BIA | BOSTON INSTITUTE OF ANALYTICS ®

power consumption prediction

Presented By : vishal Ramesh Nishad

Batch : AND-JUN2024-DSAI-1

Date : 10TH JANUARY 2025

Place : Mumbai

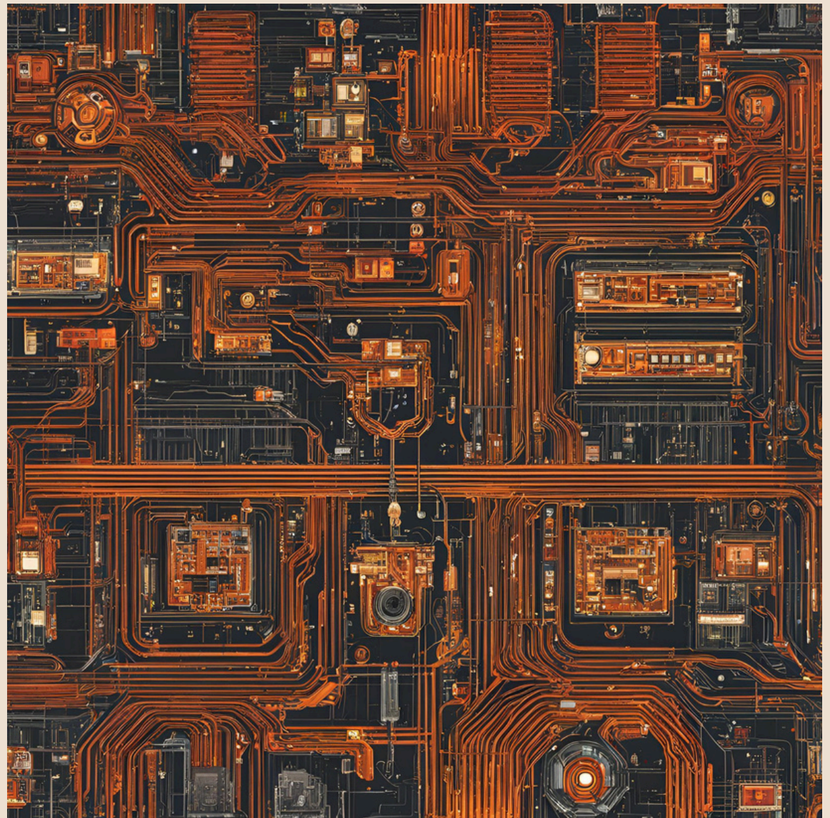
AGENDA

- Introduction
- Data Overview
- Exploratory Data Analysis (EDA)
- Model Selection
- Model Evaluation
- Feature Importance
- Results
- Conclusion

BIA | BOSTON INSTITUTE OF ANALYTICS ®

INTRODUCTION

Today, we will explore how machine learning can predict power consumption patterns based on factors such as weather conditions, time of day, energy usage trends, and external variables. This analysis offers valuable insights to help optimize energy usage, improve efficiency, and guide decision-making in power management.



BIA BOSTON INSTITUTE OF ANALYTICS ®

GOAL OF THE PROJECT

The primary goal of this project is to predict power consumption categories based on various factors such as time of day, weather conditions, usage patterns, and historical consumption data. By building a machine learning model, the project aims to classify power consumption into categories (e.g., "low," "moderate," or "high") based on performance metrics like energy usage and external variables.

HOW MACHINE LEARNING HELPS ?

Machine learning analyzes factors like time of day, weather conditions, historical usage, and external variables to identify patterns influencing power consumption. It enables energy providers to make quick, data-driven decisions for optimizing energy usage and improving efficiency.

