AI-Driven Analysis of Trolley Service Gaps in Miami

1. Introduction

Accessible public transportation is vital for urban equity and sustainability. In Miami, the trolley system offers free transit services across various neighborhoods. However, questions arise regarding the system's coverage adequacy, especially concerning key landmarks and densely populated areas. This project employs geospatial analysis and machine learning techniques to identify service gaps in Miami's trolley network and propose data-driven recommendations for service expansion.

2. Data Sources

The analysis utilizes the following datasets:

Miami Trolley Routes: GeoJSON files detailing current trolley routes.

Landmarks: GeoJSON and CSV files listing significant landmarks, including parks, schools, and community centers.

Neighborhood Boundaries: GeoJSON files defining Miami's neighborhood boundaries.

Census Tracts & Population: Shapefiles providing demographic information at the tract level.

NAIP Aerial Imagery: RGB composite images from the National Agriculture Imagery Program (used in exploratory analysis).

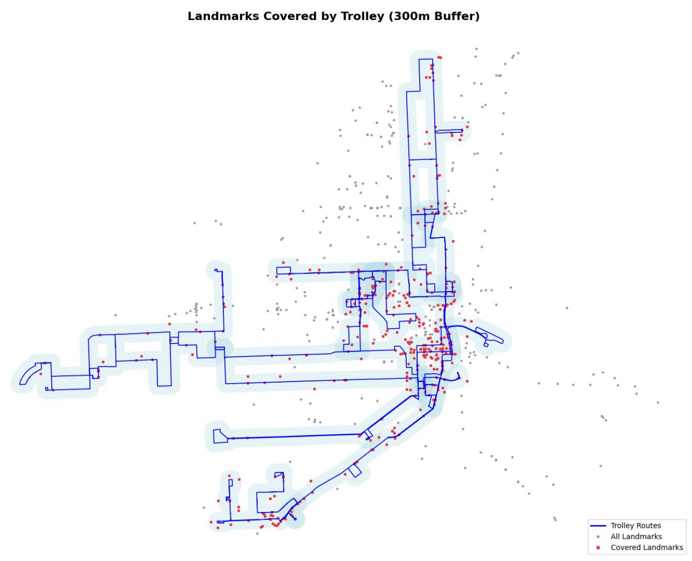
All spatial datasets were projected to EPSG:3857 for consistency and processed using Python libraries such as geopandas, scikit-learn, and matplotlib.

3. Methodology

3.1. Coverage Analysis

A 300-meter buffer was generated around existing trolley routes to simulate reasonable walking distances. Landmarks falling outside this buffer were classified as "uncovered," indicating potential service gaps.

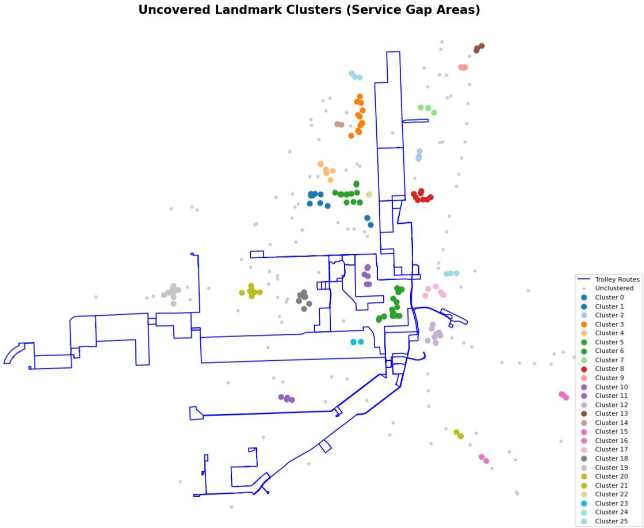
Figure 1



3.2. Clustering Uncovered Landmarks

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm was applied to the uncovered landmarks to identify clusters representing concentrated areas lacking trolley access.

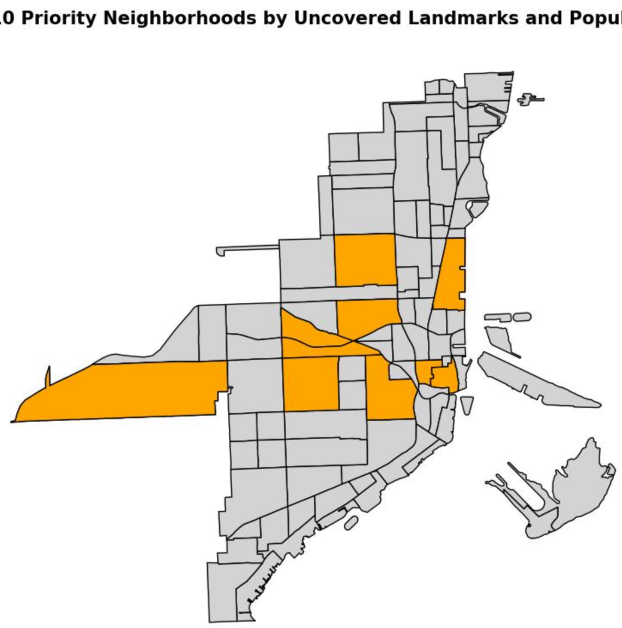
Figure 2



3.3. Neighborhood Prioritization

Uncovered landmarks were spatially joined with neighborhood boundaries to count the number of service gaps per neighborhood. This count was then weighted by neighborhood population to prioritize areas where trolley service expansion would benefit the most residents.

