

Can machine learning models help us read CTG results?

CAPSTONE THREE PRESENTATION
BY MORGAN SNELGROVE

Did you know...

Every day, 800 women die from preventable causes related to pregnancy and childbirth around the world.

In 2021, 5 million deaths in children under the age of 5 was recorded, mostly from preventable and treatable causes.

United Nations World Health Goals

Reduce
maternal
mortality

Reduce
newborn
and under-
5 mortality

Cardiotocography



Simple

Cost
Effective

Improve
Pregnancy
Outcomes

Human Error in CTG Interpretation

- ▶ Some measurements can't be quickly calculated
- ▶ High variability between interpretations
- ▶ Inaccurate interpretations lead to negative outcomes.

Negative Outcomes

- ▶ C-sections
- ▶ Metabolic acidosis (build up of acid in the body)
- ▶ Preterm delivery
- ▶ Low birthweight
- ▶ poor health metrics
- ▶ A stay in NICU

Computerized Analysis of CTG Results

Can a computer accurately classify fetal health?

Studies show promising results!

2016 Study by Ignatov

- ▶ Goal: Evaluate the effectiveness of Computerized CTG compared to conventional CTG
- ▶ Sample size: N =720 women in active labor
- ▶ 360 women randomly assigned to computerized CTG, remaining giving conventional CTG
- ▶ All newborns from both groups were assessed in the same way.

Their Conclusion:

The incidence of adverse perinatal outcomes was lower among women who were monitored by computerized CTG.

This ML project

Can we build a machine learning model to predict fetal health?

Our Data:

- ▶ 2126 examples of CTG readings.
- ▶ Features extracted from CTG outputs and labeled by professionals
- ▶ 3 Categories:
 1. Normal
 2. Suspect
 3. Pathological

Our Goals:

