FETAL HEALTH CLASSIFICATION

Anurag Gautam
Janak Kapuriya
Muskan Gupta
Saloni Garg

1 Abstract

Obstetricians widely use Cardiotocography(CTG) in pregnancy as a method of assessing fetal well-being, predominantly in pregnancies with increased risk of complications, it gives data on the fetal heartbeat and uterine contractions, which helps to identify whether the fetus is pathological or not. Obstetricians have traditionally analyzed CTG data artificially, which takes time and is unreliable. Due to this, creating a fetal health classification model is essential, as it may save time and medical resources in the diagnosis process.
In this study, we have taken the fetal health dataset with 22 features and attempted to classify them into normal, suspect, and pathological using different classifiers, and selecting the one that gives the best result.
It is discovered that the support vector machine produces the best result. It achieves a 90.05% recall score, which is better than all other models.

2 Problem Definition

Given a new sample that will have all the attributes value like the fetal health dataset, our task is to predict, by considering the feature values whether fetal health lies in the category of

- Normal
- Suspect
- Pathological

All these features represent the severity level of fetal health.

3 Related Work

The paper [2] takes the help of CTG data and classify fetal health as physiological, suspect and pathological on the basis of fetal heart beat and uterine contractions of mother. It uses random forest method for classification and achieved accuracy of 94.69%.
The paper [3] uses Association Based Classification Approach and was 84% accurate.
The paper [4] worked on Artificial Neural Network to classify fetal health and attained F-Score of 0.9784, 0.4514, 0.9724 for Normal, Suspect and Pathological class respectively. Along the same lines, [4] experimented with ANN and Logistic Regression to attain accuracy of 98.5% and 98.7% respectively. But this paper only considers two fetal states: normal and pathological.

4 Dataset Description

In this project study, we have used a fetal health classification dataset from the Kaggle dataset. The dataset contains 2126 records of features which were then classified into 3 classes:

- Normal:- label=1
- Suspect:- label=2

- Pathological:- label=3

These 2126 measurements are extracted from cardiotocography and classified by three expert obstetricians. The dataset contains 22 features which are named as follows:-

- Baseline value:- It is Baseline Fetal Heart Rate which helps in monitoring acceleration and deceleration which means that both attributes are computed with respect to baseline
- Acceleration:- It means the number of accelerations per second
- Fetal movement:- It is recorded per 20 minutes.
- Uterine contractionis recorded when there is pressure on the stomach
- Light deceleration:-It occurs when fetal heartbeat decreases 15 beats below baseline with uterine contraction
- Severe decelerations:- It occurs when the heart beat increases with some delay i.e. it is not increasing as a contraction moves down.
- Prolonged deceleration:- deceleration is when the heartbeat does not come back to normal for more than 15 seconds Variability is variation in lines of tocograph.
- Abnormal short term variability
- Mean value of short term variability
- Percentage of time with long term variability
- Mean value of long term variability
- Histogram width
- histogram min
- Histogram max
- Histogram number of peaks
- Histogram number of zeros
- Histogram mode
- Histogram mean
- Histogram median
- Histogram variance
- Histogram tendency
- Fetal heath

5 Methodologies

5.1 Data Preprocessing

In the data Preprocessing part, we have done the following operations:-

- **Null values detection**
  We checked if there were any values in the dataset. After checking for the same we found that the dataset contains 0 null values.
- **Duplicate rows detection**
  we checked if there were any duplicate rows in the dataset. After checking for the same we found that the dataset contains 13 duplicate rows. After that, we removed those duplicate rows from the dataset.

