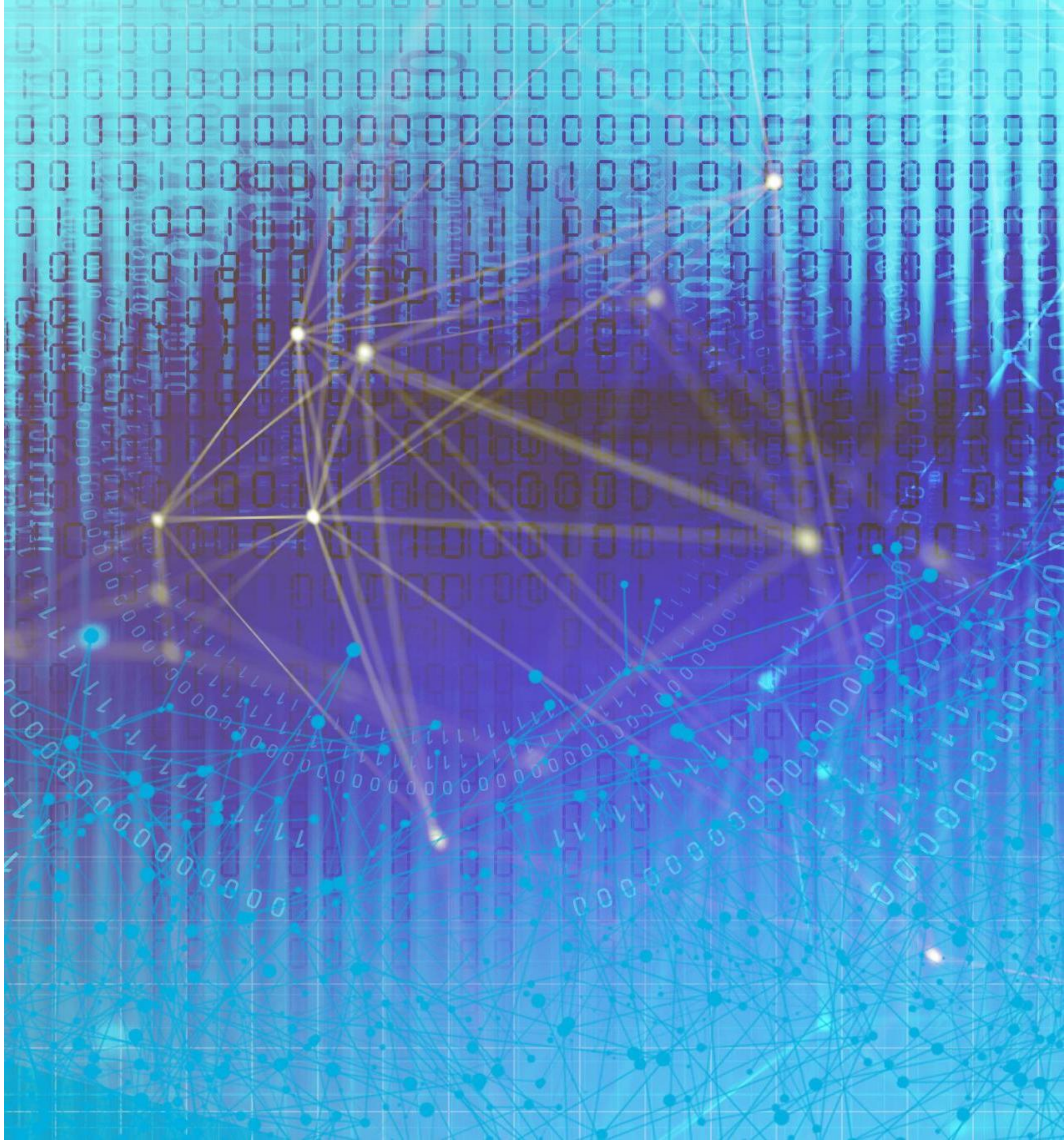
Predicting Electrical Energy Output of a Combined Cycle Power Plant: A Machine Learning Approach

