Predicting Poverty Using Demographic and Socioeconomic Features: A U.S. Based Machine Learning Study

Zhang, Qiaorui (School: Newton North High School)

Poverty is a pervasive issue worldwide, and poverty prediction is crucial for developing interventions that alleviate hardship and reduce cyclical inequality. This project focuses on poverty analysis in the United States: using supervised machine learning, I developed a more efficient and scalable alternative to conventional poverty assessment methods. Five binary classification models were trained on publicly available data from the 2022 American Community Survey Public Use Microdata Sample. The Extreme Gradient Boosting (XGBoost) model performed best, achieving an accuracy of 88.87% and an F1-score of 88.43%. Results identified employment status, educational attainment, mobility status, and relationship to householder as the most predictive features. While centered on American data, the methodology is scalable and can be applied to poverty analysis in other countries using locally relevant datasets.

Awards Won: