# Addressing Explainability in Load Forecasting Using Time Series Machine Learning Models

Mohamed Bouzid
Concordia Institute for Information Systems Engineering
Montreal, QC, Canada
mohamed.bouzid@mail.concordia.ca

Manar Amayri
Concordia Institute for Information Systems Engineering
Montreal, QC, Canada
manar.amayri@concordia.ca

Nizar Bouguila
Concordia Institute for Information Systems Engineering
Montreal, QC, Canada
nizar.bouguila@concordia.ca

## ABSTRACT

Energy management is a crucial issue in the modern world, as it affects various aspects of human life and the environment. It is a complex and challenging task that involves multiple factors and uncertainties. Machine learning has been widely adopted for improving building energy efficiency and flexibility in the past decade, as it can leverage the massive building operational data to provide accurate and reliable predictions and recommendations. However, with the increasing complexity, machine learning models are becoming black-boxes that are difficult to understand and trust by end-users. Hence, Explainable AI (XAI) has gained significant popularity in last years. In this paper, we focus on the explainability of different machine learning methods with regards to electricity consumption prediction. We apply explainability methods to interpret the results of these models and to provide insights into the factors that affect the electricity consumption. We use the Grenoble University building dataset, a non-aggregated dataset that contains electricity consumption data for different types of rooms over a period of two years. We evaluate the performance and the interpretability of the machine learning models and we discuss the implications and the limitations of our approach

## CCS CONCEPTS

• **Computing methodologies** → **Machine learning**; *Time series analysis*; • **Applied computing** → *Smart buildings*; Energy management.

## KEYWORDS

Energy management, Demand, electricity, building, Grenoble dataset, Deep learning, non-aggregated

**ACM Reference Format:**
Mohamed Bouzid, Manar Amayri, and Nizar Bouguila. 2023. Addressing Explainability in Load Forecasting Using Time Series Machine Learning Models. In *Proceedings of The 6th International Conference on Computational Intelligence and Intelligent Systems (CIIS 2023) (Conference acronym 'XX)*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX


*Conference acronym 'XX, November 25-27, 2023, Tokyo, Japan*

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

One of the most important resources in today's life is energy. Besides providing a variety of services and activities that enhance our quality of life, it plays a huge role in advancing economic growth. Specifically, the building industry counts for a significant portion of the the world's energy consumption and carbon emissions. According to the Global Status Report for Buildings and Construction [1], the sector's overall energy use and CO2 emissions rose in 2021 compared to the pre-pandemic levels. From 2020 to 135 EJ, the energy consumption of buildings grew by almost 4%, which was the biggest growth in the previous ten years. Building-related carbon dioxide emissions have peaked at around 10 GtCO2, up 5% from 2020 and 2% over the previous peak in 2019. These emissions presents threats to the planet's sustainability and ecosystem. Hence, building energy management is essential for worldwide energy savings and carbon reduction. Finding ways to use energy more wisely and effectively is therefore imperative.

Understanding and forecasting trends in energy consumption is one of the key components of smart-energy use. Predicting patterns in consumption can help building become more energy-efficient use less electricity, and better control their electricity usage through energy management systems. It can lower expenses and emissions, enhance the supply and demand for energy. However, energy consumption forecasting is a dynamic and difficult task that is influenced by a wide range of variables, including occupancy, time of the day, weather, and user behaviour.

Early methods of energy consumption forecasting consisted of using traditional rule-based techniques such as the statistical methods such as ARIMA and exponential smoothing [2]. However, energy systems often have complex patterns that are difficult to be captured by statisical models, such as seasonality, non-linearity and trends, Due to sensors and other sophisticated metering equipment being equipped in modern buildings, enormous amount of data related to energy demand. As a result, automation of building systems have gained attention in the last years giving the possibility of using such big data. A variety of data-driven techniques have been implemented to in the literature to support energy demand and learn new knowledge when it comes to load forecasting. A complete review of statistical and machine learning (ML) methods was presented in [3] and [4]. Some of the methods examined are linear regression, Support Vector Machine,etc. Moreover, another review [5] on load forecasting proved that tree-based, artificial neural networks (ANN) and deep neural networks (DNN) perform well load forecasting. Ensemble methods [6] have also been popular such as Random Forest, gradient boosted trees. Moreover, a survey have been conducted in [7] where 12 Neural networks (NN), including