Machine learning techniques for energy output prediction

Agenda for Machine Learning Approach

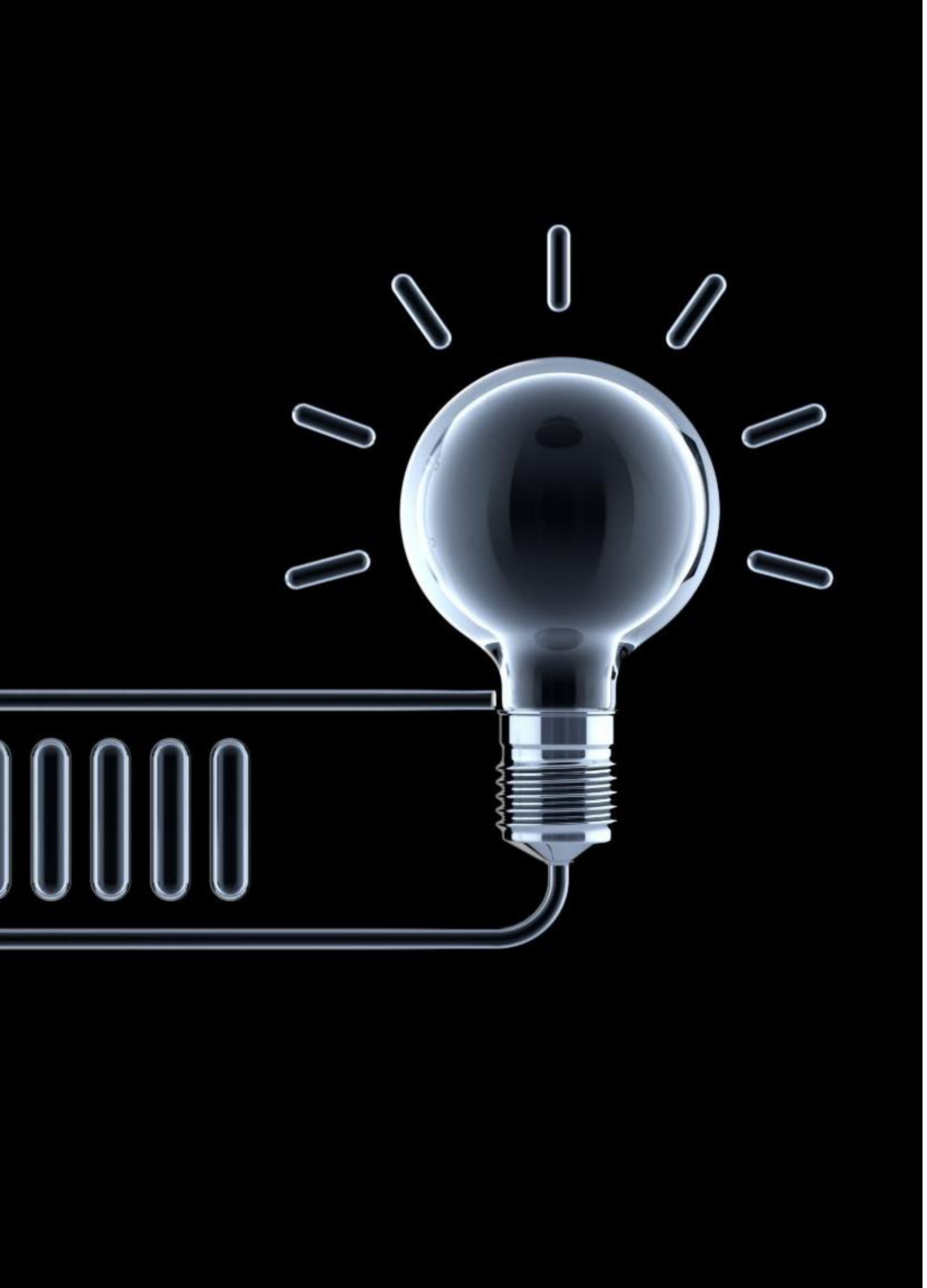
- Modeling Approach
- Model Building Process
- Model Evaluation



The Challenge: Predicting Power Output

- Forecast electrical energy output for efficient power generation.
- Optimizes operational decisions and reduces overall costs.
- Utilized a regression model built with Python for predictions.
- Analyzed data from 9,568 hourly sensor readings.
- Employed tools like VS Code and Python for effective exploration.