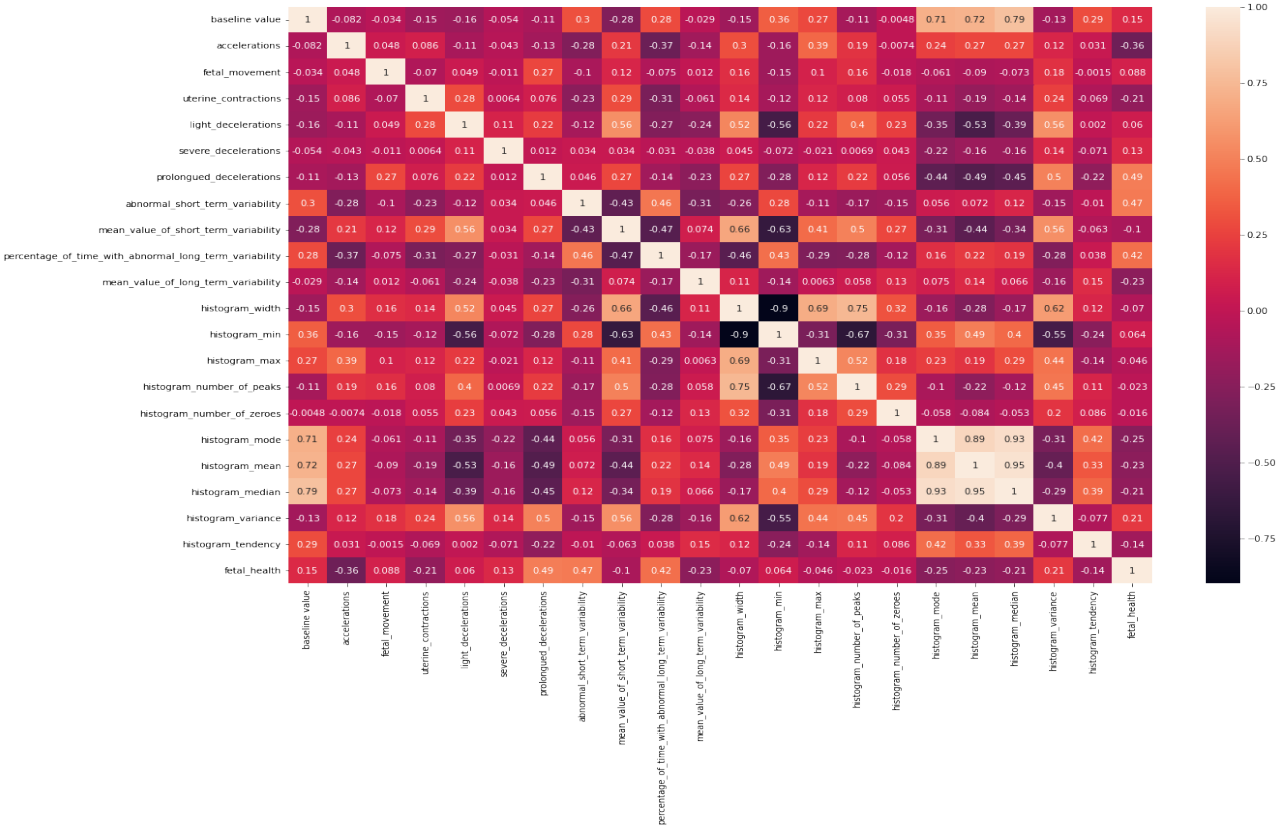
- **Correlation between the features**
  We checked which features in the dataset are highly correlated to each other by plotting a heatmap for the dataset.

  After plotting the heatmap we found the correlated features in the dataset are:-
  Histogram_mean, Histogram_mode,Histogram_median
  Therefore, we dropped all the correlated features from the dataset.

