

## FETAL HEALTH REPORT

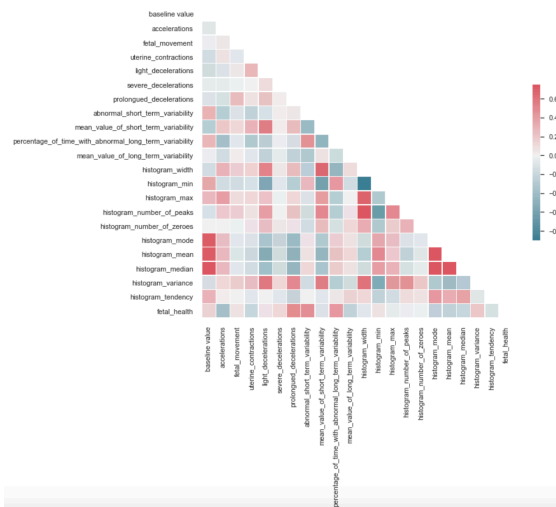
**Executive:** A Voting Classifier model, that was Grid Searched and incorporates Synthetic Minority Oversampling Technique (SMOTE), provided the best overall metrics regarding classifying the three classes; Normal, Suspect, and Pathological. The overall metrics are as follows: Accuracy: 92%, Precision: 93%, Recall: 92%, and F1: 92%.

**Introduction:** One of the important data insight extracting methods in biomedical research are classification algorithms. Classification algorithms can be divided into two categories: binary and multi-class classifiers. The interest of this project is focused on a multi-class problem identifying and classifying Normal, Suspect, and Pathological fetuses.

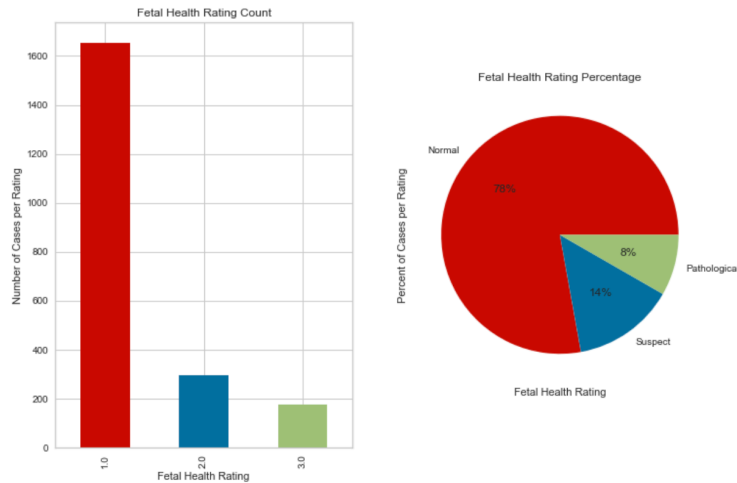
**Background:** Reduction of child mortality is reflected in several of the United Nations' Sustainable Development Goals and is a key indicator of human progress. The UN expects that by 2030, countries end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce under-5 mortality to at least as low as 25 per 1,000 live births. Parallel to notion of child mortality is maternal mortality, which accounts for 295,000 deaths during and following pregnancy and childbirth (as of 2017). The vast majority of these deaths (94%) occurred in low-resource settings, and most could have been prevented. In light of what was mentioned above, Cardiotocograms (CTGs) are a simple and cost accessible option to assess fetal health, allowing healthcare professionals to take action in order to prevent child and maternal mortality. The equipment itself works by sending ultrasound pulses and reading its response, thus shedding light on fetal heart rate (FHR), fetal movements, uterine contractions and more.

**Purposes:** The purpose of this analysis is to satisfy Coursera's/IBM Supervised Learning: Classification course final project requirements.

**Attributes:** This dataset contains 2,216 records/observations of patient Cardiotocogram exams, which were then classified by three expert obstetricians into one of the following three classes: 1) Normal, 2) Suspect, and 3) Pathological. Moreover, the data was tidy and did not necessitate cleaning. Additionally, there was a significant class imbalance, but that was remedied by utilizing the Synthetic Minority Oversampling Technique (SMOTE) on the training set. Furthermore, the dataset contained 22 feature columns including; baseline values, fetal movements, histogram features, etc.



