UCWIT2024 Korean Information Science Society Symposium

Development of an AI Model for Optimizing Boiler Efficiency and Identifying Key Influencing Factors

Presented by : Askarova Zhanylmyrza

Authors : Juyeol Park, Bomin Kwon, Heeseong Kim, Donghwan Jeon, Askarova Zhanylmyrza, Seolyoung Jeong, Chigwan Kim

Contents

INTRODUCTION

Background

- Industrial boilers play a key role in generating energy for various industries.
- Existing methods, like KS B 6205, fail to account for real-world dynamic factors and constraints.
- This limits efficiency improvements, affecting costs and carbon emissions.

한국정보과학회
Korean Institute of Information Scientists and Engineers

KNU
KYUNGPOOK NATIONAL UNIVERSITY

kit
Kumoh National Institute of Technology

Motivation

Goal # 1

Develop an AI-based model to predict and optimize boiler efficiency.

Goal # 2

Analyze real-world data(2 months 3-ton industrial boiler) to identify and improve key efficiency factors.

한국정보과학회
Korean Institute of Information Scientists and Engineers

KNU
Kyungpook National University

kit
Kumoh National Institute of Technology

Related Studies

Study #1

Nemitalah Medhat et al. explored AI and ML techniques, like neural networks and genetic algorithms, for boiler optimization.

Study #2

Vladislav Kovalenkov et al. used the Random Forest algorithm to enhance burner efficiency.

Study #3

Pauli Virtanen et al. utilized Python libraries for practical calculations in boiler-related studies.