5.2 Exploratory Data Analysis

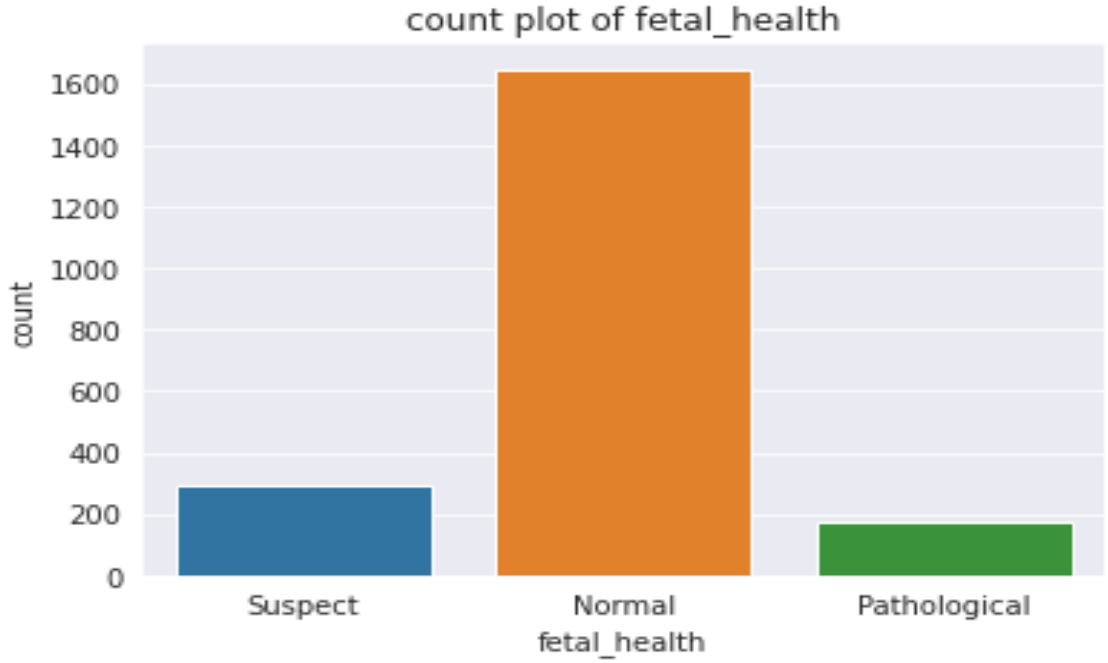
In the EDA part, we have done the data visualizations by plotting the plots and making the observations:-

- **Heat Map**

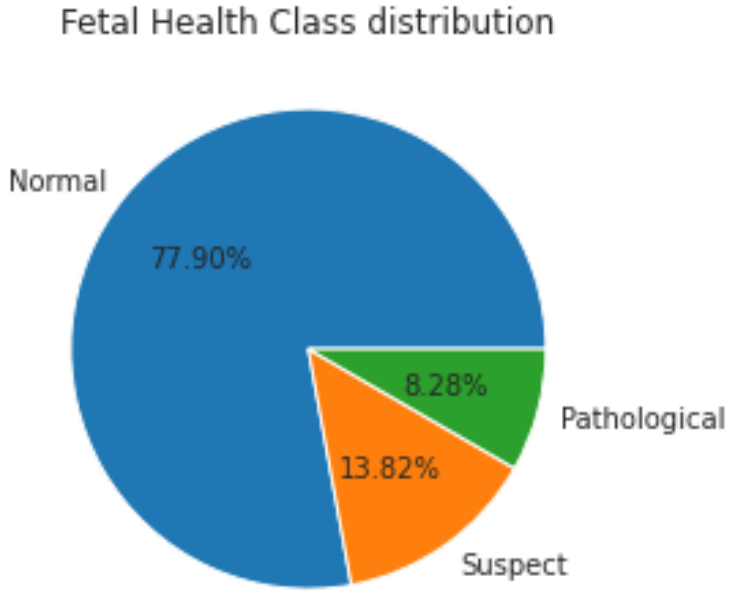


This heatmap is used in finding correlation between features of dataset.

