Predictive Modeling for Low Birth Weight Classification

Brandi Jones

Faculty Advisor: Dr. MinJae Woo

INTRODUCTION

- The purpose of this study is to explore the use of current modeling methods for infant low birth weight prediction using a variety of maternal and paternal factors.

METHODS

Dataset

- Infant dataset was obtained from the National Survey of Family Growth (NSFG) survey conducted by the Centers for Disease Control and Prevention (CDC) from 1973-2019.
  - Survey collects information on fertility, family planning, and reproductive health in the United States.
  - The sample was designed to be representative of live births in the United States using continuous interviewing/fieldwork survey methodology.
  - Dataset included 101,400 live births and 41 variables.

Data Processing

- Low Birth Weight (LBW) binary classification response variable created using 5.511557 lbs (2500 g) as threshold.
- MICE Imputation performed for missing values after handling of coded missing.
- Use of 60:20:20 ratio for train/validate/test sets for all models.

Modeling Methods Used

- XGBoost
  - Hyperparameter Tuning
  - *Code adapted from Dr. MinJae Woo DS7140 Notes*
- Naïve Bayes
- Random Forest
- Logistic Regression

Modeling Results

- AUC/ROC Curves calculated for each model.
- Confusion Matrix created for each model.
- Accuracy, F1 Score, Precision and other model performance metrics calculated.

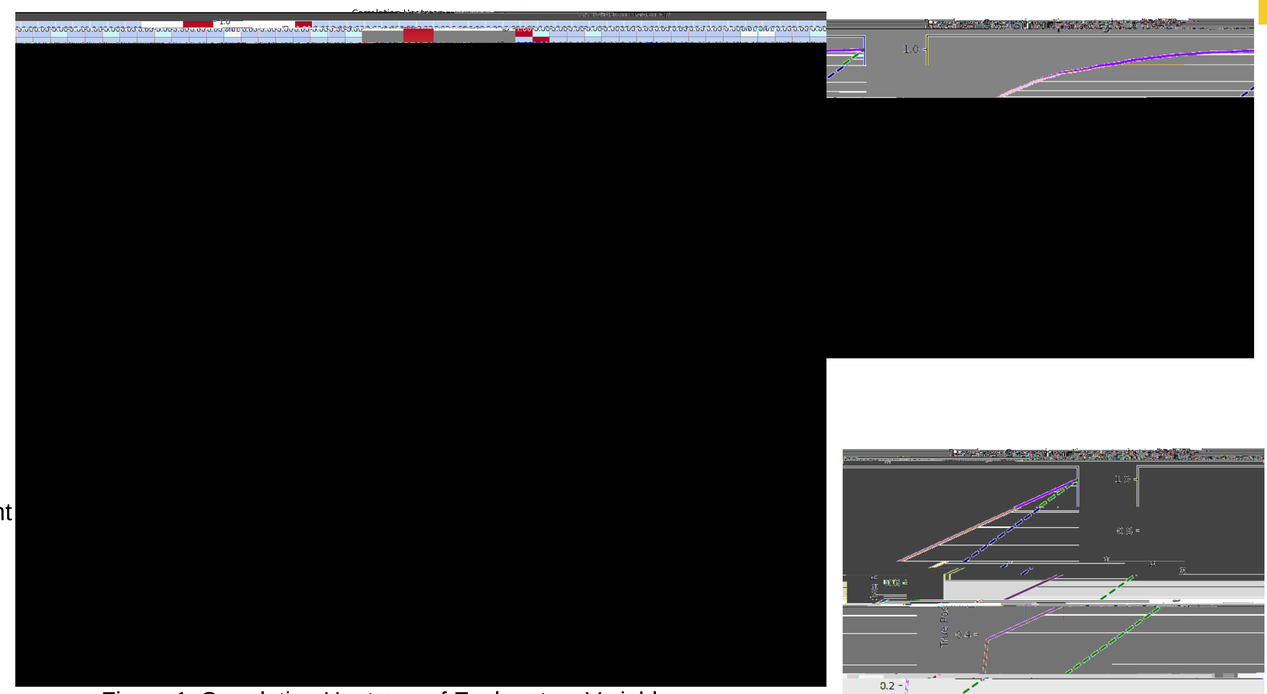
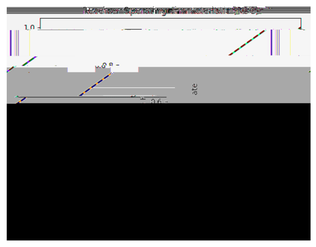


Figure 1: Correlation Heatmap of Explanatory Variables



RESULTS

Model Outcomes (Table 2)

- XGBoost had the highest AUC/ROC curve score of all models.
-

- Naïve Bayes' higher accuracy and precision preferable in development of screening programs aimed at confidently identifying LBW infants.
  - Relative model simplicity makes it ideal with limited computational resources.

Limitations

- Regardless of model performance, ability to interpret is crucial for clinicians' acceptance.
- Predictive modeling in healthcare warrants ethical considerations as regards biases in the data or algorithms.

Brandi Jones
MSDSA Graduation
Spring 2025