XGBoost Model Evaluation Report

1. Overview

This report presents the evaluation of the XGBoost Regressor model used for hourly car rental demand prediction. The model was selected for its superior performance compared to other regression algorithms tested, including Random Forest, Gradient Boosting, and Linear Regression.

---

2. Performance Metrics

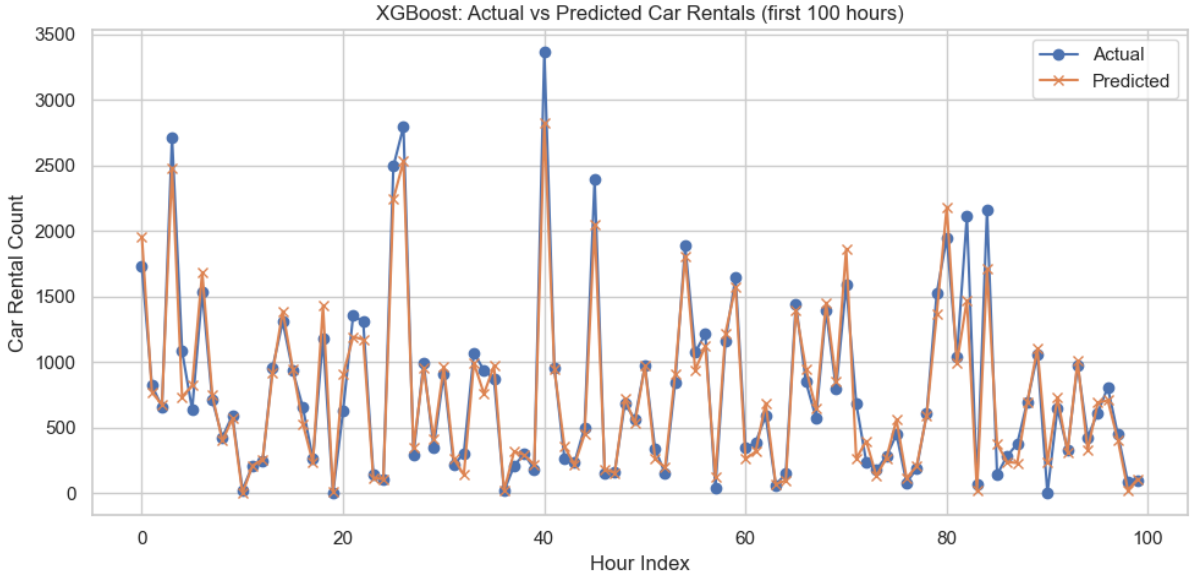
Metric	Value
RMSE	167.73
MAE	98.44
MAPE (%)	38.65
R^2 Score	0.933

- The XGBoost model explains **93.3%** of the variance in the data.
- Error metrics (RMSE and MAE) are significantly lower than the target variable's standard deviation (~645), indicating high prediction accuracy.

