Analyzing Household Energy Consumption Anomalies using Machine Learning

Ishan Biswas
Khoury College of Computer Sciences
Northeastern University
Boston, USA
biswas.is@northeastern.edu

Sri Sai Teja Mettu Srinivas
Khoury College of Computer Sciences
Northeastern University
Boston, USA
mettusrinivas.s@northeastern.edu

Divyank Singh
Khoury College of Computer Sciences
Northeastern University
Boston, USA
singh.divya@northeastern.edu

***Abstract*—Dynamic Time-of-Use (dToU) pricing scheme aims to optimize energy consumption patterns by incentivizing users to shift usage away from peak demand periods. However, the efficacy of these schemes depends on the response of consumers, which varies significantly between households. This paper presents a machine learning framework for detecting and classifying anomalous energy consumption behaviors in response to dToU pricing signals. Using data collected from smart meters in London, we first identify potential consumption anomalies using the Isolation Forest algorithm. These labeled anomalies then serve as ground truth for training supervised LightGBM and Neural Network models to classify anomalous behavior based on temporal and consumption features. Our experimental results demonstrate that both supervised approaches achieve high classification accuracy (92% and 90.3% respectively), with LightGBM slightly outperforming the Neural Network model. Analysis of detected anomalies reveals seasonal patterns with higher anomaly rates during winter months, suggesting that household responses to dToU pricing are influenced by weather-dependent consumption needs. This work provides valuable information for energy providers and policymakers looking to optimize dToU pricing strategies and improve consumer engagement with demand response initiatives.**

***Index Terms*—Anomaly Detection, SmartMeter Technology, Energy Consumption, Isolation Forest, LightGBM, Neural Networks, Time Series Analysis, dToU Pricing**

I. Introduction

Dynamic Time-of-Use (dToU) pricing represents an effort to manage energy demand by incentivizing consumers to shift consumption away from peak hours through variable tariffs [2]. Although effective for many, some households may exhibit anomalous responses that deviate from expected behavior, such as high consumption during peak prices or minimal change during off-peak periods. Identifying such anomalies is crucial to understanding the effectiveness of dToU schemes and to refine energy policies.

This project leverages machine learning techniques applied to a large dataset. Our primary goal is to identify anomalous consumption patterns within the subset of households under dToU tariffs using an unsupervised approach (Isolation Forest). We then treat these identified anomalies as ground truth labels to train and evaluate supervised approaches (LightGBM and Neural Network) capable of predicting anomalous behavior based on consumption and temporal features. This allows for a deeper analysis of the characteristics associated with the anomalies.

II. Objectives and Expected Outcomes

A. Objectives

- Visualize household energy consumption over time, focusing on the dToU subset, to identify anomalies.
- Detect anomalies using an unsupervised machine learning algorithm, Isolation Forest
- Apply supervised learning techniques LightGBM and Neural Networks to classify the identified anomalies.
- Evaluate the performance of both approaches.
- Analyze the characteristics and temporal patterns of the detected anomalies.

B. Expected Outcomes

- Quantify anomalies detected by the Isolation Forest algorithm.
- Performance metrics (e.g., accuracy, confusion matrix) for supervised algorithms.
- Visualize consumption patterns, anomaly distributions, and model results.
- Insights into the types of consumption behavior classified as anomalous by the models.

III. Literature Review

The detection and analysis of anomalous energy consumption patterns has gained significant research attention with the proliferation of smart metering technologies. This section presents a thematic and comparative analysis of relevant literature, organizing research into key themes while critically evaluating methodological approaches and findings.

A. Smart Metering and Dynamic Pricing Schemes

1) Evolution of Smart Metering Technology: Smart meters have transformed energy consumption monitoring, enabling fine-grained analysis of household responses to pricing signals. A foundational study by Faruqui and Sergici [2] analyzed 15 experiments on household responses to dynamic pricing, revealing that while consumers generally respond to price signals, the magnitude varies significantly between households. This variability in response behavior creates a natural

segmentation between households that adapt efficiently to pricing signals and those exhibiting anomalous consumption patterns.