Greater than
80% accuracy

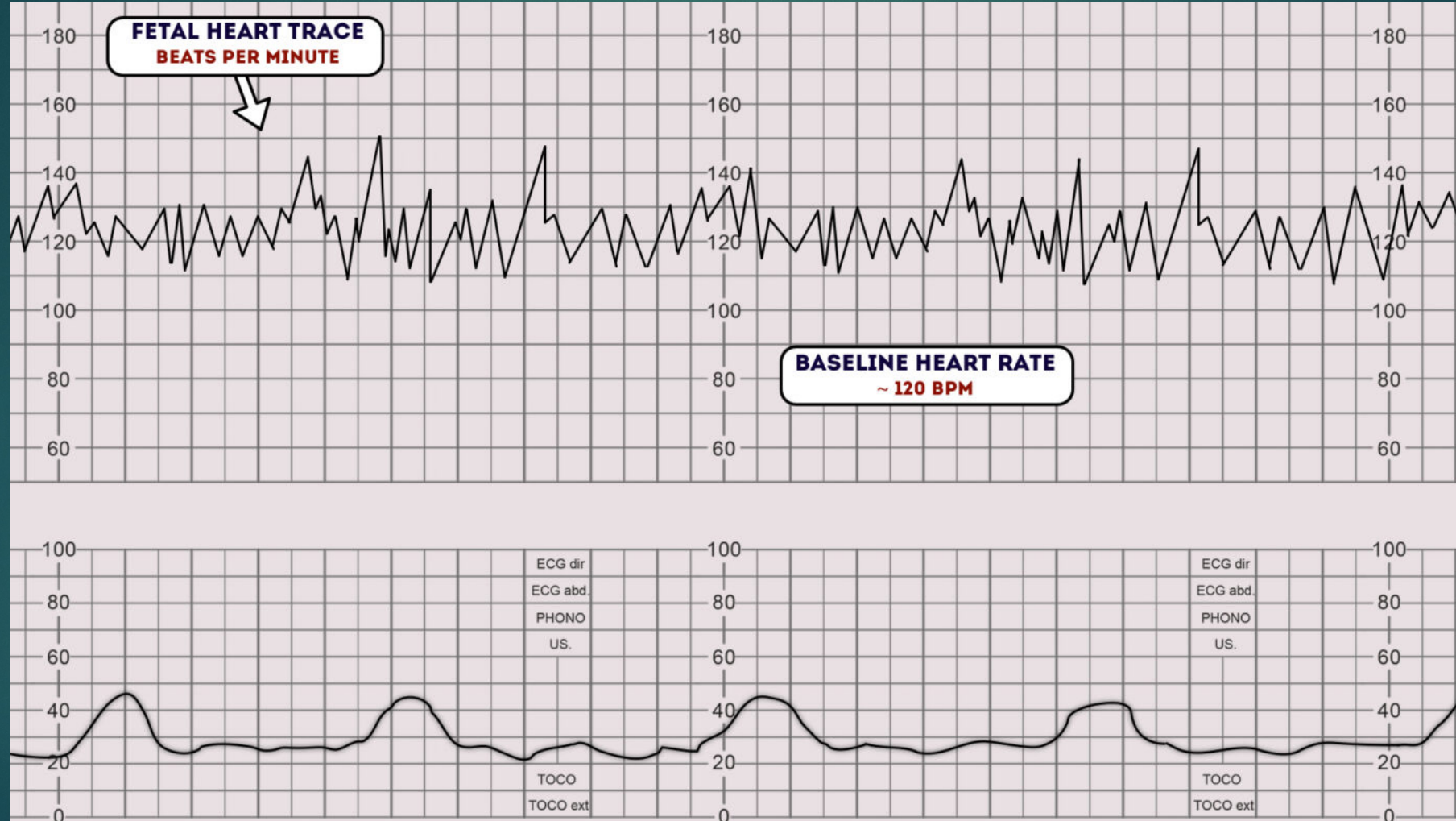
Class 3
recall greater
than 90%

Identify most
important
features

Our Features

- ▶ We have 22 columns
- ▶ First 11 columns are measurements taken from the CTG
- ▶ They deal with:
 1. Baseline Rate
 2. Variability
 3. Accelerations
 4. Decelerations