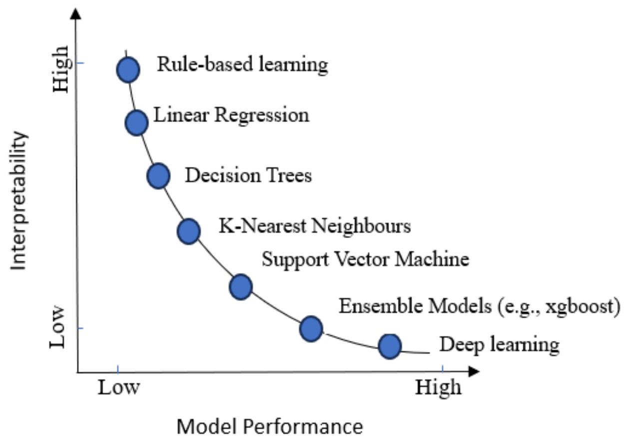



**Figure 1: Trade-off between modelś perfromance and interpretability**

Multi-layer preceptron, Recurrent Neural Networks (RNN), long short-term memory (LSTM), Convolutional-based models (CNN), have been studied for building energy forecasting. Additionally the dee autoregressive model (DeepAR) is one of the most widely used models, Finally, the attention mechanism began to become more relevant in load forecasting.

On the other hand, these advancements come at a cost. With the increasing complexity of the methods developed, the interpretability of these models decreases. The degree to which an individual can comprehend the forecast made by a machine learning model is known as model interpretability [8]. To interpret a model is to address the following question: which features have the major impact on the model's results? Despite their good performance, The inability to understand the black models of a deep learning models is a significant drawback. As the number of layers increases, it gets deeper and darker Domain experts usually want to comprehend the tool they use on a daily basis especially in sensistive fields like energy consumption. On the other hand, traditional models such as ARIMA or linear regression, which are easy-to-interpret, usually do not perform well in terms of prediction. Being unreliable or untrustworthy could affect the deployment of such model which explains why the building industry is often highly skeptical about the application of ML. Consequenntly, it is imperative to provide logical explanations for the ML model that do not sacrifice prediction accuracy or oversimplify crucial elemets. Figure 1 shows the trade-off between model's accuracy and model explainability [9].

In this paper, we focus on the usage of machine learning for the task of electricity consumption prediction. we increase the electricity load forecasting accuracy and model interpretability using different ML models and explainability techniques. We use the Grenoble University building dataset that contains electricity consumption data for different types of rooms over a period of two years and we create a prediction models models with good performance and high interpretability. Preprocessing, feature selection, model training, and performance metrics analysis and feature importance model interpretability is carried out on the dataset. It is well recognized that machine learning is an effective technique for load forecasting, which will aid in the development of energy-efficient management systems. To address this issue, we consider using four of the popular ensemble methods: lightgbm, random forest, catboost, xgboost. As the consumption dataset is not aggregated, it represents the unique behavior of each room's occupant, making this a tough scenario.

The main contributions of this paper are:

- We develop of three ML based load forecasting models and highlight that, in spite of the challenging nature of the data, they can produce satisfactory results.
- The application of feature importance technique to understand and interpret the complex nature of the ML system

The rest of the paper is organized as follows: Section 2 provides a brief review of different studies of explainability on load forecasting. In Section 3, we describe the dataset used to evaluate the models in addition to discussing data preprocessing. Next, we will review the different models in this applied on this dataset in Section 4 and compare their performance in Section 5. Finally, we dive into the explainability part and we discuss some insights from the model and we conclude in 6.

## 2 RELATED WORKS

There are various classification for explainable machine learning techniques used in time series problem. They can be divided into post-hoc and ante-hoc categories. Ante-hoc approaches incorporate explainability into the model's framework from the begining of the training process and hence they are model-specific. On the other hand, post-hoc techniques, which are applied to the model after training, are nostly model-agnostic. Several studies have been conducted on XAI approaches for load forecasting, such as decision-tree based enesmble methods and NN-based approaches. Mainly, post-hoc techniques are used mode in the litterature and the most popular techniques are Local interpretable model-agnostic explanations (LIME) [10],Shapley additive explanations (SHAP) [11], Feature importance, and visualizations plots. In [12], Zdravković et al. used LIME to interpret local features influence on the forecasting of heating demand. [11] used SHAP to examine the interpretation and relationships between three forecasting techniques for electriciy and showed that the type of building ( residential, commercial,etc) plays a big role in predictions. In [13], three explainability tools, which are LIME, SHAP, and explain it like I'm five (ELI5), have been used for solar photovoltaic forecasting. They built a random forest-based forecasting model and used those tools to explain the reasoning behind the outcomes. Feature importance was also used in different studies. In fact, tree-base methods such as Random forest, XGBoost, gradient boosting techniques can calculate each feature's importance by evaluating the impact of each feature in decreasing the impurity withing the tree. Visualization is another useful technique to analyze the complex models.The authors of [14] used t-SNE to show the latent states of autoencoders and provide an explanation for the reasons behind the predicted high or low energy usage. Finally, explainability have also been introduced to deep learning models, specifically in attension mechanism. Temporal Fusion Transformers [15] interpreted results by using the self-attention values that explains generall relationships that the model has learned.