In contrast to Faruqui's focus on aggregate responses, Zhou et al. [16] investigated individual household behavior, demonstrating that response patterns to ToU pricing vary substantially based on demographic factors and appliance ownership. Their findings suggested that anomaly detection might be more effective when contextualized within household-specific baseline consumption profiles, an approach adopted in our methodology.

2) Consumer Behavior Under Variable Pricing: Several researchers have investigated responses to Time-of-Use (ToU) pricing. Torriti [5] examined household activities in relation to ToU pricing using time use survey data, while Nicolson et al. [6] explored consumer preferences for different ToU tariff structures through controlled experiments. These studies focused primarily on the willingness to adopt ToU pricing rather than anomalous responses after adoption.

Taking a different approach, Himeur et al. [17] provided a comprehensive review of energy consumption patterns, noting that anomaly detection could prevent "minor problems from becoming overwhelming" and promote sustainable energy usage. However, they identified critical gaps including the absence of precise definitions for anomalous power consumption and annotated datasets, challenges our work aims to address through a hybrid unsupervised-supervised methodology.

B. Methodological Approaches to Anomaly Detection

1) Unsupervised Approaches: The research landscape for anomaly detection in energy consumption data reveals two dominant methodological streams: unsupervised and supervised approaches. Within unsupervised techniques, distance-based, density-based, and isolation-based methods have emerged as particularly relevant.

Isolation Forest, introduced by Liu et al. [3], has demonstrated superior performance for anomaly detection in various domains due to its ability to isolate outliers efficiently using random partitioning. This algorithm's computational efficiency (with linear time complexity) gives it a distinct advantage over density-based methods like DBSCAN and LOF, which typically scale quadratically with dataset size and struggle with high-dimensional data. Our work leverages this efficiency advantage, as smart meter datasets often contain millions of records.

Alternative unsupervised approaches include autoencoders and clustering-based methods. Sakurada and Yairi [7] proposed using autoencoders for anomaly detection by measuring the reconstruction error, while Peña et al. [8] employed clustering techniques to identify abnormal consumption patterns. These approaches, while effective for certain data characteristics, typically require more hyperparameter tuning than Isolation Forest and often lack the interpretability that is critical for energy domain applications. In our preliminary experiments, we found Isolation Forest provided more consistent results with less tuning than autoencoder-based approaches.

2) Supervised and Hybrid Classification Approaches: Supervised learning approaches for anomaly classification have evolved significantly, with ensemble methods demonstrating particular effectiveness. The LightGBM framework developed by Ke et al. [4] represents a significant advancement in gradient-boosting implementations, offering superior speed and memory efficiency compared to earlier approaches like XGBoost and traditional Random Forests. These advantages make LightGBM particularly suitable for large datasets.

Neural network architectures have also been widely applied to analysis of energy consumption. Although recurrent architectures such as LSTMs and GRUs excel at capturing temporal dependencies in consumption time series, feedforward networks with appropriate feature engineering can achieve competitive performance with lower computational requirements, as demonstrated by our study and corroborated by comparative analyses in the literature. Buzau et al. [18] demonstrated that LSTM-based classifiers outperformed traditional ML methods in the detection of electrical load anomalies, although at a significantly higher computational cost.

Hybrid approaches that combine unsupervised anomaly detection with supervised classification have emerged as a promising research direction. Wang et al. [9] employed a hybrid methodology using the entropy weight method with the Isolation Forest, demonstrating improved detection accuracy compared to single-algorithm approaches. Similarly, Lim et al. [10] combined the Isolation Forest for initial labeling with Gaussian Naïve Bayes for subsequent classification, achieving balanced accuracy scores above 89% in the detection of Smart Meter anomalies.

Our work extends these hybrid approaches by combining Isolation Forest with both LightGBM and Neural Networks, providing a comparative analysis of their effectiveness for classifying energy consumption anomalies in the specific context of dToU pricing responses.

C. Feature Engineering and Contextual Factors

The literature reveals significant differences in feature engineering approaches for energy consumption analysis. Temporal features (hour, day, month) are consistently used in studies, but their transformation and encoding vary considerably. Fernandes et al. [11] demonstrated the importance of incorporating contextual factors such as "outside temperature, solar intensity, the day of the week, the hour of the day, the building occupancy" for accurate anomaly detection in hotel energy consumption.

Weather and seasonal features have been incorporated with varying degrees of sophistication. Jiang et al. [12] showed that seasonal factors significantly impact the prevalence of anomalies, with the winter months typically showing more irregular consumption patterns due to heating requirements. This aligns with our findings on seasonal patterns in anomaly rates and highlights the importance of incorporating temporal context in anomaly detection models.

A critical gap in the existing literature concerns the specific features most indicative of anomalous responses to dToU

pricing. Although general anomaly detection features have been well studied, the characteristics that distinguish normal from anomalous responses to pricing signals remain underexplored. Our work contributes to filling this gap by analyzing feature importance in supervised models trained on the labels generated by the Isolation Forest algorithm.

D. Application Domains and Real-world Implications

Research on energy consumption anomaly detection spans various application domains, from fraud detection to equipment malfunction identification and behavioral analysis. Oprea et al. [13] applied unsupervised ML techniques to detect anomalous values in time series data for the duration of electricity fraud, while Ambat et al. [14] demonstrated the effectiveness of outlier-based techniques for monitoring energy consumption in smart homes.

Zhang et al. [15] addressed the scalability challenge by developing a lambda architecture for iterative model updates and real-time anomaly detection in large-scale smart meter data. This system-level approach enables efficient processing of streaming consumption data to detect anomalies as they occur, facilitating timely interventions.

Our work bridges these application domains by focusing specifically on behavioral responses to dToU pricing, a critical area for the development of energy policies. Unlike fraud detection studies that prioritize high recall to minimize missed cases, our approach balances precision and recall to provide actionable insights for policy refinement without generating excessive false positives.

E. Critical Evaluation and Research Gaps

Several limitations persist in the literature on the detection of anomalies in energy consumption. First, many studies rely on small-scale datasets or simulated data, limiting their generalizability to real-world smart meter deployments. Second, the evaluation methodologies vary widely, making direct comparisons between approaches difficult. Most critically, there is limited research specifically addressing anomalous responses to dynamic pricing schemes.

Our research contributes to filling these gaps by: (1) developing a specialized approach for detecting anomalous responses to dToU pricing using a large-scale real-world dataset; (2) implementing and evaluating a hybrid methodology combining Isolation Forest with LightGBM and Neural Network classifiers; and (3) analyzing seasonal patterns in anomaly occurrence to provide deeper insights into household behaviors under dynamic pricing schemes.

This review of the literature establishes the theoretical foundation and methodological context for our approach, highlighting both its grounding in established techniques and its novel contributions to the field.

IV. Methodology

A. Data Source and Preparation

The primary data source is the ”Smart Meter Energy Use Data in London Households” dataset, publicly available from the London Datastore [1]. This dataset contains half-hourly energy consumption readings (kWh/hh) for a large number of households over several years.

Initial data preparation involved several steps:

1) **Filtering:** The raw dataset (167 million rows) was processed to retain only records corresponding to households under the dynamic Time-of-Use (dToU) tariff. This was achieved by filtering out rows where the ‘stdorToU‘ column indicated a standard (’Std’) tariff, resulting in a dataset of approximately 33 million rows focused on dToU participants.
2) **Preprocessing:** The filtered dToU data underwent cleaning, including handling missing values (e.g., filling numeric NaNs with column means) and ensuring correct data types for consumption and timestamp columns. The removal of outliers was applied using a Z score threshold (e.g., z = 3) to mitigate the impact of extreme values on normalization and modeling.
3) **Normalization:** Numerical features, primarily the energy consumption (‘KWH/hh (per half hour)‘), were standardized using ‘sklearn.preprocessing.StandardScaler‘ to have zero mean and unit variance.

