Second Mini Project - Fetal Health Prediction

Marta Huhoja
December 2022

Use of Machine Learning Algorithms for Prediction of Fetal Risk using Cardiotocographic Data

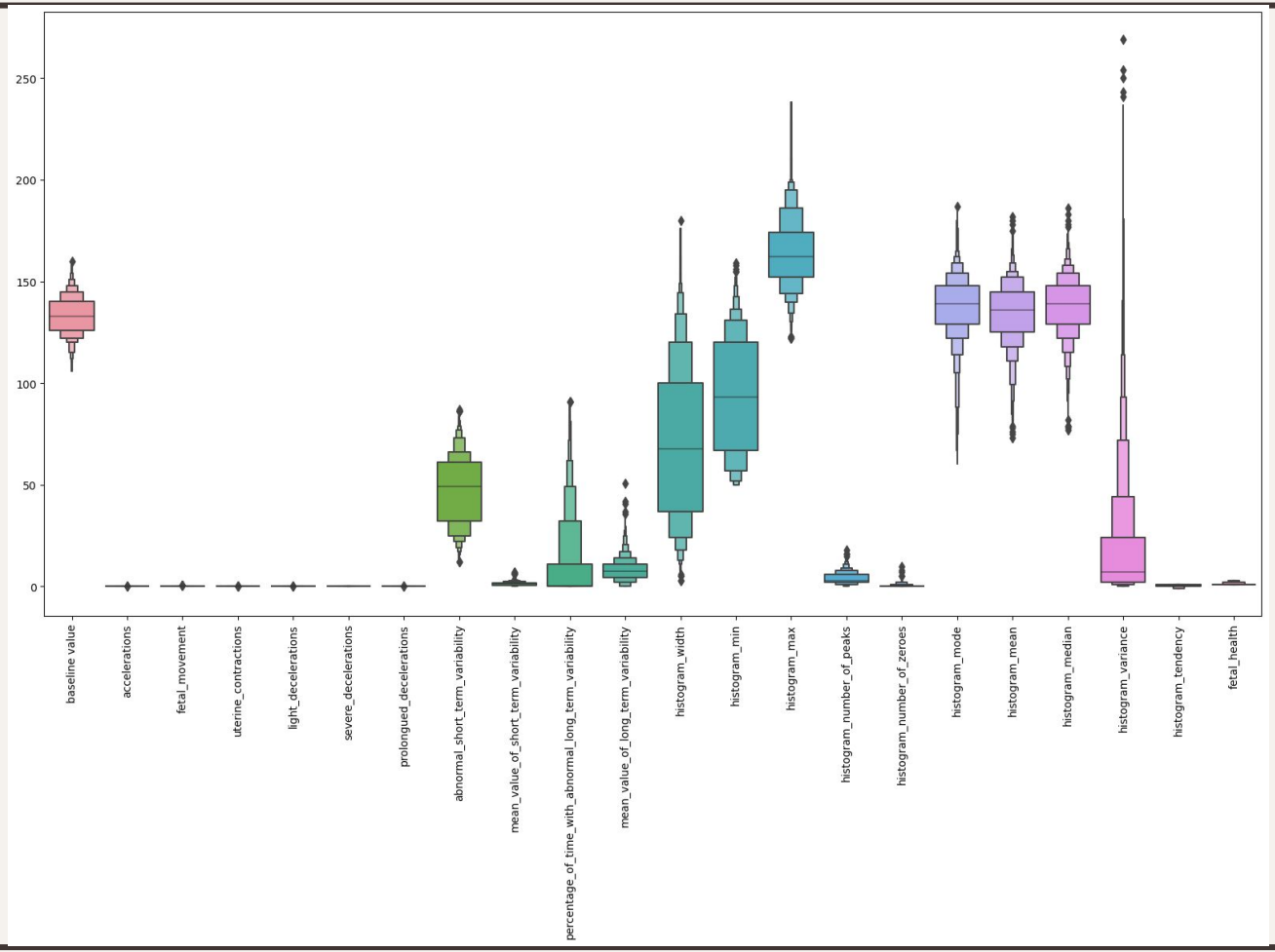
- A major contributor to under-five mortality is the death of children in the 1st month of life
- Intrapartum complications are one of the major causes of perinatal mortality
- Fetal cardiotocography (CTGs) can be used as a monitoring tool to identify high-risk women during labor
- Cardiotocography (CTG) is a continuous recording of the fetal heart rate obtained via an ultrasound transducer placed on the mother’s abdomen

