Investigating Electric Vehicle Adoption Using Correlation and Prediction Analyses

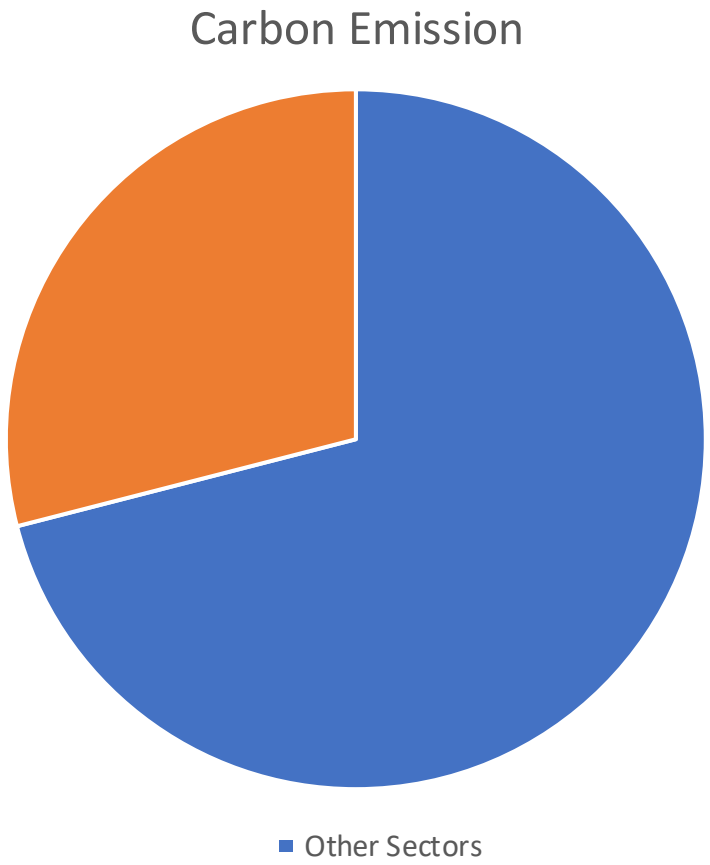
Fahad S. Alrasheedi and Hesham H. Ali
Department of Computer Science
College of Information Science and technology
University of Nebraska Omaha
hali@unomaha.edu

Presentation Agenda

- Introduction :
  - EV importance, EV Adoption barriers, Research Goals
- Methodology:
  - Data and Analysis Tools
  - Correlation Network Model
- Results:
  - Correlation Network Results
  - Statistical Prediction Results
- Conclusions
- Future Work

Introduction – EV importance

✓Rising energy consumption worsens climate change.
✓Transportation accounts for 29% of carbon emissions; light-duty vehicles contribute 59%.
✓Electric Vehicles (EVs) offer zero tailpipe emissions, helping reduce overall carbon emissions.



Introduction – EV Adoption Barriers

- EV Adoption faces several barriers such as high prices, range anxiety, lack of charging infrastructure, and others.
- Significant management efforts have been made to transition transportation toward electrification. U.S. authorities manage EV adoption through tax incentives, credits, and infrastructure programs.

Research Goals

- Despite such management efforts to promote EV usage, the EV adoption remains a complex issue that requires in-depth investigation to provide insights into how adoption rates can be increased based on the characteristics of targeted populations.
- Therefore, an essential first step in addressing the complexity of EV adoption is to examine how different counties across various states in the U.S. are working to accelerate EV adoption.
- We conducted two analyses as part of this effort: one using Graph Theory and the other employing statistical prediction analysis.

Methodology – Data Sets

- EV data : from Atlas Hub, and we found 137 counties from six states.

