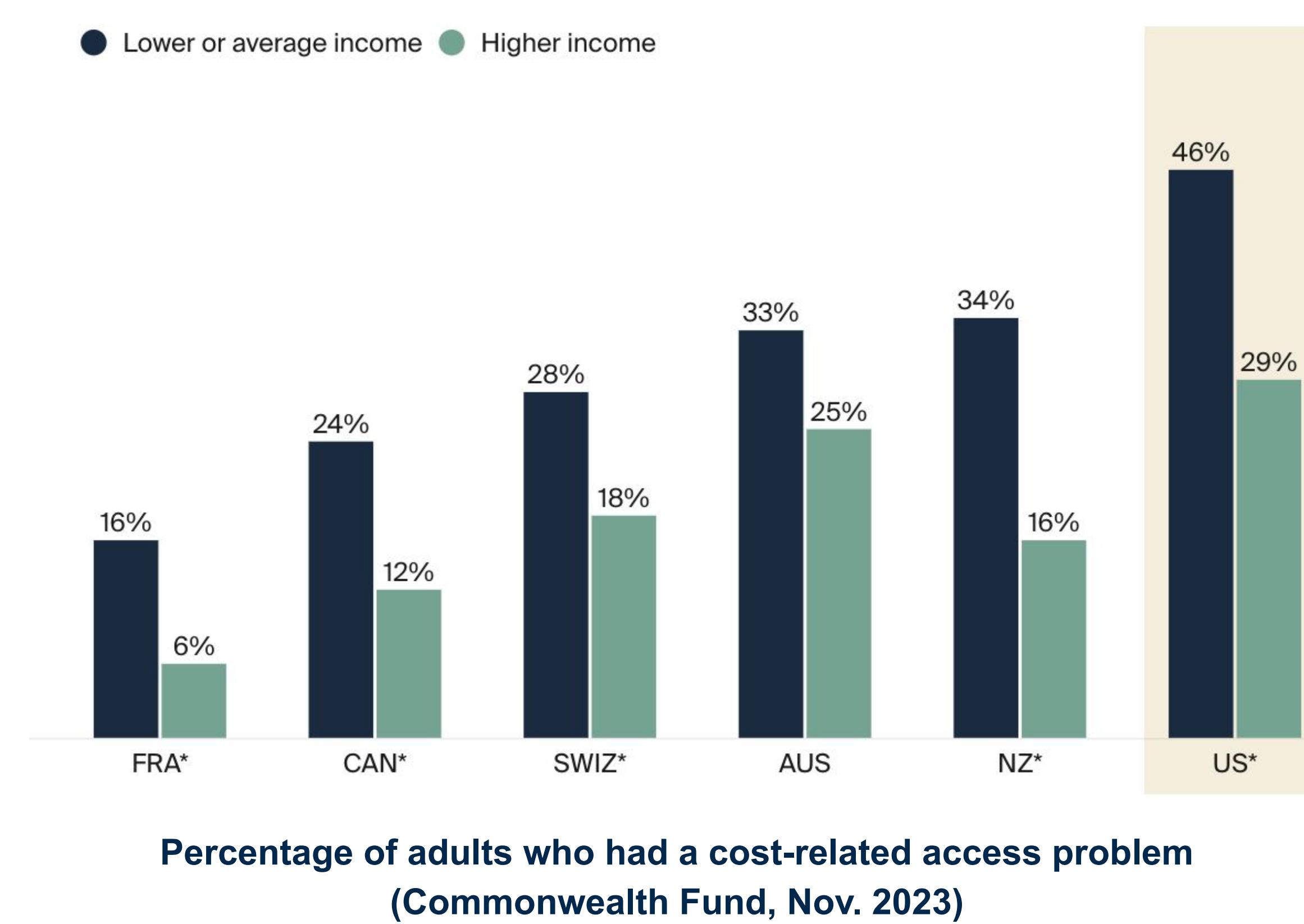


Addressing Healthcare Inequality Through Predictive Analytics

Halle Loveday, Abby Hathaway, Shao Hua Weng, Mowaniuola Olorunsola, Elizabeth Baker
 {lovedayh, hathaab, kimiweng, moloruns, lizbaker}@umich.edu

Introduction

- Healthcare inequality in the U.S. affects marginalized and low-income communities, resulting in poor health outcomes.
- Despite high healthcare spending, the U.S. ranks low among developed nations in quality of care, highlighting significant disparities.



Methodology

- Aggregated and cleaned multiple datasets, covering patient satisfaction, hospital quality, healthcare costs, infection rates, and hospice care.
- Employed feature engineering to identify key predictors of healthcare disparities, including hospital density and average copay.
- Trained a Random Forest model to classify regions by healthcare risk level (low, mild, moderate, high).

	precision	recall	f1-score	support
0	0.75	1.00	0.86	382
1	0.97	0.87	0.92	1952
2	0.91	0.91	0.91	1462
3	0.88	0.99	0.93	660
accuracy			0.91	4456
macro avg	0.88	0.94	0.91	4456
weighted avg	0.92	0.91	0.91	4456

Evaluation metrics of random forest classification model.

Social Impact

- Unequal healthcare access leads to higher mortality rates, especially in rural and socioeconomically disadvantaged areas.
- Financial barriers, like high medical costs and inadequate insurance, disproportionately impact low-income individuals.

Technical Solution

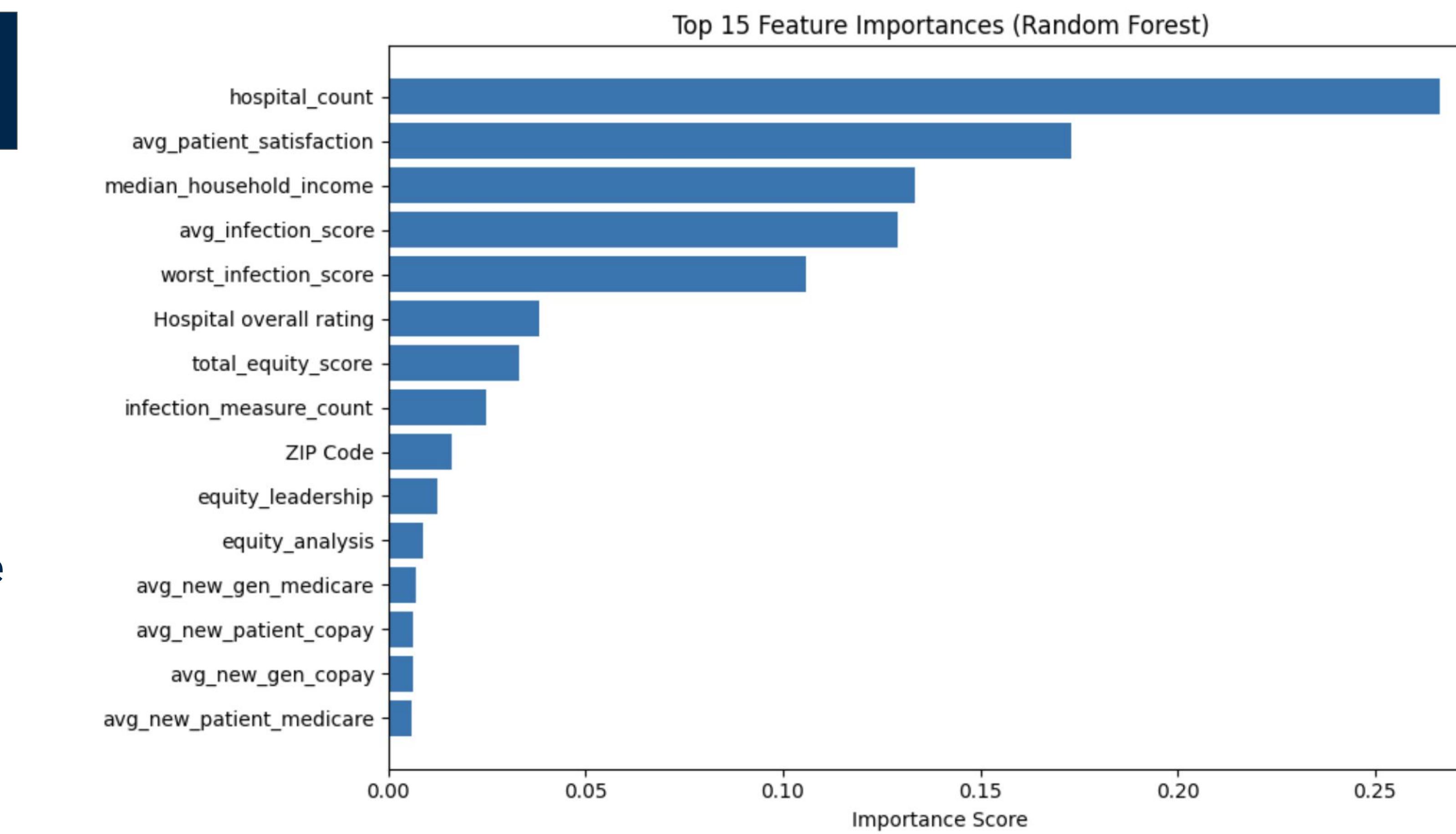
- Leverage predictive models trained on diverse healthcare datasets—patient surveys, financial reports, hospital effectiveness metrics, infection rates, and hospice care quality—to quantify and predict ZIP-level healthcare access.
- Provide an interactive U.S. dashboard for users to visualize and identify high-risk ZIP codes with significant healthcare disparities.

Datasets

- HRSA facility locations: geo-coordinates for hospital and clinic access.
- Patient satisfaction and hospital quality metrics (HCAHPS-Hospital).
- Healthcare equity assessments (Health_Equity_Hospital).
- Cost data from Family and General Practice records.
- Infection rates and hospice care quality information.

Experiments and Results

- Model accuracy: 91% over 4,456 samples.
- Effective identification of high-risk regions crucial for targeted interventions.
- Hospital count and socioeconomic factors emerged as top predictors.



Feature Importance Bar Chart identifying influential predictors. This supports the hypothesis that geographical location is heavily predictive of healthcare quality because rural areas typically have fewer healthcare providers in comparison with urban areas.

Future Work

- Interactive UI: allow zoom, filter, and on-click details per ZIP.
- Enhance the interactive dashboard with real-time data updates.
- Expand predictive models to include broader geographical and socioeconomic datasets.
- Benchmark against simpler models (e.g., logistic regression) for performance validation.

References

- Elizabeth J. Brown, Daniel Polksy, Cristina M. Barbu, Jeffrey W. Seymour, and David Grande. 2020. Patterns in geographic access to health care facilities across neighborhoods in the United States based on data from 2000 to 2014. *JAMA Network Open*, 3(5):e205105.
- Anisa Butler, Kyla Covington, and Bridget Parsh. 2021. Identifying and tackling racial disparities in healthcare. *Nursing* (Jenkintown, Pa.), 51(9):40–43.
- Michelle M. Doty, Roosa Tikkainen, and Melinda K. Abrams. 2023. The cost of not getting care: Income disparities in the affordability of health care. Accessed: 2025-04-02.
- Garth N. Graham. 2016. Why your zip code matters more than your genetic code: Promoting healthy outcomes from mother to child. *Breastfeeding Medicine*, 11:396–397. Epub 2016 Aug 11.
- Jessica Ochs and others. 2024. 9.4 Geographical Disparities. OpenStax, Houston, Texas.
- Michael G. Zywiel, Tiffany C. Liu, and Kevin J. Bozic. 2017. Value-based healthcare: The challenge of identifying and addressing low-value interventions. *Clinical Orthopaedics and Related Research*, 475(5):1305–1308.