Capstone 3 Proposal: Wind Power Forecasting

Problem Statement:
How much wind power one windmill will generate over the next 15 days?

Context:
Renewable and clean energy sources continue to be a controversial topic today. Specifically regarding wind power, those in opposition to wind farms have concerns regarding the loss of land use, the expense of production and maintenance of windmills, and the longevity of windmills. It is questioned whether or not the windmills produce enough electricity to be resource-effective. Data regarding the benefits of wind energy may be helpful in effectively making decisions regarding energy sources moving forward. Data for a specific windmill was recorded from January 2018 until March 2020 in 10-minute intervals. The data includes 22 features documenting weather information, power generated, and wind turbine and rotator information.

Criteria for Success:
Forecast the power one windmill will generate over the next 15 days. Determine the average power produced daily, and the cumulative power produced by the windmill over the course of the 15-day period.

Scope of Solution Space:
- MAE
- MSE
- RMSE
- MAPE
- MASE

Constraints:
Declining operation due to a potential lack of maintenance or cessation of operation due to repairs may affect the actual wind power generated over the forecasted time. Weather may also affect the actual wind power generated since weather naturally has a high variability. Extreme weather could also alter the wind power generated over the forecasted time period. Local land use could change and reduce wind access and power generated (for example, new construction near the wind farm). Finally, external factors such as governmental regulatory requirements, environmental restrictions, or land use policies may impact the actual wind power generated.

Stakeholders:
The stakeholders are the land owners where wind farms are developed, energy companies, and government systems (both local and national) making energy regulations.

Data Sources:
Data for this project is downloaded from Kaggle.

- https://www.kaggle.com/datasets/theforcecoder/wind-power-forecasting

Approach:

- Data Wrangling - Download the CSV from the Kaggle website and import the file into a Jupyter Labs notebook. View the first five lines of the data set to gain an initial understanding of the data set.
- EDA - Run a Profile Report to view a summary of the data, its distribution, and any alerts regarding feature correlation, missing values, or unique values.
- Feature Engineering - Resolve any missing values, normalize data, and complete time-based aggregations as needed.
- Model Development - Run 3-5 predictive models with hyperparameter tuning.
- Model Evaluation - Create a table of model metrics comparing each model's MAE, MSE, RMSE, MAPE, and MASE in order to select the best model.
- Interpretation - Create 3-5 visualizations of the predictive modeling results.
- Recommendation - Present the final prediction of wind power generated by the specified windmill over the next 15 days for use in discussions comparing wind power to other sources of power.

Deliverables:

- Jupyter Lab code for data science method
- Metrics File
- Project Report
- Project Presentation