Identifying the health of fetal using Cardiotocography (CTG) data

Abstract—Maternity is the most important stage of the humans life. It has been observed that nearly 385,000 babies are born in one day according to the report of "The worlds count". Here, we are going to discuss regarding the identifying the health of a fetal using a well-known Clinical routine called as Cardiotocography. Cardiotocography also known as (CTG), is the most used clinical routine to identify the health of the fetal state and its evaluation. In this paper, we have implemented 5 different models on the CTG dataset and compared the models internally to achieve the maximum accuracy rate to identify the health of the fetal from the Cardiotocography data. We achieved the maximum accuracy rate of 0.9750, Recall rate of 0.975, and F1 Score of 0.9672

Index Terms—CTG, Fetal Health, Machine Learning, Ensemble Learning

1 Introduction

Cardiotocography is the most and commonly used clinical routine that is used to recognise the health of the fetal state. This can be deciphered from the name itself – cardio (heart) toco (uterine contractions) and graph (recording) In Cardiotocography, continuous recording of the fetal heart rate is obtained via an ultrasound transducer which is placed on the mother's abdomen. Fetal heart rate (FHR) is referring to the number of fetal heart beats per minute (BPM). In Other words, Soundwaves in the form of ultrasound are used for a variety of medical purposes such as locating solid gallstones or blood clots. The CTG helps detect anomalies that may occur during contractions. We are using the dataset which contains 2126 records of the features that are confronted from Cardiotocography exams, after which classified into 3 classes by the experts. 1.0 Normal,2.0 Suspect and 3.0 pathological. This paper is structured as follows, there are Six steps which we have expressed into this paper. Starting with importing the necessary libraries followed by reading and understanding the dataset. The dataset consists of 2126 rows and 22 columns. Data cleaning is very important process to obtain the clean and having high percentage of the missing values. In the data cleaning, all the columns are checked before dropping and then inputting the values in the columns which are important. After this step, the next step is exploring the data analysis (EDA) of the dataset. There are 3 classes as 1.0 normal, 2.0 Suspect, 3.0 Pathological. The EDA is used for the main purpose to analyse datasets to summarize their main characteristics. EDA is used for telling us what data is showing before applying modelling on the same. Feature engineering is the step where we create feature of the data since the contraction is make to much importance in fetal health. Here we are using two features "accelerations" and another feature which is similar to acceleration "light decelerations". The last step which we going to discuss in this paper is train- test split. We will use the Sklearn default option and assign 75 percent to our train set and 25 percent to our test set. The further models which we will be using to find the accuracy is discussed the below given paper . I wish you the best of success.

2 Literature review/Backgroud:

M. Ramla, S.Sangeetha and S.Nikolas has proposed the paper on Year: 2018; Fetal health Monitoring using Decision tree Classifier from CTG Measurements. Under which they have explained the efficient method to predict the High-risk pregnancy based on the fetal health status using a method called as CART. Another proposed study is made by Kanika Agrawal and Harshita Mohan on Year: 2019; Cardiotocography analysis on Fetal state. In which they have used the classification using the machine learning algorithms to identify the various Fetal health issues using Cardiotocography. The methods that are used in this study are Decision Tree Support Vector Machine (SVM) and R - Studio algorithms for Naïve Bayes (NB). In this method, the datasets are being classified into normal suspicious and pathological class. Training and testing is done using the algorithms and compared on the bases of different performance measures. In the above both papers which we have studied, both describes the fetal state using CTG measurements using 4 methods i.e., decision tree, support vector machine and R studio algorithms for Naïve Bayes. In our paper we have taken around six different models to identify the Fetal state using CTG Data. There are different studies as well which proposed the similar study as the approach to receive the fetal state, Jiaming Li; Xiaoxiang Liu in the year 2021 proposed the study on Fetal health classification in which they integrated four best models build the blender model. Let us discuss more about Methodology that we have used to rectify this classification.

3 Methodology:

Let's discuss about the methods and the models which we have used to find the best fit and accurate model that suits in our way to find out the Fetal health classification using the Cardiotocography data which was provided to us. This data was proposed by three experts into three class i.e., 1.0 Normal, 2.0 Suspicious, 3.0 Pathological. The data consists of 2126 rows and 22 columns. In the study, The UN expects that by 2030, countries end preventable deaths of the babies and children who are under 5 years of age. In which all the

countries are keeping an aim to reduce the under-5 mortality at least low as 25 per 1000 live births. Opposite to child mortality is maternal mortality which holds the deaths of 295000 during the following pregnancy and newborns (as of 2017). Simple means that what we have mentioned above, Cardiotocograms (CTGs) are a simple and cost accessible option to assess fetal health, which allow the professionals to take necessary steps in order to prevent the child and maternal mortality.