Modeling Approach



Task and Features

Predictive Task Overview

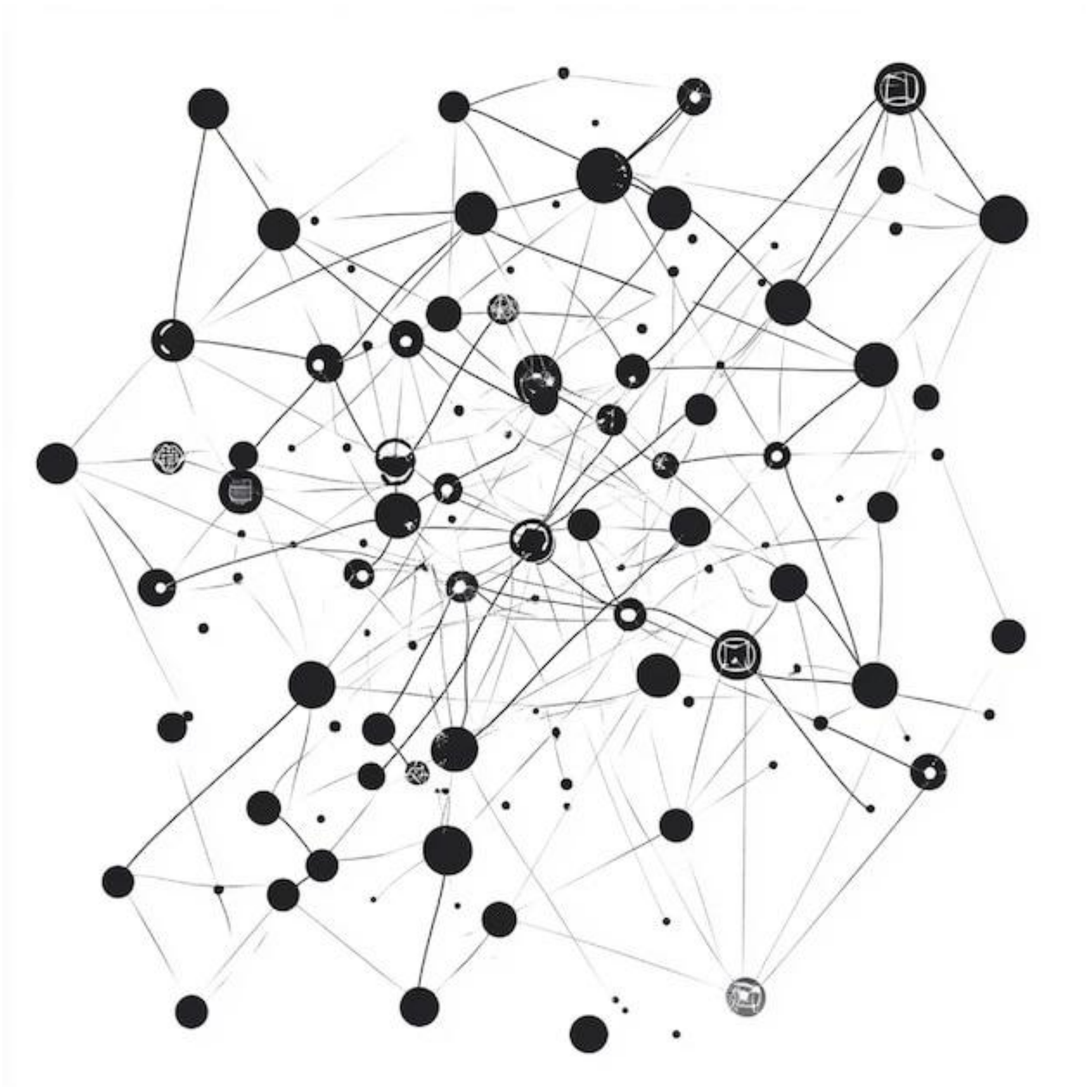
The task focuses on predicting electrical energy output using various input features, enhancing model accuracy.

Key Features

Key features include temperature, ambient pressure, relative humidity, and exhaust vacuum, each influencing energy output.

Feature Ranges

Each feature has a specific range, essential for accurate predictions in the model's performance.



Algorithms Considered

Baseline Algorithm: Linear Regression

I established a simple baseline using linear regression to evaluate the performance of our models against a straightforward predictive approach.

