Fetal Health Prediction

Topic Overview:

Every day in the U.S., about 10,000 people are born. Fetal health encompasses the physical and physiological well-being of a developing fetus during pregnancy. During the critical 40-week period of fetal development, regular screenings and tests play a vital role in identifying early warning signs of potential complications, allowing medical experts to intervene if necessary to ensure both maternal and fetal health.

The goal of this project is to work with a real-world healthcare dataset to analyze the features responsible for ensuring good maternal and fetal health during pregnancy. Students will analyze quantitative features from cardiotocograms with the aim of predicting health status from measurements of fetal movements.

Dataset Description:
This dataset contains 2,126 observations of 22 variables, capturing quantitative information on fetal health. The features are measurements from cardiotocograms, a medical test used to monitor fetal heartbeat and contractions. The test operates by a technique called **ultrasound**, where high-frequency sound waves are used to create real-time pictures of a fetus inside the body. An exam histogram of fetal heart rate and movement is produced, and then many features can be extracted from this chart for further analysis. Each row represents one cardiotocogram for one fetus.

Key Columns:

- Baseline value: Baseline fetal heart rate
- Accelerations: Number of accelerations per second
- Fetal_movement: Number of fetal movements per second
- Uterine_contractions: Number of uterine contractions per second
- Light_decelerations: Number of light decelerations per second
- Severe_decelerations: Number of severe decelerations per second
- Prolonged_decelerations: Number of prolonged decelerations per second
- Abnormal_short_term_variability: Percentage of time with 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: Range between highest and lowest histogram values
- Histogram_min: Lowest value in the histogram
- Histogram_max: Highest value in the histogram
- Histogram_number_of_peaks: Count of distinct peaks in the histogram
- Histogram_number_of_zeroes: Number of zero-frequency bins in histogram
- Histogram_mode: Most frequent value in the histogram
- Histogram_mean: Average value of histogram data
- Histogram_median: Middle value of histogram data
- Histogram_median *(duplicate)*: Same as above
- Histogram_variance: Spread of values in the histogram
  Histogram_tendency: Skew or direction of histogram distribution
- Fetal_health: 1 - Normal; 2 - Suspect; 3 - Pathological

This dataset is publicly available here:
https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification

Topics to Learn More About:

- **Cardiotocography:**
  - Cardiotocography (CTG) is a monitoring technique used during pregnancy to record the fetal heart rate and uterine contractions. It helps healthcare providers assess fetal well-being and detect signs of distress. CTG data is often used to support clinical decisions about labor and delivery.
- **Multi-class classification model evaluation:**
  - Multi-class classification model evaluation involves assessing how well a model can distinguish among three or more categories. Desired metrics should go beyond accuracy, additionally considering precision, recall, F1-score, and confusion matrices, all extended for multi-class scenarios. These evaluations help determine a model's effectiveness and guide improvements in performance.