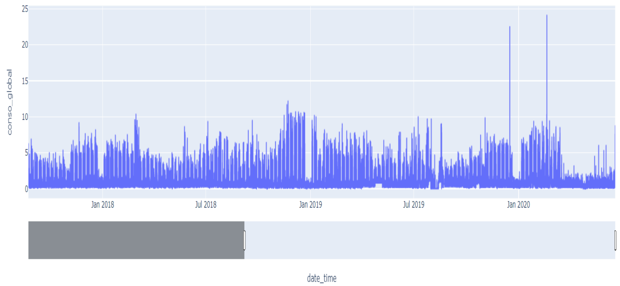



**Figure 2: Sample of Grenoble Dataset**

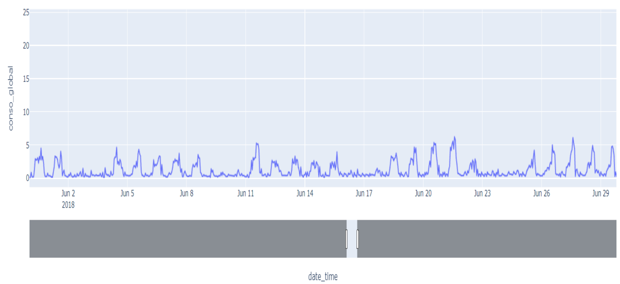



**Figure 3: Sample of 1 month of Grenoble Dataset**

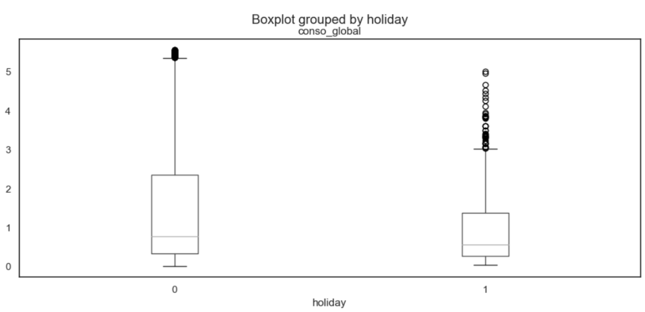



**Figure 4: Boxplot grouped by holiday**

# 3 DATASET

## 3.1 Dataset description

Measurements from different rooms at Grenoble University's Green-ER building make up the dataset [1] [16–18]. Hourly load consumption measurements were taken during a three-year period (from 2017 to 2020). While Figure 3 displays the consumption in a single month, Figure 2 illustrates the global building consumption over the course of the entire time. The noisy nature of the observations is evident, which makes handling it challenging. Boxplots of consumption for regular days and holidays are displayed in Figure 4. As can be seen, holidays use less electricity than regular days.

[1]https://mhi-srv.g2elab.grenoble-inp.fr/django/API/

## 3.2 Dataset preprocessing

Other than the hourly load consumption data, two other features were added to the dataset, occupancy and temperature, which include measurements from sensors inside the building. We refer to those characteristics as external input. Whether or not the room is occupied determines the occupancy feature's score, which is either 0 or 1. Naturally, the majority of the building's occupancy will occur on workdays and during working hours. Many preprocessing steps, such as handling outliers and missing values, were completed before the dataset was used. Our method for dealing with an outlier was to substitute its value by the observation from the same hour the week prior of that observation. Prior week observations would be more accurate than prior day observations because there may be significant differences between days within the same week (e.g., weekday and weekend), as opposed to the difference between the same day of the week from two successive weeks.

