**Fuel Consumption Prediction Models: Based on Machine Learning and Mathematical Methods** (Changed)

**Changes:**

* **Dataset**

**ML Implementation Design Framework**

**1. Introduction**

The goal of this research is to enhance the accuracy of ship fuel consumption predictions using machine learning (ML) and mathematical modeling. The framework incorporates both white-box (physics-based) and black-box (ML-based) models to provide reliable predictions while ensuring interpretability and robustness.

**Dataset**

* Dataset used: [Kaggle - Ship Fuel Consumption and CO2 Emissions Analysis](https://www.kaggle.com/datasets/jeleeladekunlefijabi/ship-fuel-consumption-and-co2-emissions-analysis?resource=download)

**2. Data Preprocessing**

* Data collected from Kaggle.
* Data cleaning performed using Kwon’s formula to remove anomalies (Z-score thresholding applied to numeric data).
* Key features selected: speed, trim, draft, wind speed, wave height, displacement, power.
* Integration of external weather data to improve prediction accuracy.

**3. Model Development**

**White-Box Model:**

* Based on ship physics and resistance calculations.
* Uses a mathematical formula to estimate fuel consumption.
* Transparent but limited in capturing real-world complexities.

**Black-Box Models:**

* Built using machine learning techniques.
* **Random Forest Model:** Ensemble-based model with decision trees and random sampling.
* **XGBoost Model:** Advanced boosting algorithm improving efficiency and accuracy.
* Hyperparameter tuning performed using Grid Search and cross-validation.

**4. Data Cleaning with Kwon’s Formula**

* Ensures model training uses only steady-state sailing conditions.
* Removes unrealistic values caused by ship acceleration and deceleration.
* Improves ML model reliability and performance.

**5. Model Evaluation**

* The models were tested using various performance metrics to ensure accuracy and reliability.
* **Visualization:** Scatter plots were generated to compare actual vs. predicted fuel consumption for different models.

**Individual Contribution**

**Tejas Vilas Kondhalkar:**

* Introduction
* Data & Preprocessing
* Black-Box Models
* Results & Evaluation
* Conclusion

**Bulbul Kumari:**

* Background & Importance
* White-Box Model
* Data Cleaning with Kwon’s Formula
* Discussion & Future Work

**Gantt Chart for Project Implementation Progress**:



