Design Engineering

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

Prediction System design for monitoring the health of developing infants from cardiotocography using Statistical Machine Learning

Aaryan Shah
Dwarkadas J Sanghvi College of Engineering
shah.aaryan27117@gmail.com

Jeet Patel
Dwarkadas J Sanghvi College of Engineering
jeetp5118@gmail.com

Devanshi Chokshi
Dwarkadas J Sanghvi College of Engineering
chokshidevanshi@gmail.com

Esha Bhave
Dwarkadas J Sanghvi College of Engineering
eshabhave@gmail.com

Drumil Joshi
Dwarkadas J Sanghvi College of Engineering
abcdrumil@gmail.com

Dr. Sunil Karamchandani
Dwarkadas J Sanghvi College of Engineering
sunil.karamchandani@djsce.ac.in

Abstract:

Today segregation models are widely used in health care, which is intended to support physicians in diagnosing diseases and reducing human error. The challenge is to use effective methods to extract real-world data from the medical field, as many different models have been proposed with varying results. Many researchers have focused on the problem of variability in real-time data sets in segmentation models. Some previous works create mechanisms that include similar graphs for information display and information acquisition. However, such methods are weak in finding different relationships between elements. The purpose of this diagnostic method is to measure the baby's heart rate and uterine contractions during the third trimester of pregnancy, when the baby's heart is fully functional. Cardiotocogram findings are usually divided into three categories: physical, suspicious, or pathological. The purpose of this work is to automatically distinguish these regions using cardiotocographic data. In this study, the Random Forest method shows that it performs very well, capable of analyzing data with 94% accuracy. Comparisons with the Separation and Reversal Tree as well as mapping methods are also provided in the corresponding research paper.

Keywords: *segregation, segmentation, cardiotocogram, reversal tree*

I. INTRODUCTION

Fetal cardiovascular diseases are caused due to a multitude of factors related to an abnormality in the number of a baby's chromosomes, single gene defects, or environmental factors and fetal conditions. These disorders are seen in 1-3% of the population, and the death rate from congenital heart disease was 81 per 100,000 live births. About 60-70% of them are detected before birth using ultrasound. Cardiotocogram is another way to measure a child's heart rate, using different signals to record a child's heart rate. Using the given guidelines by the International Federation of Gynecology and Obstetrics (FIGO), based on the child's heart rate, heart rate, acceleration, and diminution, a cardiotocograph can be classified as normal, suspicious, or abnormal. Skilled medical personnel can do this. However, the greatest variation in the fetal heart rate is observed, making it difficult for obstetricians to obtain accurate information. This involves the risk of false positives, leading to unnecessary medical interventions. To avoid this problem, machine learning algorithms can be employed for an enhanced prediction algorithm, to better predict fetal congenital anomalies, minimizing the risk of human error and misdiagnosis. This will lead to improved outcomes in cases of fetuses with structural heart disease, arrhythmias, or cardiovascular dysfunction - both, for prenatal treatment, as well as strategies for post-birth care. This will reduce the mortality rate in cases of fetal cardiovascular diseases, and lead to stability for high-risk fetuses. To produce a diagnosis, we use sophisticated mathematical algorithms and various data points from the human body. The accuracy of predicting a child's heart pattern has improved because of these mathematical examples. Medical and engineering specialists were working to translate cardiotocography automatically. The purpose of this study was to create a machine learning model that could accurately detect high risk embryos (both suspicious and pathological) accurately as highly trained medical professionals such as obstetricians, thereby eliminating false positives and minimizing ambiguity, resulting in minimal variability of outcome. separation. This method has a high degree of accuracy in predicting the pathological status of the fetus and determining the fetus's prenatal health.

II. LITERATURE SURVEY

Since [1] shows the correlation between the performance of machine learning models (SVM, random forest, retrospective, and Naive Bayes) concluded that the model retrieval model provides a maximum accuracy of 99.5 percent.

Second, [2] explains that the classification model developed using the XGBoost (Extra-Gradient Boosting) method had the highest and most accurate predictor of adverse fetal outcome and could be used by health workers in low- and middle-income countries to be screened. pregnant women in remote areas for early diagnosis.

