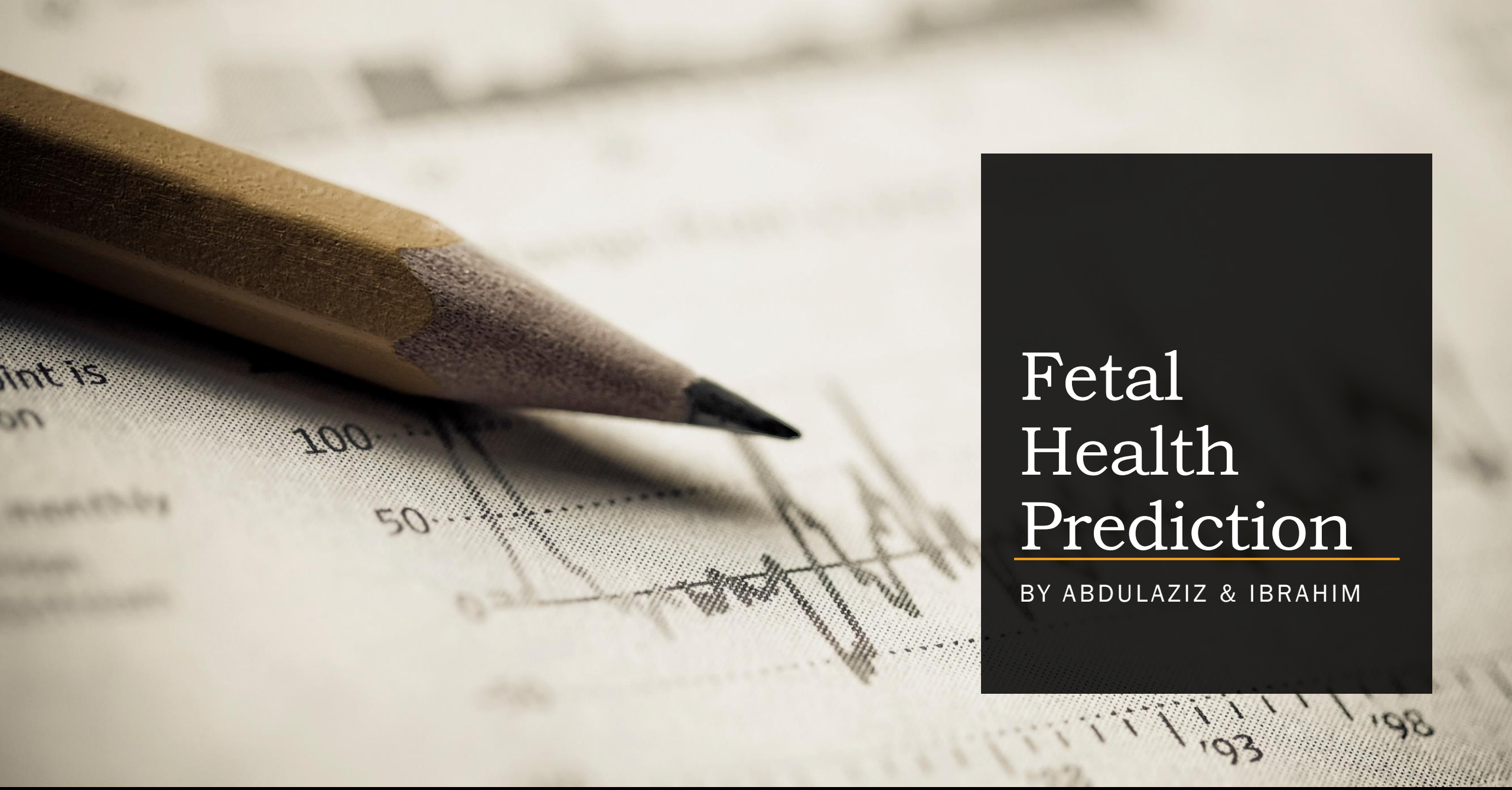


Fetal Health Prediction

BY ABDULAZIZ & IBRAHIM

Problem - Importance

- 211 death per 100,000 live birth (2017).
- Most could have been **PREVENTED.**

Problem - Goal

Classify **fetal health** in order to prevent child and maternal mortality.