Data Head:

	LCLid	stdorToU	DateTime	KWH/hh (per half hour)	hour	\
0	MAC000005	ToU	2012-06-01 10:30:00	-0.423691	10	
1	MAC000005	ToU	2012-06-01 11:00:00	-0.673404	11	
2	MAC000005	ToU	2012-06-01 11:30:00	-0.406665	11	
3	MAC000005	ToU	2012-06-01 12:00:00	-0.469093	12	
4	MAC000005	ToU	2012-06-01 12:30:00	-0.747183	12	

	day_of_week	month	is_weekend	anomaly_score	anomaly
0	4	6	False	0.033395	0
1	4	6	False	0.039046	0
2	4	6	False	0.031733	0
3	4	6	False	0.029252	0
4	4	6	False	0.034070	0

Fig. 1. Prepared Training Dataset

B. Feature Engineering

Based on the preprocessed data, temporal features were extracted from the ‘DateTime‘ column to provide context for the models:

- ‘hour‘: Hour of the day (0-23).
- ‘day_of_week‘: Day of the week (0 is Monday, 6 is Sunday).
- ‘month‘: Month of the year (1-12).
- ‘is_weekend‘: Boolean flag indicating if the day is a weekend (Saturday or Sunday).

These features, along with the normalized energy consumption, formed the input for the machine learning models.

C. Unsupervised Anomaly Detection: Isolation Forest

To identify potential anomalies without prior labels, we used the Isolation forest algorithm [3]. Isolation Forest works by randomly partitioning the data and isolating observations; anomalies are typically isolated in fewer partitions.

- **Training:** An Isolation Forest model (‘sklearn.ensemble.IsolationForest‘) was trained on the numeric features (normalized consumption

and time-based features). Key parameters included $n\_estimators = 100$ and $contamination = auto$.

- **Prediction:** The model assigned an anomaly score ('decision_function') and a binary prediction ('predict', where -1 indicates an anomaly and 1 indicates normal) to each data point. We converted the prediction to 0 for normal and 1 for anomaly for consistency.

These generated 'anomaly' labels served as the target variable for the subsequent supervised learning phase.

D. Supervised Anomaly Classification

Using the anomaly labels generated by Isolation Forest, we trained two supervised models to learn the patterns associated with these anomalies:

1) **LightGBM:** A gradient boost framework known for its speed and efficiency [4]. An 'lgb.LGBMClassifier' was trained using the engineered features to predict the 'anomaly' label. Early stopping was used during training to prevent overfitting.
2) **Neural Network (NN):** A Feed Forward Neural Network was implemented using TensorFlow. The architecture consisted of Dense layers with ReLU activation and Dropout for regularization, ending with a Sigmoid output layer for binary classification. Input features were scaled using 'StandardScaler' before training. Early stopping based on validation loss was employed.

E. Evaluation

- **Isolation Forest:** The primary evaluation was the percentage of data points flagged as anomalies and visual inspection of the anomaly score distribution and time series plots.
- **Supervised Models:** Performance was evaluated using standard classification metrics on a held-out test set: Accuracy, Confusion Matrix, Precision, Recall, and F1-Score.

V. Results and Discussion (Critical Analysis)

A. Exploratory Data Analysis (EDA)

Initial EDA focused on understanding the distribution of energy consumption and the characteristics of the anomalies identified by Isolation Forest.

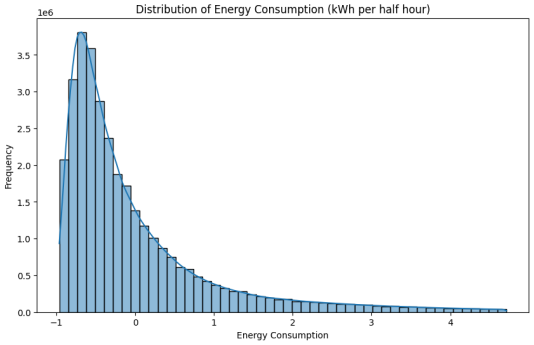

Fig. 2. Distribution of Normalized Energy Consumption (kWh/hh).

