Reducing Network Load: Predicting Reactive Power Setpoints with Fewer Features

Kanishk Varma Dendukuri, Manish Kumar Singh

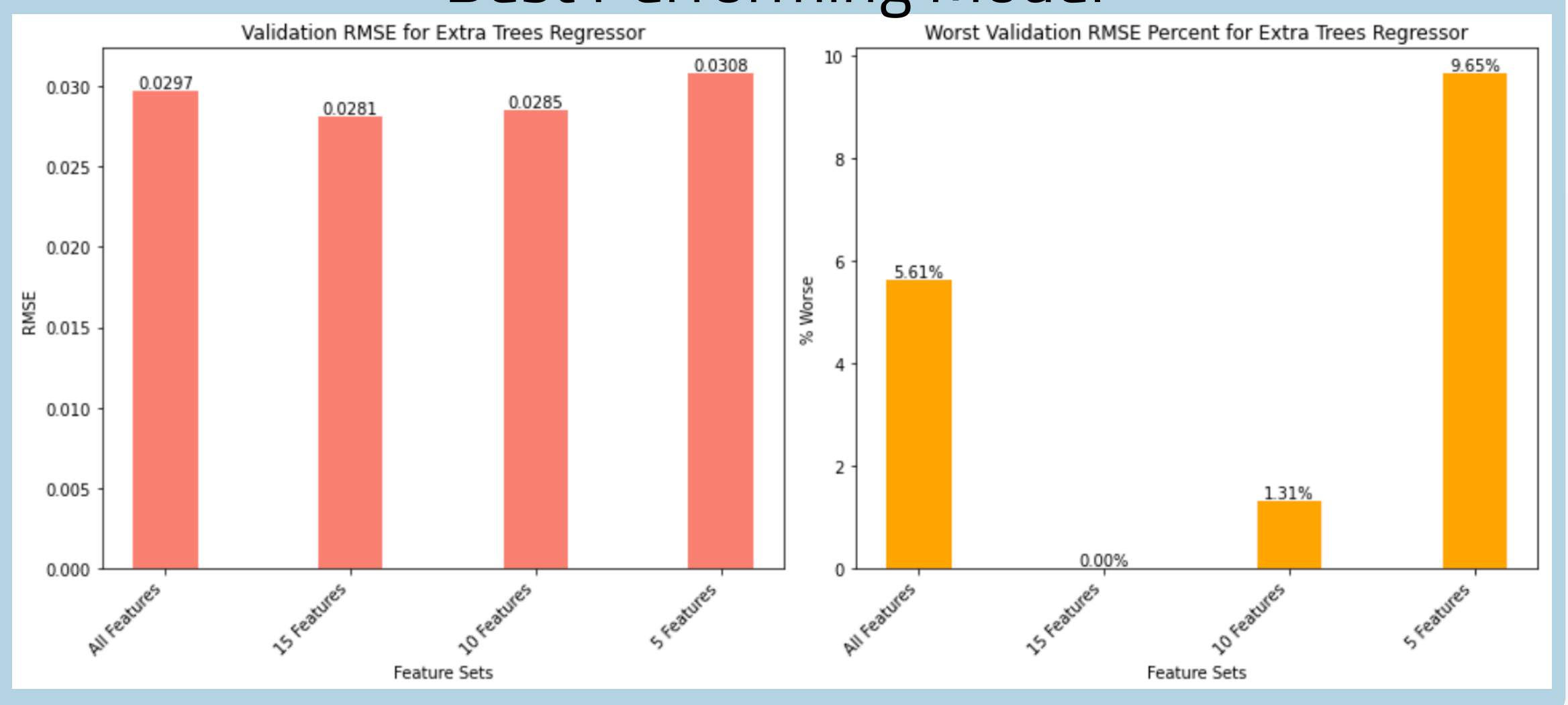
Department of Electrical and Computer Engineering

WISPO – Wisconsin Power Systems
Electric Power Systems Research Group at UW-Madison