Normal

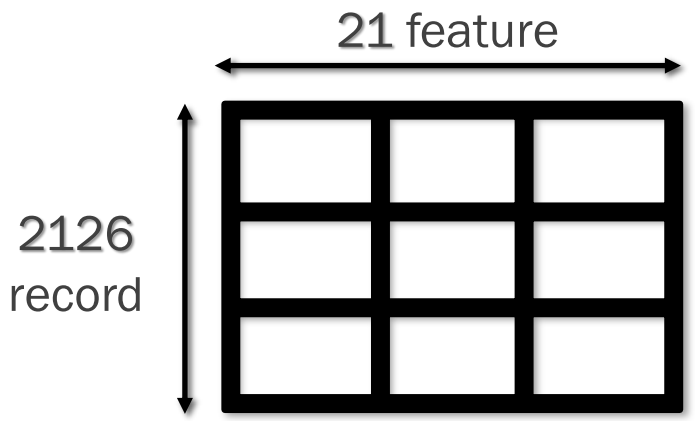
Suspect

Pathological

Data

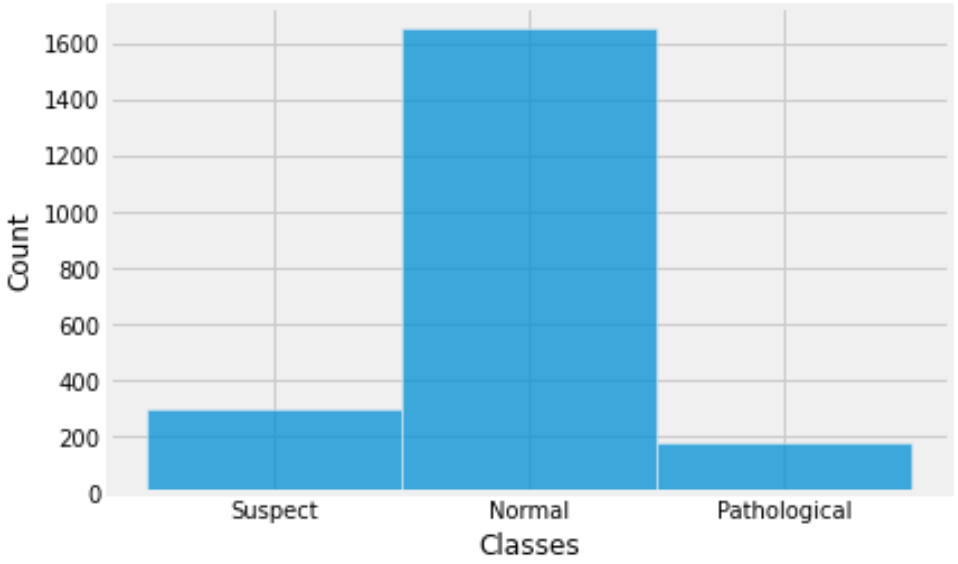
- Collected from Cardiotocograms (CTGs).
- Labeled by 3 experts.

Data

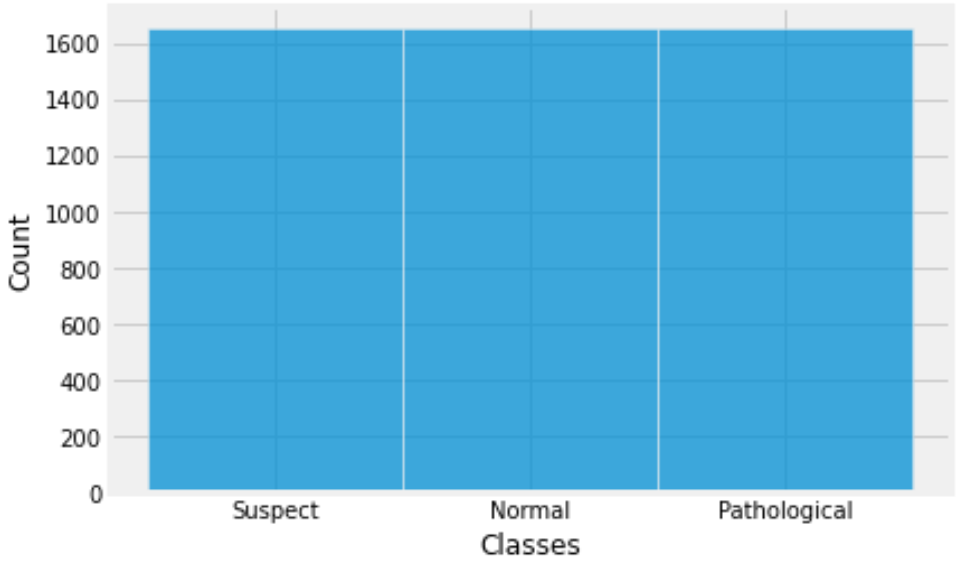


EDA – Balance Data (SMOTE)