The following article quoted [3] states that the authors calculated third, fifteenth, fiftieth, ninetieth centile curves according to the gestational age of these ultrasound measures to avoid any misconceptions about removing foreign objects, which represent international standards. by countries. With fetal growth, they recommend these international baby growth rates for clinical

Design Engineering

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

interpretation of standard ultrasound measurements.

In quote four [4] the authors propose a new legal model of segregation to help diagnose and classify embryonic, suspicious, and pathological conditions. Selection of features played an important role in the article as it describes the features that no longer work using the main variant formula of the Main Section Analysis and Line Discrimination Analysis.

In [5], in addition to the concept of machine learning, a neural network-based approach is proposed to monitor a child's health. The proposed method determines whether the child is normal or suspicious or pathological. The accuracy of diagnosing the condition is up to 99.9% which can be used effectively in child health monitoring. The parallel neural network may be at risk of overcrowding under the care of the author.

Excerpt [6] suggests an AutoML the Light Gradient Boosting Machine algorithm uses PyCaret to differentiate a child's health. The model provides 95.61% accuracy and some of the test metrics for this model are healthy and satisfying to contribute not only to the obstetrics and gynecology but also to the biomedical instrument industry.

Finally [7] the authors recommend a model to analyze the child's heartbeat records to improve the accuracy of the sections. The use of basic FHR data methods to extract features became outstanding, and SVM (Vector Support Machine Filter) and MLP were used to classify FHR records accurately.

III. METHODOLOGY

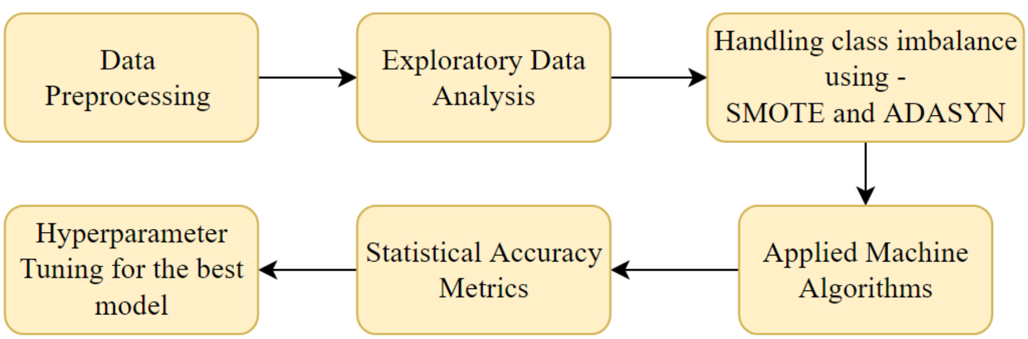

Fig 1: Block Diagram of Workflow

A. Exploratory Data Analysis

Exploratory Data Analysis is a procedure that helps to discover trends, anomalies, visualize and patterns within the data. We also get better analysis using different types of graphs like histograms, pie charts, Heatmap, Data Distribution, etc. To get an idea about the distribution of each feature in the dataset, we have used histograms to group the data into plots. Count of the number of observations in each plot created was provided by the histograms. We can easily

Design Engineering
ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

notice that the distribution is Gaussian, skewed or exponential. It overall helped us for a better understanding of the dataset and its distribution.