TL;DR

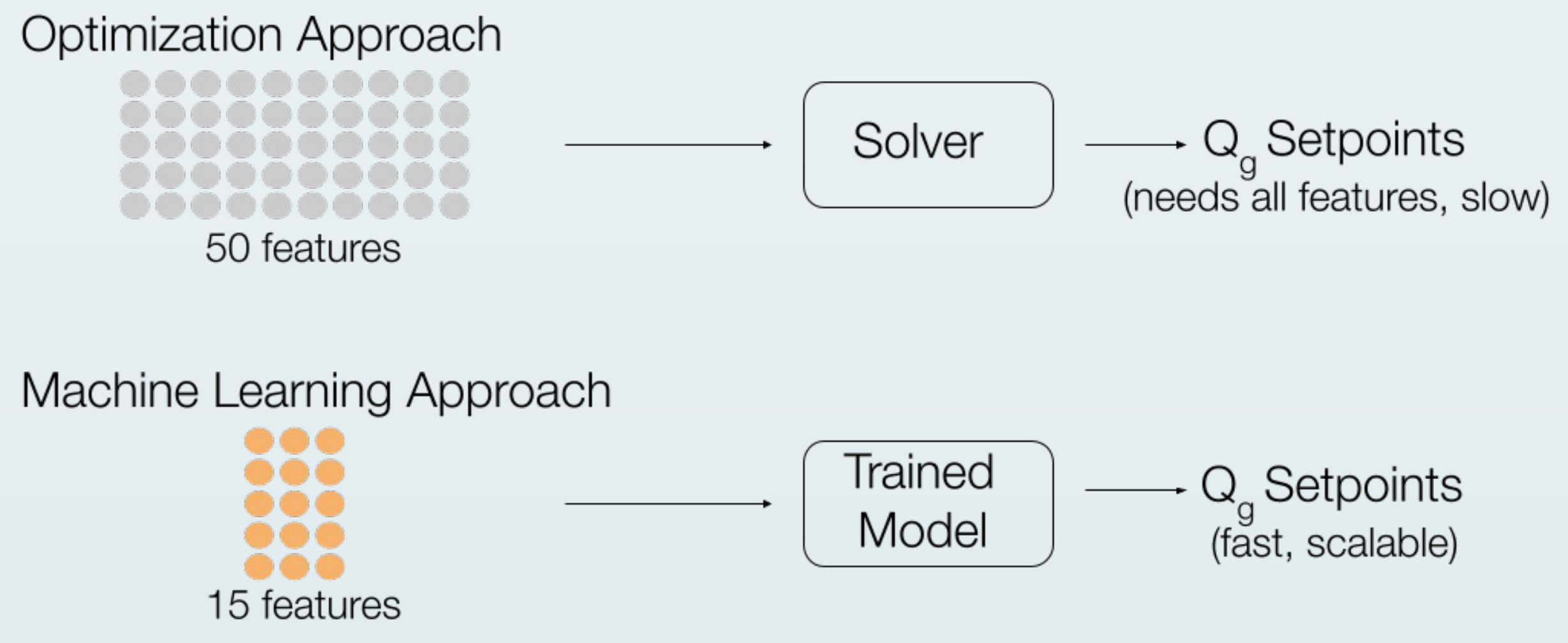
We use machine learning to predict optimal reactive power setpoints (Q_g) for solar inverters while identifying which input features matter most. By replacing traditional optimization with data-driven models, we reduce computational demands without sacrificing accuracy. Tree-based and boosting-based models showed strong performance, with Extra Trees achieving the best results when using only 15 selected features. This demonstrates that accurate predictions can be made with fewer features, offering a scalable solution for large power networks.

Best Performing Model



Introduction

- Reactive power setpoints (Q_g) help optimize solar inverter operation and minimize power losses.
- Traditional methods rely on full-network optimization, which is slow and resource-heavy.
- Our goal: Identify which features are most important for predicting Q_g and Use machine learning models to predict Q_g accurately using only those features.
- This reduces complexity while retaining predictive performance.