ORIGINAL Y



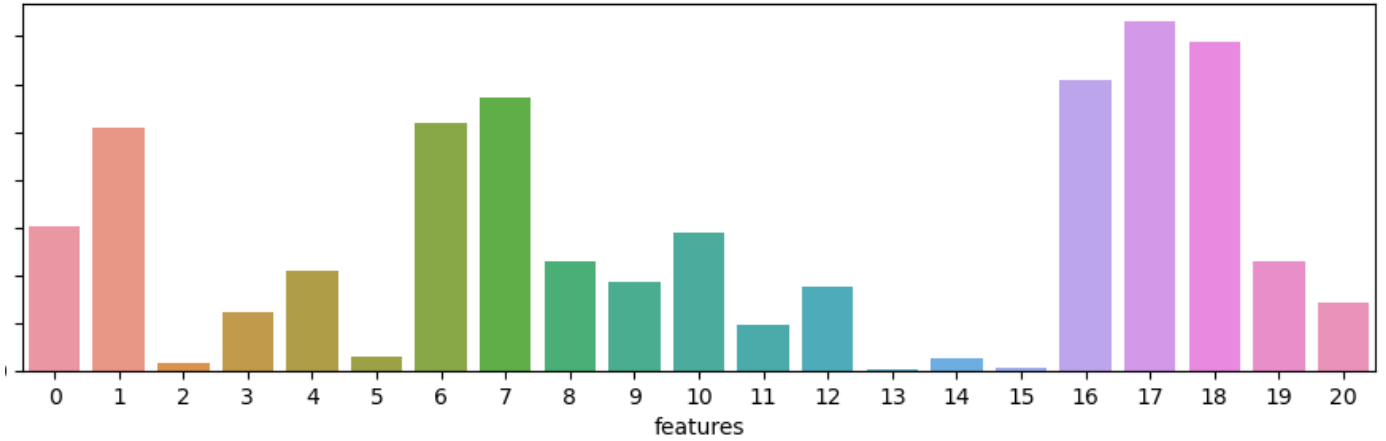
BALANCED Y



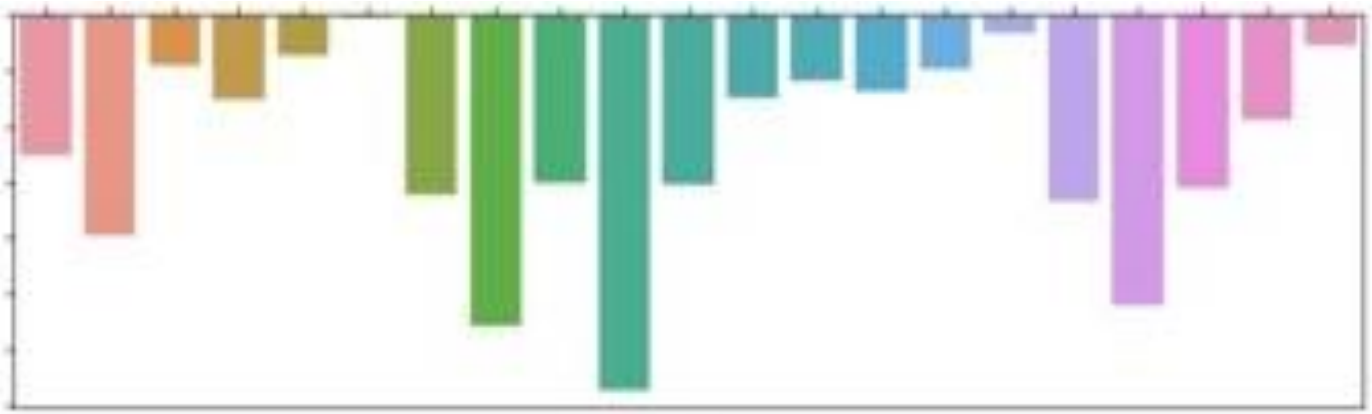
EDA – Cleaning ?