- **Count plot for goal (fetal health)**

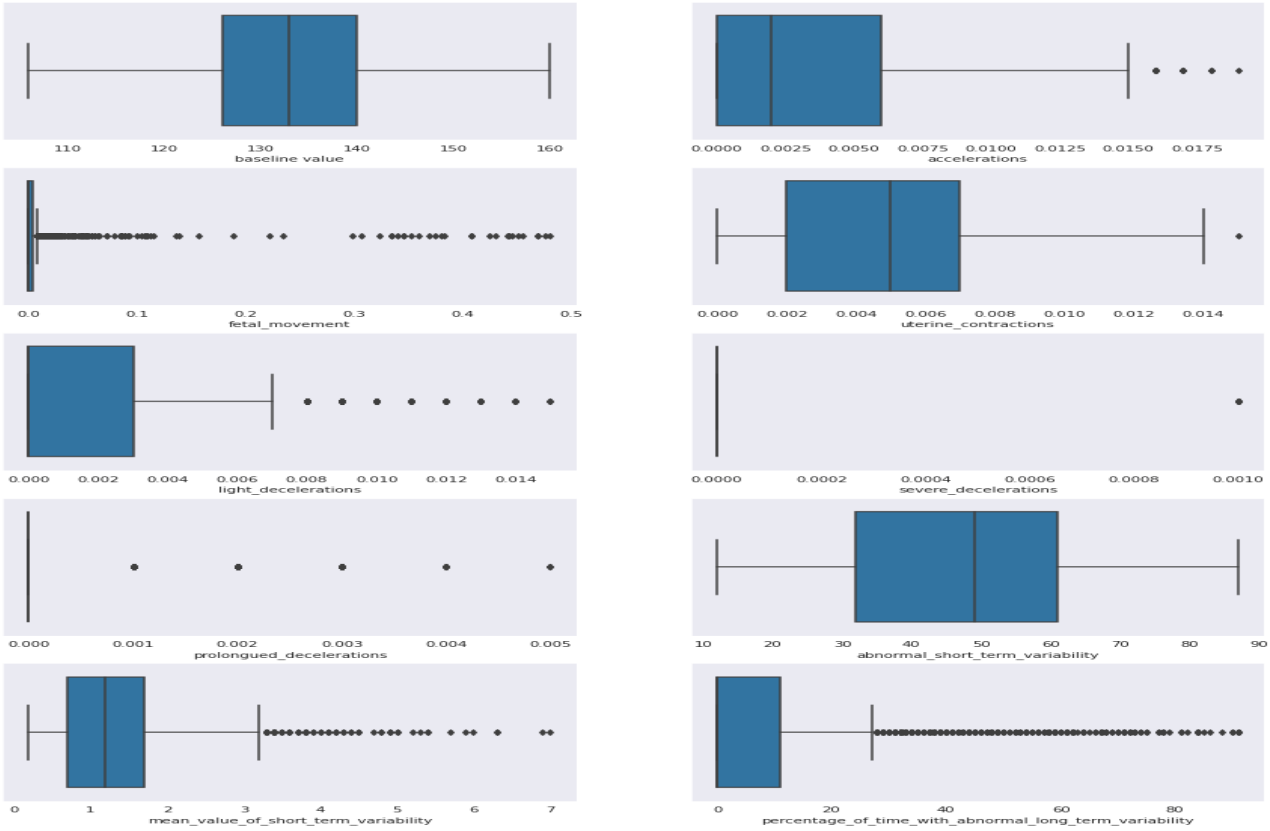


- **Pie chart**



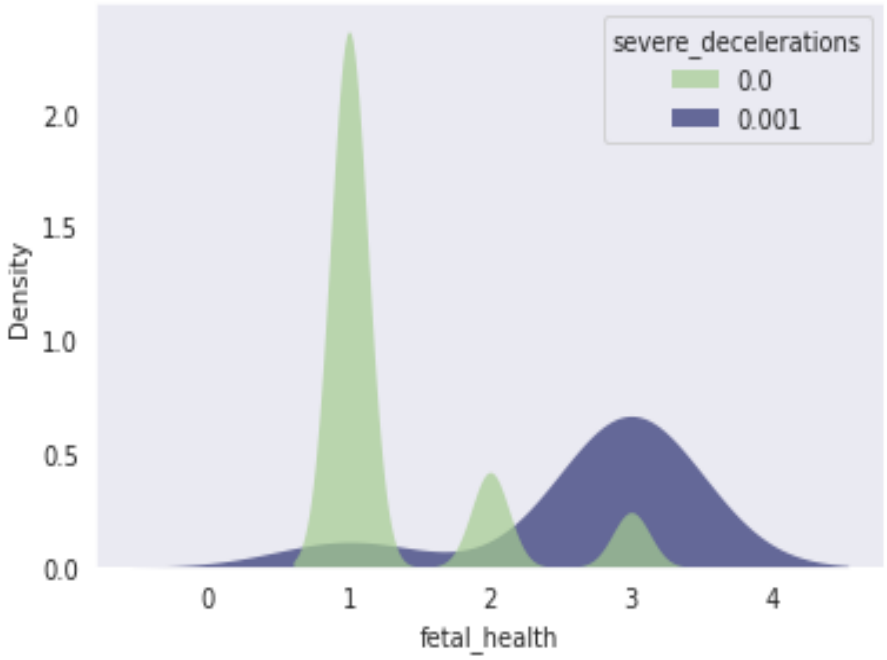
The observation made from the above two graphs is that the dataset is imbalanced. Therefore to make the dataset balance we used the weighted technique. There are no missing attributes in the dataset, and the class distribution is 1665 normal, 295 suspect, and 176 pathological.

- **Box plot**



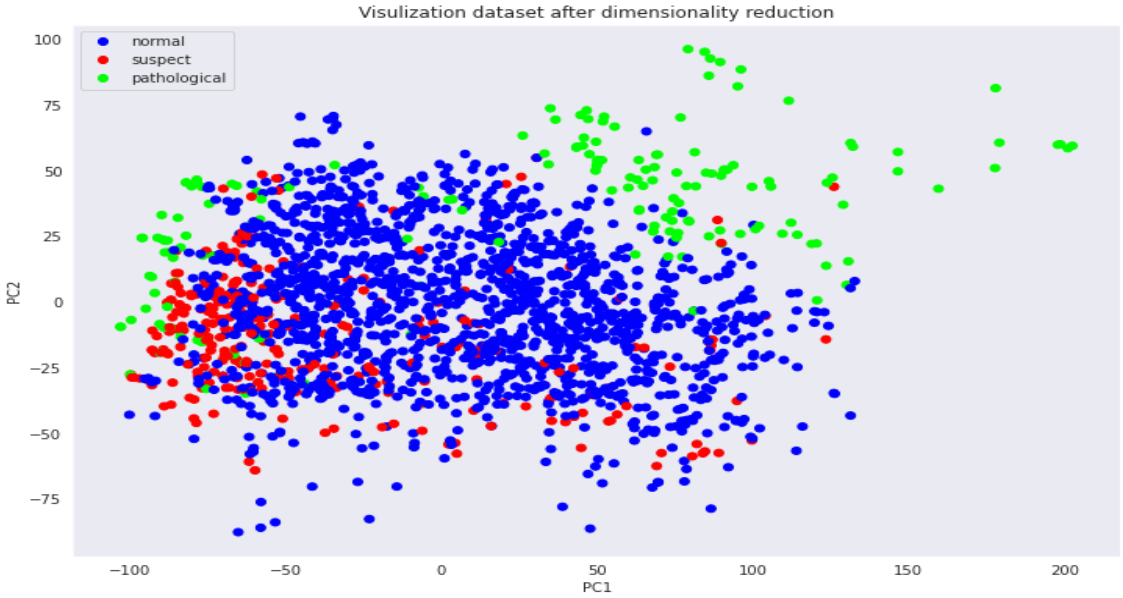
Observation made from the box plot is that only the 'baseline value' feature's data is normally distributed whereas all the remaining features' data are generally skewed and contain outliers.

- **KDE plot between features (goal and severe_deceleration)**



Observation made from the KDE plot is that most of the dataset which has fetal health= normal have severe_deceleration value= 0.0 whereas most of the other dataset having sever_deceleration=0.001 have fetal health= pathological.