No.	State	No. of Counties
1	Colorado	20
2	Minnesota	3
3	Montana	2
4	New York	48
6	Texas	30
7	Virginia	34
8	Total	137

- Charging infrastructure: from department of energy, the number of stations in 2022 (one-year lag)
- Demographic data: from Census Bureau of the United States, it has 58 features categorized into different groups: Economy, Health, education, transportation, business , geography, Income, Personal (Age, Sex, Race), Housing, Families Living Arrangement, Computer and Internet Use.

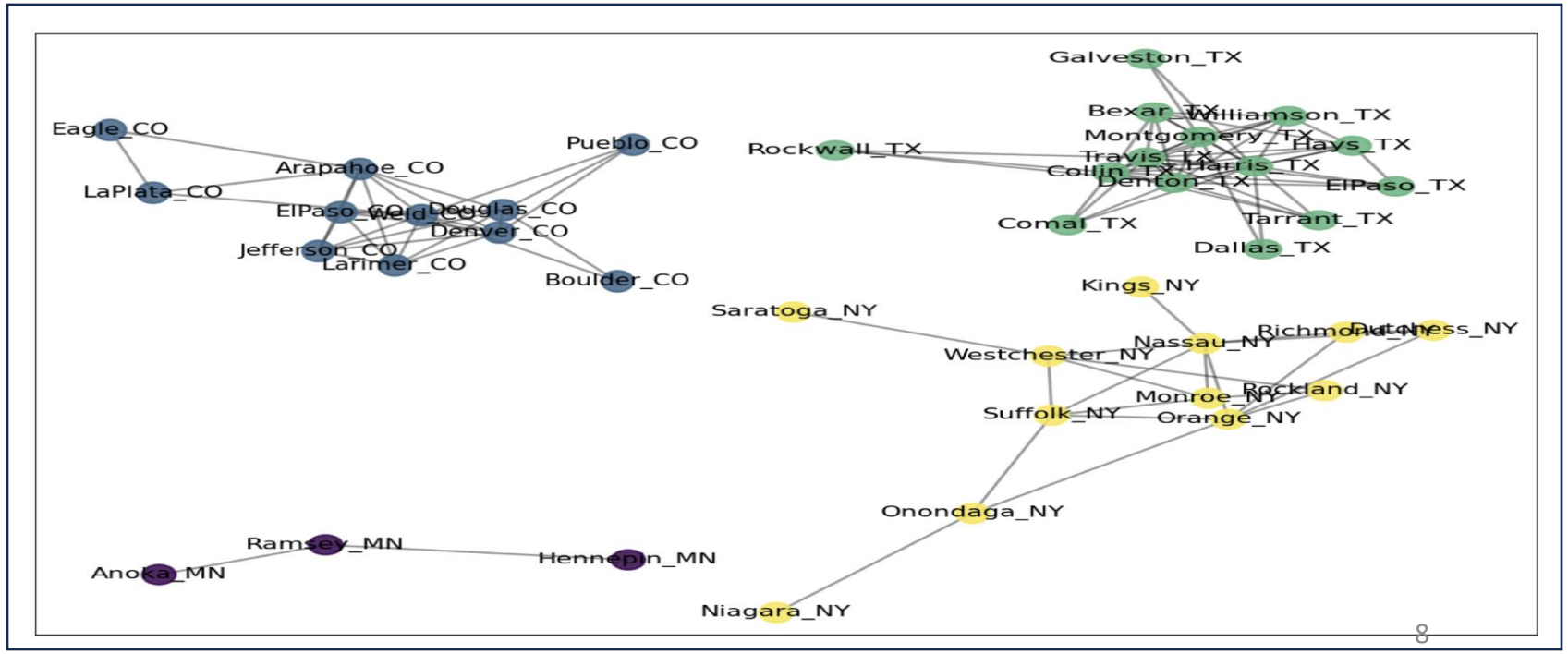
Methodology – Analysis tools

- The Correlation Network Method :
  1. We computed the month-to-month growth rates for each county
  2. Next, since our data are not perfectly linear, we calculate the Spearman correlation between counties, resulting in a 137 × 137 correlation matrix.
  3. Using this matrix, we created a correlation network where nodes represent counties and edges represent correlations that exceed a specified threshold (0.72).
- Statistical Prediction Analysis :
  1. The target variable was the number of EVs in 2023.
  2. The features included demographic data and the cumulative number of charging stations as of 2022, reflecting a one-year effect of charging stations on the number of EVs in 2023.
  3. The statistical prediction models used included Linear Regression, Random Forest, Gradient Boosting, Decision Tree, Elastic Net, Lasso, and Ridge.

