Requirement Gathering and Preliminary Analysis

Project Name : EcoTrack5G
Mentor : Dr. Charith Chitraranjan
Team Name : Group 07
Names and registration numbers of the team members:
KAJAANI B. 200279N
KARUNARATHNA J.P.N.D. 200294F
KARUNATHILAKE R.M.S.K. 200296M
Date : 16/08/2023

1. Data Requirements :

The model will estimate energy consumption for various base station products, considering architectural differences, configuration parameters, traffic conditions, and energy-saving methods. The goal is to achieve generalization capabilities across diverse products and configurations, leading to more energy-efficient 5G deployments.

Key Features/Variables Needed:

1. Base Station Information: BS name, CellName, RUType, Mode, Frequency, Bandwidth, Antennas, TXpower.
2. Cell-Level Data: Time, BS name, CellName, Load, ESMode [1-6].
3. Energy Consumption Data: Time, BS name, Energy.

2. Data Sources Identified:

Primary Data Source	Description	Rationale & Relevance to Project
Base Station Basic Information (BSinfo.csv)	Configuration parameters and hardware attributes of base stations.	Understanding base station attributes and their impact on energy consumption aligns with project goals.
Cell-Level Data (CLdata.csv)	Hourly counters related to service compliance and energy-saving methods.	Assessing load patterns and energy-saving activations aids energy-efficient model development.
Energy Consumption Data (ECdata.csv)	Hourly energy consumption measurements for specific base stations.	Essential for training the predictive model and evaluating its performance.

3. Preliminary Data Analysis:

Initial Observations:

- Datasets are complete, providing a solid foundation for analysis and model development.
- Features like 'BS,' 'CellName,' 'Mode,' and 'RUType' are categorical, needing proper categorical encoding for modeling.
- 'Time' column transformation to date type enables temporal analysis and feature engineering.

- The test set includes only 'Time' and 'BS,' requiring handling of disparity for accurate model predictions.

Challenges and Issues:

- The presence of categorical variables poses challenges for model development and generalization.
- Achieving cross-product and configuration generalization is necessary for accurate energy consumption estimation.
- The disparity between training and test sets introduces complexity, demanding effective feature incorporation.
- Handling new categorical values in the testing phase requires careful consideration and encoding strategies.

4. Stakeholder Consultation:

Stakeholders emphasized the importance of

- Developing a model with robust generalization capabilities across diverse base station products and configurations.
- A comprehensive dashboard for energy consumption patterns.
- Automatic energy-saving strategies and alerts for timely actions.

To meet stakeholder expectations, we will

- Focus on enhancing model generalization through comprehensive training on various base station products and configurations.
- Prioritize the development of an interactive and informative dashboard.
- Integrate energy-saving strategy suggestions and automated alert notifications.

5. Evaluation

- Model's accurate energy estimation for diverse base stations, cross-equipment, and cross-configuration generalization achieved.
- WMAPE and other training metrics are used to assess model performance.
- Interactive dashboard evaluated for usability.
- Effectiveness in real-world energy management, energy-saving suggestions, and timely alerts are considered.

6. Updated Project Backlog & User Stories :

Project Backlog

- Data Collection and Preprocessing
- Feature Engineering
- Model Selection and Training
- Cross-Generalization Enhancement
- Model Evaluation and Refinement
- Dashboard Development
- Documentation and Reporting

User Stories

1. Estimating Energy-Efficient Configuration Impact:
    - Users can input different configuration settings, such as RUType, Mode, Frequency, and Bandwidth.
    - The system calculates and displays the estimated energy consumption for the given configuration, aiding informed decisions for energy-efficient deployment.
2. Real-time Energy Monitoring and Alerts:
    - Users can access an interactive energy dashboard displaying real-time energy consumption trends and energy-saving action suggestions.
    - The system triggers automated alerts for unexpected energy spikes or unusual consumption patterns, enabling timely corrective actions.

3. Comparing Energy-Saving Strategies:
    - Users can input different energy-saving modes (ESMode) and evaluate their impact on energy consumption.
    - The system provides a comparative analysis, allowing users to select the most efficient energy-saving strategy for specific operational scenarios.

7. Next Steps :

Data Preparation, Exploration, and Model Development (Week 1 - Week 2):

- Encode variables, and transform 'Time' for analysis.
- Conduct EDA for insights into distributions and patterns.
- Train multiple regression models, including a decision tree and a neural network.
- Optimize models using grid search, and cross-validation.
- Evaluate using WMAPE, MAE, RMSE, and MPE.

Dashboard, and Alerts (Week 2):

- Integrate the dashboard, and gather feedback.
- Initiate alerts integration for thresholds.