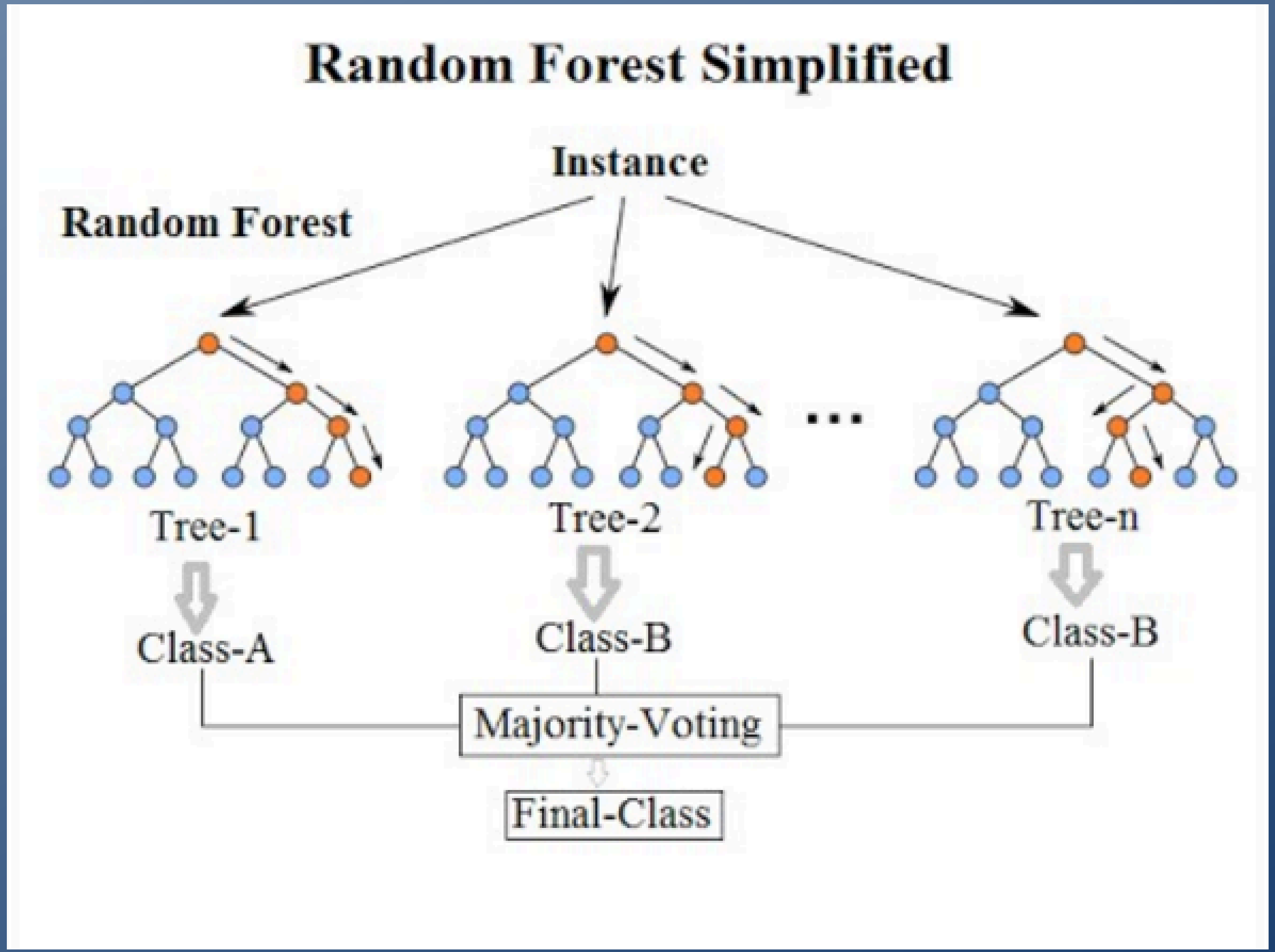
Boiler efficiency checking AI

Models

Random Forest Model

XGBoost

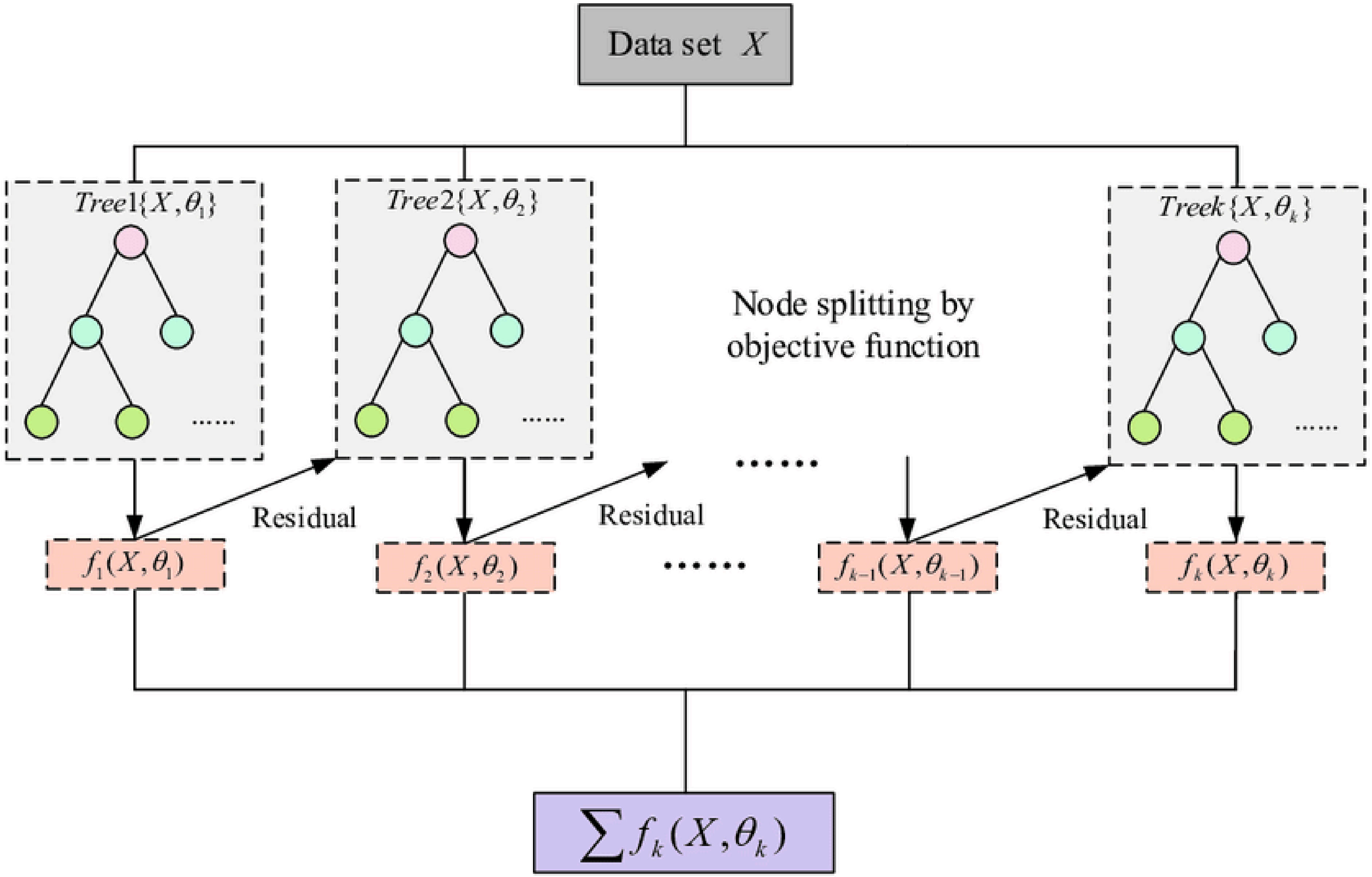
Random Forest Model



Random Forest is a bagging-based ensemble algorithm that creates **multiple decision trees** by sampling data and features. The final prediction is generated by aggregating the **results of individual decision trees**. Random Forest is effective for avoiding overfitting and can handle high-dimensional data.