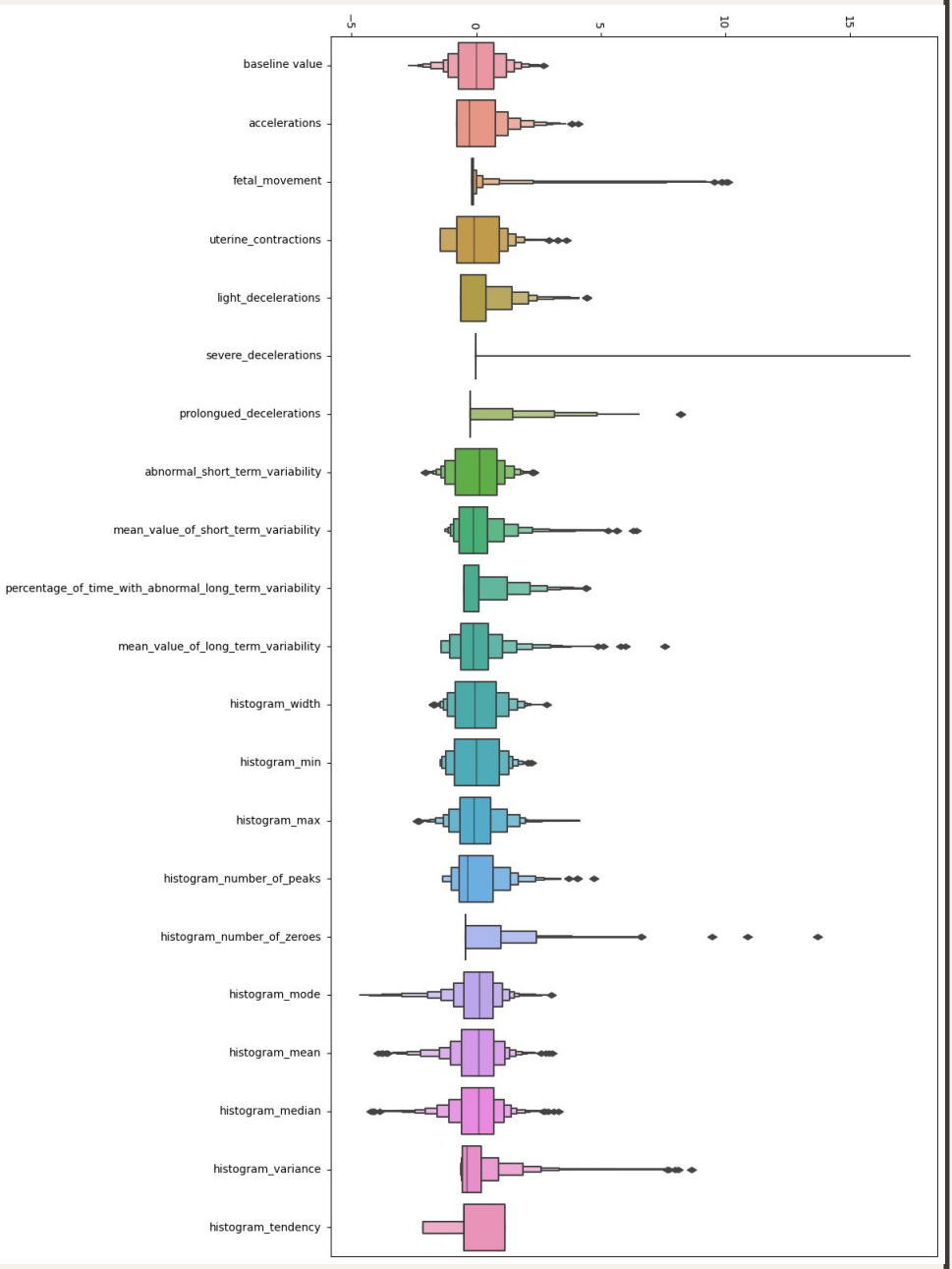
Data

- Data downloaded from Kaggle
- CTG data of 2126 pregnant women were obtained from the University of California Irvine Machine Learning Repository.

