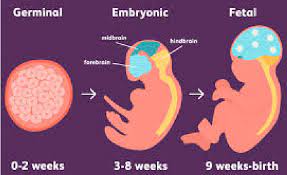
**“Fetal Health Status Prediction Using Machine Learning Techniques”**



**Contents**

[1 Introduction](#_Toc59292788)

[2 Related Work](#_Toc59292789)

[3 Methods](#_Toc59292790)

[3.1 Importing Libraries, Packages, and Dataset:](#_Toc59292791)

[3.2 Exploratory Analysis of the Data:](#_Toc59292792)

3.3 Checking for Outliers……………………………………………………………………………………………………………….

[3.4 Data Preprocessing:](#_Toc59292794)

[3.5 Machine Learning Models:](#_Toc59292795)

3.6 Evaluating Metrics………………………………………………………………………………………………………………….

4 Results…………………………………………………………………………………………………………………………………………

[4.1 Random Forest Model](#_Toc59292799)

[4.2 Naïve Bayes](#_Toc59292800)

[4.4 SVM (Support Vector Machine)](#_Toc59292801)

[4.5 KNN (K-Nearest Neighbour)](#_Toc59292802)

[4.6 Comparing accuracies of all the models](#_Toc59292803)

[4.7 Comparing the F1, precision, accuracy and recall scores of the best two performing models](#_Toc59292804)

[4.8 Comparing the auc scores of all the models](#_Toc59292805)

[5 Discussion](#_Toc59292806)

[5.1 Random Forest Model](#_Toc59292807)

[5.2 Naïve Bayes](#_Toc59292808)

[5.3 SVM (Support Vector Machine)](#_Toc59292810)

[5.4 KNN (K-Nearest Neighbour)](#_Toc59292811)

[6 Conclusion](#_Toc59292812)

[7 Contributions](#_Toc59292813)

[8 References](#_Toc59292814)

[9 Appendices](#_Toc59292815)

**INTRODUCTION:**

Health complications during the gestation period have evolved as a global issue. These complications sometimes result in the mortality of the fetus, which is more prevalent in developing and underdeveloped countries. The genesis of Machine Learning (ML) algorithms in the healthcare domain have brought remarkable progress in disease diagnosis, treatment, and prognosis. This research deploys various ML algorithms to predict fetal health from the Cardiotocographic (CTG) data by labelling the health state into normal, needs guarantee, and pathology. This work assesses the influence of various factors measured through CTG to predict the health state of the fetus through algorithms like Support Vector Machine (SVM), Random Forest (RF), Multi-Layer Perceptron (MLP), and K-Nearest Neighbors (K-NN). In addition to this, the Regression Analysis(RA) and Correlation Analysis (CA) revealed the influence of the attributes on fetal health. The results of the algorithms show that RFperforms better than its peers in terms of accuracy, precision, recall, F1-score, and support. This work can further enhance more promising results by performing suitable feature engineering in the CTG data.

Healthcare analytics is the process of evaluating current and historical industry data in order to forecast trends, enhance outreach, and even better manage disease transmission. The topic covers a broad variety of sectors and offers both macro and micro views. It has the potential to lead to improvements in patient care, clinical data, diagnostics, and business management.

As medical technology advances, the rate of preventable child mortality decreases. Lowering child mortality rates has become a key goal in advancing any society and a key in human progress as a whole. While there are many advancements that have improved the mortality rate, not all of these practices are globally available. In order to improve mortality rates in cost effective and readily available solutions need to be applied and perfected.

One such cost effective and relatively simple method would be the use of Cardiotocograms (CTG). The CTG is a non-invasive fetal monitor which is used to assess fetal health. The CTG is used to detect fetal heart rate (FHR), uterine contractions, fetal movement, and sudden changes in heart rate. Currently, doctors rely on a visual analysis of the CTG leading to erroneous interpretations of the exam. Minute changes which may be extremely detrimental to fetal health may also not be visible to the eye. Conversely some observable changes in fetal heart rate may appear to be fetal distress but are just a response to other factors like uterine contractions.

In order to improve diagnosis of fetal distress with a CTG and improve the mortality rate, the use of machine learning algorithms becomes a viable option. A well-trained model will be able to identify which variable changes to the FHR have the greatest effect on fetal health. To lower the overall risk of child mortality a ML model must not only be able to accurately classify between normal and distressed fetal health but be precise enough in its predictions to prevent any unnecessary interventive surgical procedures.