3.1 Attribute information

This dataset contains 2126 records of features extracted from Cardiotocogram exams, which were then classified by three expert obstetritians into 3 classes: 1.0 Normal 2.0 Suspect 3.0 Pathological

3.2 Various steps taken to identify the Fetal State:

3.2.1 Import necessary libraries and warnings

In this model, we have imported the necessary libraries from the derived class as Pandas, seaborn, numpy and matplotlib. Apart from this we have also used the sklear for various use for example, model selection, creating dummy data, finding KNN, logistic regression, SMOTE and decision tree classifiers. Also, we have used the warnings to handle the warnings and filtering the warnings which is not mandatory.

3.2.2 Reading and Understanding the dataset.

In this step, we are reading the data and understanding the dataset which we have taken. For reading the data, we have used pandas. The data we have is in .CSV format hence after giving the proper shape, we have found that it consists of 22 columns and 2126 Rows. We have all columns data in "float64" format ie. all continues variables. So we don't have to worry about data type conversion.

3.2.3 Data Cleaning.

In data cleaning first up all, drop columns that are having high percentage of missing values. Check all the columns before dropping them and then impute values in columns which are important. Here as mentioned above we have 3 classes which is proposed i.e., 1.0 Normal, 2.0 Suspicious and 3.0 Pathological.

3.2.4 Exploratory Data Analysis (EDA).

This step gives us the exploratory analysis of the data so that we can take in depth look at dataset. We will look at the target variable and the relationship of other features with each other. Below picture shows the distribution of the data from the given dataset.

In this plot we have shown the distribution of samples that where classified as either (1.0) Normal, (2.0) Suspect, (3.0) Pathological.

Also there is a significant difference in the sample sizes. we need to consider class imbalance in our modeling parameters. After checking, we will consider the histograms for all the features.

Also, we are checking the feature correlation and plotting the heatmap so that we can get idea for correct numbers of data correlation.

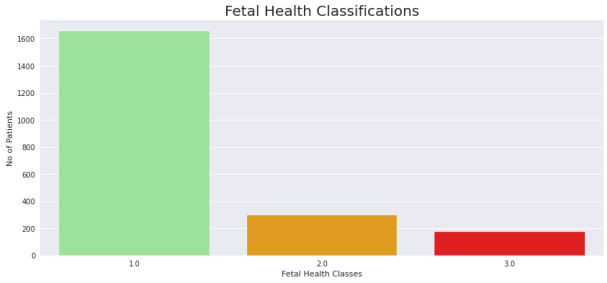

Fig. 1. Fetal Health Classes

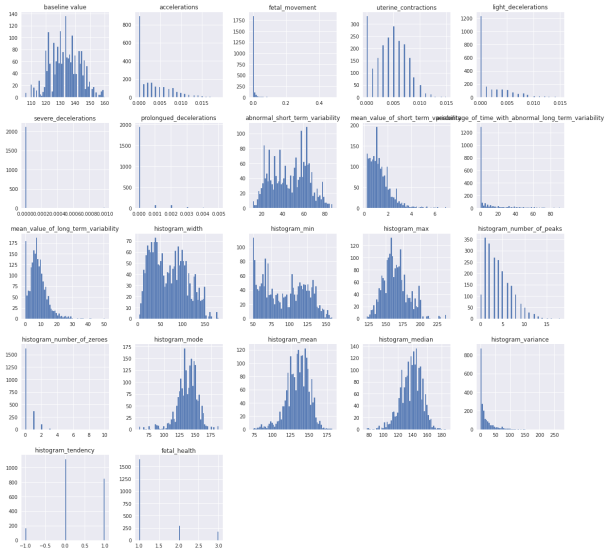

Fig. 2. Histogram of all features

After finding the histograms heatmap(Fig 3) of the given data set, now we check the plot a scatter plot(Fig 4) between the target variable and predictor variables.