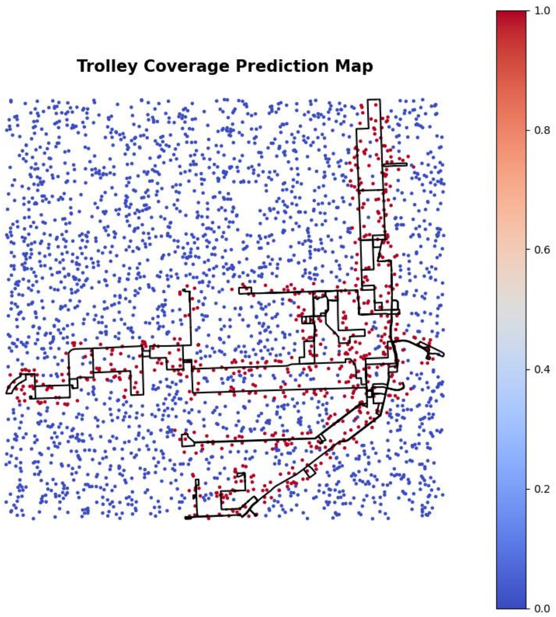
Figure 3



3.4. Predictive Modeling with Random Forest

A Random Forest classifier was trained using the coordinates of landmarks to predict coverage status. The model was then applied to a grid of randomly sampled points across Miami to estimate the probability of trolley coverage citywide.

Figure 4



4. Results

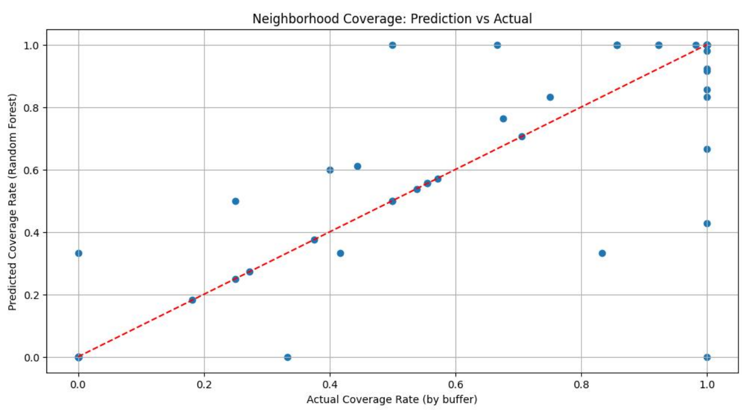
Service Gaps: Several neighborhoods, notably Allapattah and Little Haiti, exhibit high concentrations of uncovered landmarks, suggesting significant service gaps.

Cluster Analysis: DBSCAN identified multiple clusters of uncovered landmarks, highlighting specific areas where trolley service is lacking.

Neighborhood Prioritization: When accounting for population, neighborhoods like Allapattah and Little Haiti emerge as high-priority areas for service expansion.

Predictive Modeling: The Random Forest model achieved satisfactory accuracy, and its predictions align with known coverage areas, validating its utility in identifying potential service gaps.

Figure 5



5. Summary and Reflection

This project demonstrates how geospatial data, clustering techniques, and machine learning can work together to inform urban mobility planning. By identifying service gaps and ranking underserved neighborhoods, we provide actionable insights for transit equity improvement.

5.1. Limitations

The Random Forest used only spatial coordinates; richer features (e.g., built environment data) could enhance predictions.

NAIP-based land cover classification was omitted due to resolution challenges and limited training sample diversity.

5.2. Future Work

Integrate slope or walkability scores for more realistic accessibility modeling.

Explore deep learning methods for image classification and route optimization.

Collaborate with community groups to validate uncovered landmark needs.

Appendix

Figure 6

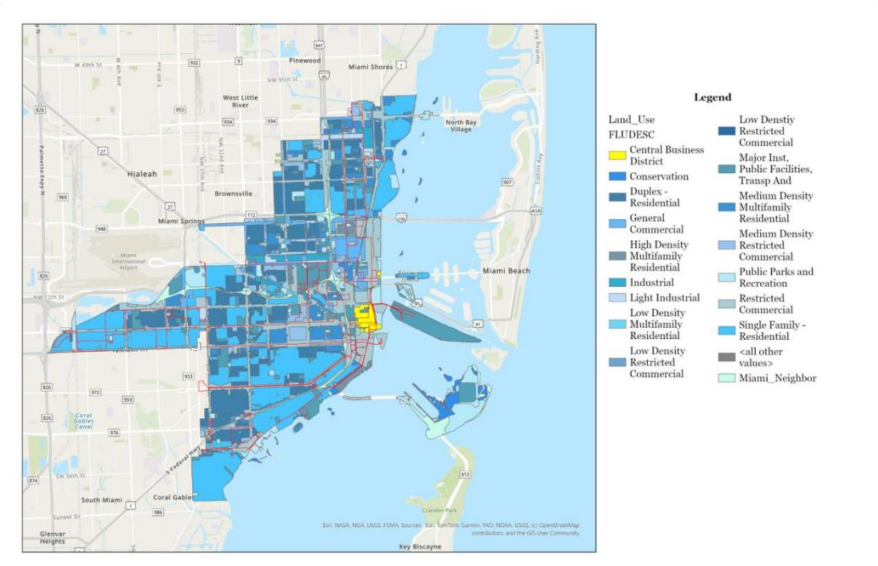


Figure 7