USING AI TO PREDICT AREAS OF HIGH UNMET HEALTH NEEDS AMONG ADULTS WITH DISABILITIES

ABSTRACT

Access to healthcare remains a critical challenge for adults with disabilities, often resulting in unmet health needs and disparities in care. This project aims to leverage machine learning techniques to predict regions with high levels of unmet health needs by analyzing indicators from the Disability and Health Data System (DHDS) provided by the CDC.

1 INTRODUCTION

Adults with disabilities often face significant barriers to accessing healthcare services, including geographic isolation, socioeconomic challenges, and systemic inequities in resource allocation. These disparities can lead to unmet health needs, poorer health outcomes, and increased vulnerability to chronic conditions. In order to address populations in need, public health agencies and policymakers require reliable, data-driven tools. This project will use data from the Disability and Health Data System (DHDS) maintained by the CDC. The DHDS provides state level indicators related to health behaviors, access to care, and social determinants of health for adults with disabilities. However, these indicators are typically used for descriptive reporting rather than predictive modeling. By applying machine learning techniques to DHDS data, this project seeks to enable proactive interventions and more equitable distribution of healthcare resources.

2 BASELINE OR INITIAL ANALYSIS

- Start by exploring the DHDS dataset: look at basic statistics, distributions, and trends across states.
- Clean the data: handle missing values, standardize formats, and make sure indicators are comparable.
- Run simple correlation checks to see which factors (like income, insurance coverage, or disability prevalence) are most linked to unmet health needs.
- Create visuals (heatmaps, bar charts, maps) to show where disparities are most visible

3 FINAL ANALYSIS

- Build predictive models (Random Forest, Gradient Boosting, Logistic Regression) to forecast which regions are most at risk.
- Use feature importance tools (like SHAP values) to explain which indicators drive the predictions.
- Check for fairness: make sure the model doesn't unintentionally disadvantage certain groups.
- Present results with visualizations: interactive maps comparing actual vs. predicted needs, plus clear charts that policymakers can understand.

4 FINAL GOALS & EVALUATION

- Deliver a working machine learning pipeline with code, visuals, and a simple dashboard.
- Show how the model can help public health agencies target resources more effectively.
- Success will be measured by:
    - Accuracy of predictions
    - Clarity of visuals and explanations
    - How easy the final dashboard is for non-technical users to navigate.

5 Related Work

- Disability Data Deserts and Inclusive AI
- Fairness Evaluation Data Set: Ethical Data Collection & Responsible AI
- Data Augmentation Techniques for Transfer Learning-Based Continuous Dysarthric Speech Recognition

6 Data & Technical Requirements

- Dataset: CDC's Disability and Health Data System (DHDS), which provides state-level indicators on health behaviors, access to care, and social determinants of health.
- Tools:Python or R, with libraries like Pandas, Scikit-learn, XGBoost, SHAP, and visualization tools
- Environment: Jupyter notebooks or a cloud platform for collaboration and reproducibility.
- References:
    - CDC DHDS: https://www.cdc.gov/dhds/about/index.html
    - Research on disability data deserts and inclusive AI frameworks.