CS6140 Machine Learning Project Proposal

Analyzing Household Responses to Dynamic Time-of-Use Pricing Signals

Ishan Biswas (002441977)
Divyank Singh (002338817)
Sri Sai Teja Mettu Srinivas (002313888)

February 21, 2025

Abstract

This project proposes to analyze whether certain London households under dynamic Time-of-Use (dToU) pricing exhibit abnormal responses to high or low tariff signals, potentially indicating inefficiency or non-adherence to price signals. This study aims at performing analysis, visualizing results to identify anomalous behavior, and train Machine Learning models on the open-source smart meter energy use dataset from London households. This proposal outlines the project objectives, data source, workload distribution, and evaluation methodology.

1 Introduction

Dynamic Time-of-Use (dToU) pricing is an innovative tariff structure designed to encourage energy conservation by charging higher rates during peak demand periods and lower rates during off-peak times. However, not all households respond optimally to these price signals, which may lead to inefficiencies or non-adherence. This project aims to use machine learning techniques to analyze electricity consumption during a timeframe in the dToU pricing scheme and identify anomalies in that duration of time.

2 Project Objectives

The primary objectives of this project are:

- To visualise energy consumption over the period of time and identify anomalies using visualisation libraries such as Pandas/Matplotlib or Tableau
- Apply Machine Learning techniques to detect abnormal responses (e.g., increased consumption during high tariffs or minimal reduction during low tariffs).
- Evaluate whether the Machine Learning models confer with the visualisation results
- Recommend actionable insights for policymakers to optimize dToU pricing strategies to consider inefficiency, non-adherence, or other behavioral factors.

3 Data Source

The data is sourced from the publicly available *Smart Meter Energy Use Data in London Households* dataset[1]. The dataset is maintained by the Greater London Authority, and widely used for research purposes. It has details on time-stamped usage data, household characteristics, and dToU pricing information. We will ensure compliance with any data usage terms and ethical considerations outlined on the website.

4 Workload Distribution

To ensure a fair and efficient distribution of tasks, the following table outlines the responsibilities of each group member:

Table 1: Workload Distribution Among Group Members

Member	**Responsibilities**
Sai Teja	1. EDA with Pandas, 2. three algorithm selection 3. Implement ML algorithm1, 3. presentation preparation.
Ishan	1. data preprocessing and wrangling, 2. implement ML algorithm2, 3. prepare report
Divyank	1. initial visualization to identify anomalies, 2. implement ML algorithm3, 3. prepare report

5 Evaluation Methodology

We will evaluate the project using the following methodology, which is suitable for the objectives and dataset:

- **Data Preprocessing**: Clean and preprocess the dataset, handling missing values and outliers, normalize consumption patterns.
- **Feature Engineering**: Extract relevant features such as time of day, day of week, tariff levels, and household characteristics to train models.
- **Model Selection and Training**: Use unsupervised learning (e.g., K-means) to identify clusters of normal and abnormal consumption patterns. Additionally, use supervised learning (e.g., Random Forest and Neural Networks) if labeled data on adherence/non-adherence is available.
- **Evaluation Metrics**: For clustering- silhouette score, for supervised learning- accuracy, precision, recall, and F1-score. Additionally, we will qualitatively analyze the interpretability of identified patterns.
- **Validation**: Cross-validation and also compare results against baseline models (e.g., simple rule-based thresholds for normal/abnormal behavior).

6 Expected Outcomes

1. Identify distinct clusters or patterns of household behavior, quantify the proportion of households exhibiting abnormal responses
2. Provide insights into potential causes- lack of awareness, technical issues, or behavioral preferences
3. Visualization using graphs (e.g., time-series plots, cluster distributions)
4. Summarise findings, model performance, and recommendations in a final report.

References

[1] Greater London Authority. (n.d.). Smart Meter Energy Use Data in London Households. Retrieved from https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households.