EDA

- CTG results described through 22 different variables
- Baseline value (BPM), movement, contractions, decelerations, accelerations, variability, CTG histogram, fetal health
- 2126 inputs
- All values are numerical
- There are no missing values
- variables are **not** in the same range - data has to be scaled


250
200
150
100
50
0
baseline value
accelerations
fetal_movement
uterine_contractions
light_decelerations
severe_decelerations
prolongued_decelerations
abnormal_short_term_variability
mean_value_of_short_term_variability
percentage_of_time_with_abnormal_long_term_variability
mean_value_of_long_term_variability
histogram_width
histogram_min
histogram_max
histogram_number_of_peaks
histogram_number_of_zeroes
histogram_mode
histogram_mean
histogram_median
histogram_variance
histogram_tendency
fetal_health

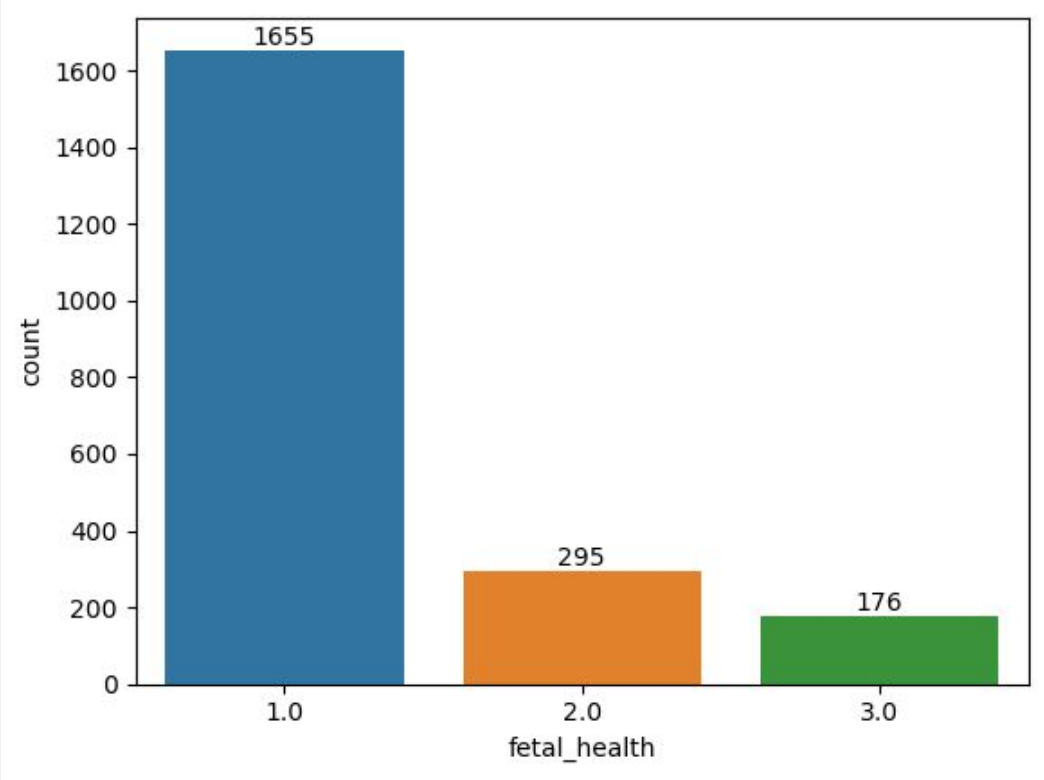

−5
0
5
10
15
baseline value
accelerations
fetal_movement
uterine_contractions
light_decelerations
severe_decelerations
prolongued_decelerations
abnormal_short_term_variability
mean_value_of_short_term_variability
percentage_of_time_with_abnormal_long_term_variability
mean_value_of_long_term_variability
histogram_width
histogram_min
histogram_max
histogram_number_of_peaks
histogram_number_of_zeroes
histogram_mode
histogram_mean
histogram_median
histogram_variance
histogram_tendency