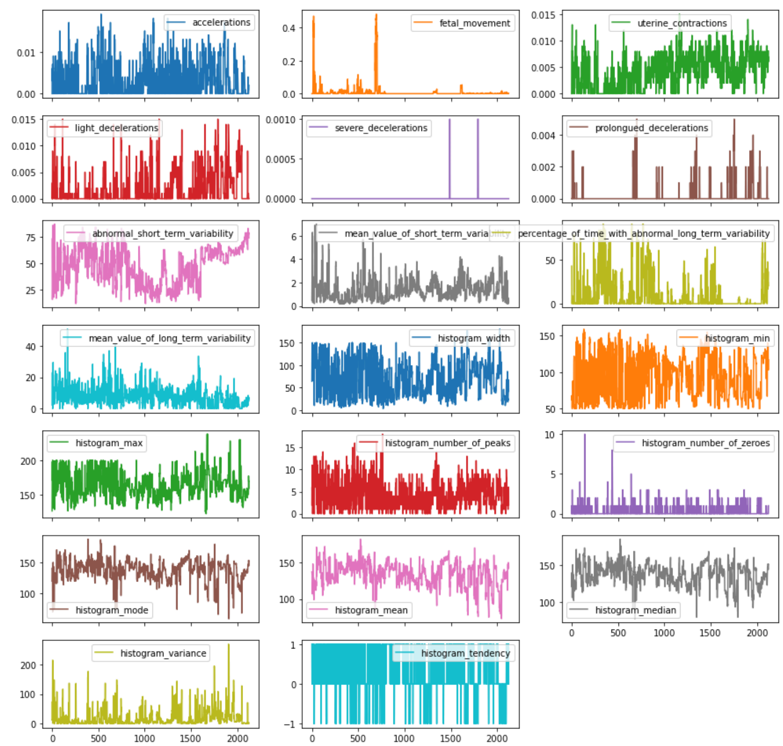

Fig. 2. Different histogram plots of features present in the dataset

A data distribution graph and pie chart have been used particularly for the better conception of the target variable (fetal health) valued 1, 2, and 3 named as normal, suspect, and pathological respectively, as this helps in the machine learning procedure.

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

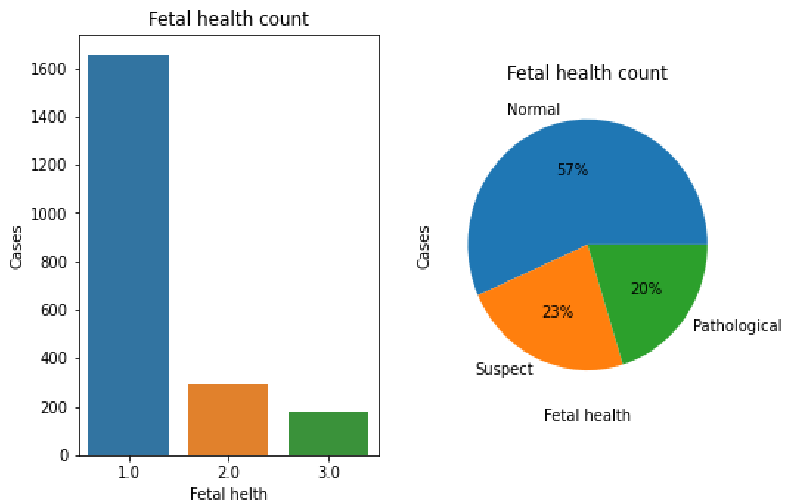

Fig. 3. Data distribution plot of target variable fetal health

We have used the lmplot() method to draw a scatter plot for accelerations and fetal movements.

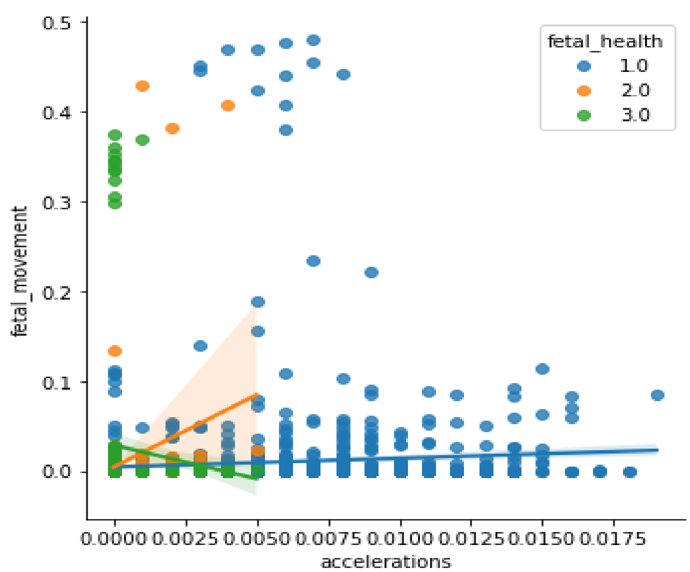

Fig. 4. Linear model plot of acceleration vs fetal movement

A heatmap essentially is a graphical presentation in which each character values of a matrix are represented by colors for a better understanding of patterns. In our project, we've used Heatmap to locate the correlation among the different independent variables.