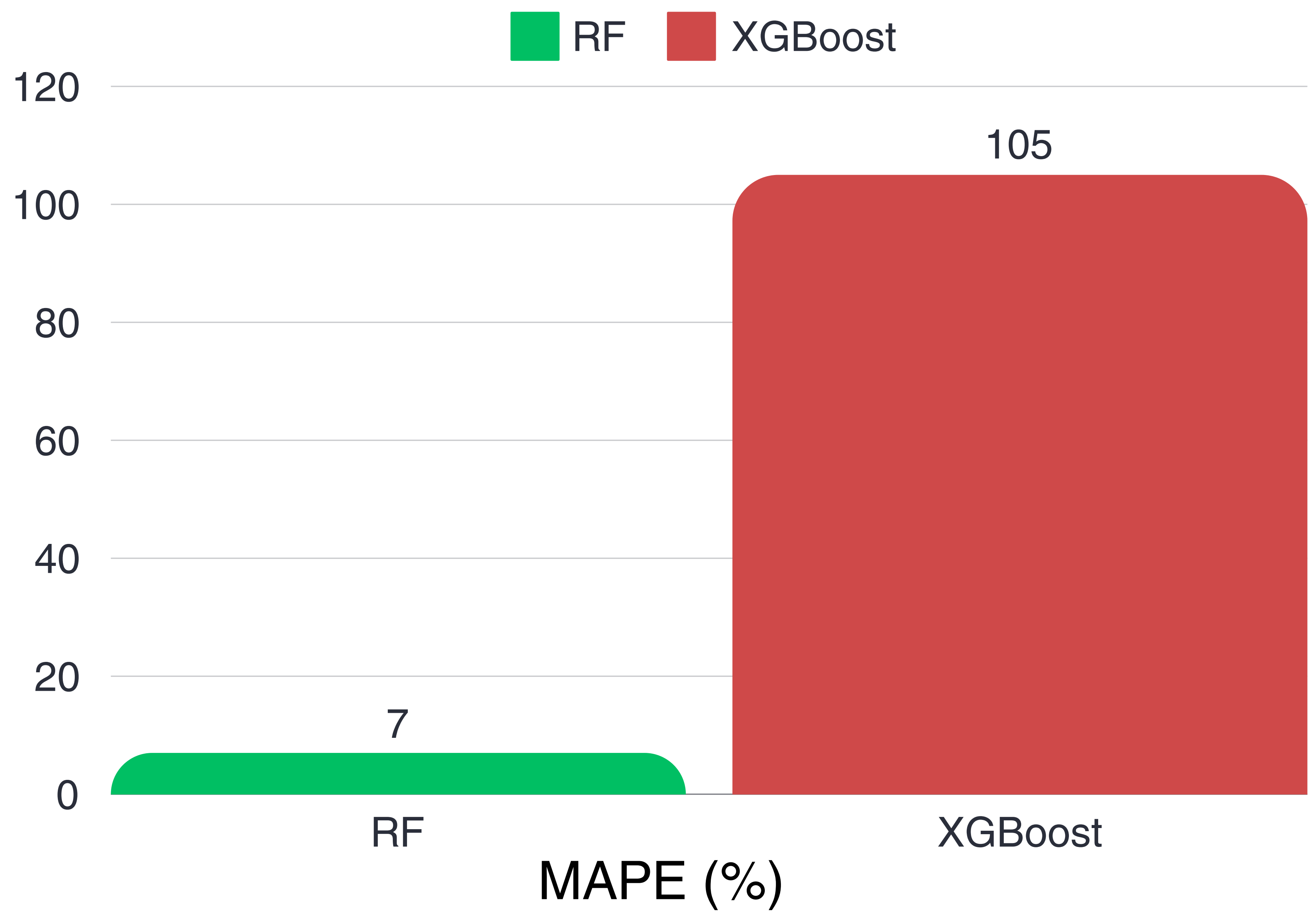
한국정보과학회
Korean Institute of Information Scientists and Engineers
KNU
Kyungpook National University
kit
Kumoh National Institute of Technology