BIA | BOSTON INSTITUTE OF ANALYTICS ®

DATA OVERVIEW

The dataset includes columns, Datetime', 'Temperature', 'Humidity', 'WindSpeed', 'GeneralDiffuseFlows', 'DiffuseFlows', 'PowerConsumption_Zone1', 'PowerConsumption_Zone2', 'PowerConsumption_Zone3', 'DayOfWeek', 'Month', 'Hour', 'Quarter'],, etc.

Dataset Dimensions

- **Number of Rows : 52416**
- **Number of Columns : 13**

Data Challenges

- **Time Series Nature**
- **Data Granularity**

Conclusion

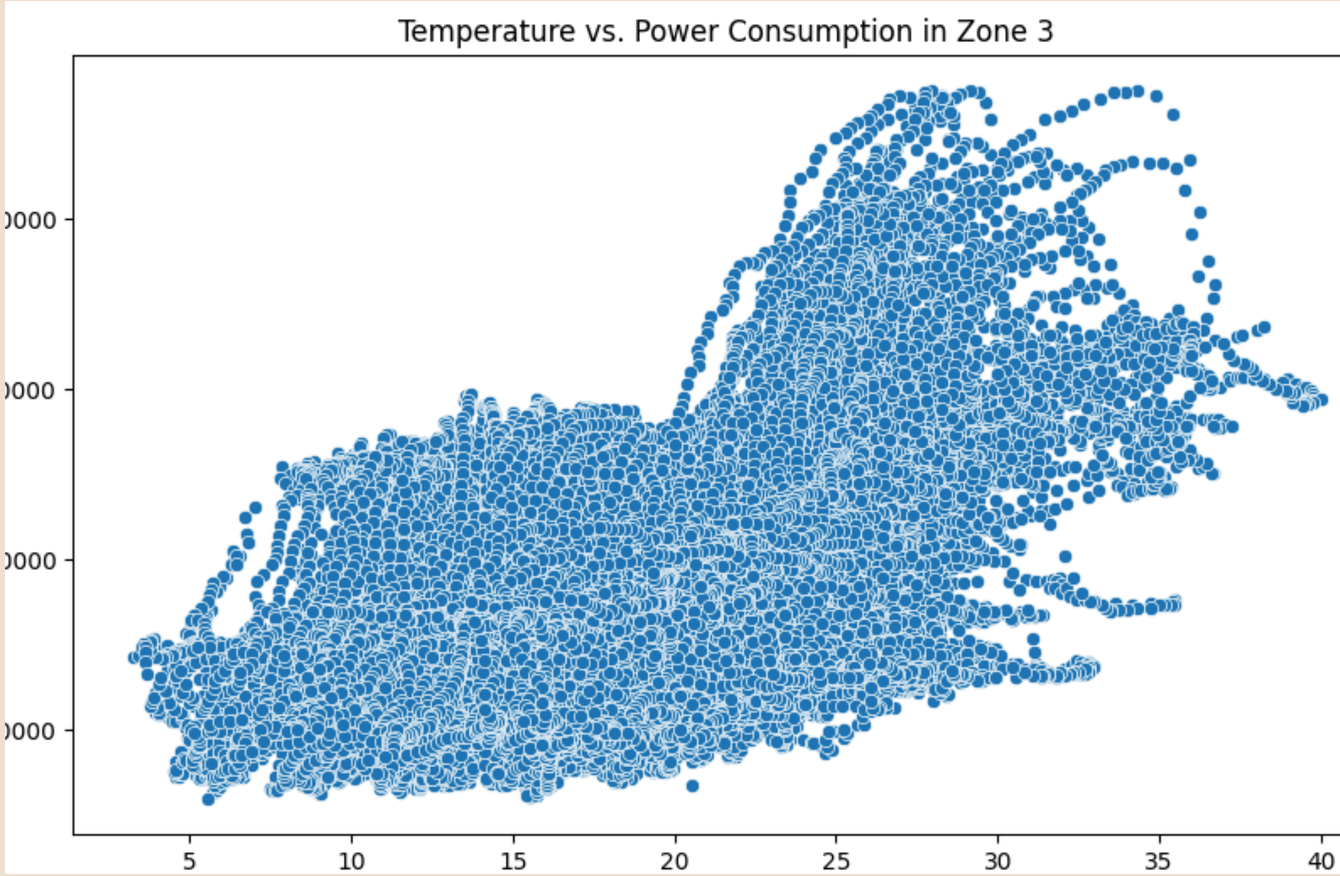
The dataset provides a robust foundation for analyzing power consumption, offering insights into critical factors such as weather conditions, time of day, and usage patterns. Its richness and diversity make it ideal for machine learning applications aimed at predicting energy consumption trends and optimizing power usage.