Moreover, since heating systems will be required on colder days, the temperature characteristic directly affects how much electricity is consumed. The temperature from the same hour the day before has been used to fill in any missing temperature data. Additionally, it was discovered through experimentation that it is preferable to utilize the daily average temperature rather than the hourly temperature because the response to temperature variations is not rapid. The occupancy feature is then examined. There were no outliers or missing results to contend with because it is specified by the school calendar rather than being recorded by a sensor. However, since there is typically no requirement for electricity consumption when there is no occupancy, occupancy is crucial for load forecasting. Global consumption also is a feature itself. Actually, lagged features of global consumption will be used where the model will be fed with the consumption of every hour of the previous day as well as the consumption of every hour of the same day from the preceding week.

# 4 MODELS

## 4.1 Light Gradient Boosting Machine (LightGBM)

Light Gradient Boosting Machine (LightGBM) is a machine learning gradient boosting technique. Its capacity to manage large volumes of data effectively has helped it become very popular; this is a feature that the data analysis industry has come to rely on more and more. It was developed by Microsoft in 2016 and it has been demonstrated that using this framework can increase training speed by up to 20 times compared to conventional gradient boosting decision trees (GBDTs). while maintaining the same level of accuracy [19]. What makes it different from GBDTs is that it builds the decision trees vertically (leaf-wise), whereas GBDTs grow them horizontally (level-wise). This enables the model to efficiently handle over-fitting and reduce the loss more quickly. Moreover, Lightgbm employs a histogram-based approach to determine the optimal split point for every feature, rather than sorting feature values. As a result, it just has to store the discrete bin values and the aggregated statistics for each bin, greatly reducing the calculation cost and memory use.

## 4.2 Extreme Gradient Boosting (XGBoost

eXtreme Gradient Boosting (XGBoost) [20] is an improved version of gradiend boosting decision tree with the intention to increasie efficiency, flexibility, and portability. In order to create a strong learner, XGBoost combines a number of weak learners. A machine learning model that performs only marginally better than random guessing is called a weak learner. On the other hand, weak learners can be merged to create a far more accurate strong learner.

Many decision trees are trained in order for XGBoost to function. A subset of the data is used to train each tree, and the aggregated forecasts of all the trees yield the final prediction. What makes XGBoost different from GBTs is that it employs a more regularized model which helps in mitigating the risk of overfitting. Another difference is that XGBoost builds trees in parallel as opposed to GBDT's sequential construction method. It employs a level-wise approach, scanning across gradient values and assessing split quality at each potential split point in the training set by utilizing these partial sums.

## 4.3 Categorical Features Boosting (Catboost)

Categorical Features Boosting is an open-source, high-performance gradient boosting library on decision trees [21], for tasks related to ranking, regression, and classification. CatBoost achieves good performance on big and complicated data sets with categorical features employing a variety of methods, including as feature engineering, decision tree optimization, and an innovative algorithm known as ordered boosting, to increase the precision and effectiveness of gradient boosting. CatBoost generates the negative gradient of the loss function in relation to the current predictions at each algorithm iteration. Next, we apply a scaled version of the gradient to the existing predictions in order to update them using this gradient. We use a line search technique that minimizes the loss function to determine the scaling factor.

CatBoost employs a method known as gradient-based optimization to construct the decision trees, in which the trees are fitted to the negative gradient of the loss function. With this method, the trees may concentrate on the feature space areas that have the most effects on the loss function, making predictions that are more accurate. Lastly, CatBoost presents a brand-new technique known as ordered boosting, which permutes the features in a certain order to optimize the learning objective function. Improved model accuracy and quicker convergence are possible outcomes of this method, particularly for feature-rich data sets.

# 5 EXPERIMENTAL RESULTS

## 5.1 Evaluation metrics

As evaluation metrics, we chose the coefficient of determination of the regression score, or $R^2$, to measure the accuracy of the predictions because we are predicting consumption values several steps in advance. Additionally, we employed root mean squared error (RMSE) and mean absolute error (MAE). Here is how the metrics are defined:

**Table 1: Results of 24-hour ahead predictions**

| Model | RMSE | MAE | $R^2$ |
|---|---|---|---|
| LightGBM | 0.04496 | 0.02804 | 0.6529 |
| XGBoost | 0.04493 | 0.02812 | 0.6533 |
| Catboost | 0.0482 | 0.02989 | 00.6008 |

$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(A_i - F_i)^2} \quad (1)$$

$$MAE = \Sigma_{i=1}^{N}|A_i - F_i| \quad (2)$$

$$R^2 = 1 - \frac{SSR}{SST} \quad (3)$$

$$SSR = \Sigma_{i=1}^{N}(A_i - F_i)^2 \quad (4)$$

$$SST = \Sigma_{i=1}^{N}(A_i - \hat{A}_i)^2 \quad (5)$$

where $\hat{A}$ is the true data average, $N$ is the number of samples, and $A_i$ and $F_i$ are the target and forecast values of the $i^{th}$ data point. Sum of squared regression is represented by SSR and total sum of squared by SST respectively.