Design Engineering

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

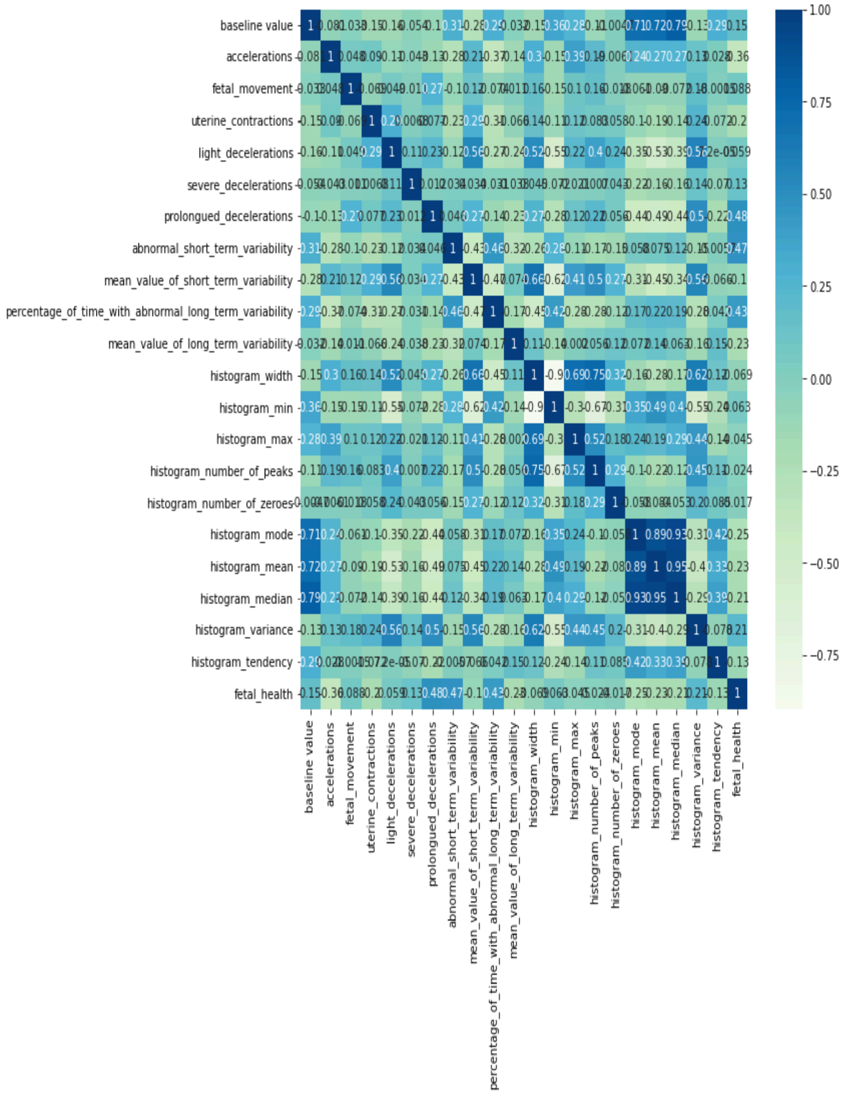

Fig. 5. Heat map plot of different features

B. Handling Class Imbalance

For our project, to overcoming the class imbalance we have used two oversampling techniques, SMOTE (Synthetic oversampling minority technique) and ADASYN (Adaptive Synthetic Sampling Approach)

SMOTE: Synthetic Minority Oversampling Technique- an exaggerated sampling method when synthetic samples were performed in a small class. This algorithm allows you to overcome the problem of overfilling caused by random sampling. It works specifically in the feature space to recruit new cases with the help of adding among the best conditions that lie together.

ADASYN: The adaptive synthetic sampling approach is a general form of the SMOTE algorithm. This algorithm also targets to oversample the minority class through generating synthetic cases for it. But the difference right here is it considers the density distribution, which makes a choice at the number of synthetic cases generated for samples which may be hard to learn. Due to this, it allows adaptively changing the decision boundaries primarily grounded on the samples tricky to learn.

