Predictive Analytics for Charging Infrastructure Development in New York

RAHUL SAI BASINA, LINGA REDDY GUDISHA, OLUWAFISAYO THEOPHILUS

Outline

Research Problem and Its Significance

Research design/framework

Evaluation

Visualization

Takeaway Messages and Conclusion

Research Problem and Its **Significance**

Focused Research Problem

Growing demand for Electric Vehicles (EVs)

Insufficient charging infrastructure is a significant barrier

Planning the future of EV charging infrastructure is crucial for sustainability

Significance of the Research

Increased EV adoption is essential for reducing greenhouse gas emissions

Inefficient infrastructure hinders EV market growth

Optimized placement and energy usage prediction will improve infrastructure efficiency

Research design/framework

- **Step 1** — Problem Definition
- **Step 2** — Data Preparation
- **Step 3** — Model Design
- **Step 4** — Model Training
- **Step 5** — Model Evaluation
- **Step 6** — Visualization

Evaluation

Electric Vehicles Dataset:

- **Rows**: 131,546, **Columns**: 20
- **Features**:
  - Includes variables such as `VIN`, `City`, `State`, `Zip`, `Model Year`, `Fuel Type`, `Reg Valid Date`, and `Reg Expiration Date`.
  - This dataset primarily tracks details about electric vehicles, including registration information, vehicle specifications, and geographic details.

EV Charging Stations Dataset:

- **Rows**: 1,000, **Columns**: 31
- **Features**:
  - Includes variables such as `station_name`, `city`, `state`, `zip`, `latitude`, `longitude`, `ev_level2_evse_num`, `ev_dc_fast_count`, `open_date`, and `ev_connector_types`.
  - This dataset provides information about charging station locations, their operational status, charger types, and station size details.

	Model	MAE	MSE	RMSE	R^2
0	LSTM	0.361309	0.184381	0.429396	-3.675525
1	Random Forest	0.342049	0.155979	0.394942	-2.955312
2	XGBoost	0.321304	0.142394	0.377351	-2.610825

2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 12ms/step

	Model	MAE	MSE	RMSE	R^2
0	LSTM	0.229488	0.091749	0.302901	-1.413115
1	Random Forest	0.174434	0.049850	0.223272	-0.311123
2	XGBoost	0.217172	0.083686	0.289286	-1.201056

Station Counts Prediction:

Input: Historical Station Count, EV Registrations, Temporal Features, Location Data, Station Properties

Output: Predicted Station CountsExample: Monthly predictions for the next 12 months.

Model Performance: Metrics comparison (MAE, RMSE, R^2) for LSTM, ARIMA, Random Forest, and XGBoost

Energy Usage Prediction:

Input: tation_size: Total chargers (Level 2 and DC fast), open_date: Operational start date for temporal patterns, ev_level2_evse_num and ev_dc_fast_count. emporal Features: month, year for seasonality.

```
Test Loss: 0.024448303505778313
Test MAE: 0.09318030625581741
```

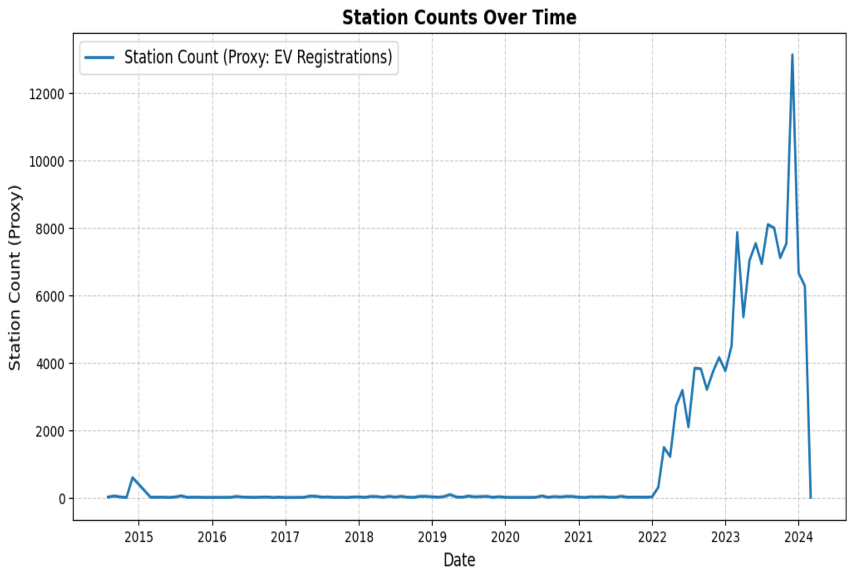
Output: Predicted Energy Usage, Model Performance Metrics