Duplicated	24 rows
Missing Values	None
Outliers	None

Feature Selection

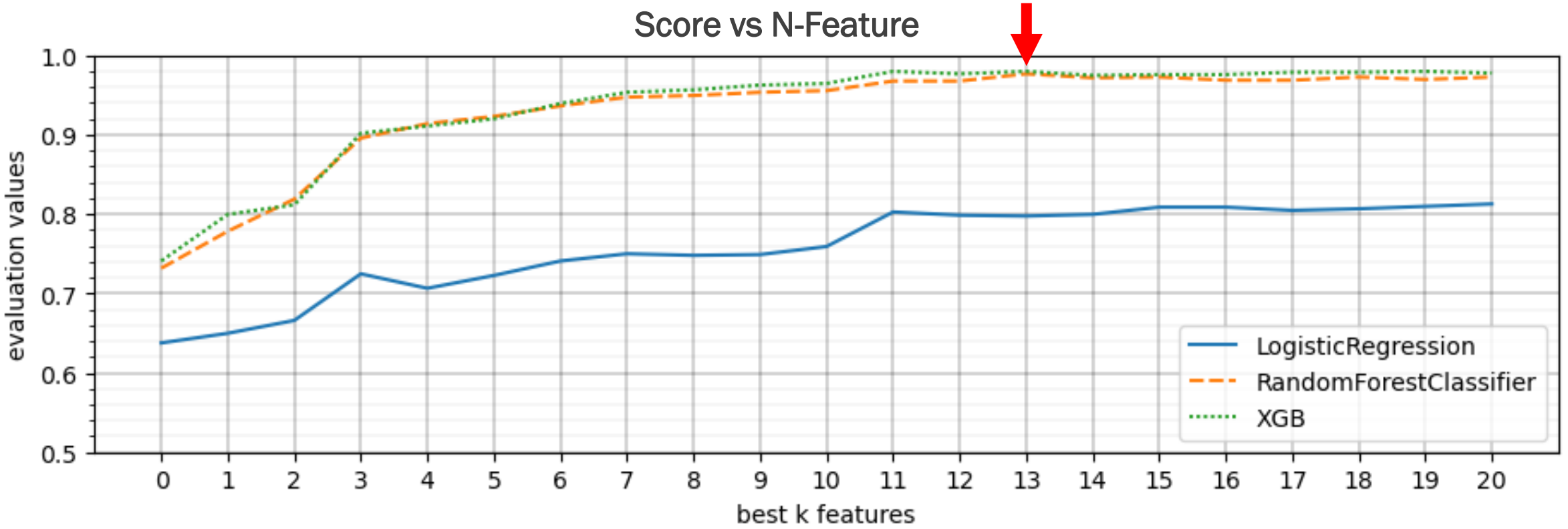


Method 1 **ANOVA**



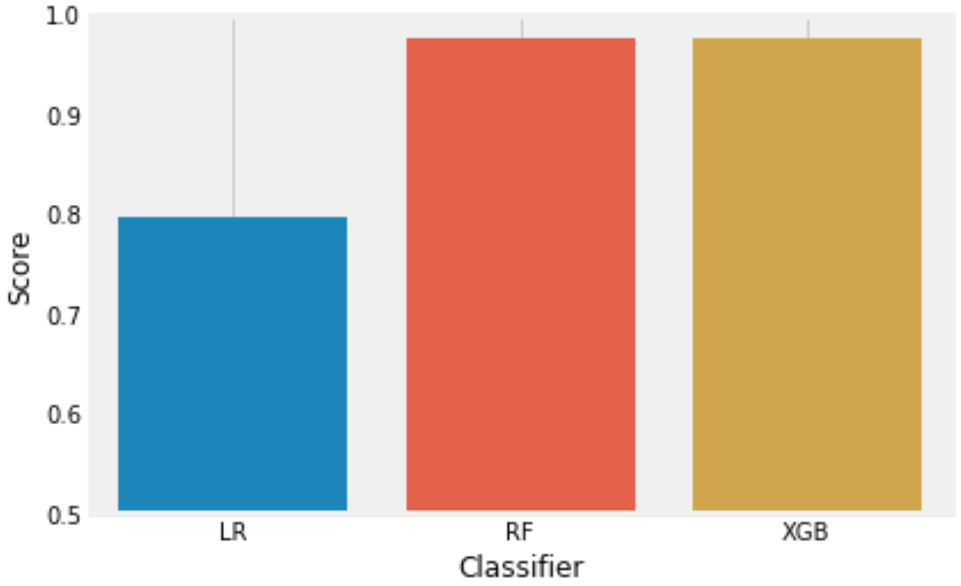
Method 2 **Random Forest**

Feature Selection

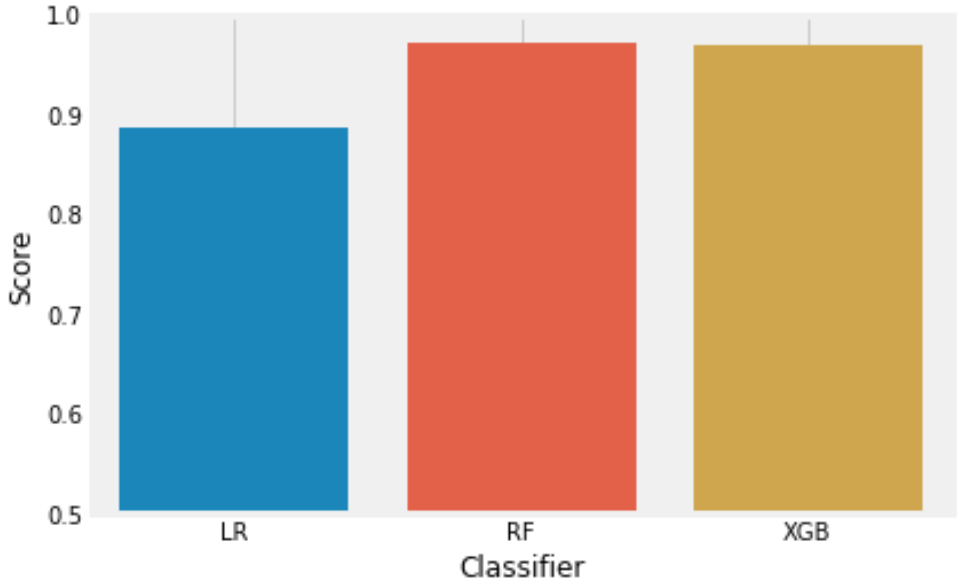


Finetuning - GridSearchCV

Before



After



Results

