

# Explainable AI (XAI) for Diagnosis Prediction Using MEPS SDOH Data

This project uses Social Determinants of Health (SDOH) data from the 2021 Medical Expenditure Panel Survey (MEPS) to predict the presence of 23 self-reported diagnoses. The focus is not only on model performance but also on the interpretability of predictions using multiple Explainable AI (XAI) techniques.

---



## Files Included

File	Description
MEPSAnalysis.py	Main script for data processing, model training, and XAI computation
model_results.csv	Summary of model performance (F1, accuracy, precision, recall) across all diagnoses
xai_summary_normalized.csv	Normalized XAI scores across models and techniques
Column_Specs_and_Names_trimmed.csv	Cleaned variable name mappings for MEPS SDOH columns

---



## Project Highlights

- **23 binary classification models**, one for each diagnosis
  - **4 model types**: XGBoost, Neural Network, Decision Tree, Ensemble
  - **6 XAI methods**:
    - SHAP
    - LIME
    - Integrated Gradients
    - Saliency Maps
    - Permutation Importance
    - Feature Ablation
  - **Normalized attribution** scores allow direct comparison across methods
  - Analysis focuses on alignment, consistency, and outliers in explanation outputs
- 



## Data Source

- **Dataset**: MEPS 2021 Full-Year Consolidated (HC-233)

- **SDOH Definitions:** Healthy People 2030 – ODPHP
- 

## Limitations

- MEPS is a self-reported survey, lacking clinical detail or longitudinal depth
  - Only SDOH variables used—no lab values or diagnoses during modeling
  - Attribution ≠ causation: some results may reflect statistical artifacts (e.g., survey month)
- 

## Future Directions

- Add EHR-based clinical variables for validation
  - Incorporate clinician review for interpretation vetting
  - Explore causal inference with SHAP-DAGs or counterfactuals
  - Expand beyond feature attribution to interactive explanation tools
- 

## Contact

Created by **Stephan Mabry**, PhD Student, University of North Dakota  
For questions or collaboration: [LinkedIn](#)

---

## Citation

If you use or build upon this work, please cite:

**Mabry, S.** *Explainable AI for Diagnosis Prediction from MEPS SDOH Data*. University of North Dakota, May 2025.

---

## Acknowledgements

- Data provided by the Medical Expenditure Panel Survey (MEPS)
- SDOH categories based on Healthy People 2030 (ODPHP)