C. Applied Machine Learning Algorithms

The goal is to construct a more accurate machine learning model. We divided our data into a 65:35 ratio for training and testing using the sklearn library. The following are the machine learning algorithms that we have considered:

a. Sigmoid Regression -

Sigmoid regression is capable to solve a specific type of problem known as classification problems. The goal is to figure out that the current object under observation belongs to which category. Probability always ranges between 0 (does not occur) and 1 (occurs). We use the logistic function to calculate probability in logistic regression. The logistic function is a simple S-shaped curve also known as a sigmoid function that is used to convert data into a value between 0 and 1. These probability ranges between 0 and 1 and using a threshold classifier, these values will be then converted into either 0 or 1.

$$Y = \frac{1}{1+e^{-(\beta 0+ \beta 1X)}}$$

'Y' is the output of logistic function, where 0Y1
'β1' is the slope
'βo' is the y-intercept
'X' is the independent variable
After applying the Logistic Regression algorithm for all three data samples we got after using the oversampling techniques, we found that the actual data sample was giving accuracy of 88% and there was a decrease in accuracy for smote and adasyn data sample giving 85% and 82% accuracy respectively.

b. Ensemble Forest -

Ensembling is a supervised machine learning algorithm that can be used to solve problems such as editing and retrieval. It is made up of different N-tree trees. Random forests create decision-making trees with randomly selected data samples, generate individual tree predictions, and vote for the best solution. Because of the variety of decision trees involved in creating models, informal forests are the most accurate and powerful method. By averaging different decision trees, it avoids the problem of overfitting in the model.

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

$$RFfi_i = \frac{\sum_{j \in all\ trees} \text{normfi}_{ij}}{T}$$

Where,
$RFfi_i$= feature i's importance
$normfi_{ij}$= characteristic that has been normalized the significance of feature i in the tree j
T = number of trees in total
After applying the random forest algorithm for all three data samples we got after using the oversampling techniques, we acquired that the smote data sample was giving accuracy of 94% and for actual and adasyn data samples, we found no such difference in their accuracies as we get 93%.

c. Decision Tree -

Decision Tree is a supervised learning method that can be used to solve planning and retrospective problems but is much better at dividing. It can be created using an algorithm that classifies the database in different ways based on the context. The Decision node and the Leaf node are two locations in this path. Decision areas are used for decision making and have several branches, and Leaf nodes are the result of those decisions and have no additional branches. Decisions or evaluations are made based on the characteristics of the data provided. The Gini Index is used to determine whether a database needs binary separation. The total value of the Gini index is 0 and the worst is 0.5 (in two-phase problems). The following figure is used to calculate the Gini index:

$$Gini(D) = 1 - \sum_{i=1}^{m} P_i^2$$

where Pi = probability that tuple in D belongs to Ci
After applying the Decision Tree algorithm for all three data samples, we found that the actual data sample was giving accuracy of 92.7%. For smote and adasyn data samples, we acquired the same accuracy of 90%.

d. SVM -

The goal of the SVM method is to find the best line or decision boundary to divide areas n-n into classes so that additional data points can be easily placed in the appropriate category in the future. Hyperplane is the name of the best decision limit. In the multidimensional space, the SVM model represents different classes in the hyperplane. SVM will generate a hyperplane in a repetitive manner to minimize error. To create a hyperplane, this selects the extreme point or vectors. Supporting vectors are extreme examples. The purpose of SVM is to classify data sets into classes to achieve the highest hyperplane (MMH).

$$decision\ function = \sum_{i \in SV} y_i\ \alpha_i K(x_i, x) + b$$

x=sample

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

αi= dual coefficient
b = intercept
After applying the Support Vector Machine algorithm for all three data samples, we learned that the actual data sample gave accuracy of approximately 92%. Whereas, smote and adasyn data sample gave an accuracy of 86% and 85% respectively compared to actual data sample.

e. Naive Bayes Classifier -

The Naive Bayes method is a supervised learning method that uses Bayes theory to solve classification problems and assist in the development of fast machine learning models that can make faster predictions. It classifies possible categories, which means making predictions based on the probability of an object. Compared to other algorithms, it works well in multiple class estimates. With prior knowledge, the Bayes' theorem is used to calculate the probability of a hypothesis. Conditional opportunities determine this. The Bayes theorem's formula is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

After applying the Naive Bayes Classifier algorithm for all three data samples, we learned that the smote data sample gave accuracy of 71%. Whereas, actual and adasyn data samples gave an accuracy of 69% and 68% respectively.