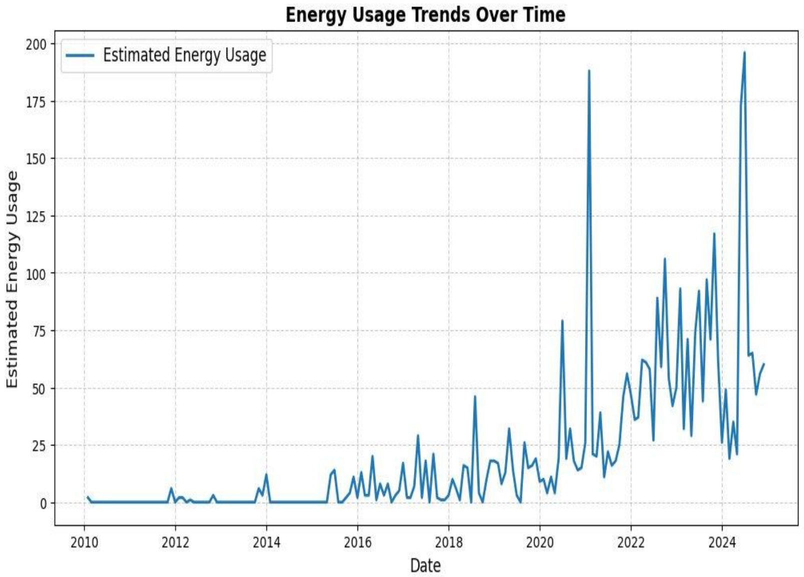
Optimal Station Locations

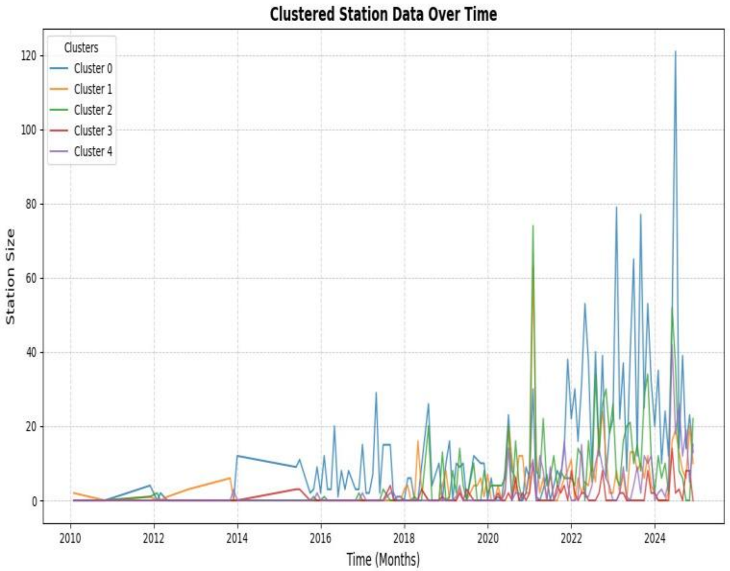
Input: Geographic Data, Station Attributes, Temporal Data, Cluster Assignment:

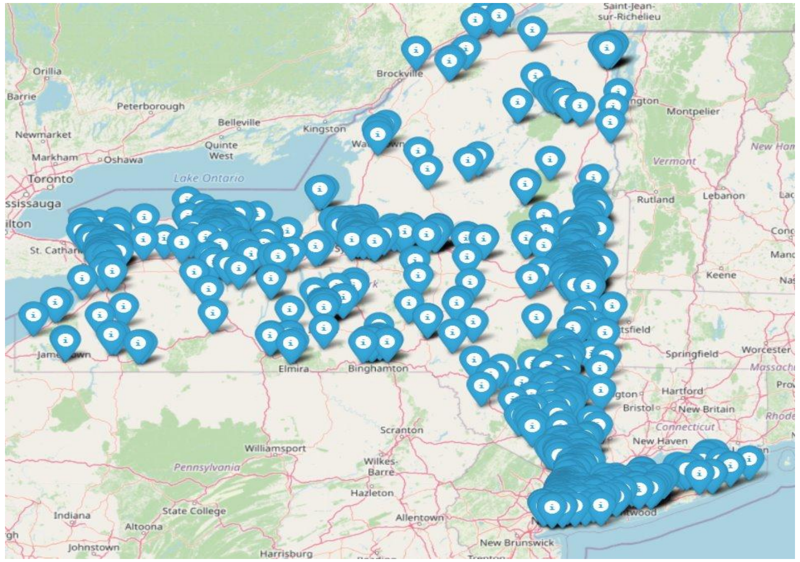
latitude and longitude: Location of charging stations, ev_level2_evse_num: Number of Level 2 chargers at the station. ev_dc_fast_count: Number of DC fast chargers at the station, open_date: Date when the station became operational.