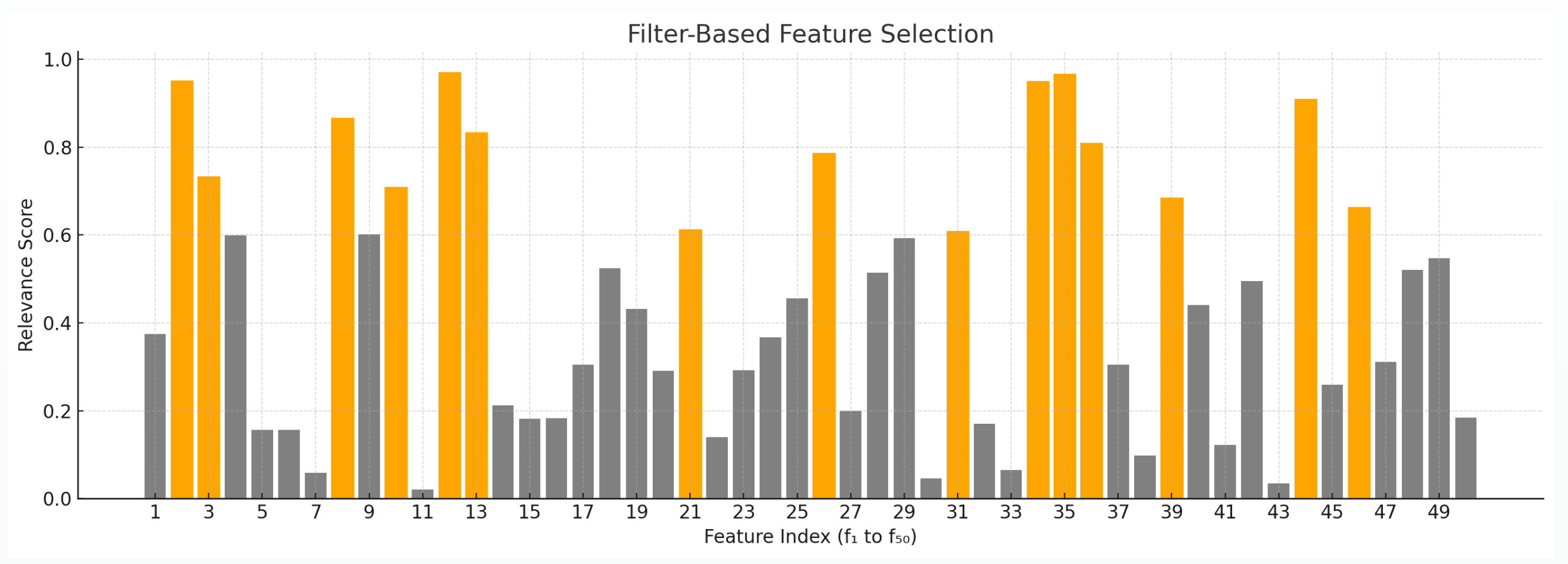


Dataset

- The dataset represents a 37-node power distribution network over 800-time instances.
- Only 25 nodes are active, resulting in 50 total input features:
  - Net Active Power Injection ($P_g – P_l$)
  - Reactive Power Consumption (Q_l)
- The output is a vector of reactive power setpoints (Q_g) for 5 solar inverters.
- All features were standardized to have zero mean and unit variance prior to model training.