D. Statistical Accuracy Measures

To evaluate the performance of a child health predictive editor, we use statistical measurements of accuracy, precision, recall, and f1-score. The accuracy of the machine learning model is the size used to determine which version differs in finding relationships and patterns between data variations based on inputs, or training data. If the higher version can access 'unseen' data, higher forecasts, and more productive data, which in turn provides more business value. The good and false cases are represented by TP and FP, respectively, while the opposing true and false cases are represented by TN and FN.

$$Accuracy = \frac{TP + TN}{TP+FP+TN+FN}$$

Precision, also known as positive predictable value, is part of the positive conditions found within the conditions obtained, while memory, also known as sensitivity, is a percentage of positive events obtained. The ratio of actual to positive to the total displayed value is known as accuracy.

$$Precision = \frac{TP}{TP+FP}$$

The number of positive instances accurately detected is referred to as recall.

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

$$Recall = \frac{TP}{TP+FN}$$

The F1-score metric makes use of an aggregate of precision and recall. In fact, the F1 score is the harmonic mean of the two.

$$F1 = \frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

E. Hyperparameter Tuning for the Best model

Machine Learning Algorithms were drafted on three different sampled data which consisted of Actual, SMOTE, and ADASYN data. The SMOTE and ADASYN are sampling techniques used to handle the problem of class imbalance making the model more biased, failing the bias-variance tradeoff. The ML Algorithms implemented to classify were Sigmoid Regression, Decision Tree Classifier, Ensemble Classifier, Support Vector Classifier and Naïve Bayes Classifier. The highest successful model was the Ensemble Classifier when SMOTE data was provided with a whooping accuracy of 94%. Then we used Grid Search CV from model selection and found out the best model parameters for the best metrics. The parameters suitable for best results are mentioned below in the table.

TABLE I

Parameters	Values
Criterion	gini
max_depth	40
max_features	sqrt
n_estimators	200
n_jobs	-1
Bootstrapping	True

IV. RESULTS AND DISCUSSIONS

A radar chart is a visual tutorial tool where multiple variables (three or more) are compared in a two-dimensional plane. We created different axes from the common central area. In SMOTE and ADASYN cases, all axes are evenly distributed and pulled evenly against each other. Sometimes, axes are also connected to each other to form different grids making it easier for us to plan a

Design Engineering

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

spider chart. In our scenario, we have plotted Accuracy, Precision, and F1-Score on the radar plot. Logistic Regression as it's based on sigmoid probability faces a backslash in all the 3 sampled data sets as the data is not in Gaussian surface. The decision tree being the base of the best performing model depicts a high accuracy but when visualized it creates a loophole for bias-variance tradeoff to come into practice. SMOTE - a synthetic minority up-sampling technique plays a vital role in all the accuracy metrics. An accuracy of range 93-95% is achieved when the data is sampled by SMOTE technique The role of precision is very under-rated when it comes to the prediction of the health status. As it is related to the life of an infant, we need to check that the precision should be as precise making the False Positive ratio negligible. Thus, to conclude the results part, the Random Forest classifier works efficiently in SMOTE technique as compared to ADASYN techniques as ADASYN considers the concept of eigenvectors, and thus a decrease of 1-2 % in accuracy is observed.