---

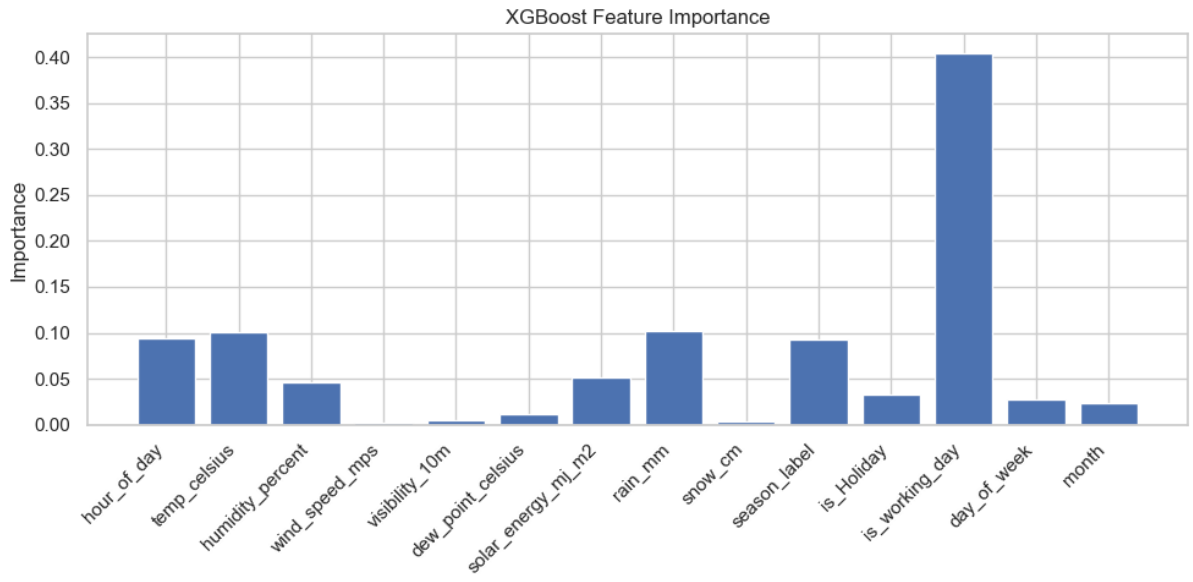
3. Graphical Evaluation

a. Actual vs. Predicted Plot

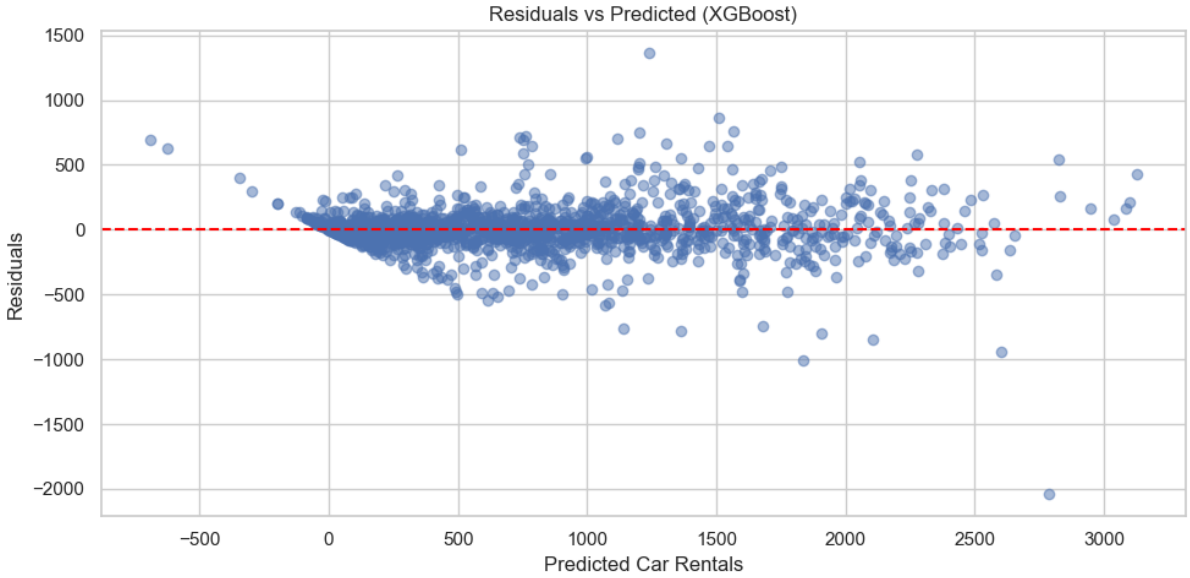


- The plot of actual vs. predicted values over the first 100 hours shows that the model captures the overall demand trend well.
- Peaks and troughs are aligned with minimal deviation, confirming strong predictive alignment.