Figure 2 shows the distribution of the normalized energy consumption. Isolation Forest is an unsupervised anomaly detection algorithm that isolates outliers by constructing random decision trees. In this project, it was applied to the smart meter data to identify households that exhibited energy consumption patterns deviating significantly from the norm. The model flagged instances where energy usage remained high during peak tariff periods or abnormally low during off-peak times, suggesting inefficiency or non-adherence to dToU pricing.

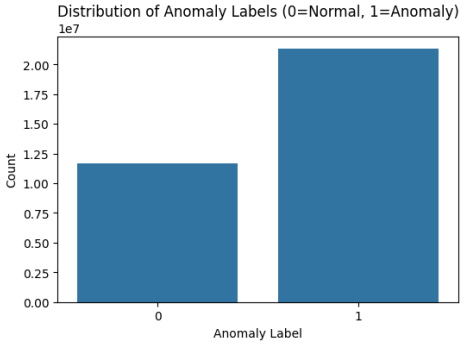

Fig. 3. Distribution of Anomaly Labels from Isolation Forest.

The Isolation Forest, with default parameters, labeled approximately 64.5% of the data points as anomalies (Figure 3). This high percentage suggests that the default 'auto' contamination setting might be too sensitive for this dataset or that many points exhibit deviations considered anomalous by the algorithm. Further tuning or alternative unsupervised methods might yield a more focused set of anomalies.

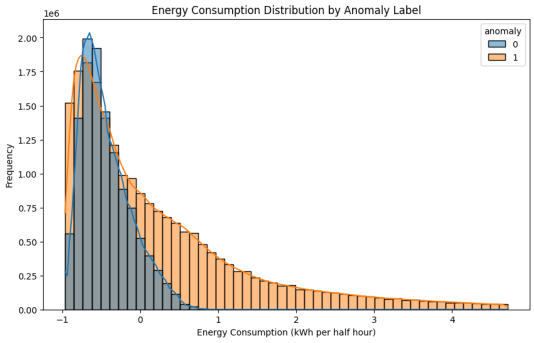

Fig. 4. Energy Consumption Distribution by Anomaly Label.

Comparing the consumption distributions for normal versus anomalous points (Figure 4) reveals that anomalous points tend to have a slightly higher average energy consumption and exhibit a wider spread in their distribution. While normal consumption values are tightly clustered around the mean, anomalies are more dispersed, indicating occasional spikes or irregular usage patterns that deviate from typical household behavior.

The daily anomaly rate plotted over time (Figure 5) indicates a clear seasonal pattern, with higher anomaly rates observed during the winter months and lower rates during

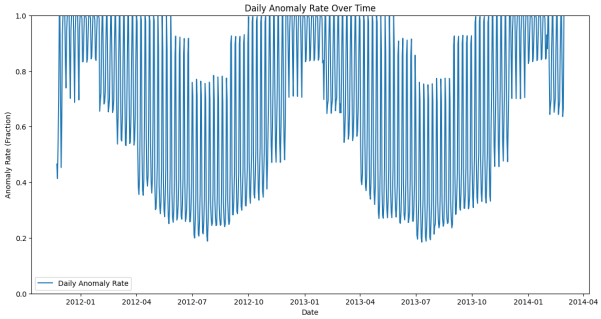

Fig. 5. Daily Anomaly Rate Over Time.

the summer. This trend suggests that energy consumption behavior is more irregular during colder periods, possibly due to increased heating requirements or varying responses to dToU pricing. In contrast, the summer months exhibit relatively stable and lower anomaly rates, reflecting more consistent energy usage patterns. This finding aligns with Jiang et al.'s [12] observation that seasonal factors significantly impact anomaly occurrence in energy consumption data.