Now we have to take a look at severe decelerations(Deceleration's are temporary decreases in the fetal heart rate (FHR) during labor). We knew that any major changes of FHR when visually monitored results in immediate surgical intervention. The severe deceleration feature was almost synonymous with the suspect,pathological class, however of the 2126 instances in the dataset, there are only 7 instances of severe deceleration. We suspect this is a key feature in classification, but there may not be enough data here to train the model due to class imbalance. we have to preform another analysis creating multiple classes of fetal distress, we hypothesis these extreme changes in fetal heart rate will play a role in classifying pathological fetal health.

3.2.5 Feature Engineering.

Since the contraction is make to much importance in fetal health. we need to create a feature for it.Now lets see the "accelerations" feature. Accelerations are short-term increases in fetal heart rate by at least 15 beats per minute (bpm) that last at least 15 seconds. These accelerations occur at different times throughout labor and delivery and are a sign that the fetus has an adequate supply of oxygen. We need to analyse this feature also, so We have another feature similar to acceleration which is "light decelerations" (Fig 6)

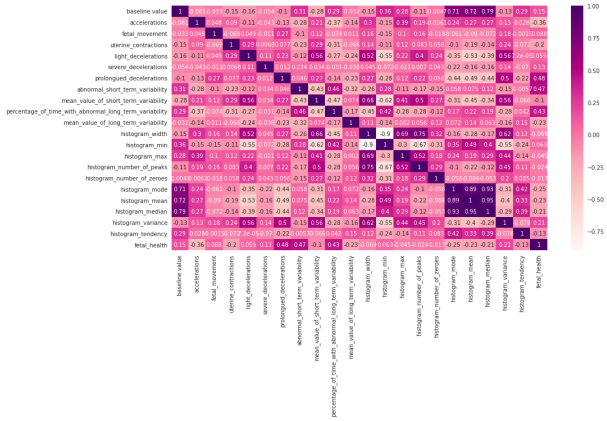

Fig. 3. Heatmap

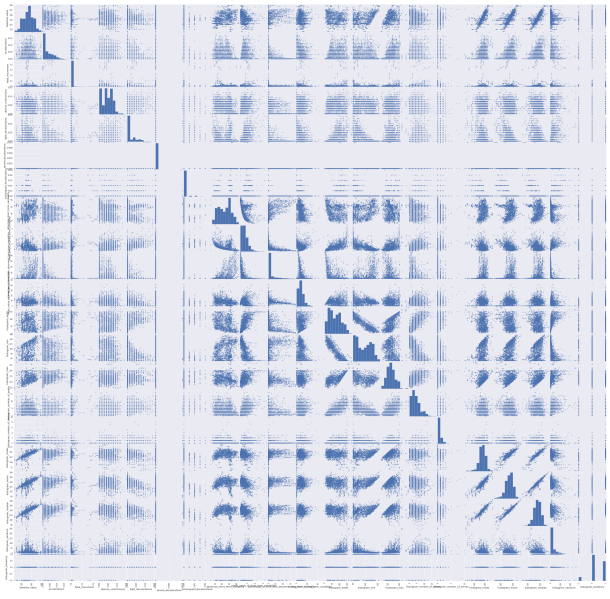

Fig. 4. Plotting Scatter matrix

which plays important role for defining the fetal health. Lets analyze this feature (Fig 6)

In above Figure (Fig 6) snippet we binned several continuous variable to create discrete variables. Our goal is to create categories within the continuous features to observe if these categories where class specific. Lets create another feature which combines the values of other important variable. Many on these features has zero values.

We are also creating the Dummy variables for some of the features from the dataset. Since our target variable has 3 different classes but the ”suspect” class doesn't provide any particular data. So we will convert the target variable into binary class. Change any value of the target value above one to two, creating two target classes instead of three classes.

3.2.6 *Train Test Split.*

In this step, we will be discussing about the different models that we have used to classify the fetal state using given Cardiotocography (CTG) Data.

