

Team Robinhood

1. Mateeh Ullah 2. Muhammad Shahroz 3. Sana Adeel

Problem Statement

- Limited access to healthcare and education facilities in underserved regions.
- Inefficient allocation of resources leads to gaps in service delivery.
- Lack of actionable insights into facility utilization and demand.



Project Objectives

- Optimize connectivity for healthcare and education facilities.
- Predict demand to enable better resource allocation.
- Provide actionable insights for policy-makers.



Solution Overview

Our Approach:

- Use AI-driven models to predict facility demand.
- Integrate geographic, ownership, and facility-type data.
- Data is obtained form John Snow labs and World Bank.
- Apply clustering and regression techniques for prioritization.



Technical Approach

Steps in Our Workflow:

- 1. Data Preprocessing: Handling missing values, encoding, and scaling.
- 2. Feature Engineering: Spatial clustering and interaction terms.
- 3. Modeling: K-Means for clustering and Gradient Boosting and Random forest for prediction.
- 4. Visualization: Geographic scatter plots for actionable insights.
- 5. Technologies Used:
- Python (Pandas, Scikit-learn, Matplotlib).
- GitHub for collaboration and version control
- Use Mistral AI for recommendations



Results and Insights

Key Metrics:

- R² Score: Gradient Boosting: ~0.96
- Consistent cross-validation scores (0.95-0.96).

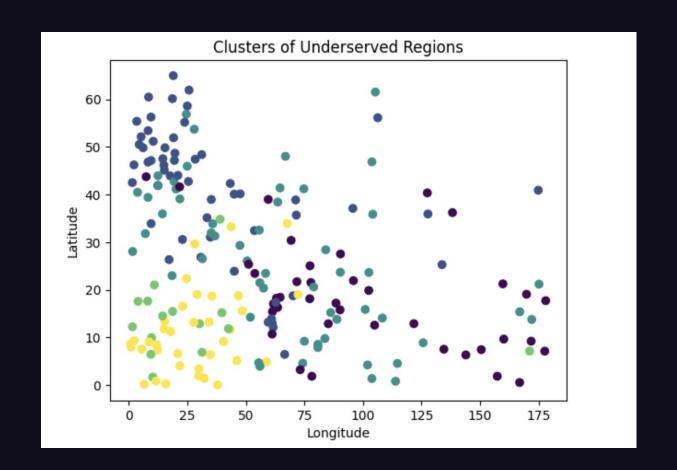
Insights:

- Facilities with the highest predicted demand.
- Regions requiring more resources.



Visualizations

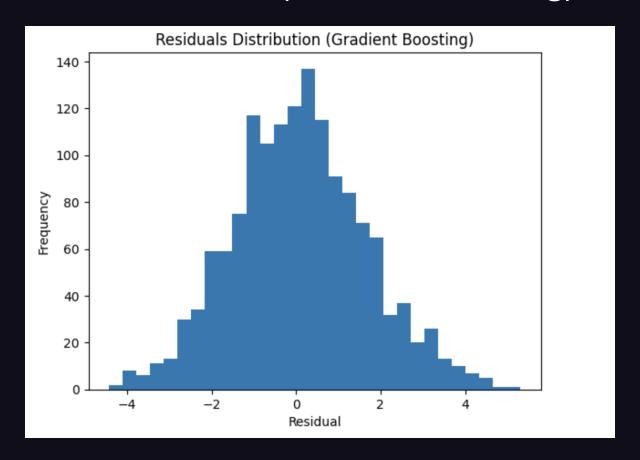
Clusters of Underserved Regions





Visualizations

Residual Distribution (Gradient Boosting)





Challenges and Learnings

Challenges Faced:

- Handling noisy and missing data.
- Initial overfitting issues with $R^2 = 1.0$.
- Incorporating diverse data sources for improved predictions.

How We Solved Them:

- Adjusted feature weights to improve geographic impact.
- Integrated Mistral AI for advanced recommendations to address resource prioritization.



Future Scope

Expansion Possibilities:

- Real-time integration with IoT and census data.
- Scaling solution to national or regional levels.
- Focused education-specific analysis for school connectivity.
- Gathering city-to-city level data for better prediction accuracy.

Next Steps:

- Collaboration with stakeholders for implementation.
- Refinement of models to support multi-region deployment.