XGBoost Model



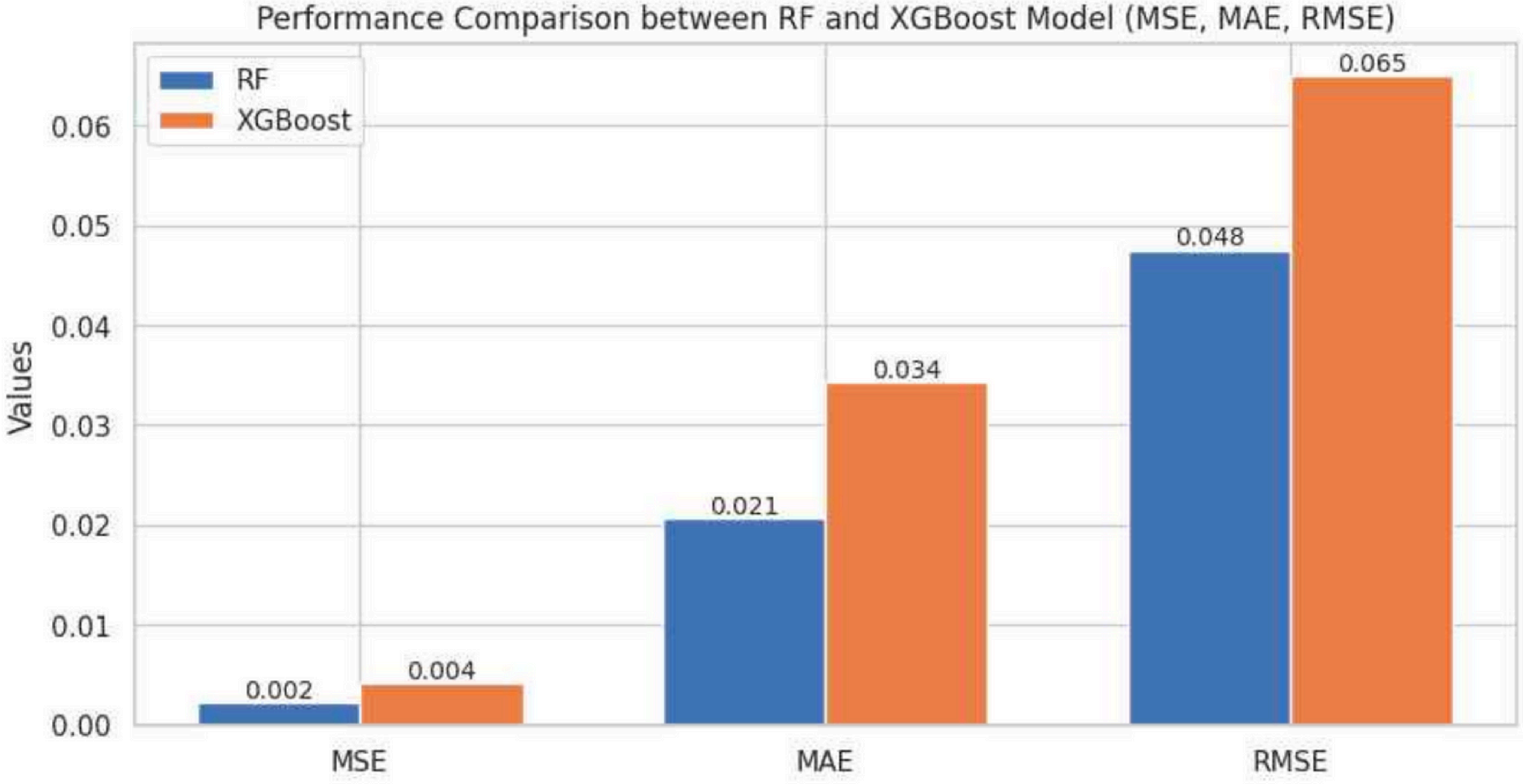
XGBoost (Extreme Gradient Boosting): XGBoost is a gradient boosting algorithm designed to improve model performance through **iterative updates**. It achieves better performance than traditional gradient boosting algorithms by introducing regularization and optimization techniques. It uses techniques like regularization, early stopping, and parallel computation to achieve higher efficiency.