B. Anomaly Visualization

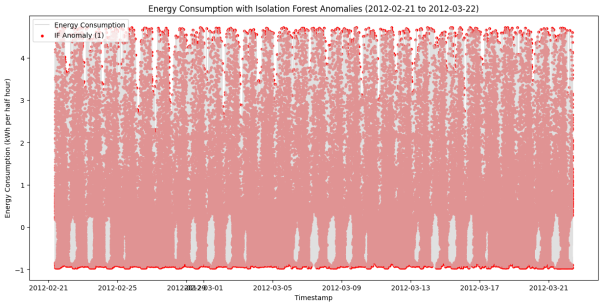

Fig. 6. Example Time Window with Isolation Forest Anomalies Highlighted.

Visualizing a subset of the time series with detected anomalies (Figure 6) provides a qualitative view of the model's behavior. The anomalies appear during periods of both high and low consumption, indicating that irregular patterns are not limited to extreme consumption values but can occur across varying usage levels. This reflects the model's sensitivity to deviations from typical consumption behavior, regardless of the absolute energy usage.

The distribution of anomaly scores (Figure 7) shows a concentration of scores near zero, indicating that most data points are considered normal by the Isolation Forest model. The distribution exhibits a long left tail extending towards lower values, corresponding to more anomalous points. These lower scores reflect instances of energy consumption behavior that significantly deviate from the learned normal patterns.

The scatter plot (Figure 8) relating anomaly scores to energy consumption suggests that highly anomalous scores (low values) are associated with both very low and very high energy consumption values. This indicates that anomalies are

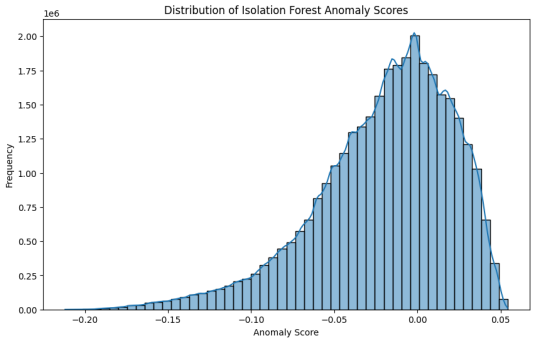

Fig. 7. Distribution of Isolation Forest Anomaly Scores.

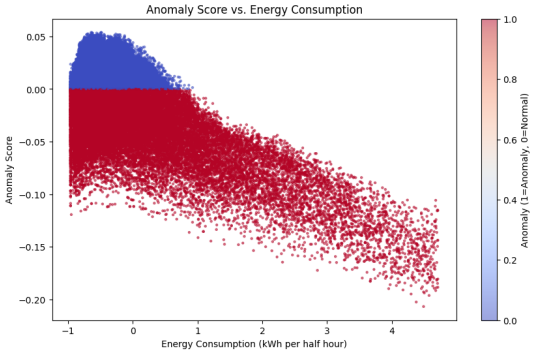

Fig. 8. Anomaly Score vs. Energy Consumption (Sampled).

not restricted to a specific consumption range but can occur at extreme ends of the usage spectrum, capturing irregular behavior such as unusually high spikes or unexpected drops in energy consumption.

C. Supervised Model Performance

The supervised models were trained to predict the binary anomaly labels generated by the Isolation Forest.

1) LightGBM: The LightGBM model achieved the following performance on the test set:

- Accuracy: 92%
- Confusion Matrix: See Figure 9
- Classification Report: The model demonstrated high precision, recall, and F1-score for both normal and anomalous classes, indicating strong performance in replicating the Isolation Forest labels.

2) Neural Network: The Neural Network model achieved the following performance on the test set:

- Accuracy: 90.3%
- Confusion Matrix: See Figure 10
- Classification Report: The model achieved high precision and recall for the normal class. The precision for anomalous points was slightly lower compared to LightGBM, but the overall F1-score was satisfactory, reflecting good learning of the anomaly patterns.

