Broadband Mapus Cellular Network Topology | Applying Data Mining and Machine Learning

FIU | Engineering & Computing

Shawna-Lee Pommells, Daidreuna Donaldson, Erika Noel, Alliantha Dervil, Dr. Herman Watson ,Dr. Diedra Hodges | Florida International University

NSF | SPECTRUMX

BACKGROUND

This research examines the cellular network performance of three major cell service providers (AT&T, T-Mobile, and Verizon) across different socioeconomic areas. By utilizing data mining and machine learning techniques, the study aims to determine if areas of similar income levels receive equitable access to quality network service. The project leverages various performance metrics to analyze the cellular network quality comprehensively.

RESEARCH QUESTIONS

1.How does the network quality compare among AT&T, T-Mobile, and Verizon across various socioeconomic regions?
2.Can a machine learning model predict the need for additional cellular infrastructure based on existing network quality data?
3.Do regions with higher economic status receive better network coverage compared to less affluent areas?

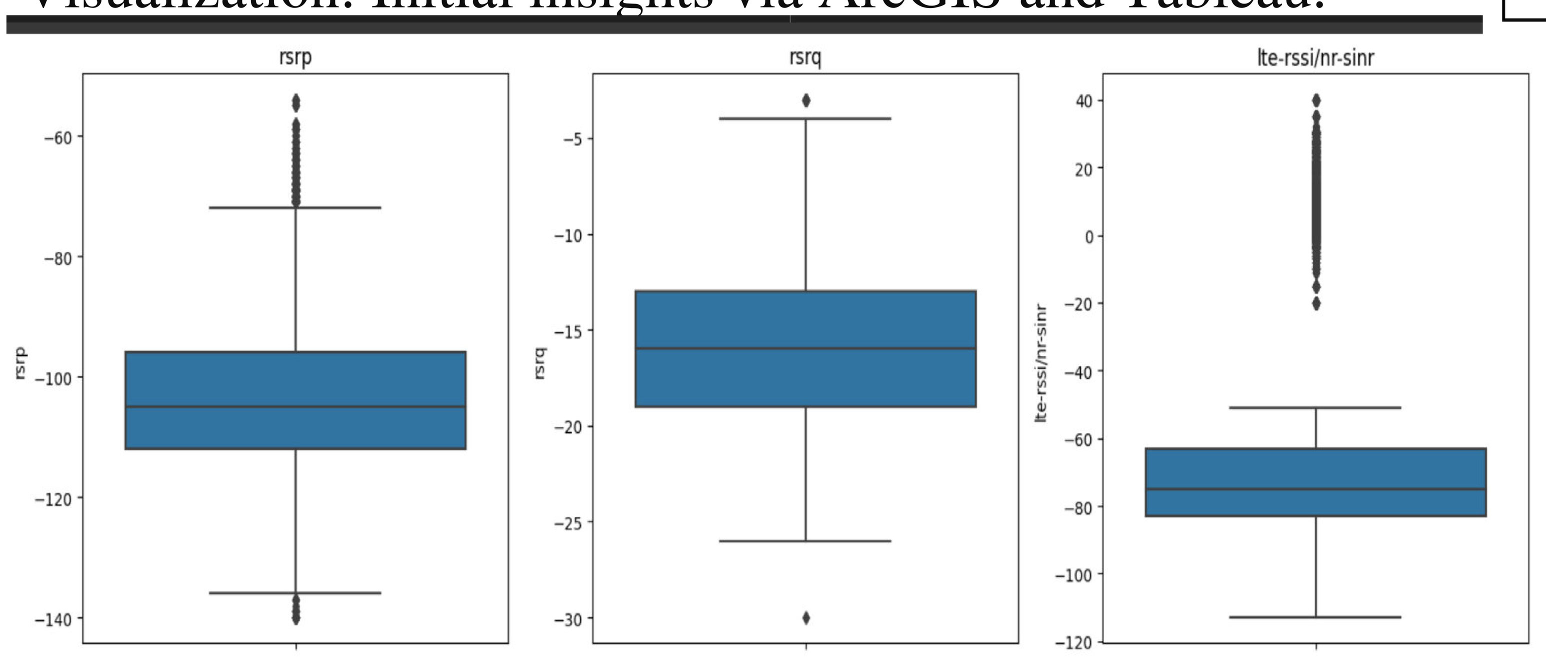
METHODS AND MATERIALS

Data Collection:

•Tools Used: FCC Speed Test, Macrodroid, Sigcap apps.
•Locations: FIU campuses, Rural, Urban, Suburban, City, Airport zones.

Data Preprocessing and Analysis:

•Preprocessing: Cleaning, normalization, and outlier handling.
•Visualization: Initial insights via ArcGIS and Tableau.