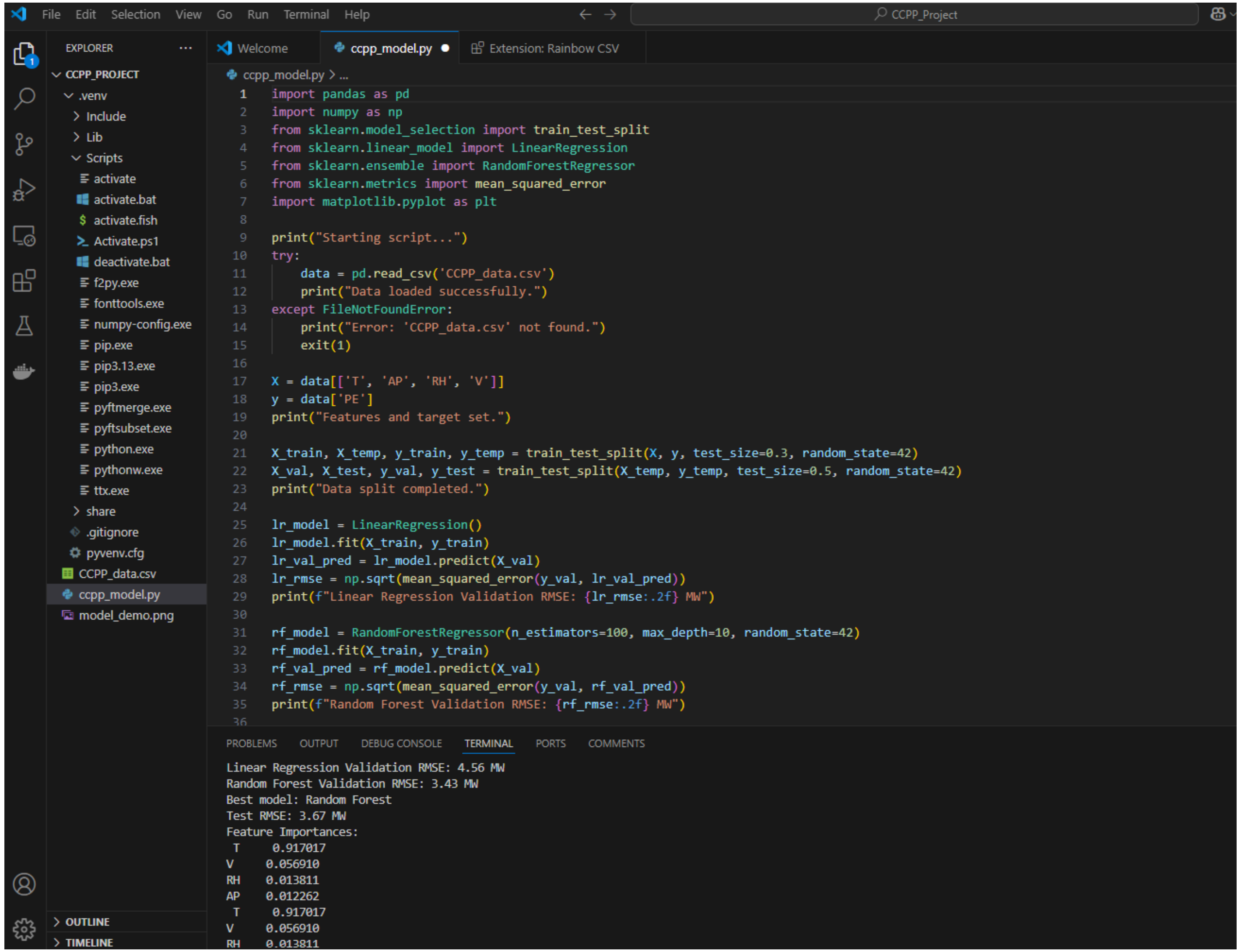
Exploration of Random Forest

I explored the random forest algorithm, which effectively captures non-linear patterns in the data, enhancing predictive accuracy.

Model Building Process

How I Built It: Python in Action

- Data handling was done efficiently using pandas.
- Training, validation, and testing sets were created using scikit-learn.
- Two models were tested: Linear Regression and Random Forest.
- Random Forest outperformed Linear Regression with lower RMSE.
- Key extensions like scikit-learn and numpy facilitated our ML process.