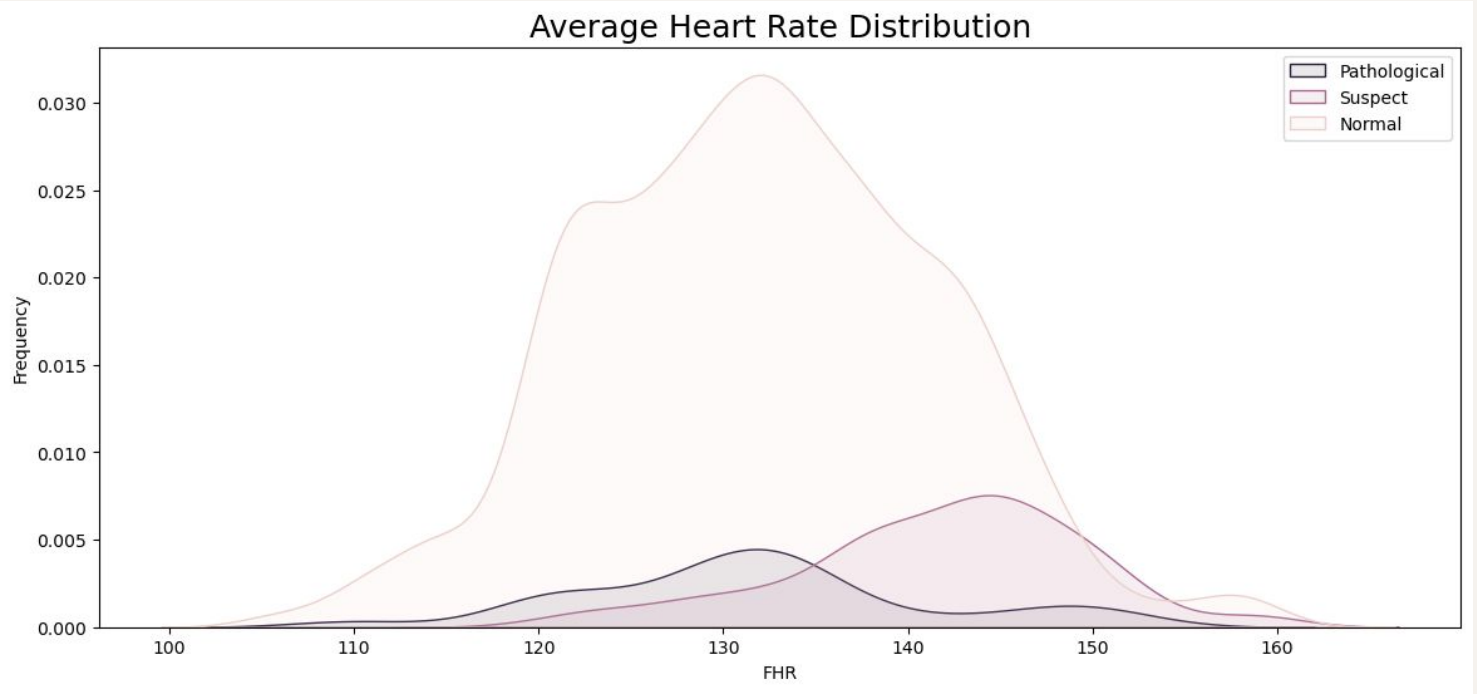
Target variable: fetal_health

3 classes:

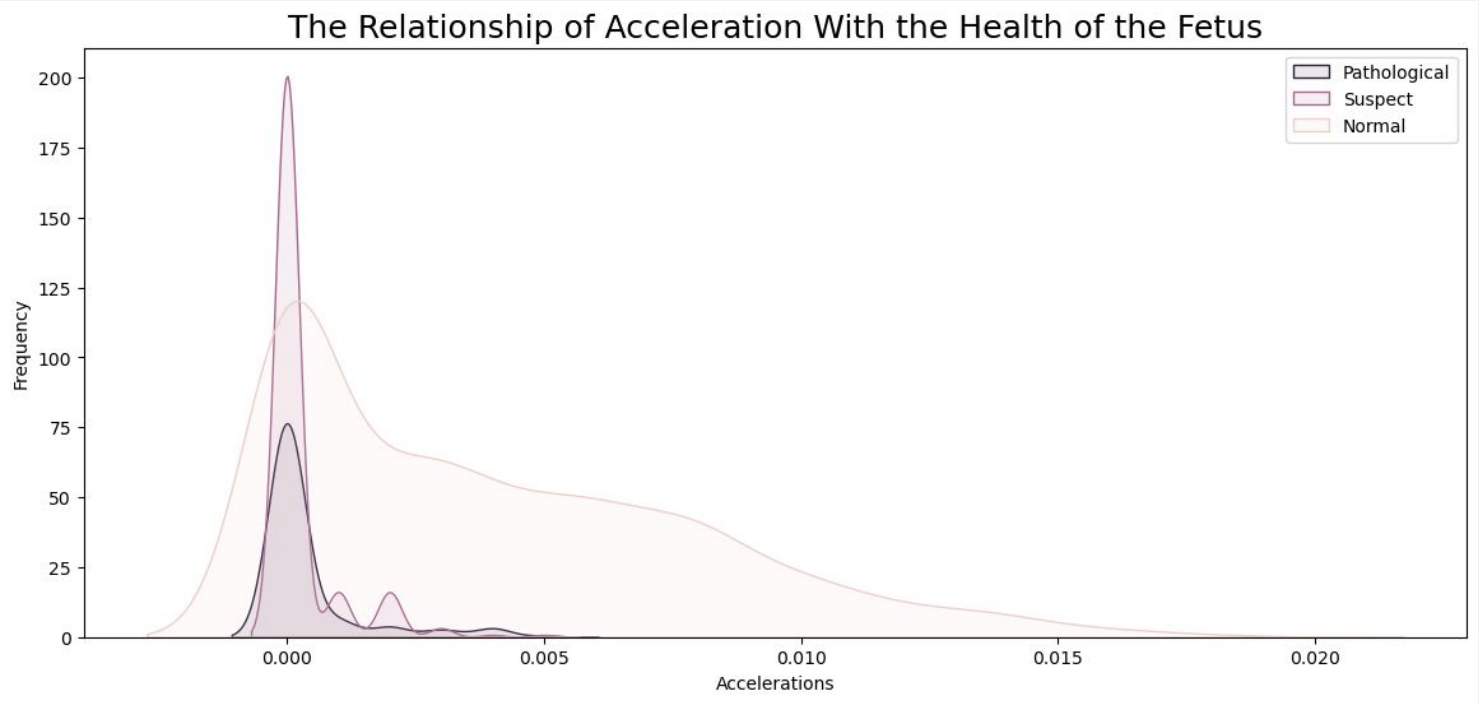
- 1 - Normal,
- 2 - Suspect,
- 3 - Pathological