Discussion: Both supervised models demonstrated a high capacity to learn the patterns identified by the Isolation Forest,

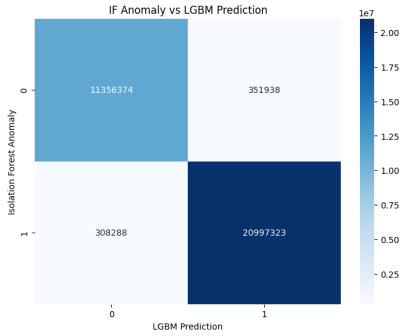

Fig. 9. LightGBM Confusion Matrix.

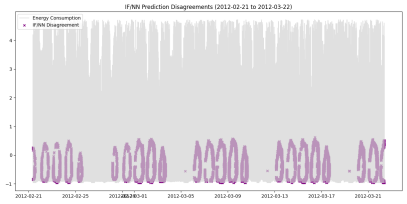

Fig. 10. Neural Network Confusion Matrix.

achieving high accuracy. LightGBM slightly outperformed the Neural Network, particularly in handling anomalous instances, likely due to its ensemble-based decision boundary modeling. However, the Neural Network also showed competitive performance, effectively capturing the key patterns in the data. The high accuracy primarily indicates that both models replicated the Isolation Forest's labels well. Given the high anomaly rate (over 60%) detected in the dataset, further validation would be necessary to confirm the real-world relevance of these anomalies.

The superior performance of LightGBM aligns with findings from Ke et al. [4], who demonstrated its efficiency in handling complex decision boundaries. This advantage is particularly relevant for energy consumption data, where anomalous patterns may involve intricate interactions between temporal features and consumption values.

VI. Conclusion and Future Work

A. Summary of Findings

This project successfully implemented a pipeline for analyzing smart meter data, identifying potential anomalies using Isolation Forest, and training supervised models (LightGBM and Neural Network) to classify these anomalies. Exploratory data analysis and visualizations provided insights into consumption patterns and the distribution of detected anomalies. The supervised models achieved high accuracy in replicating the Isolation Forest labels.

Key findings include:

- Isolation Forest detected a high proportion (64.5%) of data points as anomalies, suggesting either high sensitivity or significant irregularity in household responses to dToU pricing.
- Anomalous consumption patterns showed distinct seasonal trends, with winter months exhibiting higher anomaly rates than summer months.
- LightGBM slightly outperformed Neural Networks in classifying anomalies, with 92% versus 90.3% accuracy, suggesting ensemble methods may be particularly well-suited for this task.
- Anomalies were associated with both unusually high and unusually low consumption values, indicating that irregular behavior spans the entire consumption spectrum.

B. Limitations

Several limitations should be noted when interpreting our results:

- The high proportion of points labeled as anomalies (64.5%) raises questions about the appropriate contamination parameter for Isolation Forest in this domain.
- Without ground truth labels for anomalous behavior, the validation of our approach relies on consistency between models rather than absolute accuracy.
- The current feature set focuses on temporal factors but does not incorporate external data such as weather conditions or household demographics that might influence consumption patterns.

C. Future Work

Future research directions could address these limitations and extend our findings:

- Explore alternative unsupervised methods (e.g., Autoencoders, DBSCAN) or tune the Isolation Forest contamination parameter to potentially identify a more distinct set of anomalies.
- Incorporate external features such as weather data, pricing signals, and household characteristics to provide additional context for anomaly detection.
- Develop a domain-specific definition of anomalous behavior in response to dToU pricing, possibly through collaboration with energy economists and policy experts.
- Implement an online learning approach for real-time anomaly detection in streaming smart meter data, building on the work of Zhang et al. [15].
- Investigate the economic implications of these anomalies for both consumers and energy providers, quantifying the potential benefits of targeted interventions.

The high performance of LightGBM suggests it is a suitable and efficient model for this type of classification task on tabular time-series-derived features, making it a promising candidate for deployment in real-world energy management systems.