## 5.2 24-hours ahead forecasting

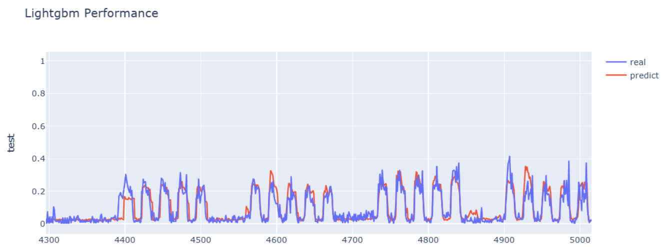
During the prediction at the test stage, the models are fed one week's worth of historical data in order to forecast the next 24 hours. The performance of the three machine learning models—LightGBM, XGBoost, and CatBoost—in predicting electricity consumption 24 hours in advance is presented and analyzed. The results, as summarized in Table 1, showcase competitive performance among the models, with key metrics including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination ($R^2$). Among the three models, XGBoost exhibits the highest $R^2$ score of 0.6533, indicating a strong correlation between the predicted and actual electricity consumption values. LightGBM closely follows with almost the same $R^2$ score of 0.6529. CatBoost, while demonstrating competitive performance, achieves a slightly lower $R^2$ score of 0.6008.

The results of the 3 models are showm in Figure 5 where 5a, 5b, and 5c represents the results of LightGBM, XGBoost, and Catboost respectively. We can see that the predictions graph have all the same trend with some minor differences.

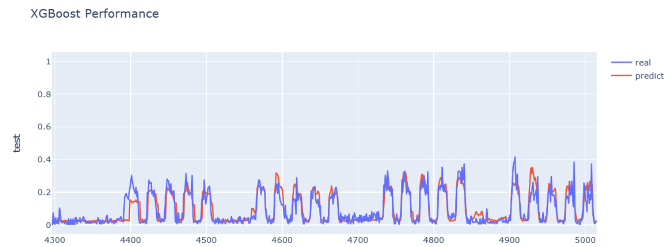
## 5.3 Explainability results

*5.3.1* ***Approach***. To explain our results, we apply the feature importance approach to add interpretabiliy into the predictions and understand which features have the most influence on the results. There are two ways to derive the feature importance scores:

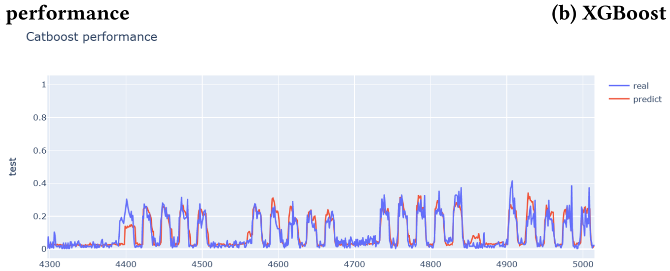
- **Gain**: When a feature is used in a tree's split, the gain technique measures how much the splitting criterion (such as impurity or information gain) improves. This improvement is used to determine the feature's importance. In fact, it quantifies the extent to which a given feature lowers the total error or raises the node purity in the trees.
- **Split**: In contrast, the split technique counts the number of times a feature is used to divide nodes among all of the




**(a) Lightgbm performance**




**(b) XGBoost performanc**




**(c) Catboost performance**

**Figure 5: Performance of the three models**

model's trees in order to determine how important a feature is. Because it is assumed that features that are used more frequently are more significant, this strategy focuses on how frequently a feature appears in the trees.

In our model, we use the split method to calculate ther scores. A higher score indicates a more important feature. The feature importance are calulated after the training of the model by using feature_importance() function that returns an array of scores for each feature after which they will be sorted and plotted.

*5.3.2* ***Results***. Before delving into the explainability results, we need to remeber that the dataset was preprocessed and occupancy data, temperate, and lagged features were added as features. If we want to predict the next 24 hours, the following features will be fed to the model to make predictions.