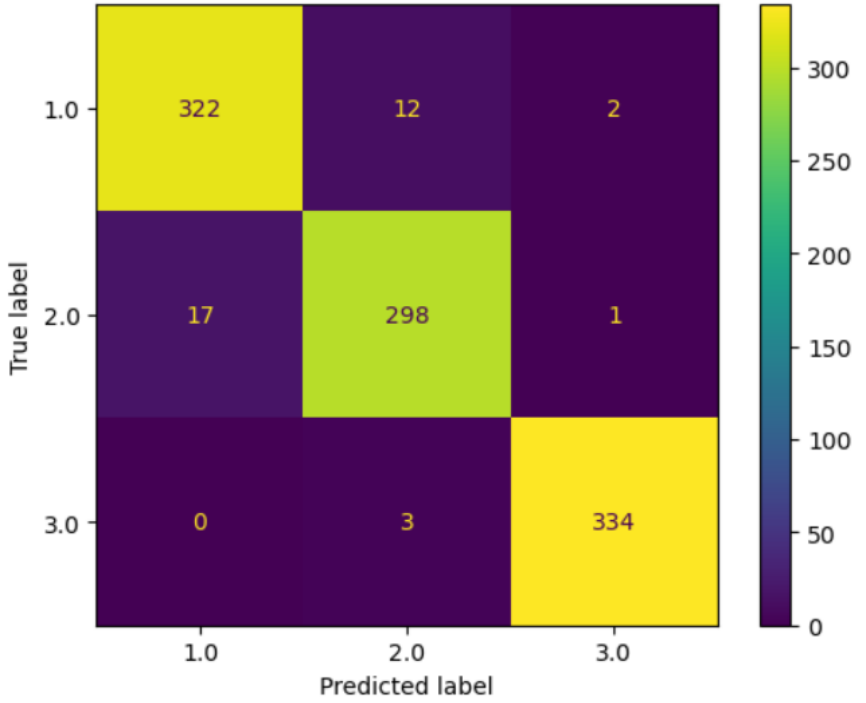
XGBoost classifier
14 Features

Precission: 96.434

Recall 96.415

Accuracy: 96.461

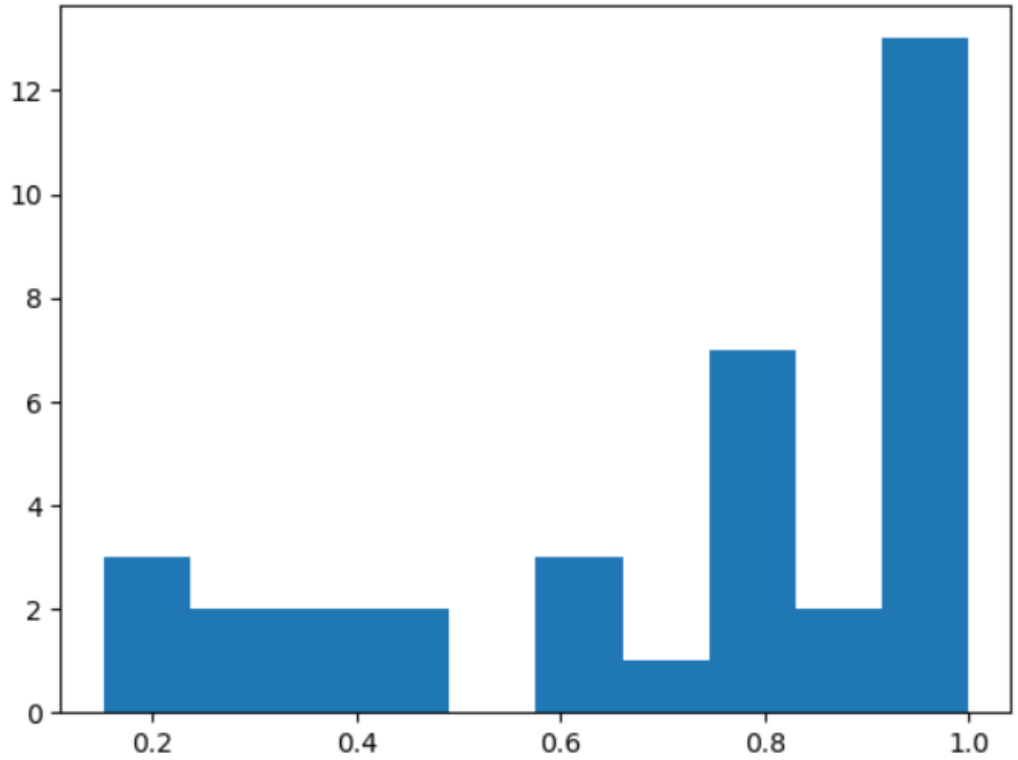
ROC over AUC: 99.622



Results

Examining Misclassified probabilities differences

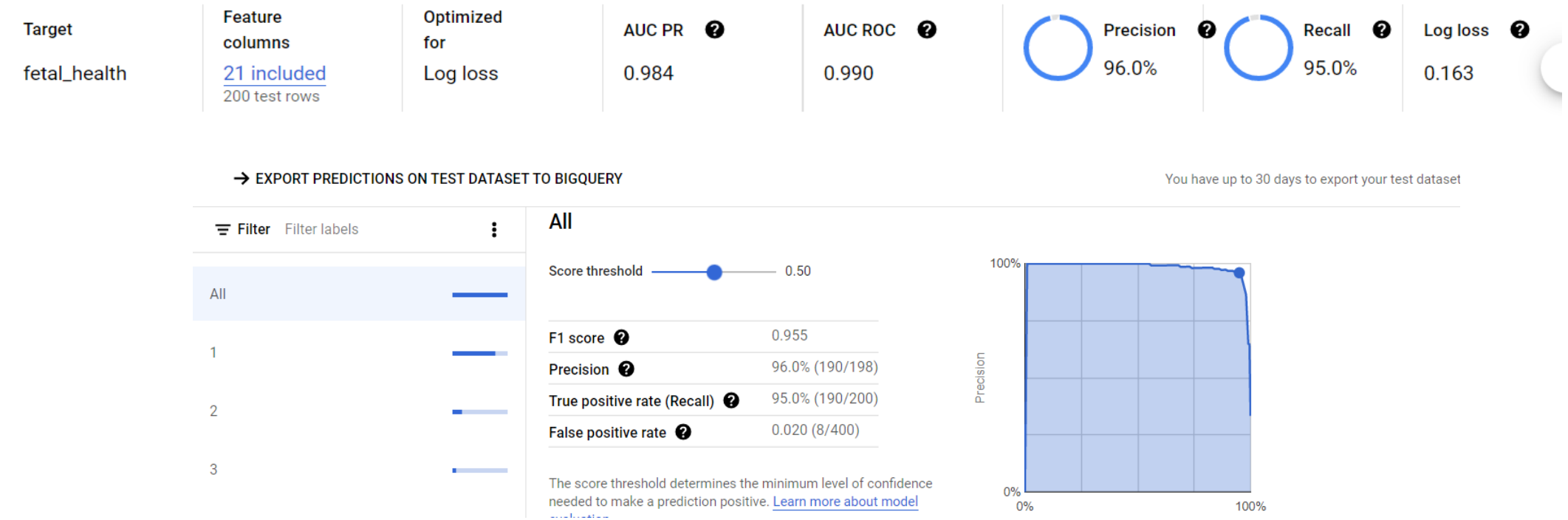