For our analysis of the we used a CTG exam dataset found on Kaggle from The Journal of Maternal-Fetal Medicine. The dataset consisted of 2126 CTG exam records which were classified into three classes (normal, suspect, pathological). In this initial analysis we decided to create a binary class of normal and distressed, combining the suspect and pathological classes into one class.

**RELATED WORK:**

The main purpose of this section is to showcase and compare the results achieved by another group. The field of work should be similar, as this is to verify that our current results of this project is in line with theirs or similar to their results of work.

A recent rush in the deployment of four ML techniques, namely NN, k-NN Classifier,SVM, and Decision Tree, which are evaluated on high dimensional data, proves that the classifier SVM dominates all other techniques in giving accurate diagnostic indices. All the Classifiers work well and use 30 ranked features to determine common risk factors in the prediction model. While using ML algorithms, feature extraction and selection are among the most used methods to select optimized features for prediction in the model. This would even help in giving top priority in feature selection. The algorithm's performance can be validated by different metrics, namely Accuracy, Specificity, Precision, and Recall. The overview of using ML for assessing fetal changes with optimizing the image acquired and the effectiveness of classifying cardiac abnormalities. Using CTG signals and their principles is described with evaluating historical data, and the process is used. The image data set has been taken with the implementation of CNN architecture in Deep Learning (DL) which motives to validate the data from CTG signals with that of the image acquisition dataset.

**Research paper 1:**

## Fetal health status prediction based on maternal clinical history using machine learning techniques:

* In this paper they have discussed about the Ultrasonographic examinations of foetuses also aid in the detection and definition of these anomalies.
* Here the system was broadly categorized algorithms were trained, tested and compared for predicting fetal health status.

They have used 5 classification models:

1. K-Nearest Neighbor (KNN)
2. Support Vector machine (SVM)
3. Decision Tree.
4. Random Forest.

|  |  |  |
| --- | --- | --- |
|  | Classifier model | Accuracy |
| 1. | KNN | 85.6 |
| 2. | SVM | 79.4 |
| 3. | DECISION TREEE | 89.5 |
| 4. | RANDOM FOREST | 84.2 |
|  |  |  |

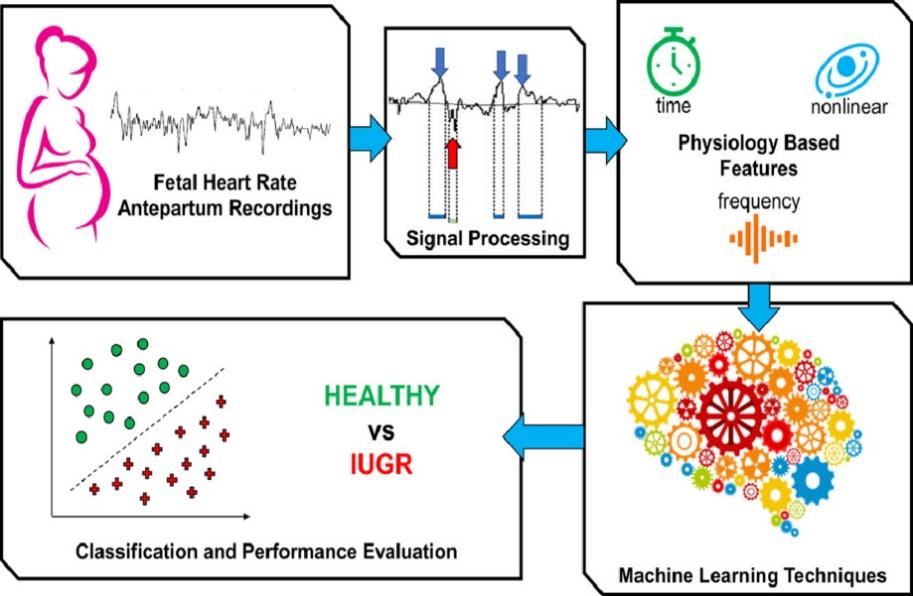
The Decision Forest model has the highest accuracy, F1-Score, and AUC, with 89.5 percent accuracy, 75 percent F1-Score, and 95 percent AUC.

