confusion matrix in Figure 8. The rulings are decided for where a data point lands within a decision boundary. F-1 score provides us with a method to monitor the precision and recall of our values.
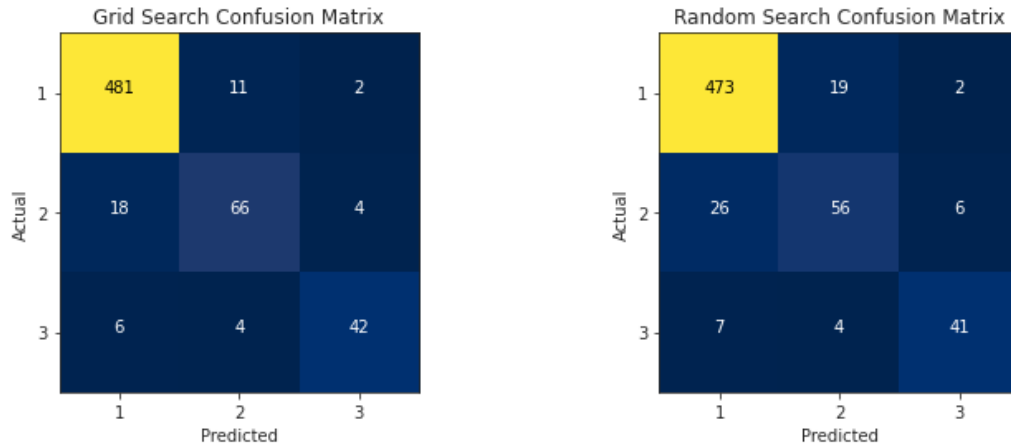

Figure 8. Confusion matrix for SVM grid and random search

2.5.2 Random Forest

Random Forest constructs decision trees from several samples and uses their majority of votes for classification and average for regression. One of the most essential characteristics of the Random Forest Algorithm is that, as in regression and classification, it can handle data sets with both continuous and categorical variables. It outperforms other algorithms in categorization tasks [12].

An ensemble method that estimates several weak decision trees and combines the mean to create an uncorrelated forest at the end. The uncorrelated forest should be able to predict more accurately than an individual tree. For this dataset Random Forest classification method gives better result than existing as shown in confusion matrix in Figure 9.
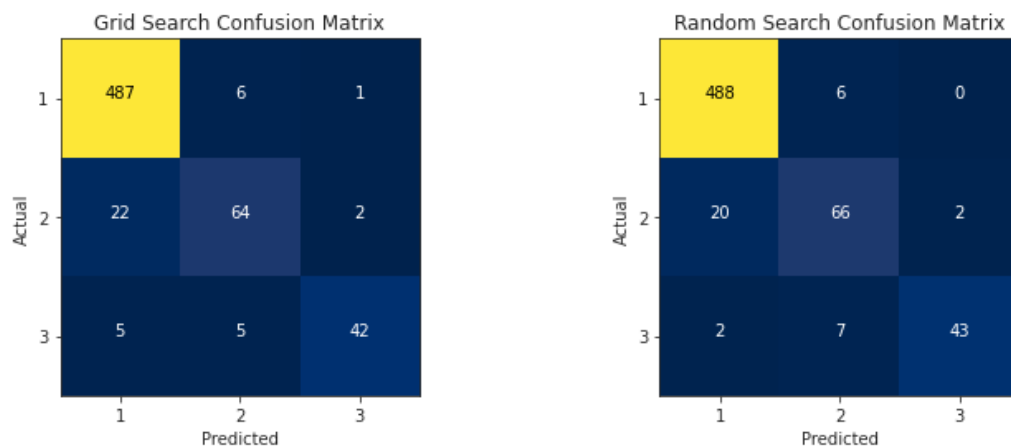

Figure 9. Confusion matrix for Random Forest grid and random search

2.5.3 Multi-Layer perceptron

Multilayer Perceptron (MLP) - A multilayer perceptron (MLP) is a feed-forward type of neural network augmentation [13-15]. Input, output and concealed these are three layers of multilayer perceptron. The input signal which is to be processed is received by the input layer. For the categorization and prediction output layer is responsible. Multi-layer perceptron is intended to approx any continuous function and can tackle issues that cannot be solved linearly Feed forward neural network. The number of nodes is determined by (2/3 * input feature count) + (number of outputs + 2). The number of layers were decided by 2/3 of the first and 1/2 the second layer. We can parametrize plenty of activator functions and set this up with the search function above. performance of multilayer