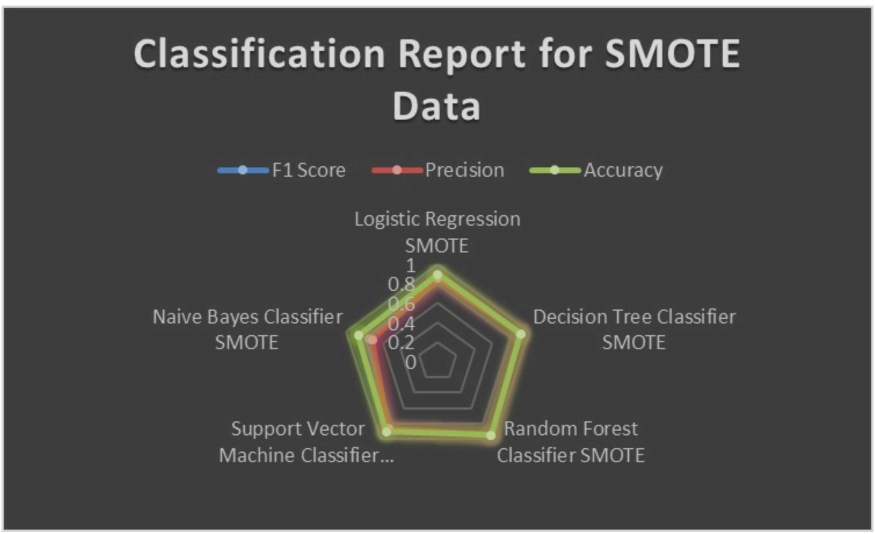

Fig. 6. Classification report of SMOTE data

Design Engineering
ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

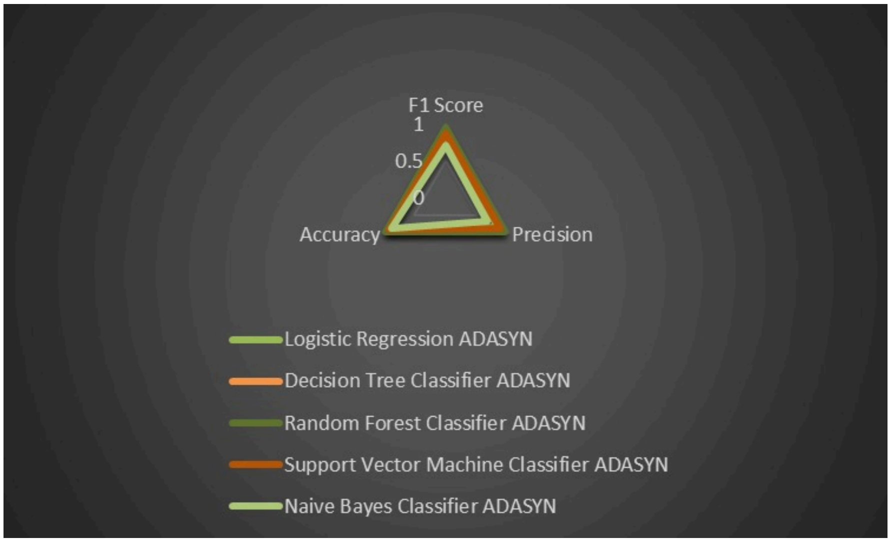

Fig. 7. Classification of ADASYN data

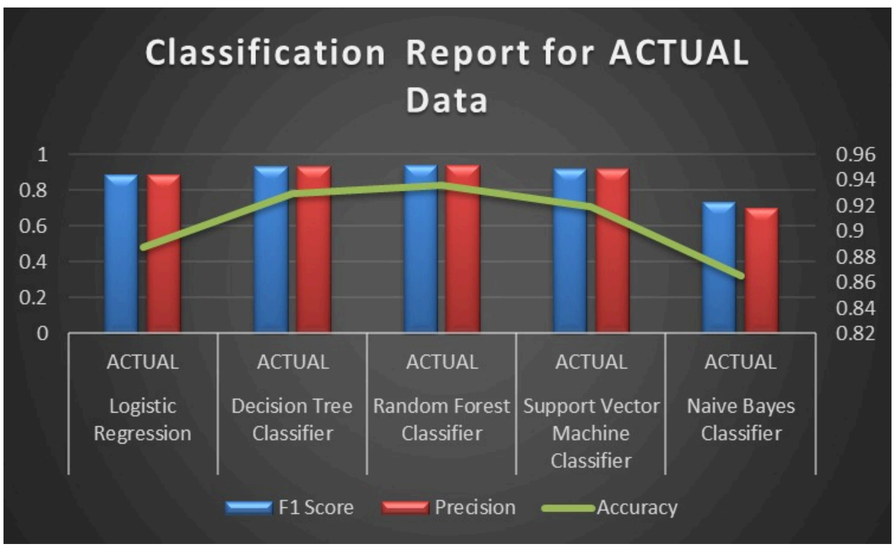

Fig. 8. Classification report for ACTUAL data

V. CONCLUSION

Cardiotocography is a low-cost means of monitoring fetal health and a method used for lowering the child fatality rate. Visual analysis errors were one of the most serious problems with CTG monitoring. Any type of interventional surgery, whether necessary or not, raises the risk of complications. Our ML model effectively differentiated normal and disturbed fetal health and was accurate enough in its predictions to avoid any unwanted interventive surgical procedures to reduce the overall risk of child mortality. In the entire research paper, we have done machine learning classification on fetal health dataset, starting with exploratory data analysis where we found the pattern and distribution of each feature in the dataset using different types of a graph

ISSN: 0011-9342
Vol: 2021 Issue: 07| Pages: 16142 - 16153

such as a heatmap, histogram, and data distribution, etc. The problem of class imbalance was resolved by using SMOTE (Synthetic oversampling minority techniques) and ADASYN (Adaptive Synthetic Sampling Approach) are data sampling techniques. A well-trained model helped in identifying which variable changes to the FHR had the most significant effect on fetal health. We used five different types of classification algorithms like Logistic Regression, Random Forest, Decision Tree, Support Vector Machine, and Naive Bayes classifier. These algorithms were applied on each of the three different sampled data which consisted of Actual, SMOTE, and ADASYN data. The statistical metrics we used to differentiate the five algorithms were accuracy, precision, recall, and f1-score. On comparing the results, the best performing algorithm was Random Forest Classifier on SMOTE data sample. It gave the highest accuracy of 94% among all the remaining classifiers. Further, we used Grid Search CV and found out the best model parameters for the best metrics.