- **T-mean-lead**: This is the average temperature of the next day (the day we are going to predict the consumption of)
- **Calendar-lead**: This is the occupancy of the next day (the day we are going to predict the consumption of)
- Cluster-lag-d7: this indicates to which cluster ( consumption profile ) corresponds the consumption of the last week from the same day (high , low, moderate)
- **Conso-lag**: these are the lagged consumption from the previous day. These are 24 features that correspond to each consumption of the last 24 hours
- **Conso-lag-d7**: these are the lagged consumption from the same day from the previous week. These are 24 features that correspond to each hour of consumption of the same day last week (24 hours).

After getting the results, feature importance method was applied to visualize the global feature importance in the time of prediction in the models. The figure 7 illustrate the sorted importance of each of the features for the LightGBM model while Figure 8 shows the feature importance for the XGBoost. The bar plot indicates the importance of each of the features. In the first one, the results provided by this figure seem very logical and convincing. Since the values are sorted, we can see that conso-lag-d7-23 is the most important feature. This is expected since this corresponds to the same hour from previous week. We predict the consumption of the next day using the same hour from last week since they are usually similar. The second most important feature is the the T-mean-lead-23. This is the average temperature of the next day that we are going to predict. It is normal that is important since temperature plays huge role in energy consumption (heating). Next important feature is conso_lag_1 which is the consumption of the last hour. This value also plays a huge role since the consumption of the first hour that we are going to predict is a continuation of the last hour of consumption since it very rare to find a brutal variability in consumption between two consecutive hours. Another important feature for prediction is the calendar_lead_23 which the occupancy in the next day. Occupancy also play a big role since occupancy means using of systems and hence electricity. Finally, in this figure, we see that Lightgbm is explainable and the explanation given above seen covinicing. On the other hand, for XGBoost, the most important feature is the same which is the conso-lag-d7-23. However, there is a small difference where calendar_lead_23 plays bigger role in XGboost (second best) than in LightGBM where T-mean-lead-23 is the second most important feature. Other features importance are similar. Finally, for Catboost, as expected, conso-lag-d7-23 is always first as them most important feature, while calendar_lead_23 and T-mean-lead-23 both play a big influence on the predictions (second and fourth)

## 5.4 Discussions

Samples from the predicted test results for LightGBM for the whole test data are displayed in Figure 11 alogside the $R^2$ score which is

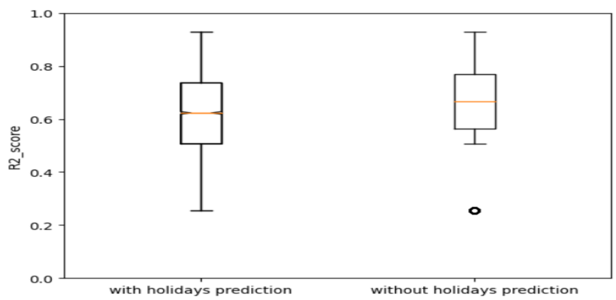



**Figure 6: Influence of vacations**

presented on top of the target and prediction values. We may observe from the figure that the model operates effectively at medium to high load consumption. Nonetheless, it typically struggles with low electricity consumption numbers, as shown by the small values $R^2$ readings in Figure 11. Due to a lack of holiday data, these models perform poorly in situations where electricity usage is low, such as during holidays or vacations. To demonstrate the impact of vacation days and weeks, we conducted an ablative study in which we eliminated forecasts made for vacation periods (forecasts with very low or negative $R^2$ values). It is confirmed by Figure 6 that predictions lose quality during vacation times. When we take away the presence of vacation periods, we can observe that the median of the boxplot of the $R^2$ score increases.

## 6 CONCLUSION

To conclude, our paper addresses the critical issue of electricity consumption prediction in the building industry, emphasizing the importance of both accuracy and interpretability in machine learning models. We have presented and compared three machine learning models—LightGBM, XGBoost, and CatBoost—in predicting electricity consumption 24 hours in advance, utilizing the Grenoble University building dataset. The results demonstrate competitive performance among the models, with XGBoost achieving the highest $R^2$ score of 0.6533.

Furthermore, we have delved into the explainability aspect, utilizing feature importance techniques to understand the model's decision-making process. The analysis revealed the significance of various features such as lagged consumption, average temperature, and occupancy in predicting electricity consumption. While LightGBM and XGBoost exhibited similar feature importance patterns, there were subtle differences, emphasizing the importance of considering model-specific interpretations.