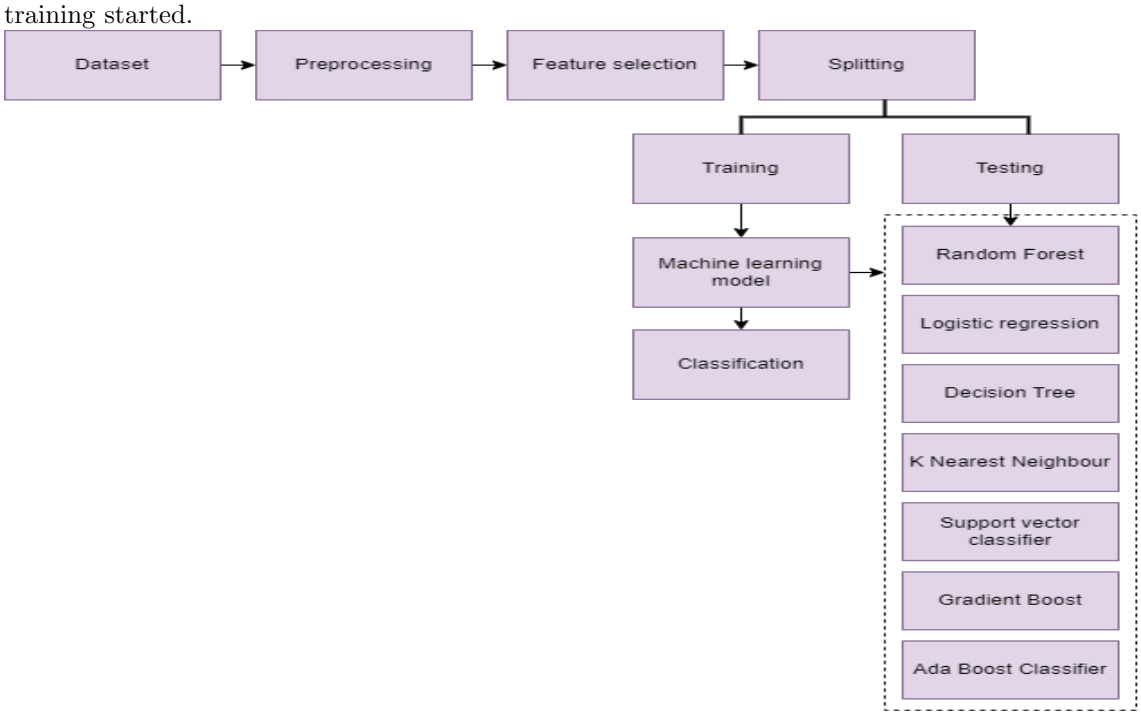
- **PCA Scatter Plot**



From the above, we observe that the data is not linearly separable.

5.3 Block Diagram of the System

The architectural diagram of the ML system is shown below.
The system makes use of the CTG dataset, which includes all of the characteristics and values. We performed the preprocessing of the data. The characteristics necessary for prediction have been assigned, and the goal value has been specified so that the model can forecast. The dataset was then separated into training and testing subsets with stratified equals to fetal health. Then, the system's training started.



5.4 Selection of Classifier

To select the classifiers for our baseline model, we ran a list of classifiers first on our dataset and recorded the values of the confusion matrix, test and train accuracy, and recall-score.

5.4.1 ML classifiers

Logistic Regression

Logistic regression is a simple and more efficient method for binary and linear classification problems. It is a classification model, which is very easy to realize and achieves very good performance with linearly separable classes.

Support Vector Machine

A support vector machine (SVM) is a machine learning algorithm that analyzes data for classification and regression analysis. SVM is a supervised learning method that looks at data and sorts it into one of two categories. An SVM outputs a map of the sorted data with the margins between the two as far apart as possible. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes.

K Nearest Neighbours

The k-nearest neighbor's algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point.