REFERENCES

[1] Fetal Health Prediction using Classification Techniques by Naveen Reddy Navuluri.

[2] Use of Machine Learning Algorithms for Prediction of Fetal Risk using Cardiotocographic Data by Zahra Hoodbhoy, Mohammad Noman, Ayesha Shafique, Ali Nasim, Devyani Chowdhury and Babar Hasan.

[3] International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project by Aris T Papageorghiou, MD Eric O Ohuma, MSc Prof Douglas G Altman, DSc Tullia Todros, PhD.

[4] Exploring Fetal Health Status using an Association Based Classification Approach by Jayashree Piri and Pushpanjali Mohapatra..

[5] An innovative method for fetal health monitoring based on an artificial neural network using cardiotocography measurements by Subrota Mazumdar, Rohit Choudhary and Aleena Swetapadma.

[6] Diagnosing the major contributing factors in the classification of the fetal health status using cardiotocography measurements: An AutoML and XAI approach by Prakriti Dwivedi, Akbar Ali Khan, Sareeta Mugde and Garima Sharma.

[7] Automatic Classification of Fetal Heart Rate Based on Convolutional Neural Network by Zhuang-Zhuang Chen, Luxiang Huang, Min Fang; Bing Li, Xianghua Fu, Huihui Wang and Qingguo Zhao

[8] Prakriti Dwivedi, Akbar Ali Khan, Sareeta Mugde, Garima Sharma. "Diagnosing the major contributing factors in the classification of the fetal health status using cardiotocography measurements: An AutoML and XAI approach"

[9] Tomáš Peterek, Petr Gajdoš, Pavel Dohnálek, Jana Krohová. "Chapter 19 Human Fetus Health Classification on Cardiotocographic Data Using Random Forests" , Springer Science and Business Media LLC, 2014

[10] Pijush Dutta, Shobhandeb Paul, Madhurima Majumder. "An Efficient SMOTE Based Machine Learning classification for Prediction & Detection of PCOS" , Research Square Platform LLC, 2021