Validation Strategy

Single Validation Set Usage

A single validation set was used to compare the performance of different models, ensuring consistent evaluation parameters.

Linear Regression Performance

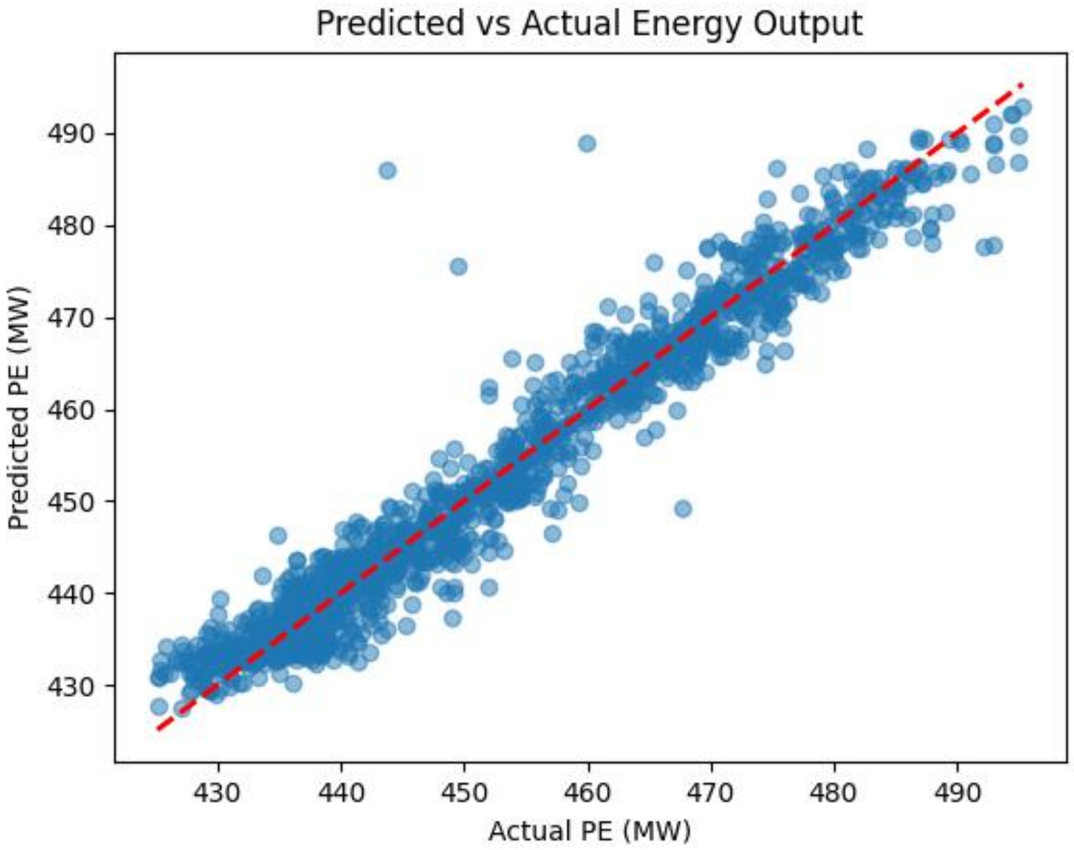
The linear regression model exhibited an RMSE of 4.56 MW, indicating its prediction accuracy on the validation set.

Random Forest Advantage

The random forest model outperformed linear regression, achieving a lower RMSE of 3.43 MW, showcasing its effectiveness.