Methodology

1. Filter-Based Feature Selection

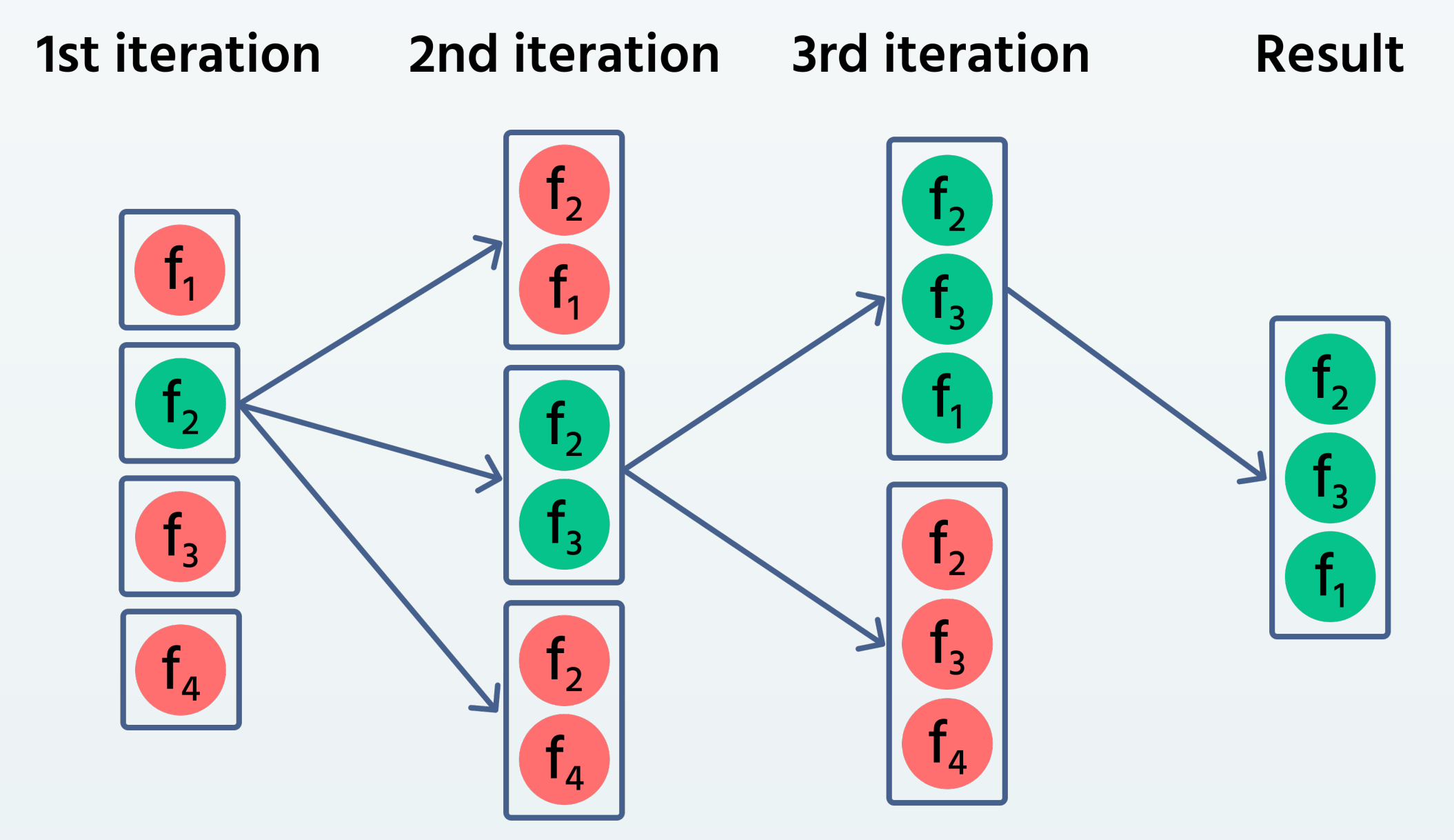


Filtering methods assign a relevance score to each feature based on its statistical relationship with the target (Q_g).

The top 5, 10, and 15 features were selected using these metrics:

- Mutual Information (MI)
- Maximal Information Coefficient (MIC)
- Distance Correlation (DCor)
- ANOVA F-value
- Pearson's Correlation

2. Wrapper-Based Feature Selection



Used Forward Sequential Feature Selection (SFS) to iteratively build subsets of the best-performing features.

At each step, SFS adds the feature that minimizes RMSE when combined with the already selected ones.

