Predicting Postnatal Depression During the COVID-19 Pandemic: A Data Mining Approach

Presenter name
University of Colorado Boulder

Abstract

This project aims to develop a predictive model for identifying individuals at high risk of postnatal depression using data collected during the COVID-19 pandemic, highlighting the use of data mining techniques and the potential impact on healthcare interventions.

Introduction

- The COVID-19 pandemic has heightened anxiety and stress among pregnant individuals, leading to increased mental health challenges due to social isolation and healthcare disruptions.
- Early detection of postnatal depression is vital for the health of both mothers and infants. Timely intervention can prevent long-term psychological issues and improve overall outcomes.
- Objectives:
  - Develop a predictive model for postnatal depression.
  - Identify key predictors from demographic, mental health, and perceived threat data.
  - Support healthcare interventions by providing actionable insights.
  - Enhance maternal and infant health through early detection and support.

Problem Statement

- Increased Stress During Pregnancy Due to COVID-19:
  - The pandemic has significantly elevated stress levels in pregnant individuals, contributing to mental health challenges.
- Goal of Predicting Postnatal Depression:
  - Develop a model to predict postnatal depression using the Edinburgh Postnatal Depression Scale (EPDS), incorporating pandemic-specific stress factors.
- Importance of Early Identification:
  - Early identification of at-risk individuals is crucial for timely intervention, reducing the risk of long-term mental health issues for both mothers and infants.

Proposed Work

- Data Cleaning: Handling missing values and converting categorical variables.
- EDA: Visualizing data relationships and identifying correlations.
- Feature Engineering: Selecting relevant features.
- Modeling: Training and evaluating regression models.
- Evaluation: Using MSE and R^2 metrics.
- Feature Importance: Identifying key features using the Random Forest model.

Evaluation Plan

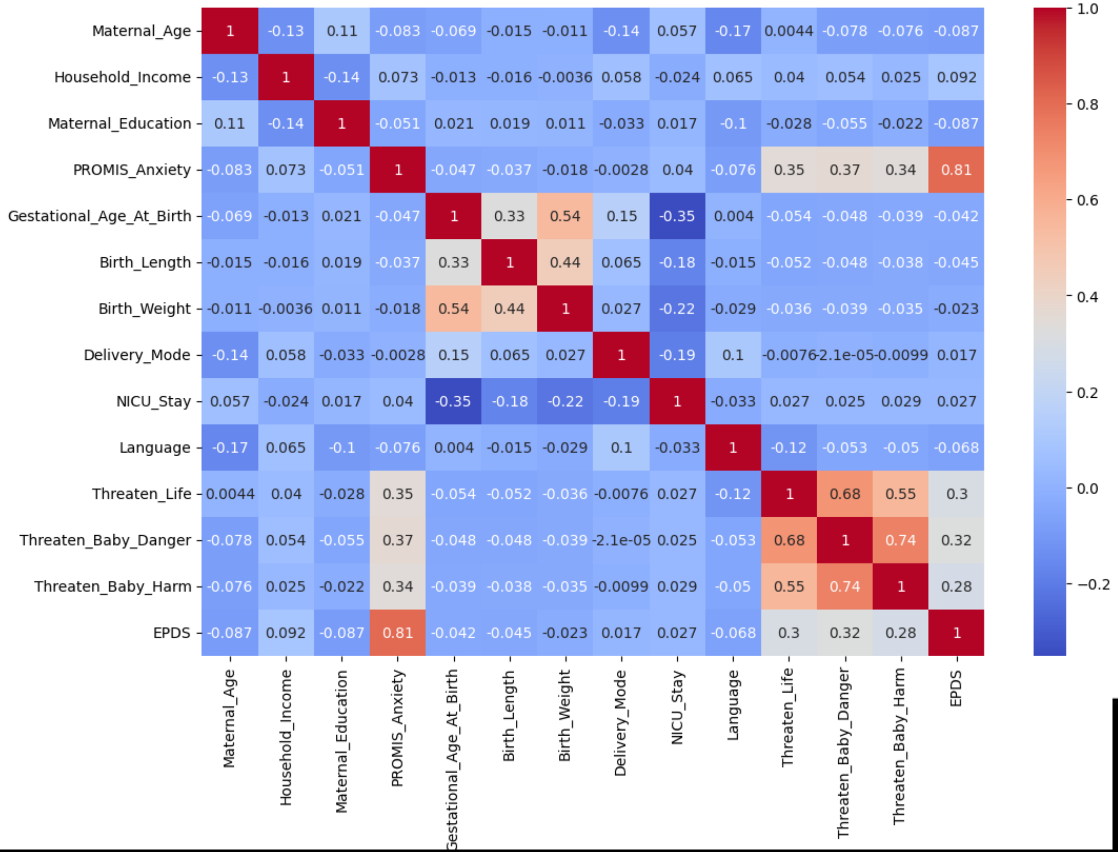
- Metrics for assessing model performance: MSE and R^2.
- Criteria for success: low MSE, high R^2, and actionable insights.
- Importance of feature importance analysis for refining the model.

Timeline

- Week 3: Data Cleaning and Preparation
- Week 4: Exploratory Data Analysis and Feature Engineering
- Week 5: Model Training and Initial Evaluation
- Week 6: Model Refinement and Feature Importance Analysis
- Week 7: Final Evaluation and Report Preparation

Progress So Far

- Completed data cleaning and preprocessing.
- Performed exploratory data analysis (EDA) and visualized relationships.
- Trained and evaluated Linear Regression and Random Forest models.
- Assessed model performance using MSE and R^2.



DATA SCIENCE
UNIVERSITY OF COLORADO **BOULDER**

Changes Made

- Updated feature selection based on initial findings from EDA.
- Improved data preprocessing steps to handle missing values more effectively.
- Enhanced model evaluation metrics to better capture performance.

References

- Gao, W., Jalal, Z., Taylor, B. K., Qian, H., Reichert, A. R., & Blank, P. R. (2023). The impact of COVID-19 pandemic on mental health in pregnant individuals. The Lancet Regional Health – Europe, 24. https://doi.org/10.1016/j.lanepe.2023.100473

- Huang, Y., Alvernaz, S., Kim, S. J., Maki, P., Dai, Y., & Peñalver Bernabé, B. (2023). Predicting prenatal depression and assessing model bias using machine learning models. medRxiv. https://doi.org/10.1101/2023.07.17.23292587