BIA BOSTON INSTITUTE OF ANALYTICS ®

EXPLORATORY DATA ANALYSIS

- **Temperature and power consumption in Zone 3.**

- "Temperature impacts power consumption in Zone 3, influencing energy demand."
- "Analyzing temperature and power consumption in Zone 3 helps optimize energy usage."
- "Temperature fluctuations drive power consumption patterns in Zone 3."
- "Understanding the link between temperature and power consumption in Zone 3 aids in better energy management."



BIA BOSTON INSTITUTE OF ANALYTICS ®

PRE-PROCESSING

Splitting

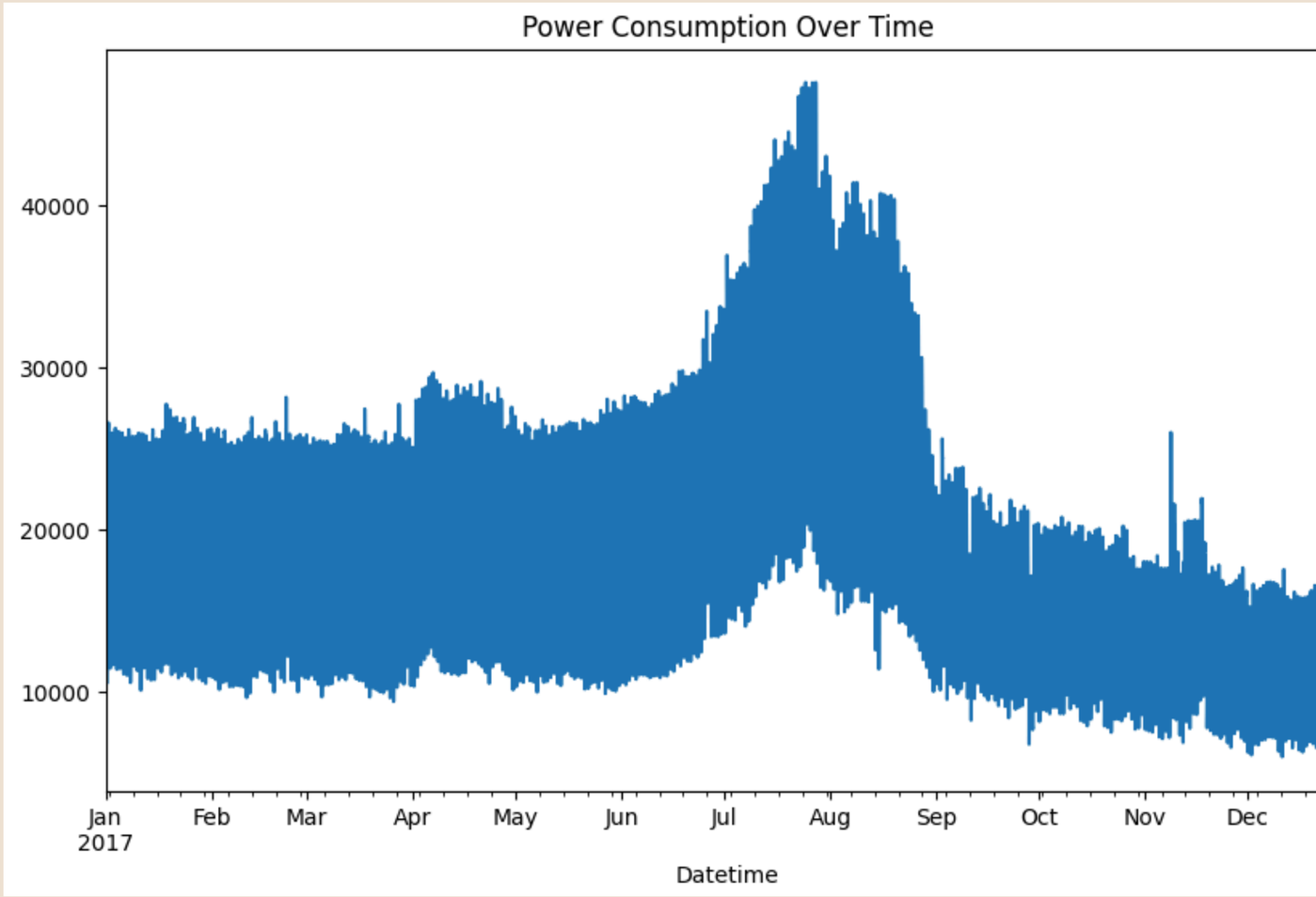
Data split into training and testing sets using train_test_split.