This was repeated until we reached 5, 10, and 15 features for:

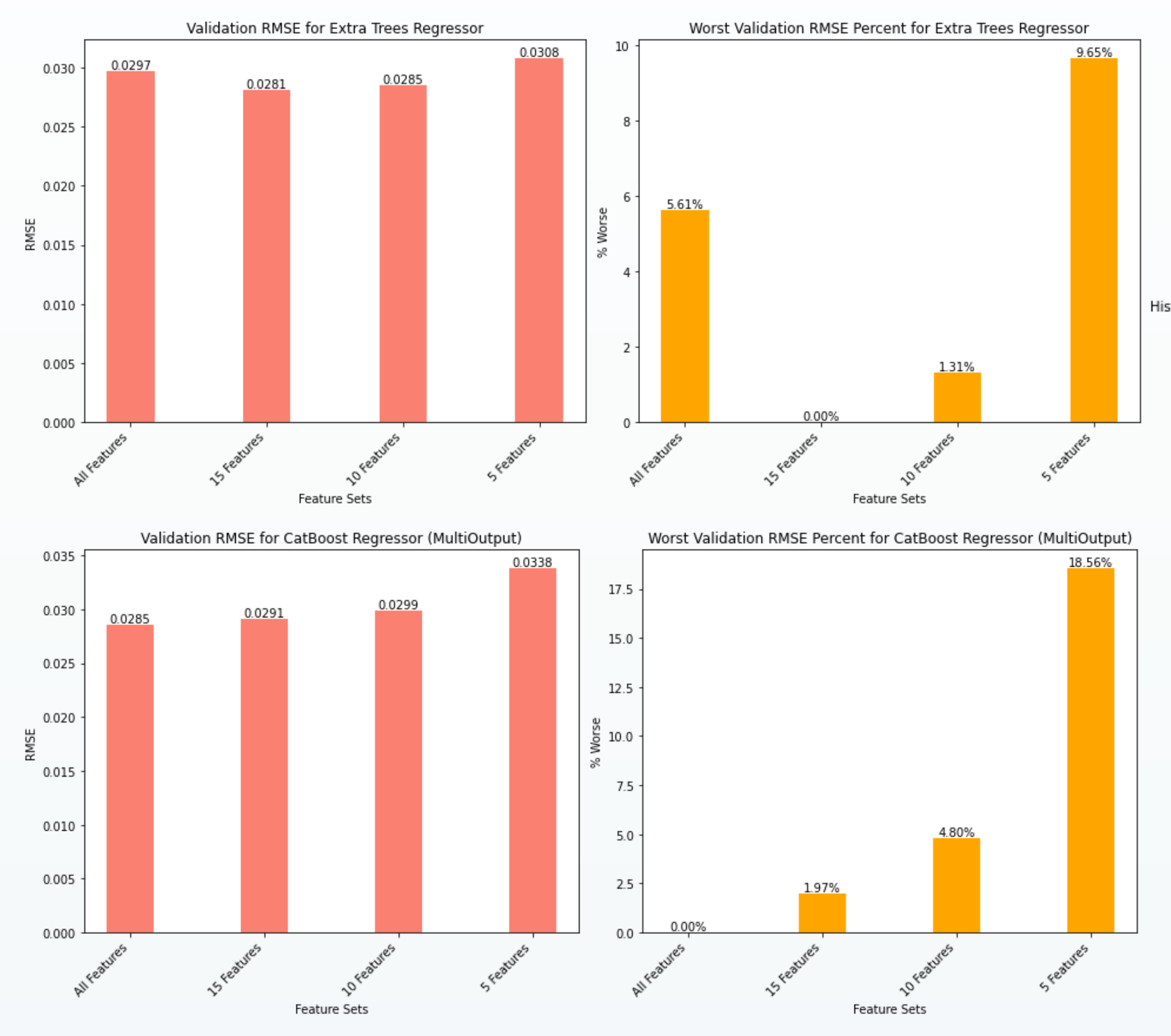
- Linear Models: Linear Regression, Ridge Regression, ElasticNet Regression
- Polynomial Regression
- Tree-Based Models: Random Forest, Extra Trees
- Neural Networks: Multi-Layer Perceptron (MLP)
- Kernel-Based Models: Support Vector Regression (SVR)
- Distance-Based Models: KNeighbors Regressor
- Boosting-Based Models: Hist Gradient Boosting, AdaBoost, CatBoost