Used STOME to balance data

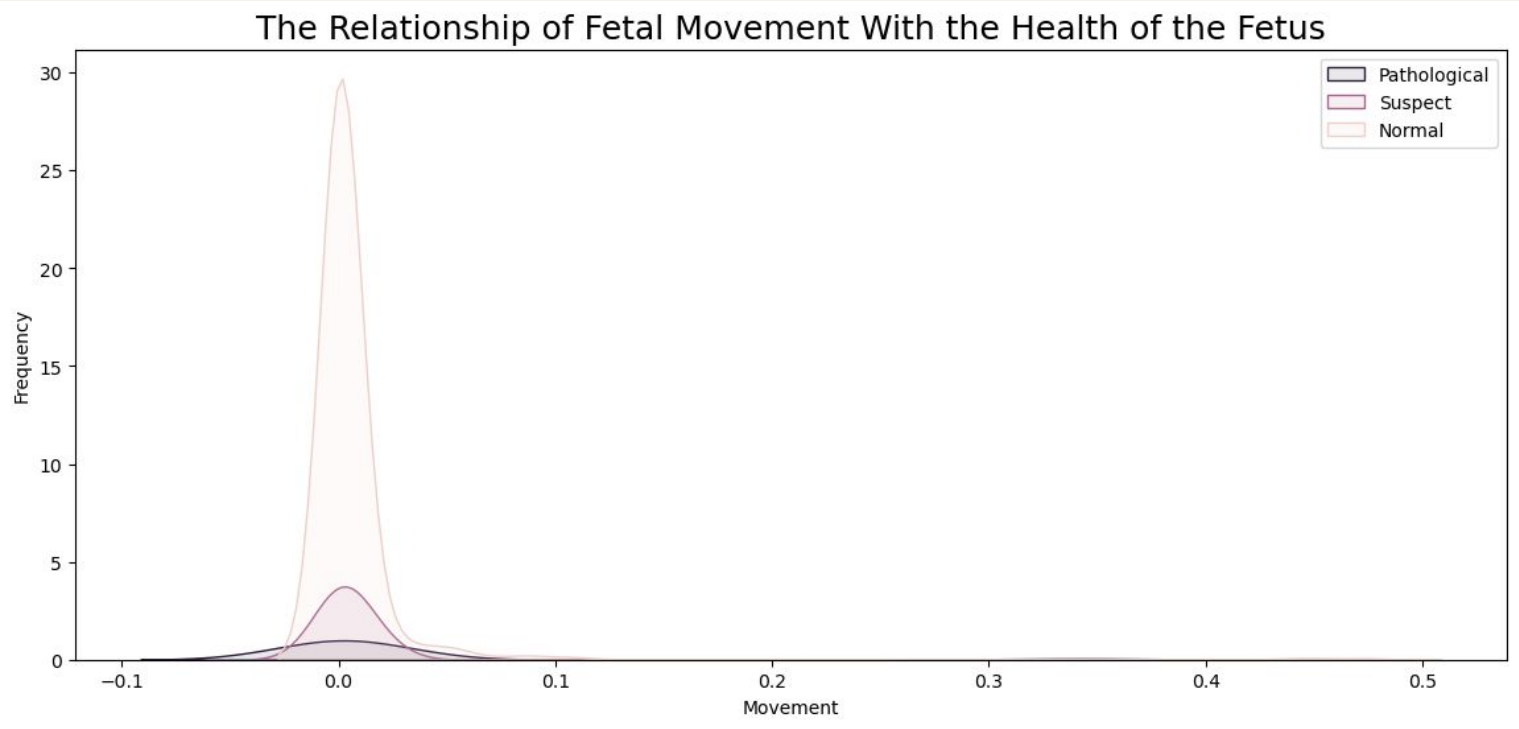


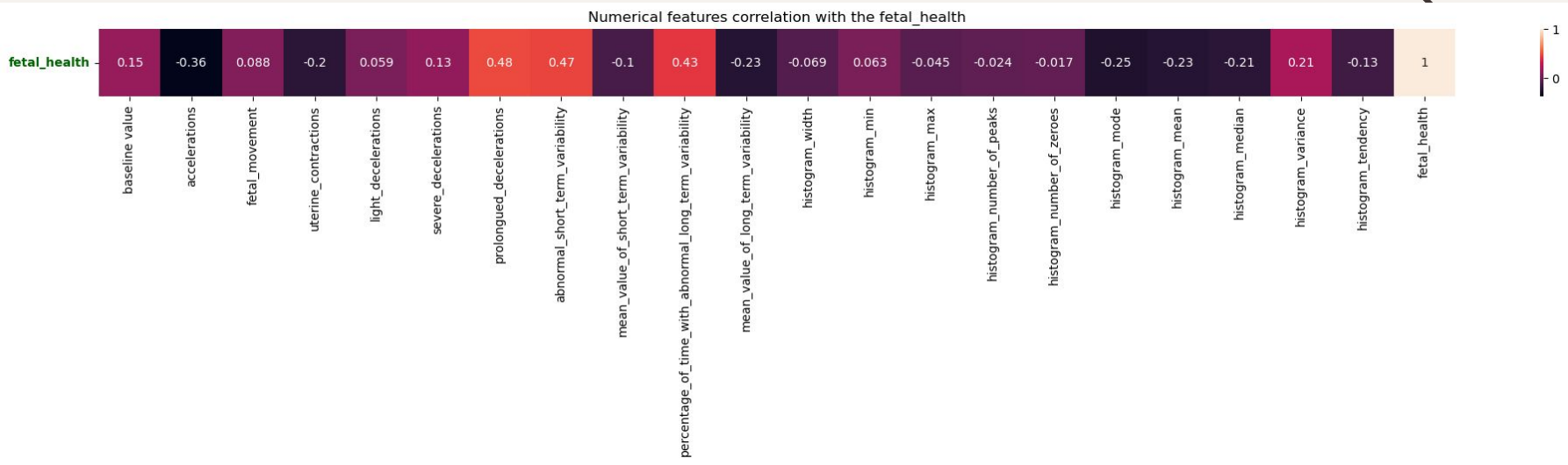


The baseline rate is the average heart rate of the fetus within a 10-minute window.
A normal fetal heart rate is between 110-160 bpm.



Accelerations are an abrupt increase in the baseline fetal heart rate of greater than 15 bpm for more than 15 seconds. The presence of accelerations is reassuring - alongside uterine contractions it is a sign of a healthy fetus.


The Relationship of Fetal Movement With the Health of the Fetus
Pathological
Suspect
Normal
Frequency
Movement
30
25
20
15
10
5
0
−0.1
0.0
0.1
0.2
0.3
0.4
0.5



Decelerations - Decelerations are an abrupt decrease in the baseline fetal heart rate of greater than 15 bpm for greater than 15 seconds.

Variability - Baseline variability refers to the variation of fetal heart rate from one beat to the next.