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

- **"The dataset is split into training and testing sets with 80% for training and 20% for testing."**
- **"We use an 80-20 split to train the model and evaluate its performance on unseen data."**
- **"The train-test split ensures that the model is tested on data it hasn't seen before for better generalization**

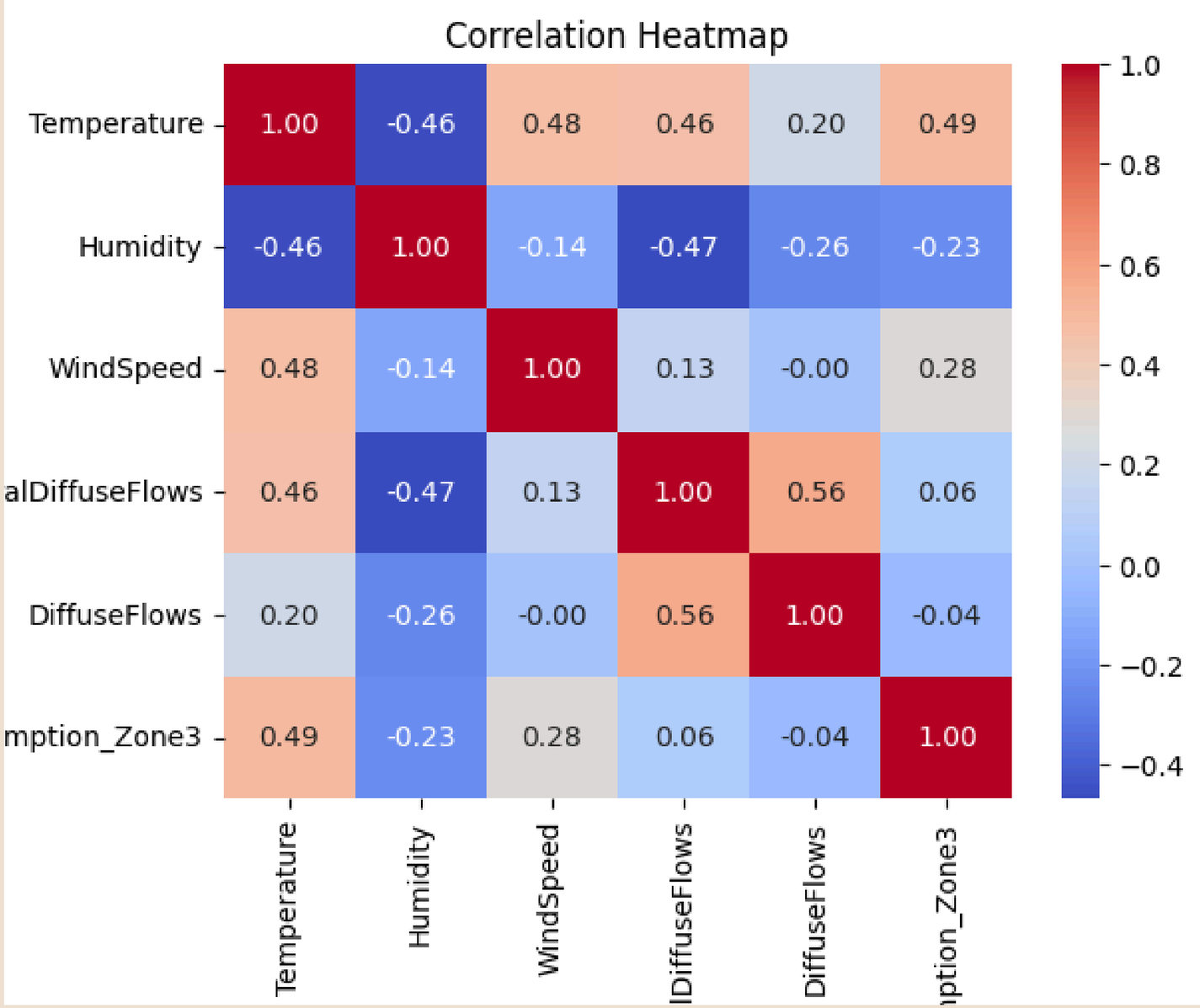
DATA VISUALIZATION

- "Power consumption trends vary over time, reflecting daily and seasonal patterns."
- "Analyzing power consumption over time helps identify peak usage periods and optimize energy distribution."
- "Time-based analysis reveals fluctuations in power consumption, aiding in demand forecasting."
- "Understanding power
- consumption over time is key to improving energy efficiency and planning."



BIA BOSTON INSTITUTE OF ANALYTICS ®

- "The heatmap shows correlations between factors affecting power consumption."
- "It highlights key variables that influence power usage patterns."
- "The heatmap reveals how temperature, humidity, and time relate to consumption."
- "It helps identify important features for predictive modeling."



BIA BOSTON INSTITUTE OF ANALYTICS ®

MODEL SELECTION

Machine Learning Model Selection

- **Random Forest :**
  - "Random Forest is used for its ability to handle complex, non-linear relationships in power consumption data."
  - "This model effectively captures interactions between multiple variables like temperature, humidity, and time."
  - "Random Forest handles large datasets well, making it ideal for predicting power consumption across different zones."
  - "The model’s robustness to overfitting and its ability to provide feature importance makes it suitable for power consumption prediction."