Performance Comparison



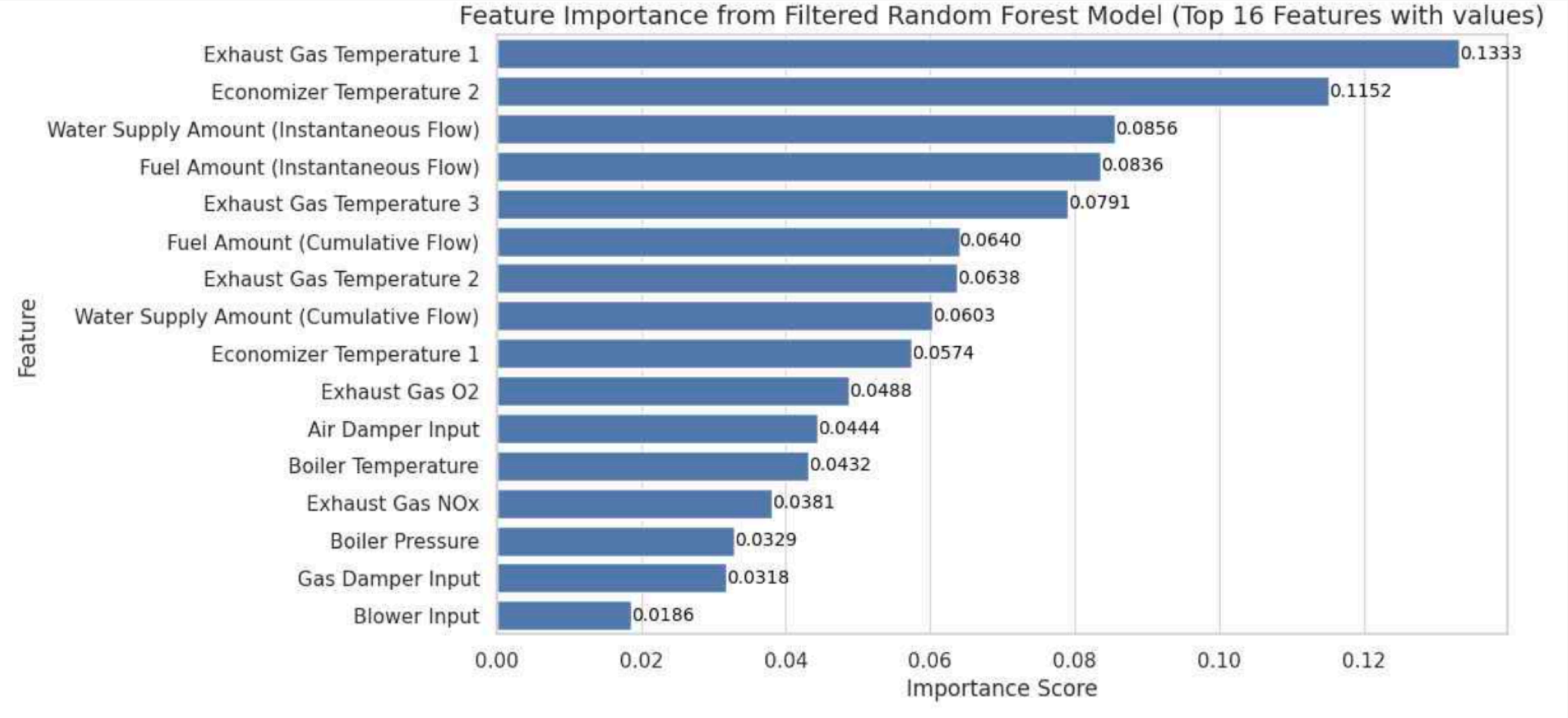
- Metrics Used: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).
- The Random Forest model outperformed XGBoost in MAPE, demonstrating superior predictive accuracy in this study's dataset.

Performance Comparison



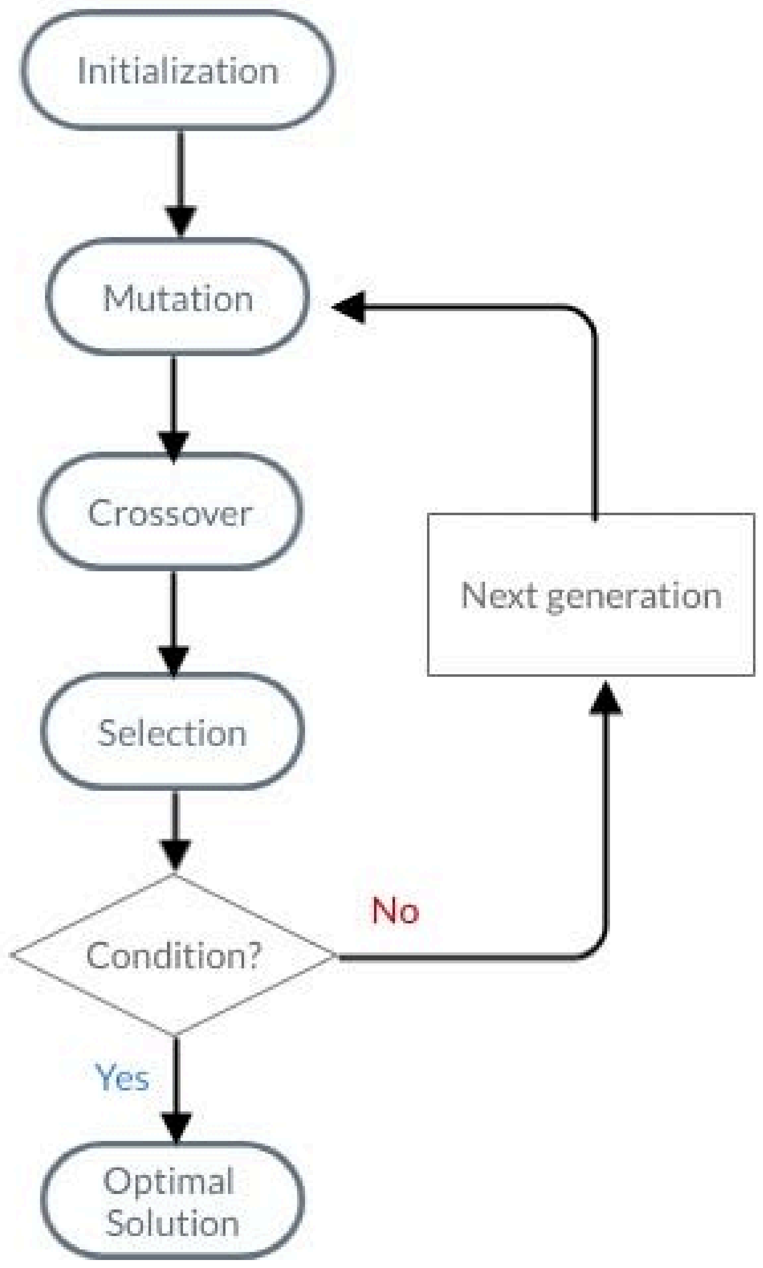
- Metrics Used: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).
- The Random Forest model outperformed XGBoost in MAPE, demonstrating superior predictive accuracy in this study's dataset.