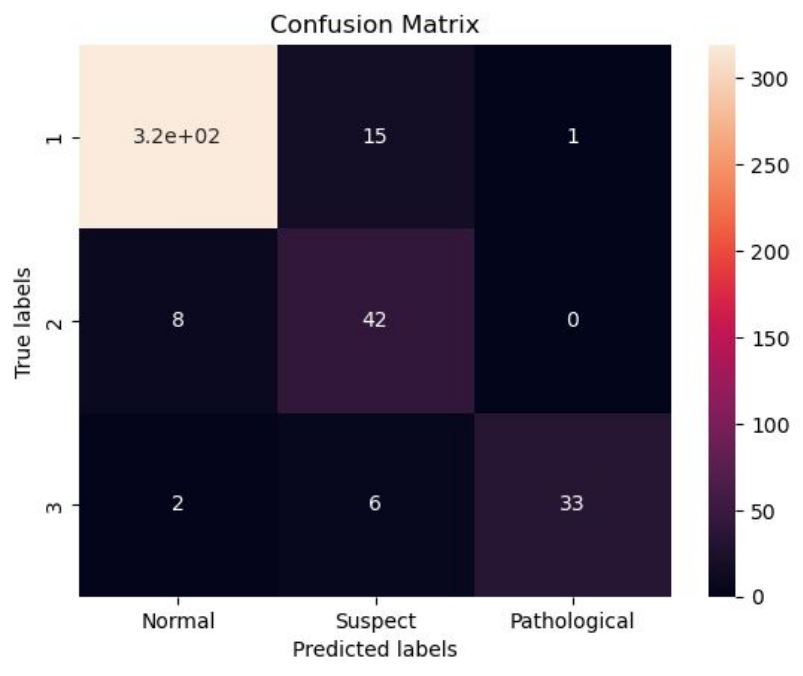
Models used

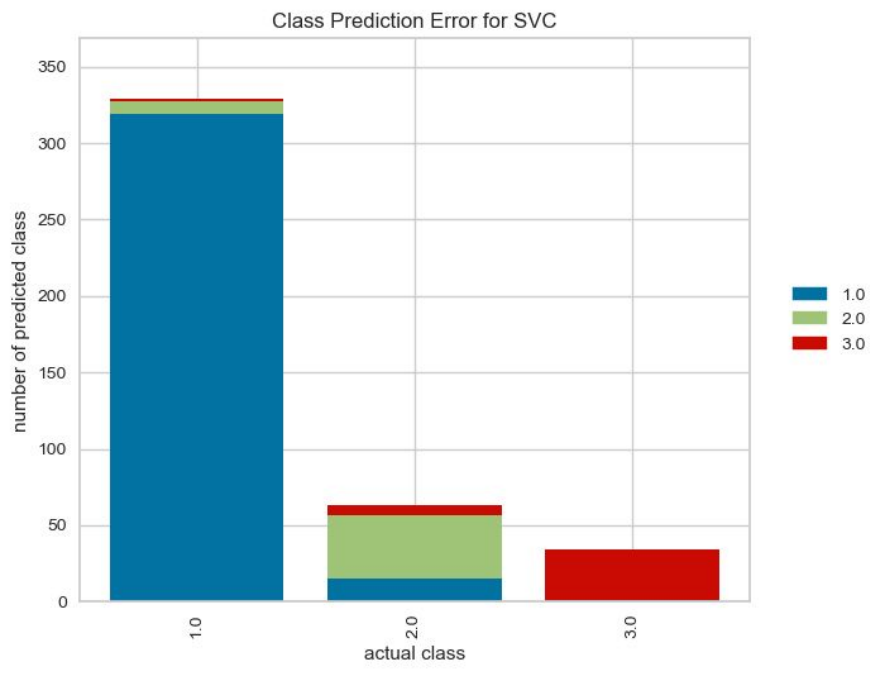
- Used 4 models: LR, SVM, KNN, NB-Gaussian
- Pipeline
- All models were fitted at the same time, doing grid search with multiple parameters
- printed out best estimator and accuracy

Results

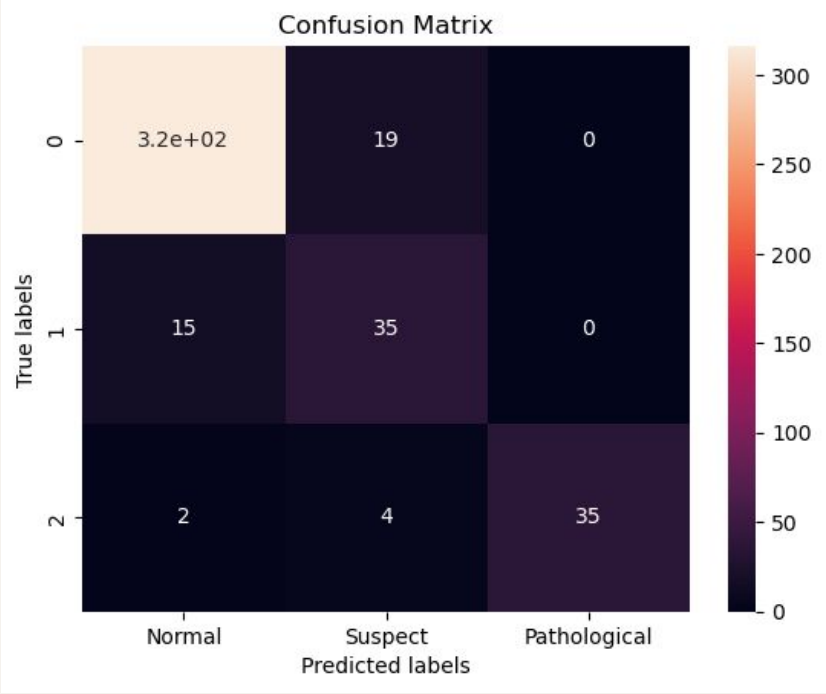
Model	Accuracy
Logistic Regression	85.21%
KNN	90.61%
SVM	**92.45%**
NB Gaussian	79.58%

