

Predicting Postnatal Depression During the COVID-19 Pandemic

Firstname Lastname
CU MSDS
University of Colorado Boulder
Boulder, Colorado, USA
fname.lname@cu.edu

Abstract

This project aims to develop a predictive model for identifying individuals at high risk of postnatal depression using data collected from pregnant individuals during the COVID-19 pandemic. The data includes demographic information, mental health scores, and perceived threat levels. By leveraging data mining techniques, we aim to provide valuable insights that can inform healthcare providers and support timely interventions.

Introduction

The COVID-19 pandemic has been a significant stressor globally, particularly for pregnant individuals. Increased stress during pregnancy can adversely affect both the mother and the infant, potentially leading to postnatal depression. This project utilizes data from the Pregnancy during the COVID-19 Pandemic (PdP) project to build a model predicting the Edinburgh Postnatal Depression Scale (EPDS) scores. The goal is to identify key factors contributing to postnatal depression and provide a tool for early identification and intervention.

Related Work

Previous studies have highlighted the impact of stress during pregnancy on maternal mental health. Research has shown that higher anxiety and perceived threats can increase the risk of postnatal depression. Various models have been developed to predict postnatal depression, but few have focused on the unique stressors presented by the COVID-19 pandemic. This project builds on existing work by incorporating COVID-19-specific stress factors into the predictive model.

Proposed Work

The proposed work involves several key steps:

1. Data Cleaning: Handling missing values and converting categorical variables to numerical form.
2. Exploratory Data Analysis (EDA): Visualizing data relationships and identifying potential correlations.
3. Feature Engineering: Selecting relevant features such as maternal age, household income, education, anxiety scores, and perceived threat levels.
4. Modeling: Training and evaluating regression models (Linear Regression and Random Forest Regressor) to predict EPDS scores.
5. Evaluation: Using Mean Squared Error (MSE) and R-squared (R^2) to evaluate model performance.
6. Feature Importance: Identifying key features contributing to postnatal depression using the Random Forest model.

Evaluation Plan

The models will be evaluated based on their predictive accuracy and interpretability. MSE and R^2 will be the primary metrics for assessing model performance. Additionally, feature importance analysis will provide insights into the most significant predictors of postnatal depression, helping to refine the model and guide future research.

Success Criteria

Success for this project will be defined by:

Achieving a low MSE and a high R^2 in predicting EPDS scores.

Identifying significant predictors of postnatal depression.

Providing actionable insights that can inform healthcare interventions.

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

REFERENCES

- [1] 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>
- [2] 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>