Our paper contributes to the growing body of literature on explainable artificial intelligence (XAI) in the context of energy consumption forecasting. The emphasis on model interpretability is crucial, especially in industries where trust and understanding of the models are paramount. As future work, exploring additional XAI techniques and expanding the analysis to deep learning models and building types can provide valuable insights into enhancing the robustness and generalizability of models for electricity consumption prediction in the building sector.

## 7 ACKNOWLEDGEMENT

The completion of this research was made possible thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC) and a start-up Grant from Concordia University.

## REFERENCES

[1] P Cdb. Global status report for buildings and construction. 2022, 2022.
[2] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3):1014–1020, 2003.
[3] Vikas Gupta and Seema Pal. An overview of different types of load forecasting methods and the factors affecting the load forecasting. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 5(IV):729–733, 2017.
[4] Baran Yildiz, Jose I Bilbao, and Alistair B Sproul. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73:1104–1122, 2017.
[5] Yongbao Chen, Mingyue Guo, Zhisen Chen, Zhe Chen, and Ying Ji. Physical energy and data-driven models in building energy prediction: A review. *Energy Reports*, 8:2656–2671, 2022.
[6] Long Cai, Jie Gu, and Zhijian Jin. Two-layer transfer-learning-based architecture for short-term load forecasting. *IEEE Transactions on Industrial Informatics*, 16(3):1722–1732, 2019.
[7] Chujie Lu, Sihui Li, and Zhengjun Lu. Building energy prediction using artificial neural networks: A literature survey. *Energy and Buildings*, 262:111718, 2022.
[8] Christoph Molnar. *Interpretable machine learning*. Lulu. com, 2020.
[9] Zhe Chen, Fu Xiao, Fangzhou Guo, and Jinyue Yan. Interpretable machine learning for building energy management: A state-of-the-art review. *Advances in Applied Energy*, page 100123, 2023.
[10] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.
[11] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.
[12] Milan Zdravković, Ivan Ćirić, and Marko Ignjatović. Towards explainable ai-assisted operations in district heating systems. *IFAC-PapersOnLine*, 54(1):390–395, 2021.
[13] Murat Kuzlu, Umit Cali, Vinayak Sharma, and Özgür Güler. Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools. *IEEE Access*, 8:187814–187823, 2020.
[14] Jin-Young Kim and Sung-Bae Cho. Interpretable deep learning with hybrid autoencoders to predict electric energy consumption. In *15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020) 15*, pages 133–143. Springer, 2021.
[15] Bryan Lim, Sercan Ö Arık, Nicolas Loeff, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764, 2021.
[16] Gustavo Nascimento, Benoit Delinchant, Frédéric Wurtz, Patrick Kuo-Peng, Nelson Jhoe Batistela, and Tiansi Laranjeira. Green-er - electricity consumption data of a tertiary building. *Mendeley Data*, 2020.
[17] Benoit Delinchant, Gustavo Martin, Tiansi Laranjeira, Muhammad Salman Shahid, Frederic Wurtz, et al. Machine learning on buildings data for future energy community services. In *SGE 2021-Symposium de Génie Electrique*, 2021.
[18] Sacha Hodencq, Benoit Delinchant, and Frédéric Wurtz. Open and reproducible use cases for energy (oruce) methodology in systems design and operation: a dwelling photovoltaic self-consumption example. In *Building Simulation 2021*, 2021.
[19] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.
[20] Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM.
[21] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. Catboost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31, 2018.