EV Correlation Network

- There are 40 out of 137 counties in the network.
- Counties from the same state in one cluster.
- The counties from the same state are also neighbors.

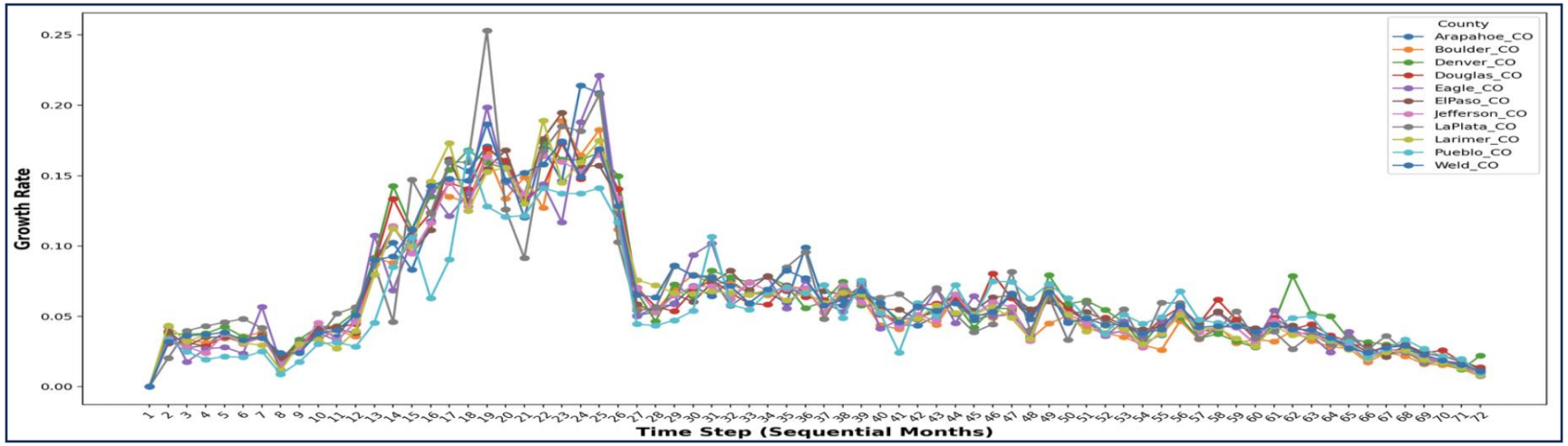
Cluster Code	No. of Counties	States
Cluster 1	11	Colorado
Cluster 2	12	New York
Cluster 3	14	Texas
Cluster 4	3	Minnesota



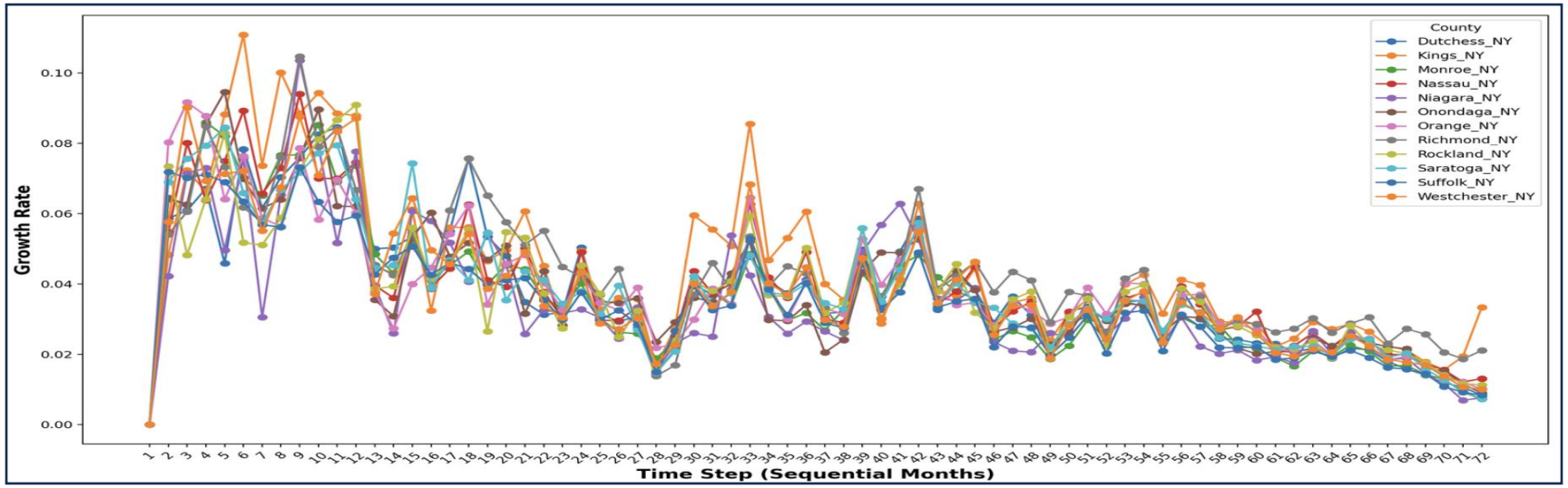
Results – Decrements in EV growth rate

- Correlated counties in cluster 1 Colorado and cluster 2 (New York).

Colorado



New York



Results - Statistical Prediction Models

The Gradient Boosting outperforms. The highest R-squared and lowest MSRE.

Model	MSRE	R-Squared
Linear Regression	101850027.5567	0.5377
Random Forest	69206289.8	0.6859
Gradient Boosting	**58108466.13**	**0.7362**
Decision Tree	141672157.654	0.357
Elastic Net	122403797.77	0.444
Lasso	98257697.207	0.554
Ridge	96591486.7415	0.5616

Results - Statistical Prediction Models

- Federal Information Processing Standards (FIPS Code)

Feature	Group	Importance
Nonminority-owned employer firms, Reference year 2017	Business	5.26e-01
Living in same house 1 year ago, percent of persons age 1 year+, 2018-2022	Families & Living Arrangements	2.02e-01
Station Counts	Station data	4.96e-2
Total annual payroll	Business	4.75e-2
Men-owned employer firms, Reference year 2017	Business	4.60e-2
Women-owned employer firms, Reference year 2017	Business	1.66e-2
FIPS Code	Geography	1.05e-02

Conclusions

- We d the EV adoption using two-pronged analysis: Correlation Network and Statistical prediction models.
- Main finding is that the geographical characteristics of counties, such as the state in which a county is located and its neighboring counties, play a significant role in EV adoption.
- This is evident in the correlation network, where counties within the same state exhibit similar EV growth rate patterns, and in the prediction model, where the FIPS feature (a geographical identifier) emerges as one of the most significant predictors in the best-performing model.
- We identified two clusters with declining EV growth rates, highlighting the need for further investigation into their underlying causes.

Future Work

- We identified two clusters with declining EV growth rates, highlighting the need for further investigation into their underlying causes.
- Future research could enhance prediction models by:
  1. incorporating political, environmental, and climatic factors.
  2. expanding the dataset to cover more counties across states.
- Lastly, distinguishing between different types of EVs in future adoption analyses may yield valuable insights.