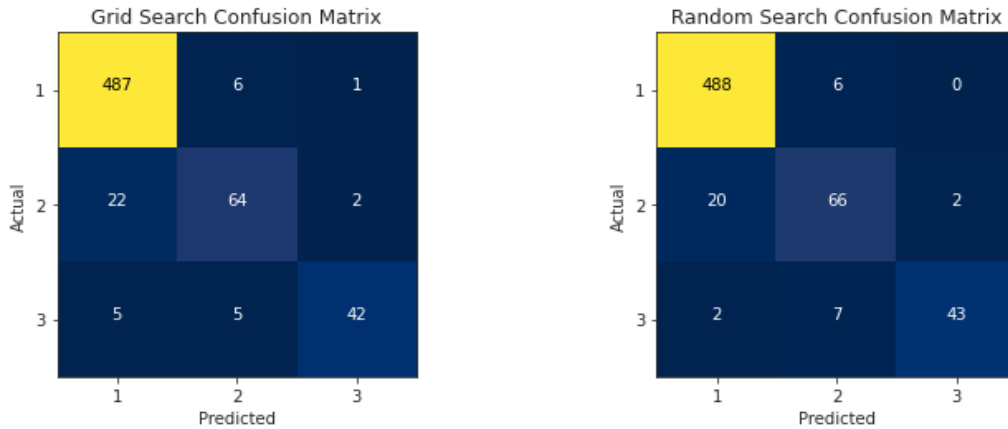perceptron is shown as confusion matrix in Figure 10.

Figure 10. Confusion matrix for multi-layer perceptron grid and random search

3. Results

3.1 Support Vector Classification (SVC)

In SCV grid search results the best parameters were: {'C':10,'degree':3,'gamma':0.1.'kernal':'rbf','random_state':1}. Classification report is shown in Table 1.
Best accuracy: 92.4%

Table 1. Classification report of SCV

Classification Report:	Precision	Recall	F1 -Score	Support
1.0	0.95	0.97	0.96	494
2.0	0.81	0.75	0.78	88
3.0	0.88	0.81	0.84	52

3.2 Random Forest

In random forest grid search results the best parameters were: {'criterion':'entropy','max_depth': 11,'n_estimator':200,'random_state':1}. Classification report is shown in Table 2.
Best accuracy: 94.3%

Table 2. Classification report of Random Forest

Classification Report:	Precision	Recall	F1 -Score	Support
1.0	0.95	0.99	0.97	494
2.0	0.85	0.73	0.79	88
3.0	0.93	0.81	0.87	52

3.3 Multi-Layer Perceptron

In multi-layer perceptron grid search results the best parameters were: {'activation':'relu', 'hidden_layer_sizes'(6,4),'learning_rate':'constant','learning_rate_init':0.001,'max_iter':1000, 'random_state': 1, 'solver': 'adam'}. Classification report is shown in Table 3.
Best accuracy: 91.5%

Table 3. Classification report of multilayer Perceptron

Classification Report:	Precision	Recall	F1 -Score	Support
1.0	0.95	0.96	0.96	494
2.0	0.70	0.72	0.71	88
3.0	0.82	0.71	0.76	52

4. Conclusion and Future Scope

Mother must take care of her health and as well as baby health monitoring. For mother fetal growth and development several tests are suggested during pregnancy. One of the tests is cadiotocogram, which is used to check the health state of the fetus in the uterus.

In this paper, CTG data is used for fetal health monitoring. This dataset consist 2,126 records of features extracted from Cardiotocogram exams, Using KBestSelection we were able to fetched the most important features from the data set. which were then classified by three classifiers namely: Support vector Machine Random Forest, and Multilayer perceptron as classifiers. We have obtained accuracy respectively Support vector Machine (92.4%), Random Forest (94.3%), Multilayer perceptron (0.91.5%). The research results show the comparison of three classifiers namely Support Vector Machine, Random Forest and Multilayer perceptron. We have observed that the random forest is the best algorithm implemented on cardiotocography data.

In future reduction techniques as a pre-processing can be apply on data. The dataset used in this paper is not too much rich; the performance may be much better and accurate if dataset is vaster. For dimensionality reduction and increase the accuracy we will use principle component analysis (PCA) and Linear Discriminant Analysis (LDA) algorithms. Both algorithms are used for retaining as much as information after the reduction of number of features in the dataset.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] "Artificial intelligence vs. machine learning vs. deep learning vs. data science," Medium, 2019, https://medium.com/@rinu.gour123/artificial-intelligence-vs-machine-learning-vs-deep-learning-vs-data-science-71bd4a8402ec.