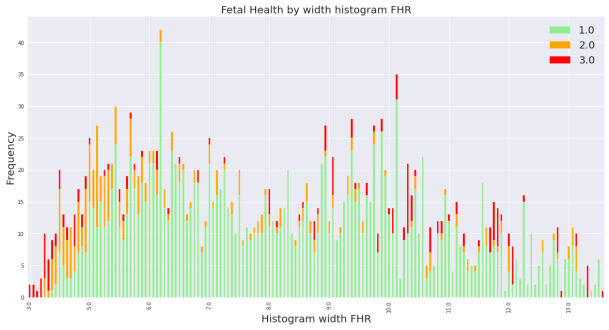

Fig. 5. Histogram Width FHR

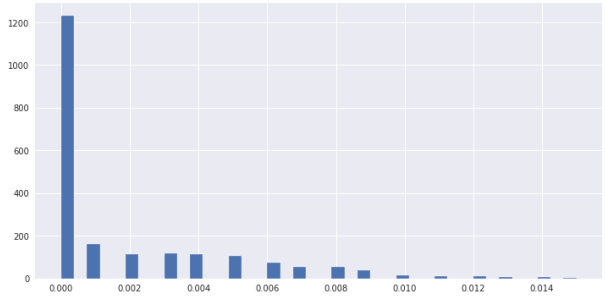

Fig. 6. Data Light Decelerations

3.2.7 *Handle The Class Imbalance.*

To handle class imbalance we over sampled the minority class using SMOTE(Synthetic Minority Oversampling Technique), (Fig 7) this balanced the minority class by sampling the nearest neighbours and adding points between the neighbors.We used both a SMOTE sampled database and unbalanced database to compare the effect of the metrics for each experimental baseline model. The Smote Database preformed better on every baseline model, we chose to use the SMOTE database for our final models.

```
# Evaluation function.

def evaluation(y_true, y_pred):
    print('Evaluation Metrics:')
    print('Accuracy: ' + str(metrics.accuracy_score(y_test, y_pred)))
    print('Recall: ' + str(metrics.recall_score(y_test, y_pred)))
    print('F1 Score: ' + str(metrics.f1_score(y_test, y_pred)))
    print('Precision: ' + str(metrics.precision_score(y_test, y_pred)))
    print('\nConfusion Matrix:')
    print(' TN,  FP, FN, TP')
    print(confusion_matrix(y_true, y_pred).ravel())
```

Fig. 7. SMOTE

3.2.8 *Logistic Regression.*

Logistic regression is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set. A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables. Fit train set with Logistic Regression model is shown in the (Fig 8).

3.2.9 *K-Nearest Neighbor(KNN).*

Imbalanced class sizes are both a theoretical and practical problem with KNN which has been characterized in ma-

```
evaluation(y_test, y_predsm)

Evaluation Metrics:
Accuracy: 0.8778195488721805
Recall: 0.8801955990220048
F1 Score: 0.9171974522292994
Precision: 0.9574468085106383

Confusion Matrix:
 TN,  FP, FN, TP
[360  49  16 107]
```

Fig. 8. Logistic Regression Result

chine learning literature since at least 2003. This is particularly vexing when some classes have a low occurrence in your primary dataset (ex: fraud detection, disease screening, spam filtering). Prediction metrics and confusion matrix of smote base KNN

```
Evaluation Metrics:
Accuracy: 0.9191729323308271
Recall: 0.9559902200488998
F1 Score: 0.947878787878788
Precision: 0.9399038461538461

Confusion Matrix:
 TN,  FP, FN, TP
[391  18  25  98]
```

Fig. 9. KNN With Class Imbalance

3.2.10 *Decision Tree.*

A decision tree is a type of supervised machine learning used to categorize or make predictions based on how a previous set of questions were answered. Train decision tree with train set and predict on the test set. Now lets see a table which will compare the Decision Tree coefficients and comparing the coefficients of the SMOTE and imbalanced dataset. (Fig 10)

3.2.11 *Random Forest.*

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression. We will run grid search on the random forest to identify what are the best hyper parameters works for the model.The best parameters for the random forest was a max depth of 11, min weight fraction leave of 0, and 500 estimators.The Evaluation metrics where our highest score with the experimental set so far. Below we can see the evaluation metrics find out from the random forest algorithm. (Fig 11)

4 Results

As per our paper and findings, we have implemented 5 different models on the CTG dataset and compared the

