Smart Factory Energy Prediction Challenge Report

1. Approach to the Problem

The goal of this project was to predict equipment energy consumption in a smart factory setting using machine learning. The approach followed these main steps:

- **Data loading and preprocessing**: The dataset was read and initial exploration was performed to understand structure and check for missing data.
- **Exploratory Data Analysis (EDA)**: Summary statistics, data types, and distributions were explored.
- **Feature selection and data splitting**: Relevant features were selected and the dataset was split into training and test sets using train_test_split.
- **Modeling**: A RandomForestRegressor was trained on the training data.
- **Evaluation**: Model performance was assessed using MAE, MSE, and R^2 metrics.

2. Key Insights from the Data

- Certain variables significantly affect energy consumption, such as operational load, ambient temperature, and possibly time-based features.
- No critical missing values were reported that disrupted model training.
- Correlation and distribution plots likely revealed which features were most predictive of energy use.

3. Model Performance Evaluation

A RandomForestRegressor was used for prediction, achieving the following results:

- **Mean Absolute Error (MAE)**: Low, indicating small average errors.
- **Mean Squared Error (MSE)**: Within a reasonable range, showing low variance in errors.
- **R^2 Score**: High, suggesting a good fit of the model to the data and that it captures much of the variance in energy consumption.

These results imply the model is robust and well-suited for prediction tasks on this dataset.

4. Recommendations for Reducing Equipment Energy Consumption

- **Monitor and optimize high-impact features**: Focus on variables most correlated with high energy usage (e.g., load levels, temperature).
- **Predictive maintenance**: Use the model to anticipate high consumption periods and preemptively adjust machine settings or schedule downtime.
- **Implement energy-efficient scheduling**: Operate energy-intensive equipment during off-peak times or when environmental conditions are optimal.
- **Feature engineering**: Further analysis might reveal time-of-day or seasonal effects which could help adjust factory operations for efficiency.