**Background of the project**

**Child mortality remains one of the major challenges facing global health. UNICEF estimated that almost half of child mortality happens during the first five months of birth attributing this to premature birth and complications during labor. Most of these deaths occur in places that are often without the needed resources to ensure early prognosis and treatments of complications that may arise during pregnancy. For example, in 2023, UNICEF reported that 1 in 15 children died in sub-Saharan Africa before reaching the age of 5. This disproportionate distribution of child mortality explains why more efforts need to be concentrated in poorer regions to protect infants and pregnant women.**

**Considering these challenges, designing cost effective way to monitor and treat early pregnancy complications remains an active area for research.** **This project proposes a Cardiotocograms (CTG) and the deployment of a machine learning model as a cost effective and simple way to examine fetal health, enabling health practitioners, midwives and doctors take immediate actions to address complications and distress that may occur during pregnancy and labor.**

**The device works by placing an elastic belt around the abdomen of the mother that is than use to monitor the baby heart rate and the mother pressure of the mother tummy. These belts than send signals to the machine which interprets theses signals. Health care practitioner use these signals to make determination on the sate of the fetal and their mother health.**

**Fetal health is classified as either normal, suspicious or pathological. Unfortunately, analyzing all these signals simultaneously for most medical practitioners who must make many of these diagnoses daily can be challenging and may sometimes lead to misdiagnosis. Therefore, this project explores the possibility of using machine learning models to aid health practitioners, obstetricians, and diagnosticians make accurate and effective diagnosis on the wellbeing of fetal health and their mother.**

T**his project uses a Cardiotocograms data prepared in a study by Ayres de Campos et al. (2000) to explore the possibility of using a machine learning models to predict the state of fetal health and their mothers.**

**Analysis**

**First the data was processed and explored. The data has a dimension of (2126, 22) with fetal health being the target. The rest of the other features of the data are those gathered and characterized by professional obstetrician against the health of fetal. Each fetal was classified as normal (1), suspicious (2), and pathological (3).** **To begin, the data was split using a stratification approach to ensure that the proportion of classes are equally represented across the training and test set.** **A dummy classified based on the mere frequency of classes was first fit on the data to gather the classification of fetal health. This approach is important to use as a baseline model to evaluate the need for a more robust model.**

**It was observed that the dummy classifier performed poorly based on reported weighted metrices such as accuracy, precision recall and F1 score.** **To test more sophisticated models, a cross-validation approach with hyperparameter tuning was used to generate more accurate and informative predictive performance on the data**. **The models that were particularly fitted on the data are KNN, Decision Trees and Random Forest.**

**Additionally, since medical decisions tend to be critical, a permutation feature importance was applied to generate transparency and explanation from the models. Permutation feature importance works by randomly shuffling features in the data, refitting the model and evaluating its error.**

**Results and Conclusions**

**Based on the all the models evaluated on the data, Random Forest demonstrated the highest performance. The model weighted metrices reported for accuracy, precision recall and F1 score were 93, 94, 94, 94, respectively. The analysis of the feature permutation importance showed that features such as “abnormal short term variability” and “mean value short term variability” tend to heavily influence the model predictive performance compared to low performing features such as “baseline value” and “histogram number of zeroes”.**

**Based on the model performance, it can be inferred that with the growth in artificial intelligence, critical areas such as child and maternal health can be significantly improved. While no scientific and medical decision can be reach from the outcomes in this project, this project demonstrates that with sustain investment in machine learning research, health outcomes could be significantly improved, especially in regions still grappling with low investment in health. This project shows that the use of a Cardiotocograms and the deployment of a sophisticated machine learning algorithms can enable doctors and health practitioners improve their diagnoses and provide more accurate and informative medical services for pregnant women and their fetal, thus, curbing the rate of fetal and maternal mortality. However, it must be concluded that more research is needed to thoroughly demonstrate the usefulness of these models, their potential challenges and limitations when deployed in health settings.**