Model Evaluation

Interpretation of Results



Average Prediction Error

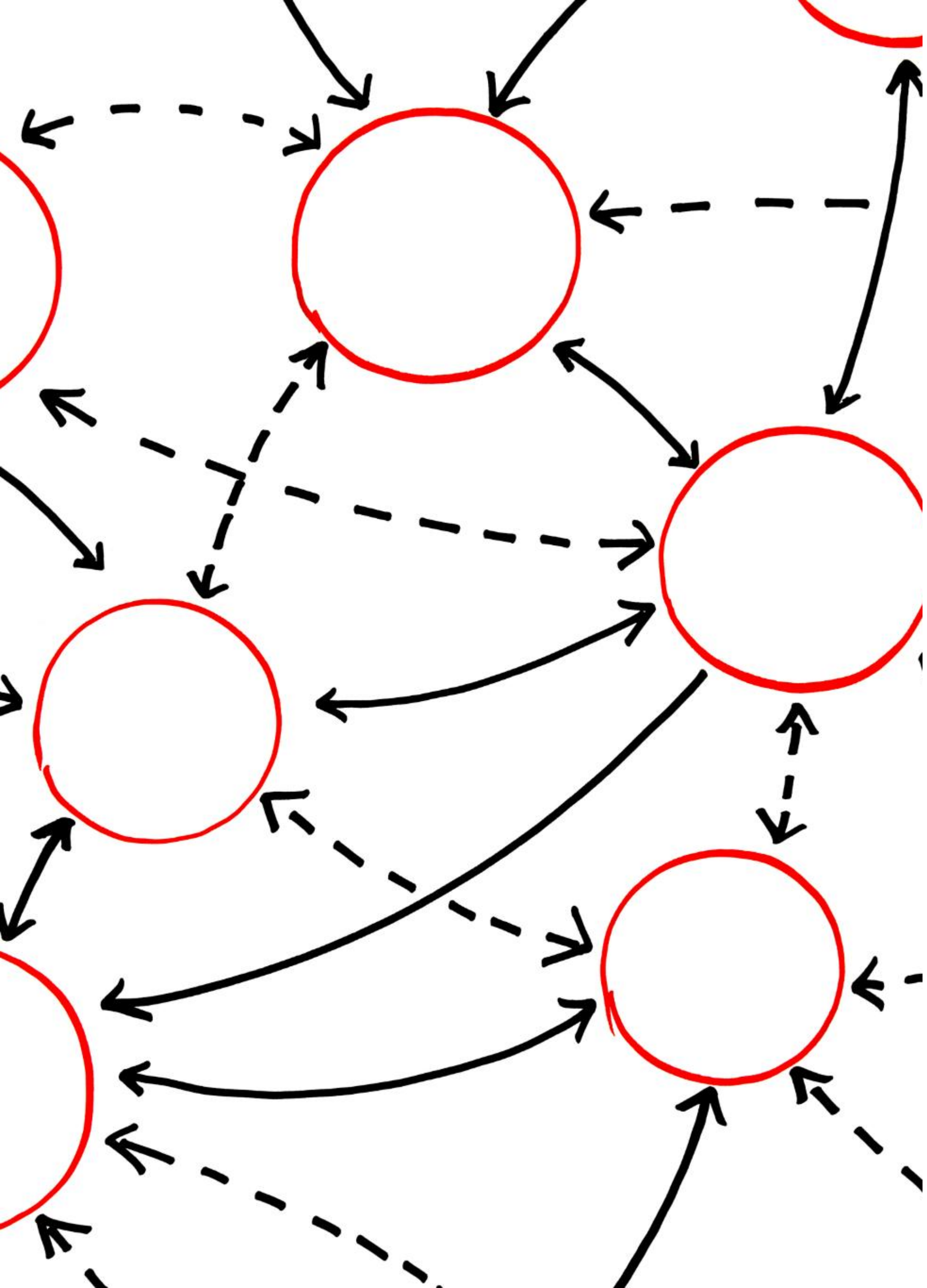
The average prediction error of 3.67 MW demonstrates the model's accuracy in forecasting energy output.

Output Range Comparison

With an output range of 420-495 MW, the small error showcases the model's reliability for energy output prediction.

Model Reliability

The low error margin indicates that the model can be trusted for future energy output predictions.



Model Selection

Model Comparison

The random forest model showed a lower validation RMSE compared to linear regression, indicating superior predictive performance.

Random Forest Advantages

Random forest offers greater accuracy and robustness in predicting electrical energy output for power plants.

Predictive Performance

Choosing the right model is crucial for accurate predictions in energy output, enhancing operational efficiency.

Conclusion

Machine Learning Application

I successfully utilized machine learning techniques to enhance the prediction of energy output in power generation.

Random Forest Model Success

The random forest model showed efficient performance with low error rates, proving its effectiveness in energy output predictions.

Energy Output Predictions

This model serves as a valuable tool for accurately predicting electrical energy output, aiding in operational efficiency.