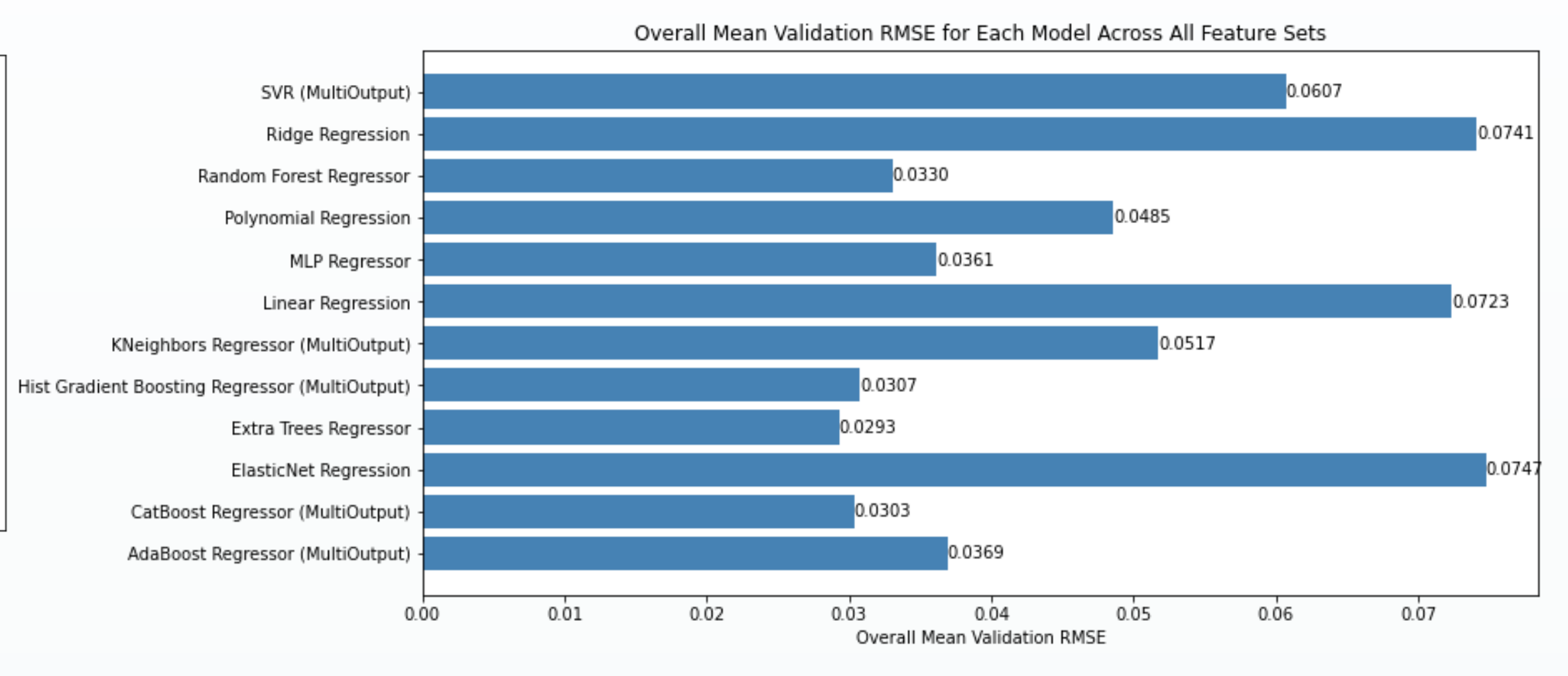
3. Evaluation Metric

We used Root Mean Squared Error (RMSE) as the evaluation metric to compare model accuracy. Validation RMSE was computed as the average across 5-fold cross-validation. Models were evaluated using 5, 10, 15, and all 50 features to assess performance across different feature subsets.