Feature Importance Analysis



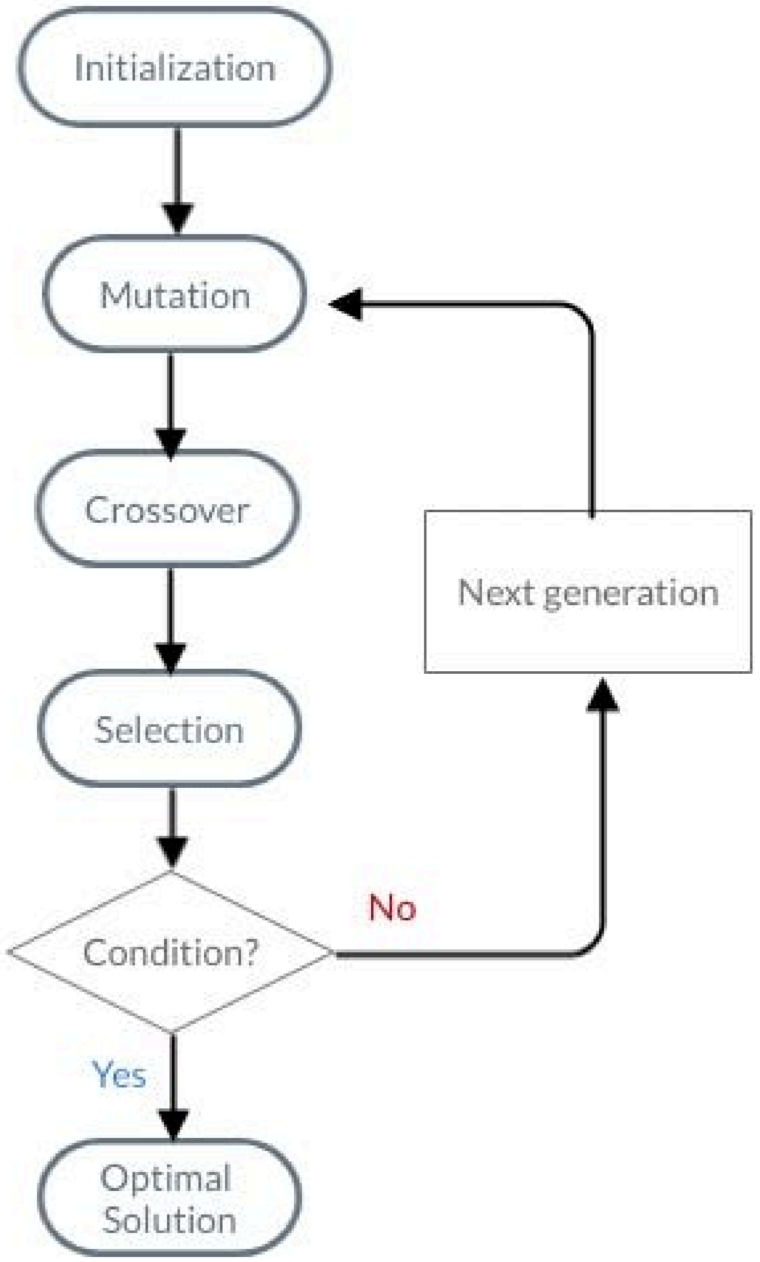
- To identify the most important factors we optimized the RF model through **hyperparameter tuning using Randomized Search**.
- After selecting and using only the top 16 features we improved model's accuracy due to reduced overfitting, noise from irrelevant features

Differential Evolution Algorithm-Based Optimization



The Differential Evolution (DE) algorithm is a population-based optimization technique, particularly effective for solving continuous numerical problems. **The DE algorithm** is inherently designed to **minimize** a given function.Its goal is to find the set of input parameters (in this case, boiler operating parameters) that result in the smallest possible value of the objective function, also called the loss function.