An example of a CTG printout



Baseline Rate

The average
fetal heart
rate

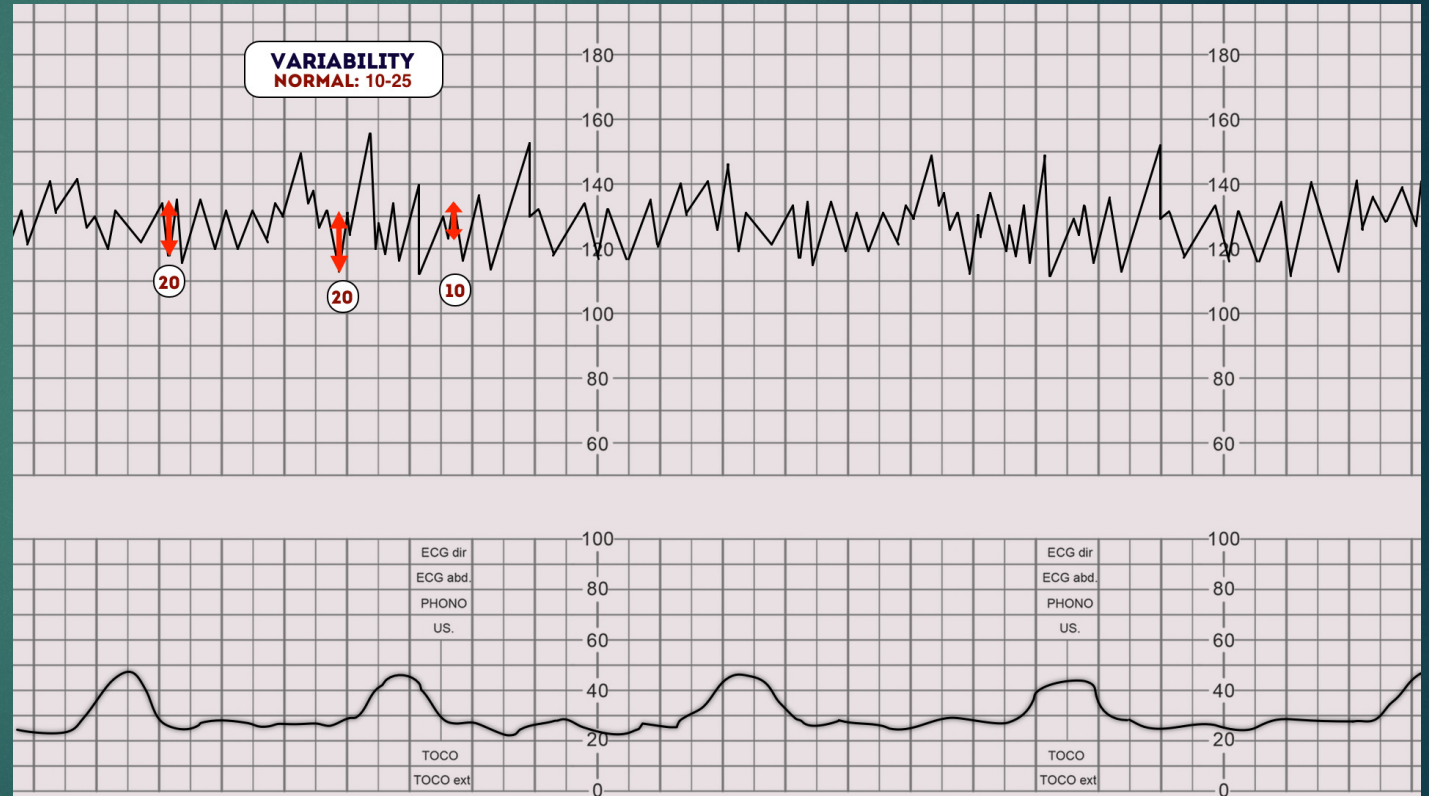
Normal is
between 110
to 160 bmp

Variability

Fluctuations in
fetal heart rate

Normal is
between 5 to 25
bpm

Measured from
peak to
following trough

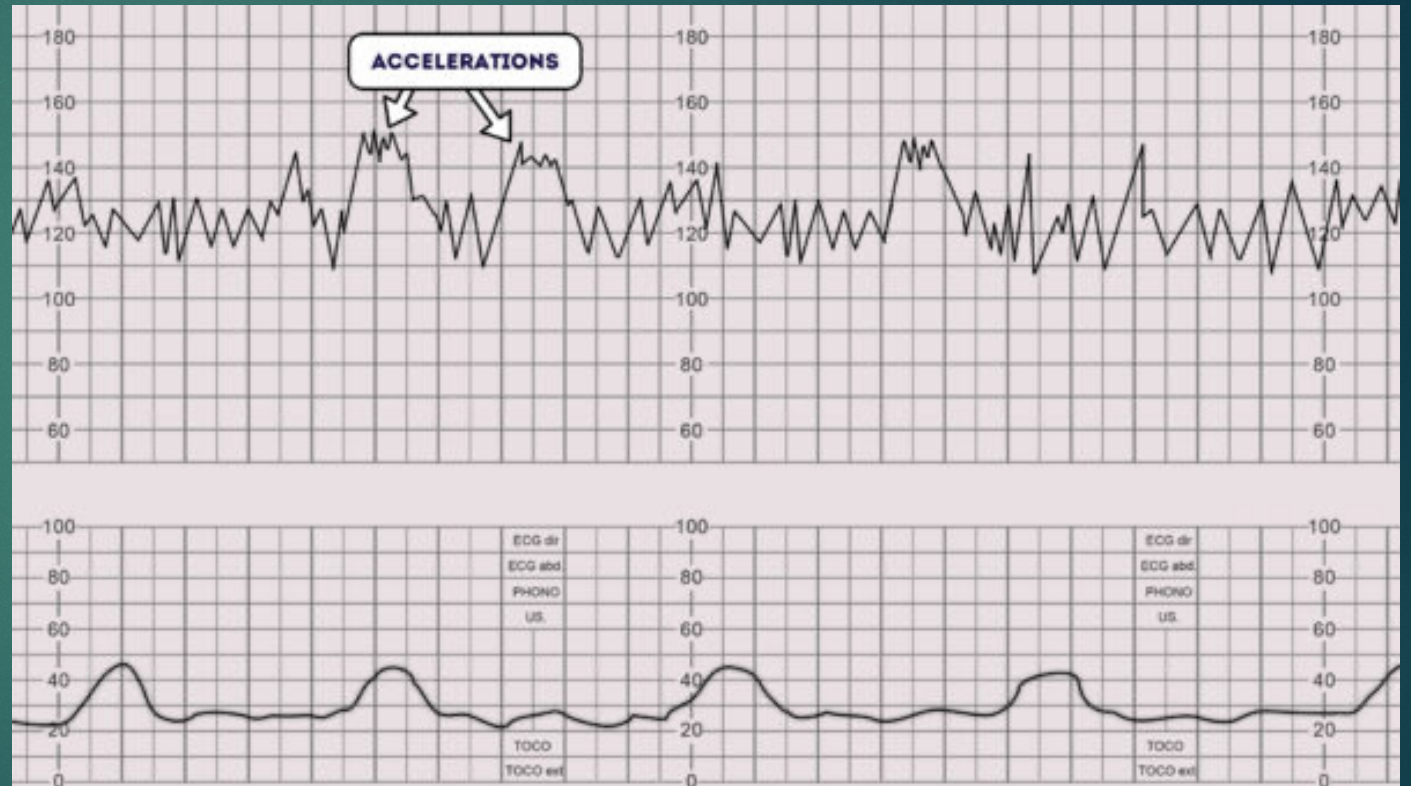


Accelerations

Short increases of
FHR above the
baseline

Above the
baseline by >15
bpm for > 15 sec

A sign of good
fetal health

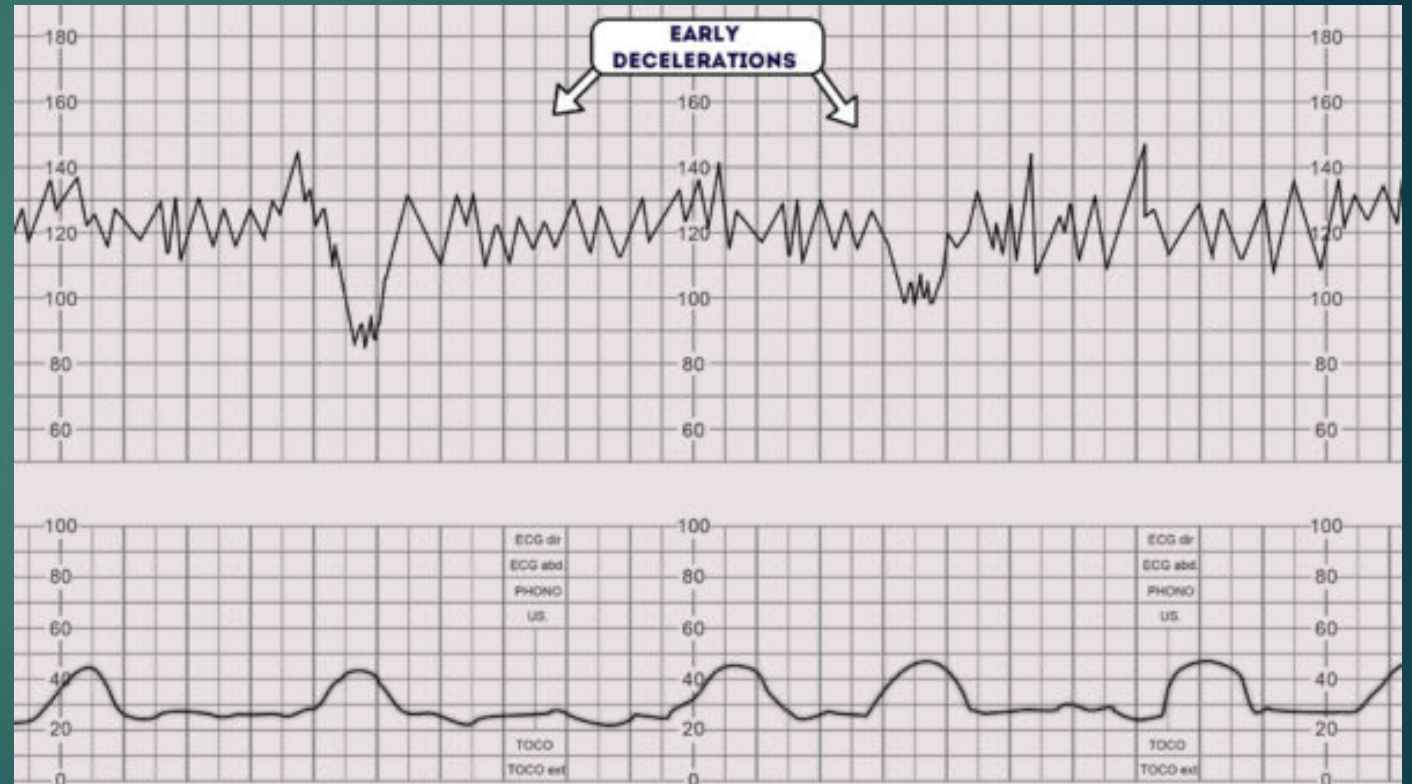


Decelerations

Short decreases of
FHR below the
baseline

Below the
baseline by 15
bpm for 15 sec

Many different
types



Histogram of Fetal Heart Rate

- ▶ Remaining feature columns extracted from a histogram of FHR
- ▶ Measures things like:
 - ▶ Mean
 - ▶ Median
 - ▶ Mode
 - ▶ Number of Peaks
 - ▶ Width
 - ▶ Max
 - ▶ Min

Positive Correlation with Target

- ▶ Prolonged deceleration
- ▶ Abnormal short-term variability
- ▶ Percentage of time with abnormal long-term variability