About 13 predictions were incorrectly classified with a very high confidence



Minimum difference of probability to other classes

AutoML

AutoML is google cloud service that automatically find the best model and hyper parameter to use.



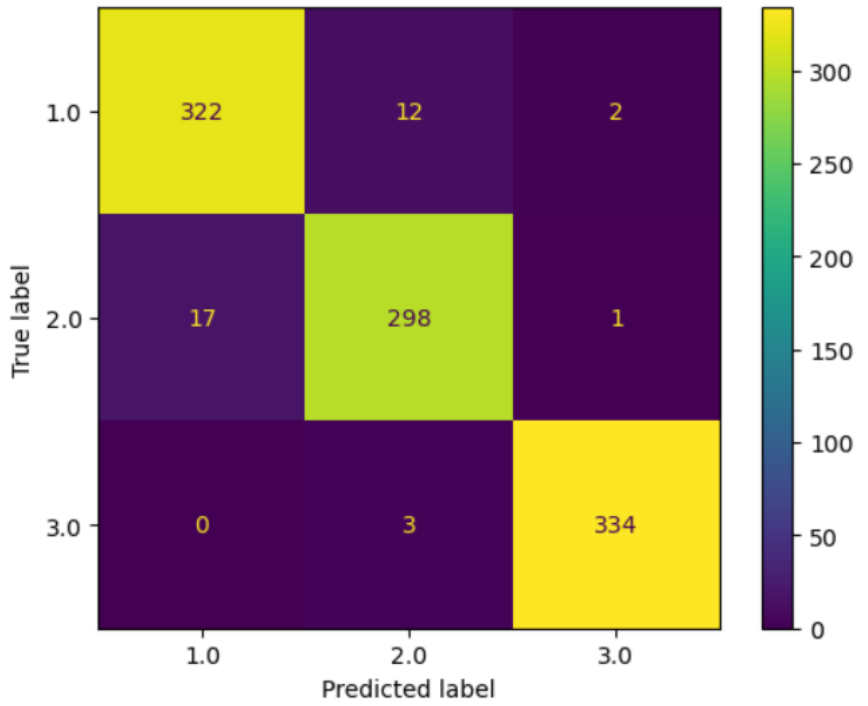
AutoML

Confusion matrix

A confusion matrix summarizes how successful a classification model's
columns are labels predicted by the model. Scan each row to determine
accurate predictions for the label (true positives or negatives) while the
matrix cells to learn more.

True labels \ Predicted labels	1	2	3
1	98%	1%	1%
2	15%	85%	-
3	8%	-	92%

Google Cloud AutoML



Our XGBoost model

Thank You