[2] A. Mehbodniya, A. J. P. Lazar, J. Webber, D. K. Sharma, S. Jayagopalan, K. Kittusamy, P. Singh, R. Rajan, S. Pandya, and S. Sengan, "Fetal health classification from cardiotocographic data using machine learning," *Expert Syst.,* vol. 39, no. 6, 2021. https://doi.org/10.1111/exsy.12899.

[3] A. K. Pradhan, J. K. Rout, A. B. Maharana, B. K. Balabantaray, and N. K. Ray, "A machine learning approach for the prediction of fetal health using CTG," In 2021 19th OITS International Conference on Information Technology, (OCIT 2021)*,* Bhubaneswar, India, December 16-18, 2021, IEEE, pp. 239-244. https://doi.org/10.1109/OCIT53463.2021.00056.

[4] J. Piri and P. Mohapatra, "Exploring fetal health status using an association based classification approach," In 2019 International Conference on Information Technology, (ICIT 2019), Bhubaneswar, India, December 19-21, 2019, IEEE, pp. 166-171. https://doi.org/10.1109/ICIT48102.2019.00036.

[5] N. R. Navuluri, "Fetal health prediction using classification techniques," *Int. J. Eng Res. Technol.,* vol. 10, no. 11, 2021.

[6] M. Ramla, S. Sangeetha and S. Nickolas, "Fetal health state monitoring using decision tree classifier from cardiotocography measurements," In 2018 Second International Conference on Intelligent Computing and Control Systems, (ICICCS), Madurai, India, June 14-15, 2018, IEEE, pp. 1799-1803. https://doi.org/10.1109/ICCONS.2018.8663047.

[7] J. Li, Y. Wang, B. Y. Lei, J. Z. Cheng, J. Qin, T. F. Wang, S. L. Li, and D. Ni, "Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting," *IEEE J. Biomed. Health,* vol. 22, no. 1, pp. 215-223, 2018. https://doi.org/10.1109/JBHI.2017.2703890.

[8] P. Dwivedi, A. A. Khan, S. Mugde and G. Sharma, "Diagnosing the major contributing factors in the classification of the fetal health status using cardiotocography measurements: An AutoML and XAI approach," In 2021 13th International Conference on Electronics, Computers and Artificial Intelligence,

(ECAI 2021), Pitesti, Romania, July 01-03, 2021, IEEE, pp. 1-6. https://doi.org/10.1109/ECAI52376.2021.9515033.

[9] M. L. Huang and Y. Y. Hsu, "Fetal distress prediction using discriminant analysis, decision tree, and artificial neural network," *J. Biomed Sci. Eng.,* vol. 5, no. 9, pp. 526-533, 2012. http://dx.doi.org/10.4236/jbise.2012.59065.

[10] S. Jadhav, S. Nalbalwar, and A. Ghatol, "Modular neural network model based foetal state classification," In 2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops, (BIBMW 2011), Atlanta, GA, USA, November 12-15, 2011, IEEE, pp. 915-917. https://doi.org/10.1109/BIBMW.2011.6112501.

[11] R. Afridi, Z. Iqbal, M. Khan, A. Ahmad, and R. Naseem, "Fetal heart rate classification and comparative analysis using cardiotocography data and known classifiers," *Int. J. Grid Distrib.*, vol. 12, no. 2, pp. 31-42, 2019. http://dx.doi.org/10.33832/ijgdc.2019.12.1.03.

[12] V. Chudacek, J. Spilka, L. Lhotska, P. Janku, M. Koucky, M. Huptych, and M. Bursa, "Assessment of features for automatic ctg analysis based on expert annotation," In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (IEMBC 2011), Boston, MA, August 30-September 3, 2011, IEEE, pp. 6051-6054. https://doi.org/10.1109/iembs.2011.6091495.

[13] D. Ayres-de-Campos, J. Bernardes, A. Garrido, J. Marques-de-sá, and L. Pereira-leite, "A program for automated analysis of cardiotocogram," *J. Matern-Fetal Med.*, vol. 9, no. 5, pp. 311-318, 2009.

[14] A. Batra, A. Chandra, and V. Matoria, "Cardiotocography analysis using conjunction of machine learning algorithms," *In* 2017 International Conference on Machine Vision and Information Technology, (CMVIT *2017), Singa*pore, February 17-19, 2017, IEEE, pp. 1-6. https://doi.org/10.1109/CMVIT.2017.27.

[15] H. B. Zhou and G. W. Ying, "Identification of ctg based on bp neural network optimized by pso," In 2012 11th International Symposium on Distributed Computing and Applications to Business, Engineering Science, (DCABES 2012), Guilin, China, October 19-22, 2012, IEEE, pp. 108-111. https://doi.org/10.1109/DCABES.2012.97.