Vestas: Predicting Long-Term Corrected Wind Speed

In this assignment, we worked with real-world data from Vestas to optimize methods for generating Long-Term Corrected (LTC) wind data using regression models. The objective was to evaluate whether regression techniques could replace neural networks for predicting wind speeds and directions over long time spans, providing a cost-effective solution.

Key Tasks Performed

1. **Data Preprocessing**:
    - Worked with mast and meso time series data to align and clean datasets.
    - Resampled mast data from 10-minute intervals to match the hourly meso data.
    - Corrected timestamps by converting mast data to UTC and addressed daylight saving time adjustments.
    - Interpolated meso data to match the height of mast measurements for consistent comparison.

2. **Exploratory Data Analysis**:
    - Analyzed wind speed and direction distributions to identify trends and patterns.
    - Calculated and compared NaN percentages for key variables, ensuring clean and usable data for modeling.
    - Explored Weibull distributions to understand wind speed behavior across datasets.

3. **Regression Model Development**:
    - Developed regression models to predict mast wind conditions using meso data for overlapping time periods.
    - Trained and validated models using techniques like cross-validation to ensure robustness.
    - Compared regression outputs to the actual mast data to assess the accuracy of LTC predictions.

4. **Model Evaluation**:
    - Evaluated models using metrics like Mean Squared Error (MSE) and R-squared (R^2) to quantify performance.
    - Compared the predicted and actual wind speed distributions to assess the effectiveness of the regression approach.

5. **Documentation**:
    - Documented the process thoroughly to ensure reproducibility and clarity.
    - Discussed potential limitations and opportunities for further improvement.

Through this project, we gained hands-on experience working with large datasets, preprocessing time series data, and applying regression techniques in a real-world scenario. This assignment highlighted the potential of regression models as a time-efficient and cost-effective alternative to neural networks for wind data prediction.