Chosen Metrics for Model:
RSRP, RSRQ, LTE-RSSI/NR-SINR, Network Quality, Longitude, Latitude

Machine Learning Models	Definition
LASSO Regression	Minimizes coefficient size to enhance model simplicity and prevent overfitting.
Ridge Regression	Similar to Lasso but penalizes the square of the coefficients, promoting model stability.
Decision Tree Regressor	Splits data into branches based on feature values, handling non-linear patterns effectively.

RESULTS

Model Performance Evaluation

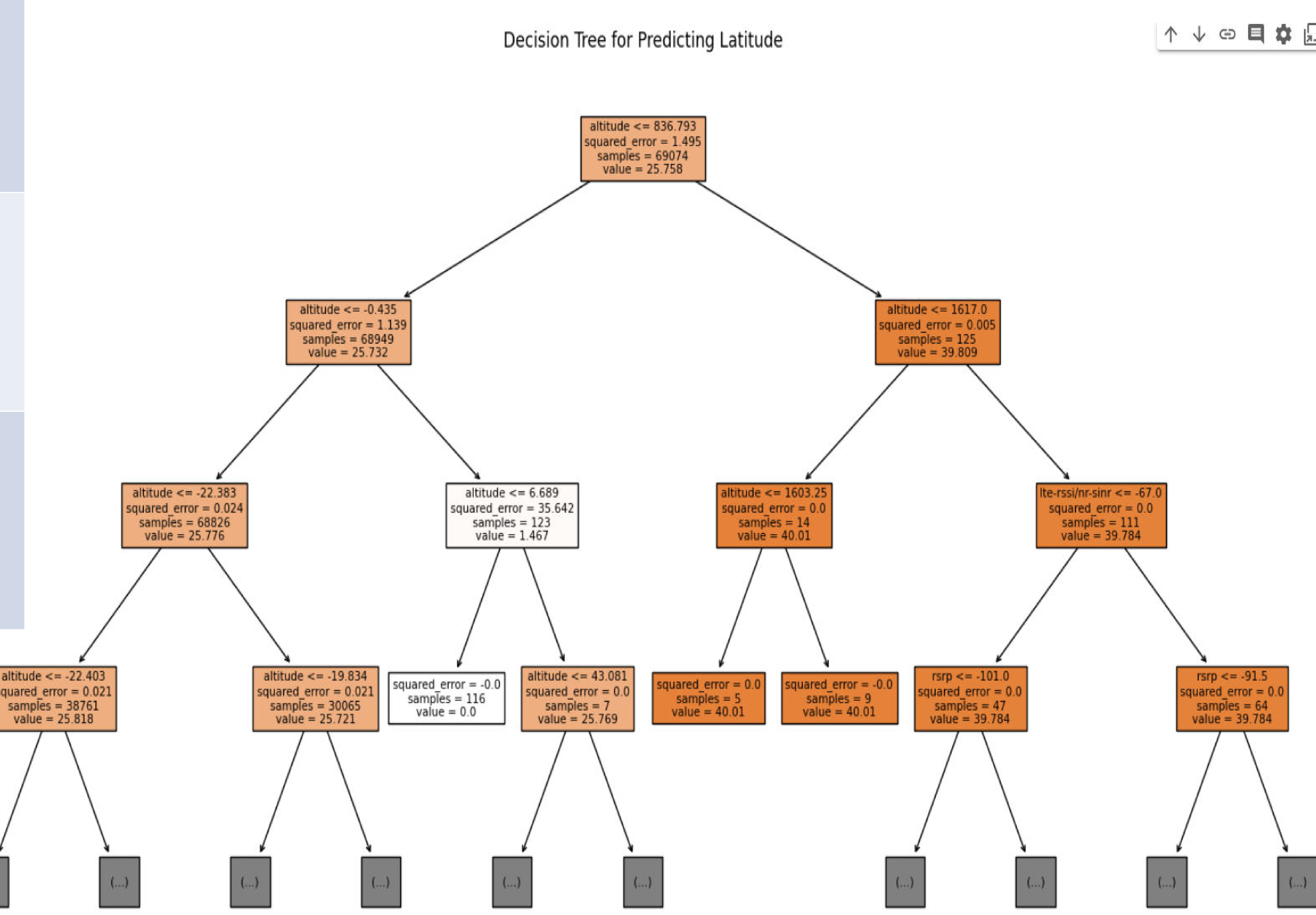
Model	Mean Squared Error	R^2 Score
DTR (Longitude)	0.004570273	0.999607165
DTR (Latitude)	0.010162263	0.993159768
Lasso (Longitude)	10.531439776	0.094775387
Lasso (Latitude)	1.117788158	0.247615399
Ridge (Longitude)	10.529471376	0.094944579
Ridge (Latitude)	1.117633468	0.247719521

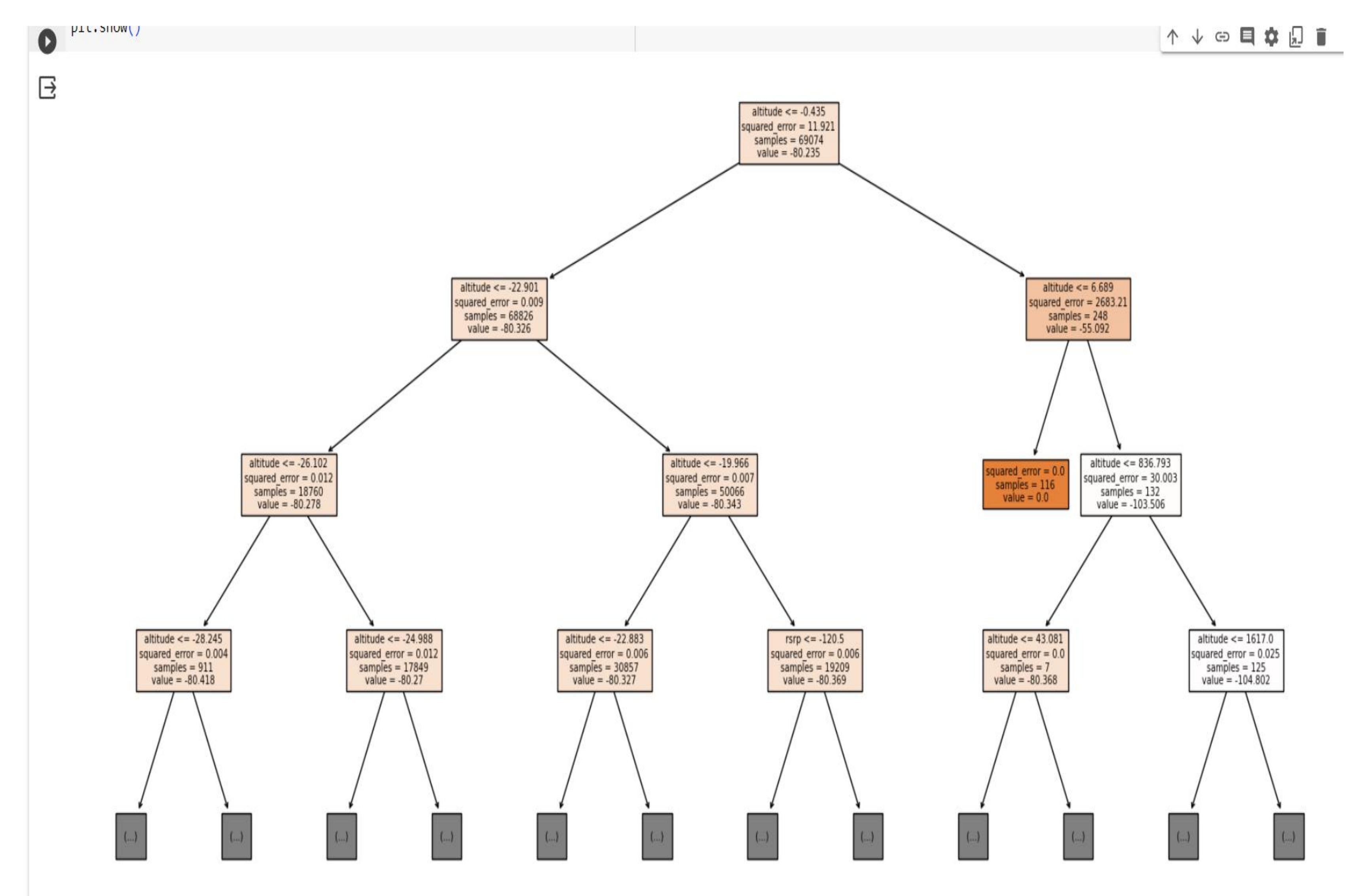
Cross-Validation R^2 Scores

Fold	Longitude	Latitude
Fold 1	0.998603	0.99000315
Fold 2	0.9990416	0.96891166
Fold 3	0.99956096	0.98310811
Fold 4	0.99746701	0.95641913
Fold 5	-0.34594171	-0.05907206
Average R^2	0.729746174	0.767873999

CONCLUSION

The Decision Tree Regressor (DTR) proved to be the most effective model for predicting cellular network performance, particularly in areas of varying socioeconomic status. By accurately identifying locations needing infrastructure improvements, the DTR highlighted the disparities in network quality—areas with higher economic status often have better coverage. This research supports the deployment of a DTR-based application that could guide the strategic placement of new cell towers and optimize network resources, ensuring equitable network access across all regions.





ACKNOWLEDGEMENTS

Our special thanks goes to Dr. Deidra Hodges for her guidance and instruction. We would also like to give thanks to Muhammad Akbar, for his technical expertise.