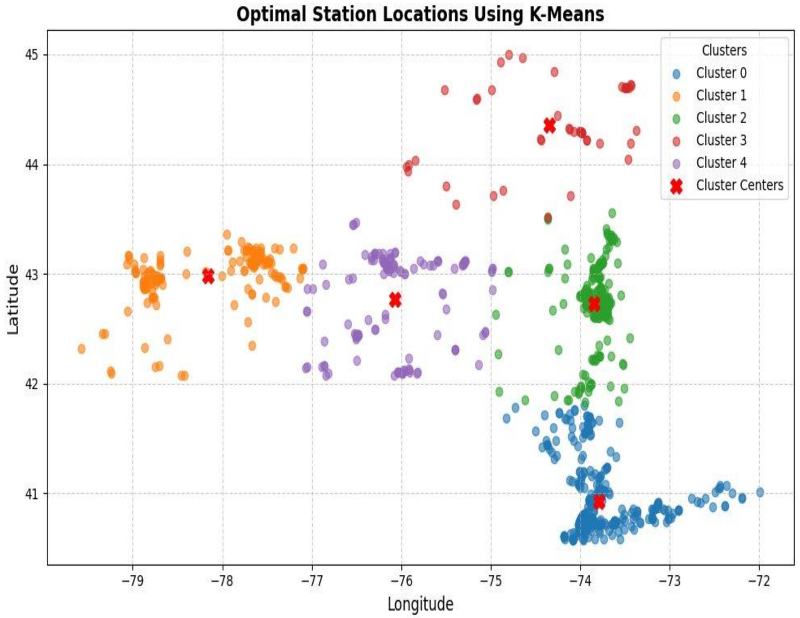
Output for Optimal Station Locations: Cluster Assignments, Cluster Centers, Station Demand Trends

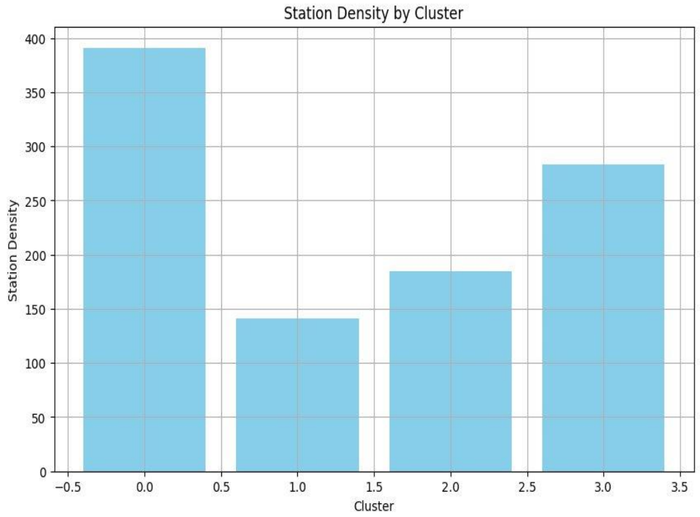

Station Counts Over Time
Station Count (Proxy: EV Registrations)
Station Count (Proxy)
Date


Energy Usage Trends Over Time
Estimated Energy Usage
Estimated Energy Usage
Date


Clustered Station Data Over Time
Clusters
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Station Size
Time (Months)


Orillia
Barrie
Peterborough
Belleville
Kingston
Brockville
Newmarket
Markham
Oshawa
Toronto
Lake Ontario
St. Catharines
Montpelier
Vermont
Rutland
Lebanon
Keene
Springfield
Worcester
Hartford
New Britain
Connecticut
New Haven
Elmira
Binghamton
Scranton
Williamsport
Pennsylvania
Wilkes-Barre
Hazleton
Allentown
State College
Altoona
Indiana
Johnstown
Harrisburg
New Brunswick


Optimal Station Locations Using K-Means
Clusters
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster Centers
Latitude
Longitude


Station Density by Cluster
Station Density
Cluster

Takeaway Message and Conclusion

During this analysis, significant insights were gained across three key areas.

Station Counts Prediction: LSTM models effectively captured temporal trends in station growth.

Energy Usage Prediction: Analyzed station size and usage patterns to model rising energy demands.

Optimal Station Locations: K-Means clustering identified high-demand areas for future EV charging stations.

Conclusion:

This study highlights how predictive analytics can enhance EV charging infrastructure in New York. Using LSTM and K-Means, we forecasted station counts, energy usage, and optimal locations. Findings emphasize data-driven planning for efficient deployment, energy management, and equitable access. These insights support policymakers and stakeholders in accelerating EV adoption, optimizing resources, and promoting sustainability in transportation and energy systems.

Thank you