**References:** <https://www.sciencedirect.com/science/article/abs/pii/S0169260718302384>

**Research paper 2:**

Integrating machine learning techniques and physiology-based heart rate features for antepartum fetal monitoring

* In this paper they have discussed about intrauterine growth restriction (IUGR) is a fetal condition defined as the abnormal rate of fetal growth.
* Here they tested the performance of five techniques in discriminating healthy versus IUGR foetuses. The various models were trained with a set of 12 physiology-based heart rate features extracted from a single antepartum Cardiotocographic (CTG) recording. The reason for the utilization of time, frequency, and nonlinear indices is based on their standalone documented ability to describe several physiological and pathological fetal conditions.
* We validated our approach on a database of 60 healthy and 60 IUGR foetuses. The machine learning methodology.
* Here they got the best performance for Random Forest with highest accuracy of 91.1%

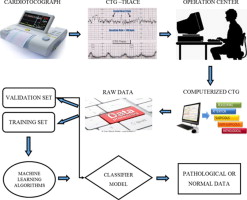


**References:** <https://www.sciencedirect.com/science/article/abs/pii/S0169260719308107>

**Research Paper 3:**

# Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques

* The main aim of this research is evaluating the classification performances of eight different machine learning methods on the antepartum cardiotocography (CTG) data.
* The aim of the machine learning methods is by using attributes of data obtained from the uterine contraction (UC) and fetal heart rate (FHR) signals to classify as pathological or normal. The dataset contains 1831 instances with 21 attributes, examined by applying the methods.
* In the paper, the highest accuracy displayed as 92.3%



**References**:[https://www.sciencedirect.com/science/article/abs/pii/S1568494615002653#fig0030](https://www.sciencedirect.com/science/article/abs/pii/S1568494615002653%23fig0030)

**METHODS:**

In this section we provide a brief description of process we went through in order to perform the tasks like EDA, pre-processing, detection outliers, developing model etc.

Installing packages, importing libraries, and loading data set:

we have installed packages and imported libraries which we thought are required in order to run all chunk of codes in the python notebook. After importing all the libraries, we have loaded our data set and stored in in a data frame.

* **EDA and Pre-processing:**
* EDA was performed in order to get a better understanding on our data set, as we went through few simple steps like performing head, tail functions, data info in order to check for the data types of all the variables weather they are in numeric format or not and going through some other features like null value count, man value, min value, standard deviation and quartiles.

Graphical user interface, application, table, Excel

Description automatically generated

Using the null value function, we checked for the null values and came to know that there are no null values in our data set.

Text

Description automatically generated

After we came to know that there are no null values, we plotted correlation heat map for the dataset in order to check the correlations of every other variable

A picture containing calendar

Description automatically generated

Chart, box and whisker chart

Description automatically generated

In the above plot fetal health 2.0 has poor fetal heart rate which indicate babies health is not good

1.0 : Normal

2.0: Suspect

3.0: Pathological

Chart, box and whisker chart

Description automatically generated

There seems to more commonly be movements among records classified as pathological(3.0).

Chart, box and whisker chart

Description automatically generated

There doesn't seem to be a lot of variability among uterine contractions. There are more higher values for normal records.

## **Checking for Outliers:**

Outliers are data points which is located far from the other data points or observations. Identification of Outliers can help us to find out the bad data points associated with our feature columns. The data might be placed incorrectly, but also due to a variation by random can help us to identify something interesting. Based on this, we checked for outliers in our columns to understand the distribution of data points

* we wanted to look for correlation of every variable to the target variable, so we used the specific code.

corr\_matrix**=**df**.**corr()

corr\_matrix['fetal\_health']**.**sort\_values(ascending**=False**)

Graphical user interface, text, application

Description automatically generated

Here we got an idea of which variables are positively correlated and which are negatively correlated, and we dropped few columns which has high negatively correlation in response to our target variable (fetal health)

## **Machine Learning Models:**

Machine Learning is the study of algorithms which can be improved as time goes by. It is a field of Artificial Intelligence which is used for building a model that can process sample data, Using the train test split function initially we divided the data into 2 parts train (20%) and test (80%), We have used 4 machine learning models in this project they are:

1. KNN
2. RANDOM FOREST
3. NAVIE BAYES
4. SVM

These are the different models we have used for our dataset. The models which we have used for this dataset could be different from the previous work we have considered. In order to achieve the best accuracies.

## **Evaluation Metrics:**

An evaluation metric quantifies the performance of a predictive model. This typically involves training a model on a dataset, using the model to make predictions on a dataset not used during training, then comparing the predictions to the expected values in the dataset.