MODEL PERFORMANCES

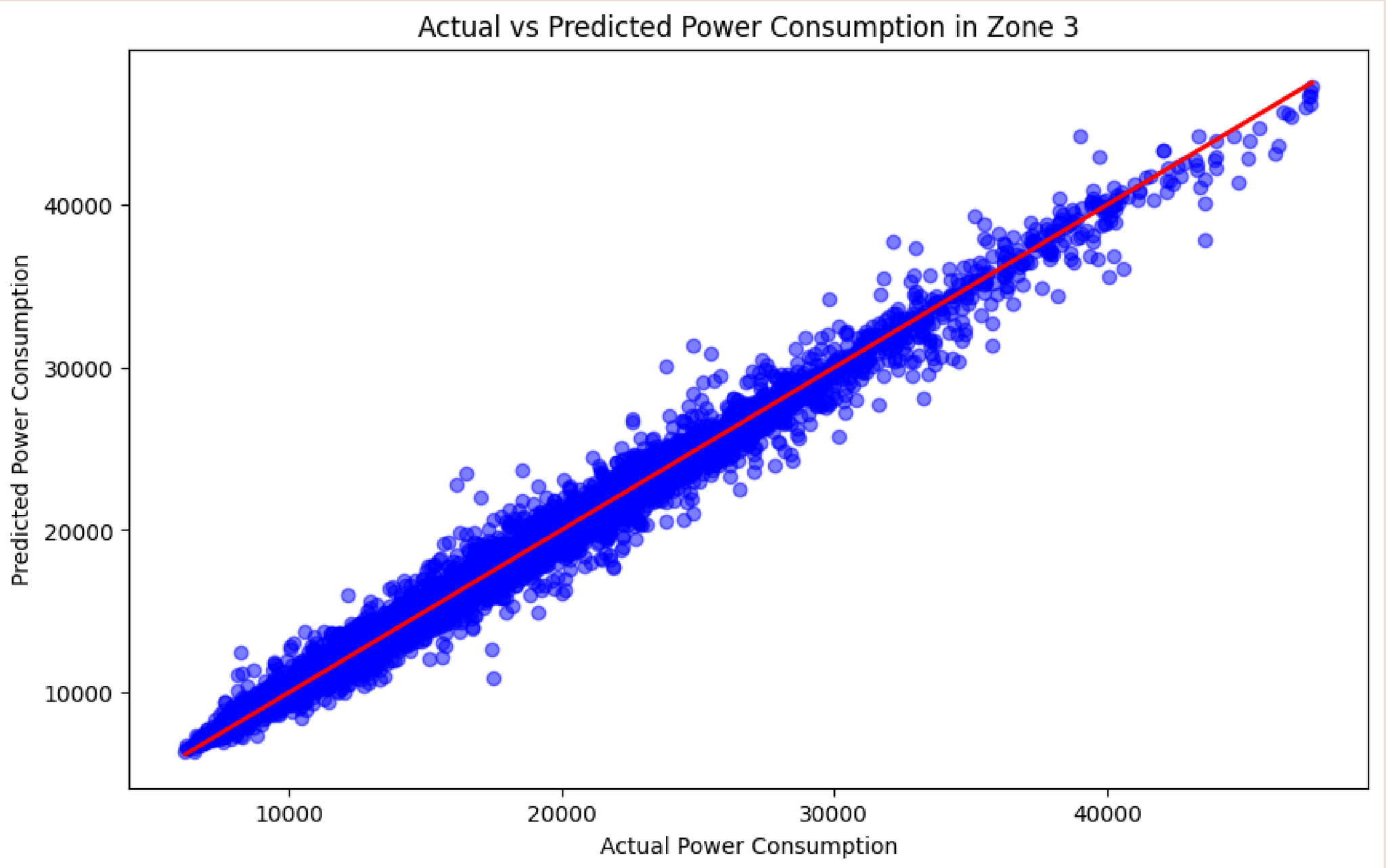
Insight:

Random Forest achieved the best performance with the highest of Model Accuracy: 98.71%

Metric	Value
MAE	**468.49**
RMSE	**749.74**
R^2	**0.9871**
Model Accuracy	**98.71%**