b. Feature Importance Plot

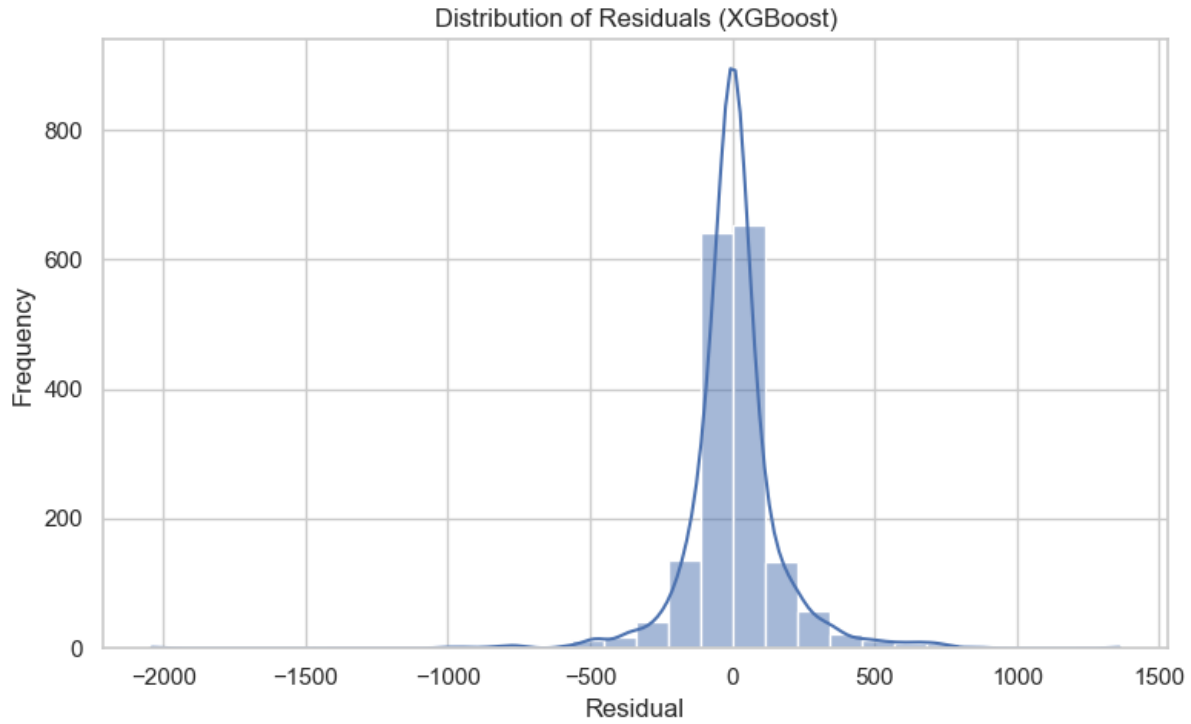


- Top contributing features:
  - temp_celsius
  - hour_of_day
  - is_working_day
  - solar_energy_mj_m2
- Insights: Both **temporal** and **weather-based** features are critical drivers of rental demand.
- **c. Residuals vs. Predicted Plot**



- Residuals are symmetrically distributed around zero with no visible patterns.
- Indicates absence of systematic over- or under-prediction.

d. Residuals Histogram



- Residuals form a bell-shaped distribution centered at zero.
- Suggests prediction errors are small, unbiased, and mostly random.

---

4. Technical Strengths

- **Generalization:** High R^2 and low error values on the test set confirm minimal overfitting.
- **Nonlinearity Handling:** Tree-based structure captures complex feature interactions better than linear models.
- **Robustness:** Model performs consistently even during demand spikes or irregular patterns.

---

5. Challenges and Mitigations

Challenge	Solution
Capturing complex, nonlinear patterns	Utilized XGBoost's boosting mechanism to model nonlinear relationships.
Identifying relevant features	Used feature importance analysis to guide feature engineering.
Avoiding overfitting	Applied careful data splitting, residual analysis, and multi-metric evaluation.

---

6. Conclusion

The XGBoost Regressor proved to be the most effective model for forecasting car-sharing demand, offering:

- High predictive accuracy
- Strong interpretability through feature importance
- Robust performance across diverse demand scenarios

These qualities make XGBoost a reliable model for practical deployment in car-sharing platforms to aid in **real-time demand forecasting and operational decision-making**.