Decision Tree

Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

Random Forest

A random forest is a meta estimator that fits a number of decision tree classifiers on various subsamples of the dataset and uses averaging to improve the predictive accuracy and control overfitting.

Gradient Boost

This algorithm builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions.

Ada Boost

AdaBoost, short for Adaptive Boosting, is a statistical classification meta-algorithm used in conjunction with many other types of learning algorithms to improve performance.AdaBoost is adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers.

5.4.2 Final Models

Dataset was first filtered using various filters and preprocessing techniques to reduce the noise and then fed to the models.
Recall Score is our main evaluation metric for the models.
Higher value of Recall score lowers the False Negative and vice-versa. The final Models with their accuracy and classification report are:-

- **Logistic Regression**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
0.8420118343195266	0.8203309692671394	0.8001990346817932

  The confusion matrix build through the model.

	1	2	3
1	272	41	17
2	3	50	5
3	1	9	25

- **Support Vector Classifier**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
0.9242603550295858	0.8912529550827423	0.900532417773797

The confusion matrix build through the model.

	1	2	3
1	292	37	1
2	3	54	1
3	2	2	1

- **K Nearest Neighbour**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
1.0	0.9267139479905437	0.8730507040851868

The confusion matrix build through the model.

	1	2	3
1	316	13	1
2	12	45	1
3	2	2	31

- **Decision Tree Classifier**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
0.9491124260355029	0.9148936170212766	0.7894561377319999

The confusion matrix build through the model.

	1	2	3
1	322	6	2
2	16	41	1
3	10	1	24

- **Random Forest Classifier**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
0.9177514792899408	0.900709219858156	0.8808876946807981

The confusion matrix build through the model.

	1	2	3
1	301	24	5
2	8	49	1
3	1	3	31

- **Gradient Boost Classifier**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
0.9781065088757397	0.9408983451536643	0.8781509678061402

The confusion matrix build through the model.

	1	2	3
1	323	6	1
2	14	43	1
3	2	1	32

- **Ada Boost Classifier**
  The model's accuracy and recall score

Training accuracy	Testing accuracy	Recall Score
1.0	0.9361702127659575	0.8600636910981739

The confusion matrix build through the model.

	1	2	3
1	322	8	0
2	13	45	0
3	3	3	29

6 Experimental Results

In this section, we analyzed the results obtained using different classifiers on the preprocessed dataset and found that the model for the dataset with the highest accuracy and recall score is the Support vector classifier.
Therefore, we applied Grid SearchCV to find the best parameters for the support vector classifier. The model, built after finding the best params using Grid SearchCV, accuracy and recall score

Testing accuracy	Recall Score
0.9125295508274232	0.8262775538637608

We observe here that the accuracy and recall score for the SVC after applying Grid SearchCV is not better than before. Therefore, We continued with our previous parameters.

7 Analysis

We have observed from the distribution of class labels that the dataset was imbalanced.
Recall Score is our main evaluation metric here because we need less False Negative as less as possible from model prediction. Because in the medical domain problem we need fewer False Negative as they would have a bad impact on Fetus's health.

8 Contribution of Each Member

Phases of the project are:-

- Data Preprocessing
- Exploratory Data Analysis
- Model Training and Evaluation

All the project memebers have contributed in each of the phase of the project.

9 References

[1] https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification

[2] Peterek, J. (2014). Human Fetus Health Classification on Cardiotocographic Data Using Random Forests. In Intelligent Data analysis and its Applications, Volume II (pp. 189-198). Springer International Publishing.

[3] Piri, J., Mohapatra, P. (2019). Exploring Fetal Health Status Using an Association Based Classification Approach. In 2019 International Conference on Information Technology (ICIT) (pp. 166-171).

[4] Chinnasamy, S., Chitradevi, M., Geetharamani, G. (2012). Classification of Cardiotocogram Data using Neural Network based Machine Learning Technique. International Journal of Computer Applications, 47, 19-25.