Negative Correlation with Target

- ▶ Accelerations
- ▶ Uterine Contractions
- ▶ Mean Value of Long-Term Variability

Early Models

Dummy Model for Baseline

- ▶ Just predict the most common class every time
- ▶ Accuracy 78%
- ▶ Class 3 Recall of 0

First Real Models

- ▶ Decision Trees
 - Accuracy: 89%
 - Class 3 Recall: 77%
- ▶ Random Forest
 - Accuracy: 93%
 - Class 3 Recall: 89%

Best Multiclass Model

Gradient Boosting Classifier

Accuracy of 94%
Over 20% increase!

Class 3 Recall of 93%

Assessing the Models

- ▶ All the models struggled the most with class 2 – Suspect.
- ▶ Would combining Classes 2 and 3 into one category help?
- ▶ New Categories:
 - ▶ 1 = Suspect or Pathological
 - ▶ 0 = Normal

Best Binary Model

Gradient Boosting Classifier

Accuracy of 96%

Class 1 Recall of 88%

Which is Better?

- ▶ Suspect Category is hard to pick out
- ▶ Worse Case: Pathological Classified as Normal
- ▶ The Multiclass Model misclassified:
 - ▶ 2 Pathological as Normal
 - ▶ 19 Suspect as Normal
- ▶ The Binary Model misclassified 14 cases as Normal
 - ▶ We don't know if these were Suspect or Pathological

Keep all
three
classes

Conclusion

Machine Learning has a lot to offer in this area!

- ▶ With the aid of computerized CTG we could:
 1. Reduce the under-5 mortality rate
 2. Improve decision making strategies for labor and delivery
 3. Enable medical professionals to help more patients
 4. Decrease the occurrences of negative outcomes for newborns

Questions and Continued Study

- ▶ Since we are questioning the labels, it would be interesting to do a clustering analysis of the data without the labels.
- ▶ It would be very helpful to have the negative outcomes for each birth.
- ▶ How many of the Suspect cases were healthy babies?
- ▶ How many of these healthy class 2's did the model misclassify? In the end, who was right?
- ▶ We need more studies with larger sample sizes.