**📝 Meeting Report**

**Date:** April 24, 2025  
**Subject:** Project Progress – Power Consumption Model  
**Attendees:** Project Team

**1. Data Adequacy Review**

* The current dataset sourced from Kaggle was evaluated for completeness and relevance.
* It was found to be **inadequate** in terms of granularity and volume, particularly for building a reliable machine learning model.
* As a result, it was concluded that **additional synthetic data** must be generated to move forward effectively.

**2. Project Objective**

* The primary goal of the project is to **develop a model that identifies households with higher energy consumption**.
* The analysis will be based on features such as:
  + Date and Time
  + House ID
  + Total Power Consumed (kV)
  + Meter Readings
  + Location
  + Breakdown of consumption across categories:
    - Entertainment
    - White Goods
    - Lighting
    - Air Conditioners
    - EV Items

**3. Plan for the Upcoming 2 Days**

* **Generate Data:** Create detailed, realistic power consumption data for **50 individual houses**.
* **Configure House Profiles:** Each house will be assigned unique characteristics (location, usage patterns, etc.) to simulate real-world diversity.
* **Prepare Dataset:** The synthetic dataset will be structured and stored in CSV format for model training and testing.

Date: April 28, 2025  
Topic: Data Generation Code Improvement

This is a brief summary of our meeting held on April 28, 2025 regarding improvements to the data generation process.

Key Update:  
We discussed and implemented changes to the entire codebase responsible for synthetic data generation. The primary improvement was in changing the data types and logic used in the script to enhance the quality and realism of the generated data.

**Meeting Report – April 29**

**Key Updates:**

1. **Enhancements to Dataset Structure:**
   * Added new columns including **date range** to support more robust model training and evaluation.
   * Ensured the dataset now better reflects real-world usage intervals and tracking needs.
2. **Data Quality Improvements:**
   * Addressed and **rectified issues** in the previously generated data (e.g., timestamp inconsistencies, logical errors in usage simulation).
   * Cleaned and validated appliance usage patterns for improved reliability.
3. **Dataset Optimization:**
   * Refined column organization and added **category-based summaries** (e.g., white\_goods, entertainment, lighting, etc.).
   * Enhanced compatibility with downstream **analytics and machine learning workflows**.

**Next Steps:**

* Begin integration with model training pipeline.
* Continue validating outputs with domain benchmarks.

**Meeting Report – May 1, 2025**

**Date:** May 1, 2025  
**Participants:** [Your Team Name or Attendees]  
**Duration:** [Optional – e.g., 30 mins]  
**Facilitator:** [Optional – e.g., Your Name]

**✅ Key Updates & Actions**

1. **GitHub Repository Created**
   * A new GitHub repository has been set up to centralize project collaboration and version control.
2. **Improved Data Quality**
   * Enhancements were made to increase the level of information in the dataset, leading to improved data quality and context.
3. **Organized Project Structure**
   * Separate folders were created for scripts, data, and configurations to streamline file management and development workflows.
4. **Data Generator Code Updated**
   * Modifications were made to the data generation script to support:
     + More accurate energy consumption logic
     + Proper handling of time-based usage
     + Structured CSV outputs with per-category and per-device detail

**Meeting Report – Key Points**

**Date:** 5 may  
**Topic:** Updates on Training Dataset Preparation & Configuration Integration

**✅ 1. Refactoring Training Data Preparation**

* **Time Feature Engineering & Encoding**
  + Moved processing of time (conversion to seconds, cyclical features) and date\_range encoding directly into the training dataset preparation script.
  + Ensures cleaner separation of concerns and simplifies the training pipeline.

**🔄 2. Configuration Enhancements**

* **Month-wise Variations Added**
  + Introduced month\_variations in house configuration files to reflect realistic seasonal electricity consumption trends.
  + Example: Higher usage in winter (January: 1.2), lower in summer (May: 0.8).
* **Lighting Usage Adjustments**
  + Refined lighting schedules in configurations to better reflect actual household behavior.
* **New Seasonal Configuration File**
  + Added a dedicated file for seasonal changes, enabling flexible simulation control based on the time of year.

**🛠️ 3. Generator Code Modifications**

* **Adapted Simulation Logic**
  + Updated the generator script to read and apply month\_variations from the configuration.
  + Ensures monthly adjustments are correctly applied during energy consumption simulation.

**Meeting Report – Power Consumption Forecasting System**

**Date:** *6 May*  
**Topic:** System Updates and Enhancements

**✅ Key Points Discussed**

1. **📦 New LSTM Model Integration**
   * A new LSTM-based model was successfully created and trained.
   * It uses sequential time-windowed data with cyclical features for improved temporal learning.
   * Evaluation showed significant improvement in predictive performance (R² up to ~0.89).
2. **🧩 Configuration Enhancements**
   * Seasonal configuration logic was finalized and corrected for Indian seasonal mapping.
   * Two new appliances were added to the simulation: **mortar** and **water heater**.
   * Updated power ratings were integrated into the generator.
3. **⚙️ Generator Code Update**
   * The data generator was updated to:
     + Reflect the new device types.
     + Incorporate seasonal multipliers and peak-hour adjustments per device.
   * Enhanced time handling and seasonal logic now simulate data for the entire year.
4. **📈 Training Pipeline Adjustments**
   * The trainer module was modified to accommodate:
     + New feature columns from the updated data format.
     + Normalization of time and power-related inputs.
     + Output labels now reflect the expanded device category list.

Meeting Report

Date: *9 May*

Key Points Discussed

1. Model Selection
   * Comparative evaluation of models: LSTM and Random Forest Regressor.
   * LSTM was favored for its sequential learning capability with time-based features.
   * Random Forest performed well with tabular features and achieved a strong R², indicating robust generalization.
   * Decision made to include both models in the documentation, highlighting their pros and cons.
2. College Documentation
   * Final structure for the written report was discussed:
     + Abstract, Problem Statement, Data Pipeline, Model Architecture, Evaluation, and Results.
     + Detailed explanation of preprocessing, seasonal adjustments, and appliance simulation logic.
     + Screenshots and diagrams (like the system pipeline) to be included.
   * Assignment of sections and deadlines for drafting was agreed upon.
3. Project Deployment
   * Discussed possible deployment options: local server vs cloud
   * Next steps involve setting up:
     + A prediction endpoint using the trained model.
     + A basic front-end interface for testing.
   * Deployment will be part of the final demonstration and submission package.