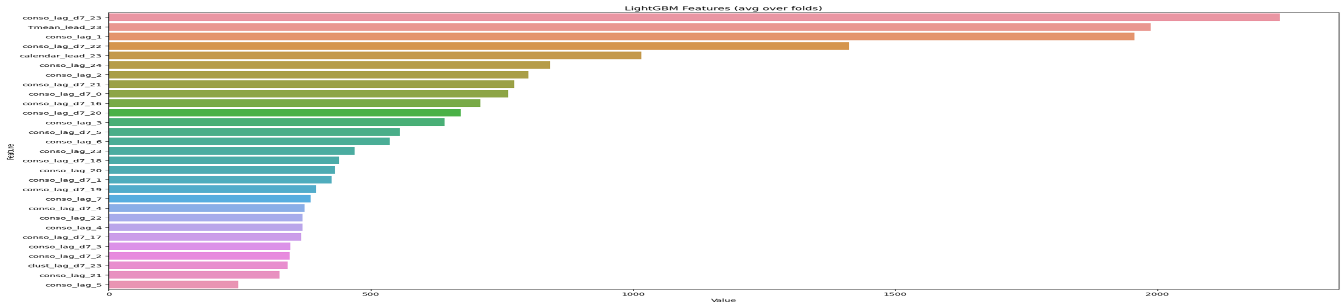



**Figure 7: LightGBM feature importance results**

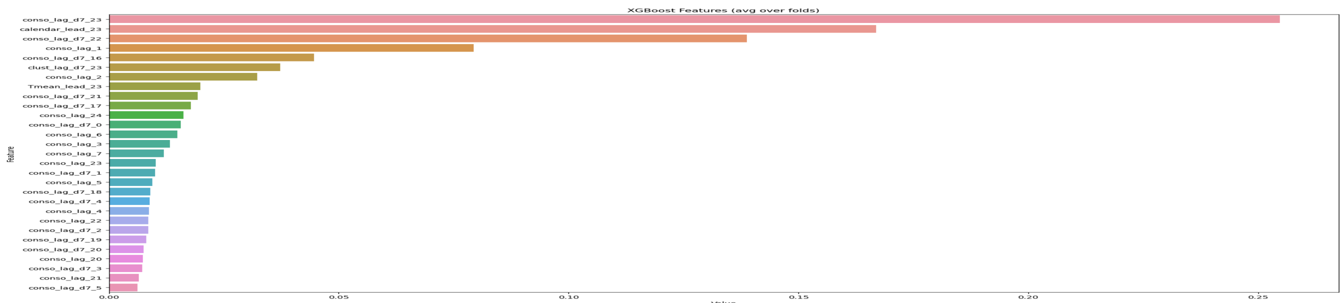



**Figure 8: XGBoost feature importance results**

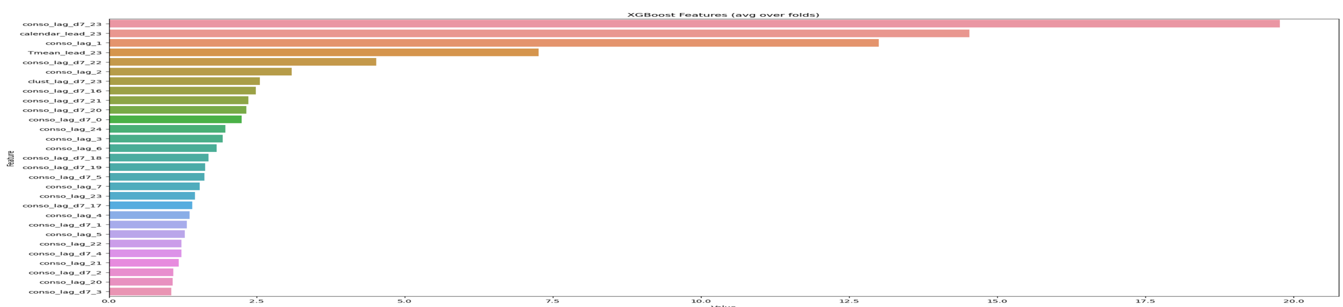



**Figure 9: Catboost feature importance results**

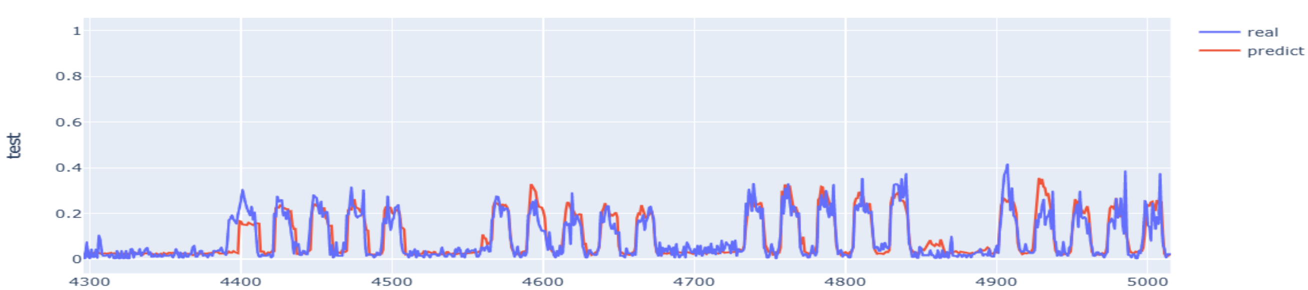



**Figure 10: LightGBM 24-ahead predictions**


**Figure 11: R2score analysis**