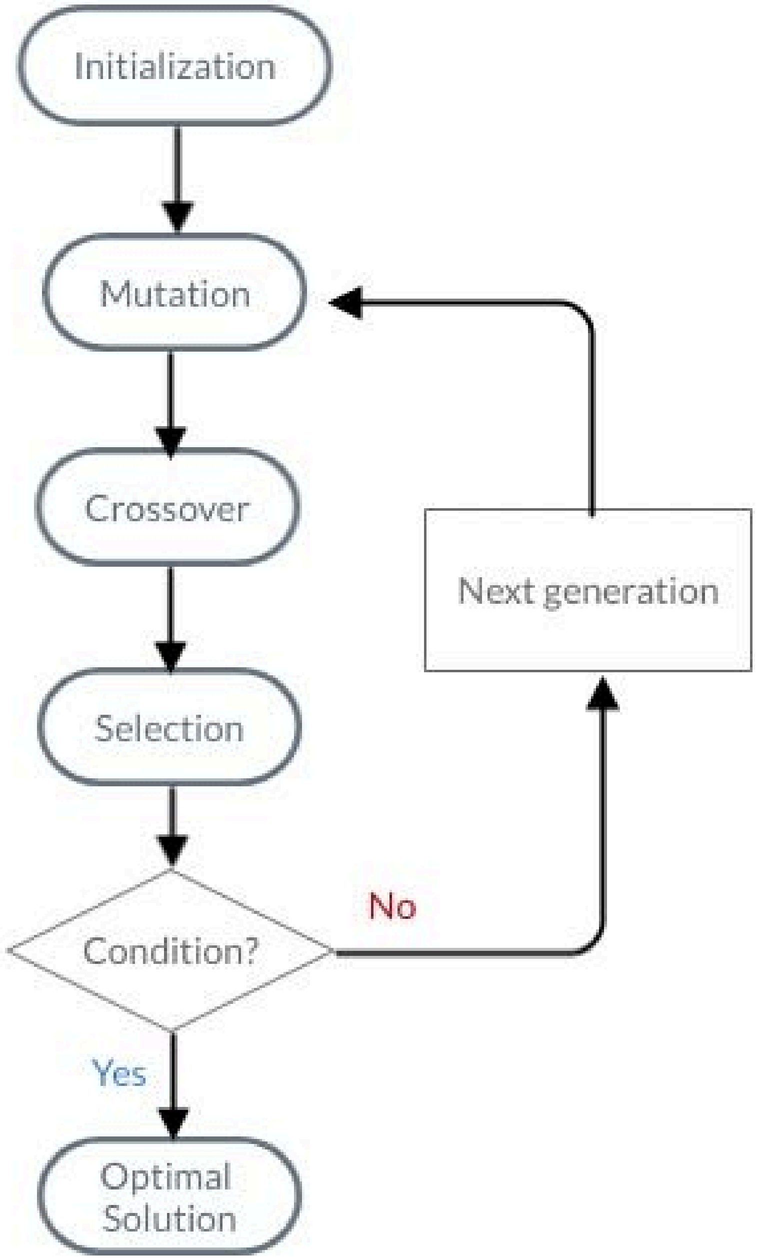
Differential Evolution Algorithm-Based Optimization



It operates through **three key** steps:

1. Mutation: Generating new solutions by combining differences between randomly selected candidates.
2. Crossover: Mixing mutated candidates with existing ones to maintain diversity.
3. Selection: Retaining only the best-performing solutions for the next iteration.

How does it handle minimization?



To handle maximization problems, the goal is simply reframed:

- **Instead of directly maximizing a function f(x), the algorithm minimizes −f(x)**
- This transforms a maximization task into a minimization one, allowing DE to work without modification.
- In our project, the DE algorithm minimizes a loss function calculated as the negative efficiency values, which is equivalent to maximizing positive efficiency values.

If the efficiency of the RF model is 95%, the DE algorithm reverses it to -95%, turning it into a minimization problem. The algorithm then tries different parameter combinations to make -95% even smaller, in other words even higher efficiency.