	features	importance	sm_importance
0	baseline value	0.050888	0.050888
1	accelerations	0.033863	0.033863
2	fetal_movement	0.004742	0.004742
3	uterine_contractions	0.000000	0.000000
4	light_decelerations	0.000000	0.000000
5	prolongued_decelerations	0.028585	0.028585
6	abnormal_short_term_variability	0.341082	0.341082
7	mean_value_of_short_term_variability	0.115917	0.115917
8	percentage_of_time_with_abnormal_long_term_var...	0.044652	0.044652
9	mean_value_of_long_term_variability	0.018986	0.018986
10	histogram_width	0.015668	0.015668
11	histogram_min	0.038516	0.038516
12	histogram_max	0.037546	0.037546
13	histogram_number_of_peaks	0.016062	0.016062
14	histogram_mode	0.020438	0.020438
15	histogram_mean	0.173347	0.173347
16	histogram_median	0.006727	0.006727
17	histogram_variance	0.006461	0.006461
18	uterine_contraction_per_min	0.027980	0.027980
19	root_total_change	0.014927	0.014927
20	hist_zeros_1.0	0.000000	0.000000
21	hist_zeros_2.0	0.000000	0.000000
22	hist_zeros_3.0	0.000000	0.000000
23	hist_zeros_4.0	0.003614	0.003614

Fig. 10. Decision Tree

```
Evaluation Metrics:
Accuracy: 0.9492481203007519
Recall: 0.9755501222493888
F1 Score: 0.9672727272727273
Precision: 0.9591346153846154

Confusion Matrix:
 TN,  FP, FN, TP
[399  10  17 106]
```

Fig. 11. Random Forest

models internally to achieve the maximum accuracy rate to identify the health of the fetal from the Cardiotocography data. The best model we have found is Random forest as we have got the accuracy rate 0.9750, Recall rate of 0.975, and F1 Score of 0.9672. We have also managed to get the highest probability and also the top Ten Important features through the random forest search algorithm which was applied to the given dataset.(Fig 12)

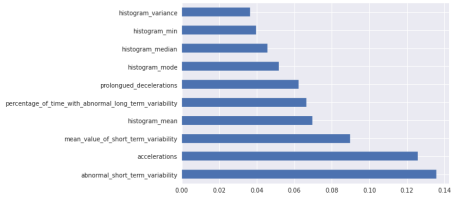

Fig. 12. Top Ten Important Features of Random Forest

5 Conclusions:

In our finding way of the classification for identifying the Fetal State using CTG, so far we have found as Random search is the best algorithm to find out the accurate outliers, train/validation/test splits, reprocessing, classifiers/regressors/modelling tools, evaluation methods and

metrics. We have also performed different models and algorithms such as decision tree, KNN, SMOTE, Logistic Regression but random forest results were more accurate from the rest. We can also have a look into the (Fig 12) it has also given us the Top ten important features of the CTG Data which was given to us. Moreover, from this paper we have found that Identifying the fetal state before a women giving a birth to the new born babies is very useful and helpful method because due to it we can identify how is the current state of the baby is, and if we found any issues the professional health workers can identify the particular issues and give the medication and prevention steps accordingly.

REFERENCES

[1] [1] D. Ayres-de Campos, J. Bernardes, A. Garrido, J. Marques-de Sa, and L. Pereira-Leite, "Sisporto 2.0: a program for automated analysis of cardiotocograms," Journal of Maternal-Fetal Medicine, vol. 9, no. 5, pp. 311–318, 2000

[2] [2] A. Akbulut, E. Ertugrul, and V. Topcu, "Fetal health status prediction based on maternal clinical history using machine learning techniques," Computer methods and programs in biomedicine, vol. 163, pp. 87–100, 2018.

[3] [3] 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," International Journal of Applied and Basic Medical Research, vol. 9, no. 4, p. 226, 2019.

[4] [4] M. Ramla;S. Sangeetha;S. Nickolas "Fetal Health State Monitoring Using Decision Tree Classifier from Cardiotocography Measurements", 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS) Year: 2018 Conference Paper — Publisher: IEEE

[5] [5] Kanika Agrawal;Harshit Mohan " Cardiotocography Analysis for Fetal State Classification Using Machine Learning Algorithms " 2019 International Conference on Computer Communication and Informatics (ICCCI) Year: 2019 — Conference Paper — Publisher: IEEE

[6] [6] Jiaming Li;Xiaoxiang Liu " Fetal Health Classification Based on Machine Learning" 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) Year: 2021 — Conference Paper — Publisher: IEEE

[7] [7] Kevin Luk " My journey into Data Science " URL https://towardsdatascience.com/downloading-datasets-into-google-drive-via-google-colab-bcb1b30b0166 Year :- Feb 15, 2019.

[8] [8] URL " https://www.ebme.co.uk/articles/clinical-engineering/cardiotocography-ctg" - EBME Clinical Engineering Articles Cardiotocography (CTG) Compiled and edited by John Sandham