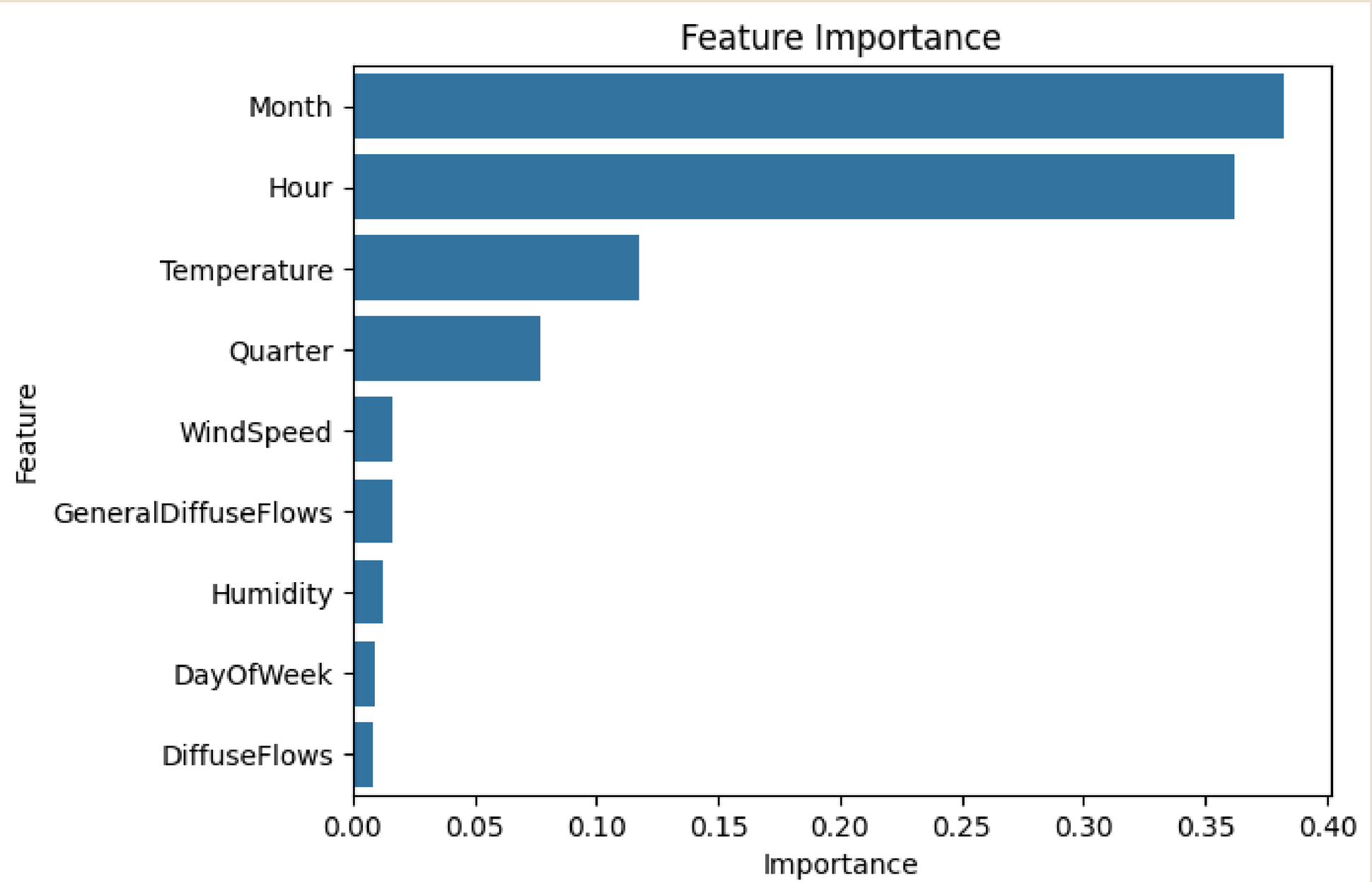
BIA BOSTON INSTITUTE OF ANALYTICS ®

SCATTER PLOT - MODEL PERFORMANCES



BIA BOSTON INSTITUTE OF ANALYTICS ®

FEATURE IMPORTANCE - RF MODEL



BIA BOSTON INSTITUTE OF ANALYTICS ®

CONCLUSION

Best Model: 

Random Forest emerged as the most accurate, achieving the best accuracy.

Random Forest is the recommended model for predicting power consumption, leveraging its robustness and interpretability to provide reliable and accurate predictions.

BIA BOSTON INSTITUTE OF ANALYTICS ®

BIA | BOSTON INSTITUTE OF ANALYTICS ®

THANK YOU

Presented By : Vishal Ramesh Nishad