We have used 4 different metrics to verify which model performs best and to be used to get the best results, they are:

* Accuracy
* Precision
* F1 score
* Recall

After all the parts are done, we have printed the accuracies of each model comparing each model which has the highest accuracy. All the results can be seen in the results section.

**RESULTS:**

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy** | **F1 Score** |
| Random forest | 94.3 | 97 |
| KNN | 89.4 | 81.1 |
| NAVIE BAYES | 79.1 | 69.2 |
| SVM | 84.7 | 67.4 |

Out of the 4 models we used we got highest accuracy in random forest model and least in naive bayes. When comparing all models with all metrics we see that

A screenshot of a computer

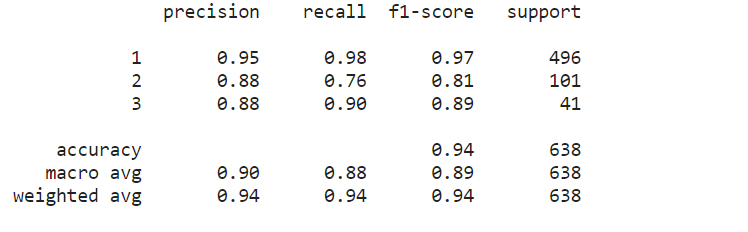
Description automatically generated

Form this chart we can say that random forest has the best score in all the metrics when compared to other models that we have used till now.

**DISCUSSION:**

Following are the precision and recall scores from classification report and its interpretation of each of the machine learning models:

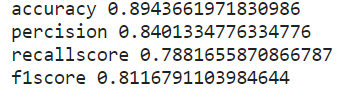
**Random forest model:**



The accuracies yielded by **random forest model** are of 94.3% and to evaluate the performance of the model we created confusion matrix. the model’s precision is high 95%

The recall scores us 98% which means that the model has correctly. Similarly, the model has f1 score as 97% and precision as 95%.

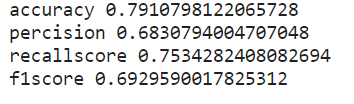
**KNN (k-nearest neighbors):**



The accuracies yielded by **KNN** are of 89.4% and to evaluate the performance of the model we created confusion matrix. the model’s precision is high 84%

The recall scores us 78.8% which means that the model has correctly. Similarly, the model has f1 score as 81.1%.

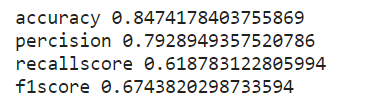
**Naïve bayes:**



The accuracies yielded by **Naïve bayes** are of 79.1% and to evaluate the performance of the model we created confusion matrix. the model’s recall is high 75.3%

The precision score us 68.3% which means that the model has correctly. Similarly, the model has f1 score as 69.2%.

**SVM (Support Vector Machine):**



The accuracies yielded by **SVM are** of 84.7% and to evaluate the performance of the model we created confusion matrix. the model’s precision is high 79.2%

The recall score us 61.8% which means that the model has correctly. Similarly, the model has f1 score as 67.4%.

**CONCLUSION:**

Cardiotocographs are a low-cost means of monitoring fetal health and are being utilised to reduce child death rates. 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. We used numerous modelling approaches to predict fetal health class in this work, and we tried to achieve the best precision in our prediction to limit the number of false positive classifications.

From the models like Random Forest, KNN, Naïve bayes and SVM. Random forest has performed the best.

The weighted average of Precision and Recall is used to calculate the F1 score, where Precision is the ratio of correctly predicted positive observations of all positive predicted positive observations and Recall is the ratio of correctly predicted positive observations of all true class observations. F1 score is more valuable than accuracy score, especially when the class distribution is uneven, and thus it is best if Precision and Recall must be balanced.

**Finally, we have concluded that Random Forest model performs the best which has the highest accuracy of 94.3% and even has the high precision score which is 96%.**

**REFERENCES:**

**Dataset:** <https://www.kaggle.com/andrewmvd/fetal-health-classification(kaggle)>

**Fetal health:**

* <https://www.sciencedirect.com/science/article/abs/pii/S0169260718302384>
* <https://medium.com/ai-techsystems/fetal-health-classification-on-cainvas-1129886daa19>
* <https://ieeexplore.ieee.org/document/9389902>

**Machine learning:**

* <https://r4ds.had.co.nz/exploratory-data-analysis.html>
* <https://www.javatpoint.com/classification-algorithm-in-machine-learning>
* <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>
* <https://github.com/IanTirok/KNN-and-Naive-Bayes-practice>
* <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>