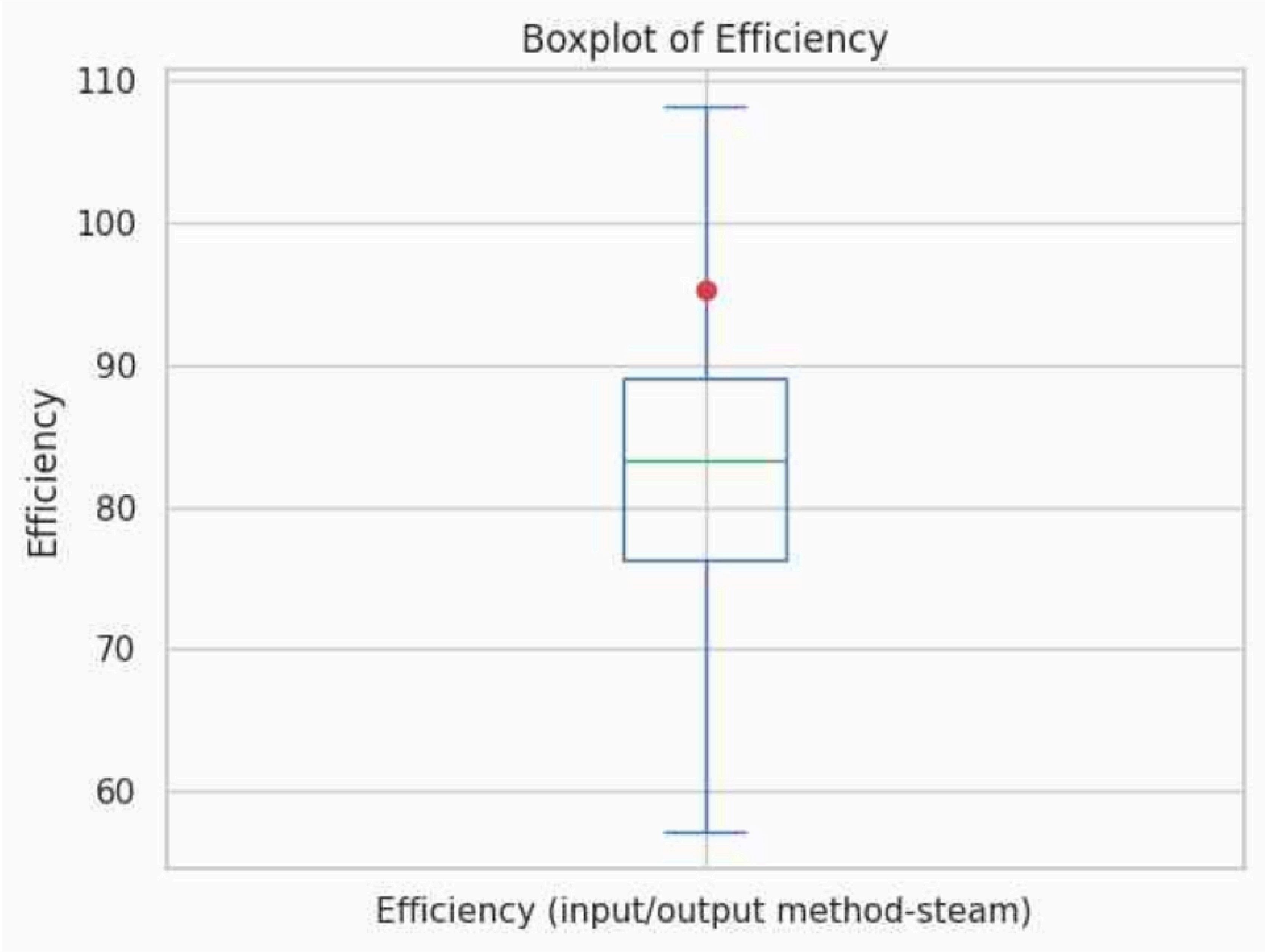
This unique approach ensured the algorithm identified the parameter combinations **that maximized efficiency.**

Reverse Processing

Reverse Processing and Recovery

- Optimized efficiency values are scaled back to their original range for user interpretation.
- Process Overview:
  - Data initialized with zeros to ensure clean processing.
  - Optimized values are inserted into respective columns and rescaled.
  - Results are restored to the original scale for user-friendly analysis.

Result of a boiler efficiency checking model



- Results:

The optimized settings achieved approximately 95% efficiency.
The DE algorithm's integration with the Random Forest model provides a robust framework for achieving practical improvements in boiler efficiency.

Conclusion

Key Findings:

1. The Random Forest model identified more diverse efficiency factors than the KS B 6205 formula.
2. Key Factors:
   - 16 significant variables were identified as primary contributors to boiler efficiency.
3. Optimization:
   - Differential Evolution (DE) model discovered combinations that could yield higher boiler efficiency.
4. Limitations:
   - Real-world physical constraints such as equipment capabilities may prevent achieving the full potential of the optimized results.

Thank you!

Q&A