Results

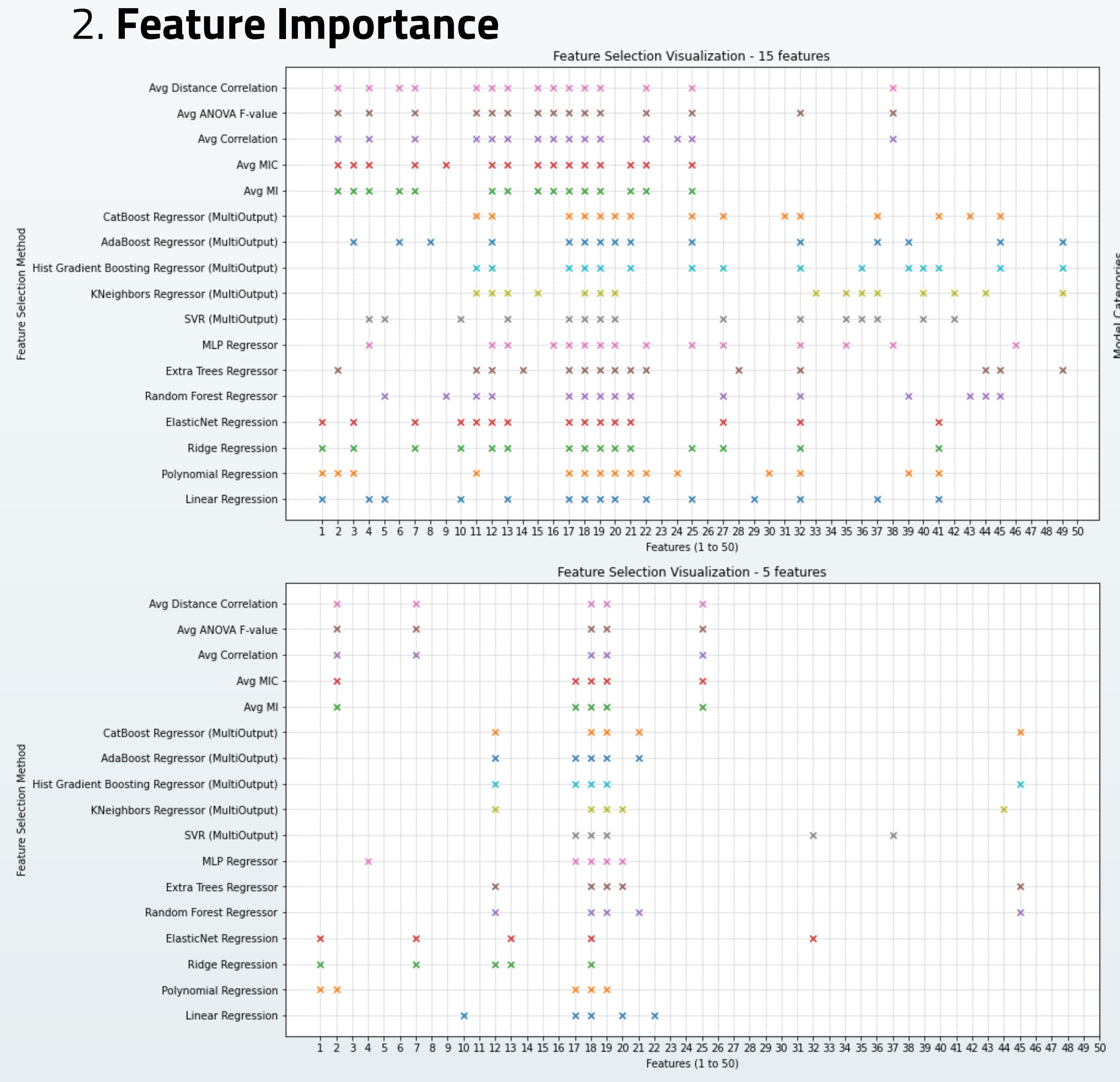
1. Model Performance

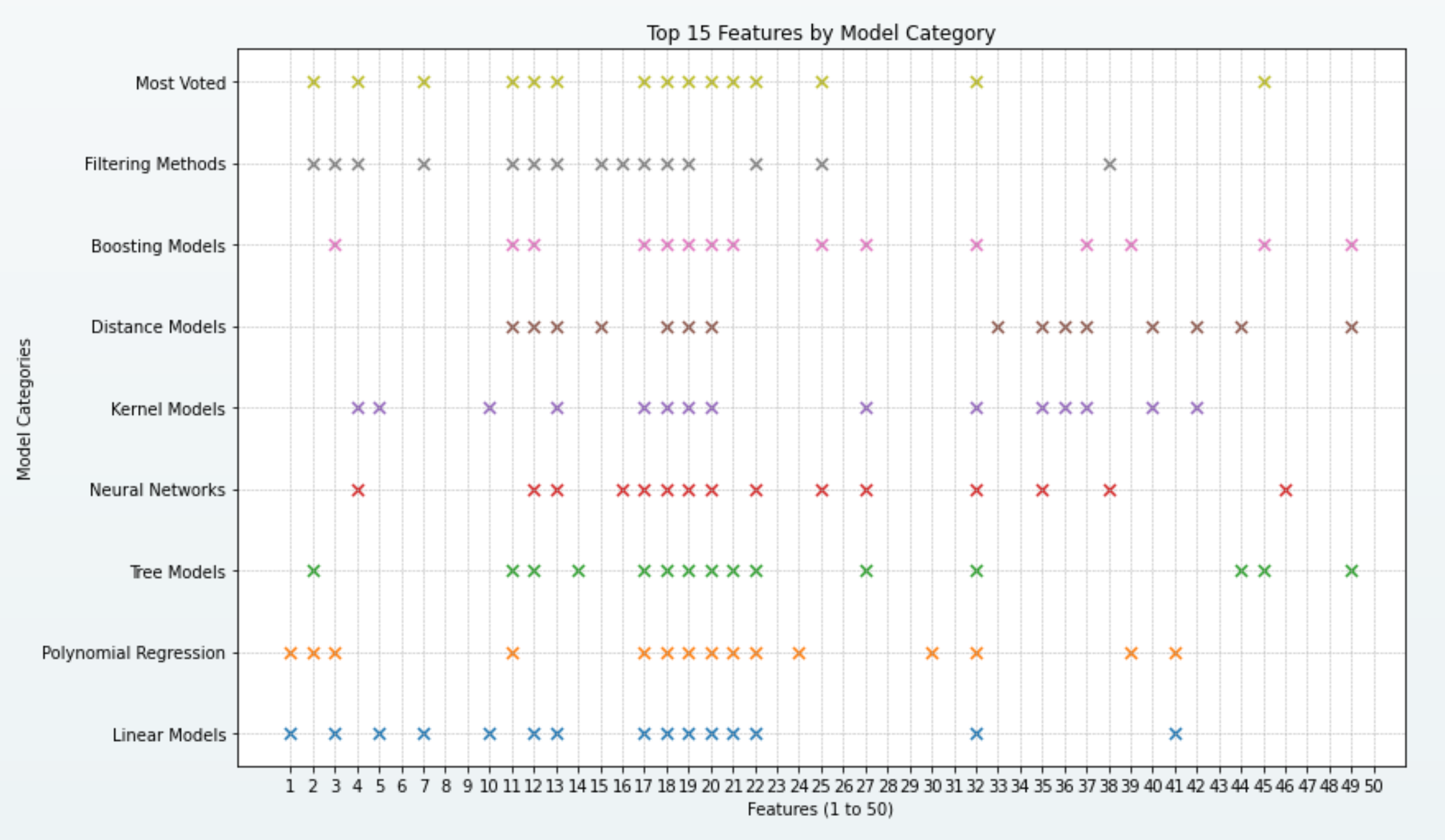




- Tree-based and boosting-based models performed best.
- Extra Trees Regressor achieved the lowest RMSE with just 15 selected features, outperforming even the full feature set.
- Hist Gradient Boosting and CatBoost were close behind.

2. Feature Importance





- These charts show the top features selected by each feature selection approach.
- The Extra Trees Regressor, our best model, selected 11 of the 15 most voted features, reflecting a 73% overlap.

Conclusion

- We successfully used machine learning to predict reactive power setpoints (Q_g) using only a subset of network features.
- Tree-based and boosting-based models, especially Extra Trees, performed best with just 15 selected features.
- Our results show that high accuracy can be achieved with fewer features, reducing computational demands and making the approach scalable for large networks.

Acknowledgements

I would like to thank Dr. Manish Singh for his guidance and mentorship, and the Alvarado Sustainability Scholars Program for supporting this research opportunity.

QR Code for GitHub:

QR Code for Report: