AI/ML FOR ELECTRIC DISTRIBUTION GRID OUTAGE ANALYSIS

STUDENTS: Laxman Balamurugan, Shiyu Chen, Andrew Nguyen, Sreeram S, Louis Serchuk, Yifei Zhao, Kaiyuan Zheng

Abstract

Traditionally, power outage management relies on preventive maintenance, robust infrastructure, and emergency responses. However, these methods often fail to prevent unexpected failures or adapt quickly to changing conditions. Integrating AI into power systems enhances these methods. AI and ML models analyze data from weather, grid usage, and equipment status to predict failures before they occur.

Project Requirements:
• Collect and preprocess outage and weather data
• Test and train the cleaned data using a variety of models
• Analyze results

Data Collection

Data was collected from three different sources:
• Power outage data was purchased from PowerOutage.us
• Weather data was collected from the National Centers for Environmental Information (NCEI) and OpenWeatherMap and combined to form a comprehensive database.

Model Plan

Problem Objective: Develop a Machine Learning model that will predict the percentage of power outages in a Florida county based on outage and weather data.

Figure 1: High level block chart of design process

Data Processing

• Categorical features were numerically encoded.
• New features of Holiday/Weekday were introduced to inspect the effect of public holidays and weekends on Outage %.
• Correlation matrices were presented to identify redundant features in the database.
• DBSCAN Clustering method was used to remove outliers from the dataset.

Figure 2: Visualization for how DBSCAN clustering method operates.

Figure 3: Box plots of notable weather features showing the result of the DBSCAN outlier cleaning process.

Successful Training Models

We followed two types of ML training pipeline: One is Time-Series modeling and Independent modeling. Our AI solution Roadmap for Power outage prediction is shown below:

Figure 4: Time Series Model Flow Chart

Figure 5: Transformer Model Flow Chart

Other Models Attempted

Before arriving at our Time-Series and Transformer based models, we experimented the Outage % prediction with foundational/trivial and advanced models like Linear Regression, SVM, Tree based methods, SVM, by following Occam’s Razor, then proceeding with higher level models upon realizing the performance and limitations of the primary models attempted.

Figure 10: Mean Square Error comparison between the six primary models attempted compared to the baseline MSE.

Figure 11: Regression Plots - Actual percentage of power outages in Broward County for the six primary models attempted.

Figure 12: Confusion Matrices for the six models examined: Transformer Model, Time Series and Independent modeling. Our AI solution Roadmap for predicting power outages for major metropolitan areas in different states, trying more methods, or collecting more data to try and predict where the outage occurs during transmission.

Conclusion

• The goal for this project was to find the best fitting model for predicting power outages based on weather data.
• The best fitting model based on R² values is the Time Series Model.
• The next highest was the Transformer Model.

Future Work

Future work could include expanding to different states, trying more advanced methods, or collecting more data to try and predict where the outage occurs during transmission.

References and Acknowledgments

We would like to thank our sponsor GE Vernova and our advisors Armina Foroutan and Baosen Zhang for their help with our project.

Wang, Ranfeng; Vanga, Venkata Leela MO; Zaiken, Zachary B.; and Bennett, Jonathan (2022) “Analysis of the electric power outage data and prediction of electric power outage for major metropolitan areas in Texas using Machine Learning and Time Series Methods,” SMU Data Science Review: Vol. 8 : No. 1, Article 5.