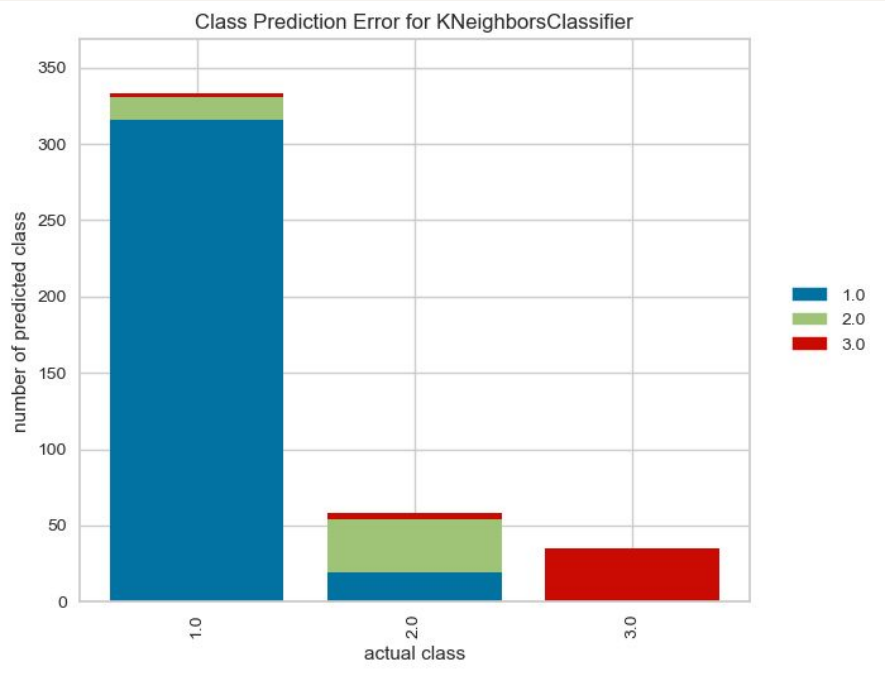
SVM


Confusion Matrix
3.2e+02
15
1
8
42
0
2
6
33
True labels
1
2
3
Normal
Suspect
Pathological
Predicted labels
300
250
200
150
100
50
0


Class Prediction Error for SVC
number of predicted class
350
300
250
200
150
100
50
0
1.0
2.0
3.0
actual class
1.0
2.0
3.0

KNN


Confusion Matrix
3.2e+02
19
0
15
35
0
2
4
35
True labels
0
1
2
Normal
Suspect
Pathological
Predicted labels
300
250
200
150
100
50
0


Class Prediction Error for KNeighborsClassifier
number of predicted class
350
300
250
200
150
100
50
0
1.0
2.0
3.0
actual class
1.0
2.0
3.0

Conclusion

- Machine Learning models can be useful to help predict fetal health and possible complications - possible use by health professionals
- Limited data - no sociodemographic data, information about pregnancy or mother health
- Future improvement of ML algorithm using different models, larger dataset and feature selection

References

- **Dataset**: https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification
- **Grivell RM, Alfirevic Z, Gyte GM, Devane D. Antenatal cardiotocography for fetal assessment.** Cochrane Database Syst Rev. 2015 Sep 12;2015(9):CD007863. doi: 10.1002/14651858.CD007863.pub4. PMID: 26363287; PMCID: PMC6510058.
- **Hoodbhoy Z, Noman M, Shafique A, Nasim A, Chowdhury D, Hasan B. Use of Machine Learning Algorithms for Prediction of Fetal Risk using Cardiotocographic Data.** Int J Appl Basic Med Res. 2019 Oct-Dec;9(4):226-230. doi: 10.4103/ijabmr.IJABMR_370_18. Epub 2019 Oct 11. PMID: 31681548; PMCID: PMC6822315.
- https://geekymedics.com/how-to-read-a-ctg/