1. **Introduction and Approach**   
   Understanding and optimizing equipment energy consumption is critical for reducing operational costs and environmental impact. In this project, I leveraged time-series data to build a predictive model that estimates energy consumption using historical patterns, operational variables, and temporal indicators.   
   My approach followed a structured pipeline:   
   **1. Data Loading and Cleaning:**   
   • Imported timestamped sensor data from CSV.   
   • Parsed datetime features and handled missing or corrupted entries.   
   • Removed uninformative variables and encoded categorical fields.   
   **2. Feature Engineering:**   
   • Extracted features such as hour of the day, day of week, month, and weekend indicators.   
   • Created lag features (lag1 energy, lag2 energy) and rolling averages to capture temporal dependencies.   
   **3. Modeling:**   
   • Scaled data using StandardScaler and split into training and test sets without shuffling (preserving time order).  
    • Tuned a Random Forest Regressor using TimeSeriesSplit and GridSearchCV.   
   • Developed a Stacking Regressor with Random Forest, XGBoost, and Gradient Boosting as base models, and Ridge regression as meta-model.  
   **4. Evaluation and Visualization:**   
   • Evaluated the model using RMSE, MAE, and R2 metrics.   
   • Generated feature importance and prediction vs. actual plots for interpretation.
2. **Data Insights**   
   From the processed dataset, several patterns emerged:  
   **2.1 Feature Importance**   
   The Random Forest model revealed the most impactful predictors:   
   • Lag features: Previous energy values (lag1.energy, lag2.energy) were highly predictive  
   • Rolling mean: The 3-hour rolling average helped smooth short-term fluctuations.   
   • Temporal variables: Hour of day and day of week played critical roles in characterizing operational cycles.  
   **2.2 Temporal Influence**  
   • Energy consumption showed clear trends tied to the hour of the day, with peak usage typically observed during business hours.   
   • Weekday vs. weekend behavior differed significantly, with weekends showing lower energy use.   
   • Monthly trends suggested seasonality, possibly due to climate-driven operations or maintenance schedules.
3. **Model Performance Evaluation**   
   I evaluated the ensemble model on a hold-out test set using the following metrics:   
   • Root Mean Squared Error **(RMSE): 86.39**   
   • Mean Absolute Error **(MAE): 28.49**• R2 Score: **0.768**   
   These results indicate that the model captures the variance in energy consumption well, with minimal overfitting due to the use of cross-validation and model ensembling.