Energy Production Prediction Using Environmental Variables

A Regression Analysis Project

Your Name

January 2025

Objective

- To model energy production as a function of environmental variables using regression techniques.

- Key Target: Predict continuous energy production based on environmental conditions.

Dataset Overview

- Dataset Name: solarpowergeneration.csv
- Size: 2920 instances (rows) and 10 variables (columns)
- Target Variable: power_generated (in Joules per 3 hours)
- Features: Environmental variables such as temperature, wind speed, and humidity.

Features Description

- 1. distance_to_solar_noon: Radians
- 2. temperature: Degrees Celsius
- 3. wind_direction: Degrees (0-360)
- 4. wind_speed: Meters per second
- 5. sky_cover: Scale 0-4
- 6. visibility: Kilometers
- 7. humidity: Percentage
- 8. average_wind_speed: Meters per second
- 9. average_pressure: Mercury inches

Target Variable

- Variable: power_generated
- Unit: Joules
- Measurement: Energy produced every 3 hours.

Regression Problem

- Type of Analysis: Regression
- Goal: Predict continuous values of power_generated.
- Approach: Evaluate relationships between the target variable and independent features.

Data Preparation

- Steps:
- – Data Cleaning: Handle missing values, outliers.
- – Feature Scaling: Normalize or standardize variables.
- – Encoding: Convert categorical variables (if any) into numerical.
- Splitting: Train-test split for model evaluation.

Exploratory Data Analysis (EDA)

- Insights on feature correlations with power_generated.

- Visualizations:
- – Scatter plots
- – Heatmap of correlations
- – Histograms for variable distributions

Modeling Approach

- Models Considered:
- - Linear Regression
- - Decision Trees
- - Random Forests
- - Gradient Boosting Models

- Evaluation Metrics:
- - Mean Absolute Error (MAE)
- - Mean Squared Error (MSE)
- - R-squared (R^2)

Results

- Summarize model performance.
- Highlight the best-performing model based on evaluation metrics.

Conclusion

- Importance of environmental factors in energy prediction.
- Potential applications for optimizing solar energy production.

Future Work

- Incorporate more granular data (e.g., hourly measures).
- Explore advanced models like Neural Networks.
- Investigate additional environmental factors.

Acknowledgments and Questions

- Acknowledge contributors or sources.
- Invite questions from the audience.