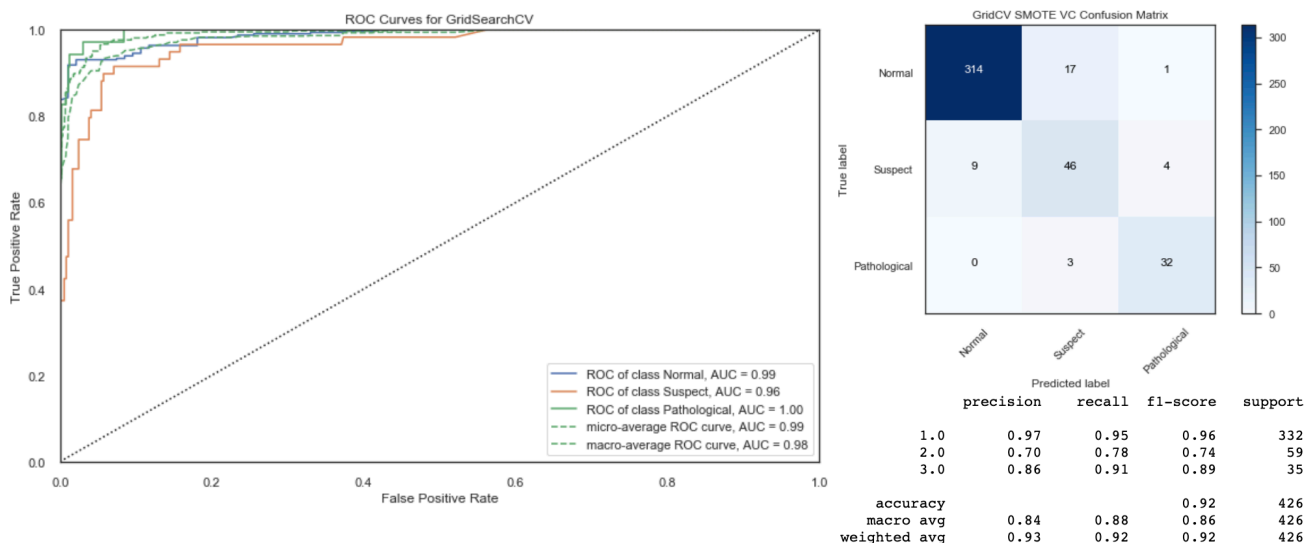
**Limitations:** A glaring limitation with this dataset is the small amount of observations, 2,126, available. To give context the UN estimates that roughly 385,000 children are born daily, thus the observations in this dataset account for 0.55% of possible observations on a daily basis. Therefore, with more data it is possible that new trends could be uncovered or that particular features become more/less important in reference to what was observed during my analysis.



**Methods:** The following models were fit to classify the data into the three aforementioned classes: Logistic Regression, K-Nearest Neighbors, Decision Tree, Random Forest, XGBoost, ADABOost, and a Voting Classifier.

**Results:** Important to note that while the Random Forest model, that was Grid Searched and incorporated Synthetic Minority Oversampling Technique (SMOTE), provided the best raw metrics for Accuracy, Recall, and F1 Score the model still resulted in classifying 2 Pathological fetuses as Normal. Given that this is healthcare related data a False Negative, or type 2 error, is egregious and not acceptable when performing diagnostic testing of this nature. Thus, the following visualizations will provide a ROC-AUC graph, Confusion Matrix, and Accuracy/Precision/Recall/F1 scores for the best overall model. Moreover, this information for the rest of the models can be found on my [GitHub](#) repository.

1) Voting Classifier:



**Recommendations:** For future analysis it would be interesting to compare Principle Component Analysis and Linear Discriminant Analysis dimension reduction techniques in attempt to further separate the data. Then train the models again on dimensionally reduced data for comparison. Furthermore, as previously mentioned obtaining more data could result in uncovering new trends and or resulting in particular features become more/less important.

**References:** Ayres de Campos et al. (2000) SisPorto 2.0 A Program for Automated Analysis of Cardiotocograms. J Matern Fetal Med 5:311-318 ([https://onlinelibrary.wiley.com/doi/10.1002/1520-6661\(200009/10\)9:5%3C311::AID-MFM12%3E3.0.CO;2-9](https://onlinelibrary.wiley.com/doi/10.1002/1520-6661(200009/10)9:5%3C311::AID-MFM12%3E3.0.CO;2-9))