VII. Team Contributions and Acknowledgments

A. Team Contributions

All team members contributed equally to this project across various phases including research, implementation, data analysis, and documentation:

- **Ishan Biswas:** Contributed to data preprocessing, feature engineering, and implementation of supervised models. Assisted with model evaluation and documentation.
- **Sri Sai Teja Mettu Srinivas:** Led the implementation of the Isolation Forest algorithm, contributed to visualization development, and assisted with data analysis and result interpretation.
- **Divyank Singh:** Focused on dataset preparation, supervised model training, and results analysis. Collaborated on presentation materials and report documentation.

Each team member was involved in all aspects of the project, from initial concept development through final documentation and presentation, with tasks distributed to ensure equal participation and contribution.

B. Acknowledgments

We would like to express our sincere gratitude to Dr. Uzair Ahmad for his valuable guidance, feedback, and support throughout this project. His insights and expertise significantly contributed to the success of our work.

C. Code Repository

The complete code implementation for this project is available on GitHub at:
https://github.com/singhdivyank/Anomaly-Detection

The repository contains all code, documentation, and visualizations related to this work, including data preprocessing scripts, model implementations, and evaluation metrics.

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

[2] Faruqui, A., & Sergici, S. (2010). Household response to dynamic pricing of electricity—A survey of 15 experiments. *Journal of Regulatory Economics*, 38, 193–225.

[3] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. 2008 Eighth IEEE International Conference on Data Mining, 413-422.

[4] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30.

[5] Torriti, J. (2012). Price-based demand side management: Assessing the impacts of time-of-use tariffs on residential electricity demand and peak shifting in Northern Italy. Energy, 44(1), 576-583.

[6] Nicolson, M., Huebner, G., & Shipworth, D. (2018). Are consumers willing to switch to smart time of use electricity tariffs? The importance of loss-aversion and electric vehicle ownership. Energy Research & Social Science, 37, 123-133.

[7] Sakurada, M., & Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 4-11.

[8] Peña, M., Biscarri, F., Guerrero, J. I., Monedero, I., & León, C. (2013). Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach. Expert Systems with Applications, 40(4), 1394-1403.

[9] Wang, J., Guo, C., & Liu, K. (2022). Anomaly electricity detection method based on entropy weight method and isolated forest algorithm. Frontiers in Energy Research, 10, 984473.

[10] Lim, J. Y., Tan, W. N., & Tan, Y. F. (2022). Anomalous energy consumption detection using a Naïve Bayes approach. F1000Research, 11, 64.

[11] Fernandes, A., Ramos, C., Faria, F., Abreu, J., & Aranha, J. (2023). Anomaly Detection of Consumption in Hotel Units: A Case Study Comparing Isolation Forest and Variational Autoencoder Algorithms. Applied Sciences, 13(1), 314.

[12] Jiang, Z., Zhan, H., Yao, X., & Xu, W. (2022). High-Dimensional Energy Consumption Anomaly Detection: A Deep Learning-Based Method for Detecting Anomalies. Energies, 15(17), 6139.

[13] Oprea, S.V., Bâra, A., Puican, F.C., & Radu, I.C. (2021). Anomaly Detection with Machine Learning Algorithms and Big Data in Electricity Consumption. Sustainability, 13(19), 10963.

[14] Ambat, S.K., Venkatesan, S., & Tharakan, M. (2023). Anomaly detection and prediction of energy consumption for smart homes using machine learning. ETRI Journal, DOI: 10.4218/etrij.2023-0155.

[15] Zhang, B., Zhang, C., & Yi, X. (2018). Scalable prediction-based online anomaly detection for smart meter data. Information Systems, 77, 34-47.

[16] Zhou, K., Yang, S., & Shen, C. (2016). Time-of-use pricing model based on power supply chain analysis for price-based demand response. Journal of Modern Power Systems and Clean Energy, 4(2), 236-245.

[17] Himeur, Y., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Applied Energy, 287, 116601.

[18] Buzau, M., Tejedor-Aguilera, J., Cruz-Romero, P., & Gómez-Expósito, A. (2018). Detection of non-technical losses using smart meter data and supervised learning. IEEE Transactions on Smart Grid, 10(3), 2661-2670.