Explainable Lightweight Federated Learning for Load Forecasting with Smart Meter data

Le Vu Hoang Duc[1], Ngo Van Uc[1], Kim Duc Tran[2,1], Ludovic Koehl[2], and Kim Phuc Tran[2]

[1]International Chair in DS & XAI, International Research Institute for Artificial Intelligence and Data Science, Dong A University, Danang, Vietnam
[2]Univ. Lille, ENSAIT, ULR 2461 - GEMTEX - Génie et Matériaux Textiles, F-59000 Lille, France

December 2, 2024

Abstract

The purpose of forecasting electricity demand is to achieve a balance between supply and demand between providers and customers, which is a crucial factor in the energy field. This study aims to employ an approach using XAI and FL in a DNN model, combined with data-driven techniques to deliver accurate predictions based on data from current smart meter systems in smart buildings towards a case study based on load forecasting for households. The paper continues the work on federated learning to ensure the privacy of individual meter data, allowing several parties, commonly referred to as clients, to keep their data distributed and not centralized in the training process. Additionally, the paper highlights the benefits of the XAI approach in providing clear and reasonable explanations for the model's predictions. This not only supports energy providers in analysis, decision-making, and optimizing operations but also enhances the trust and understanding of stakeholders regarding the forecasting model, thereby optimizing the efficiency of energy distribution and usage.

Keywords: Deep neural network, Explainable Artificial Intelligence, Federated learning, Energy Distribution

1 Introduction

The advancement of smart grid systems has led to the adoption of smart meters (SM) devices in power systems as an alternative to humans to record electricity consumption indicators. The large amount of data from SM brings many benefits to electricity suppliers, such as the data related to electricity consumption demand for each region and specific time. Therefore, predicting the demand for electricity consumption paves the way for energy services and data-based business models[3].

Many advanced deep learning models have been deployed in load forecasting. However, modern models often operate as "black boxes", making it difficult for people to understand and trust their forecasting mechanism[13]. Some complex models, such as deep neural networks, make it difficult to interpret the prediction results and lead to a lack of model transparency[8]. These challenges require an innovative solution that meets the requirements of user data security and privacy but also provides accurate load forecasting performance, while explainability and transparency of outputs are key to sustainability.

Therefore, in this study, we use Explainable Artificial Intelligence (XAI) techniques to explain the prediction results on a lightweight DNN model in the FL framework published in[5]. By using XAI, we explain how data features influence model output, ensuring transparency and trustworthiness, while the federated learning technique helps secure user data. The main contributions of this research include:

- Re-implementing a lightweight deep learning model within the FL framework to forecast the load with the same or higher accuracy than other advanced models, the purpose is to optimize the computational resources while the model must be suitable for resource-constrained environments.
- Integrating XAI techniques into the model, helping to explain load forecast results, thereby increasing transparency, accountability, and ethics in forecasting applications. This helps us better understand the model results and increase trust in the system.

This study includes three parts, in Related work we review a series of studies that have been done to address the load forecasting problem, then outline the current context and reasons why XAI is needed to address this problem in the era of artificial intelligence. Next, in the Explainable Lightweight Federated Learning for load forecasting section, we briefly discuss the lightweight DNN model within the FL and XAI framework. Finally, we apply XAI to the test data and explain how the model predicts the output.

2 Related work and background

The need for load forecasting in power systems is becoming increasingly urgent, especially in the context of the strong development of smart meter systems[3]. Traditional and modern methods have been widely applied to solve this problem. The Deep Neural Network (DNN)-based method has been proven to be effective in its ability to extract features from large and complex data[11]. A. M. Pirbazari et al. (2020) proposed using Long-short Term Memory (LSTM) and other deep learning models for load forecasting, and the results showed that the model can generalize satisfactorily and produce accurate results if they are provided with sufficient data[10]. In addition, the study by A. Duttagupta et al. (2024) used a federated learning approach using a lightweight DNN for load forecasting. The results showed that the model is simpler but has comparable accuracy to other state-of-the-art models while still ensuring the privacy of individual smart meter data[5].

The mentioned studies using the DNN method have shown impressive results in the load forecasting task. However, as models become more complex, the results of these models become difficult to understand and question their reliability due to their "black box" nature. It is necessary to understand and explain the reasoning behind the results of AI models[2]. Therefore, this study aims to apply XAI to a lightweight federated learning model, thereby explaining the outputs. This research focuses on providing a clear and reliable explanation for the predictions of lightweight DNN models in the FL framework, which helps improve transparency and accountability in smart energy systems, making an important contribution to safe and sustainable systems in the context of rapid development of distributed computing and big data from smart metering devices for smart buildings.

3 Explainable Lightweight Federated Learning for load forecasting

In this section, we present the concept of federated learning and how it can be applied in load forecasting. Subsequently, we conduct data preparation to adapt our FL simulation with the lightweight deep neural networks model. Then, XAI is applied to improve the trustworthiness and transparency of the model.

3.1 Federated learning

In the context of growing concerns about data privacy, Federated Learning (FL) enables training ML models across devices without sharing raw data, safeguarding user privacy[7]. Figure 1 shows the FL setup, where each smart building trains a local model on its edge data.

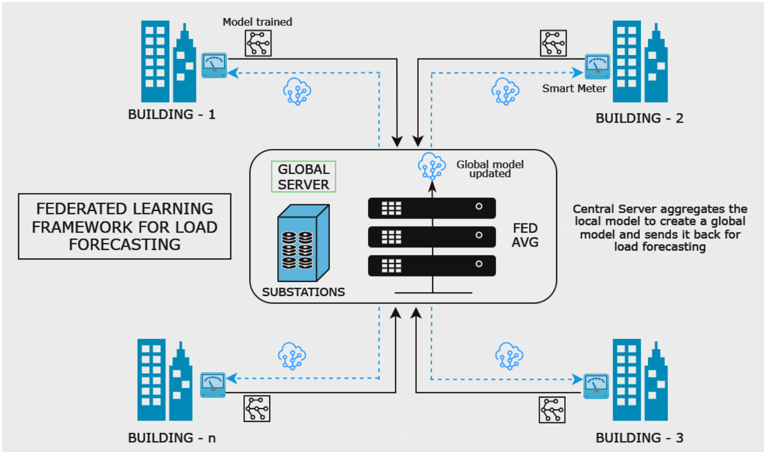

Figure 1: Application of Federated Learning in load forecasting analysis.

Algorithm 1: Federated Learning Algorithm

Input: Total device set M, selected device set N, batch size B, learning rate η, local epochs E

Output: Global model W

Initialization: Randomly initialize global model W;

while *not converged* **do**

  foreach *device n in selected device set N* **do**

    Retrieve local data P on device n;

    Train local model w_n on data P for E epochs with batch size B and learning rate η;

    Share local model w_n with central server;

  Aggregate models w_n from all devices on a central server using federated averaging;

  Update global model W as the average of all local models w_k;

  Share global model W with all devices;

The FL algorithm flow used in this study, as detailed in[5] and derived from[6], begins with selecting a subset of user devices (N) from the total population M and initializing a global model W. This model is distributed to the selected devices for parallel training on their local data P, using parameters such as batch size B, learning rate η, and local epochs E. The locally updated models w_n are sent back to the server, which aggregates them into a new global model W via weighted averaging. This updated global model is redistributed to the devices, and the process repeats until the global loss function (GLF) shows no significant improvement, indicating it has reached a minimum.

3.2 Propose method

In figure 2 is our proposed method, first the data will be collected and pre-processed. Then we cluster the data into 18 clusters using cluster algorithm. We use 60 percent of the clusters for training purpose and the remaining 40 percent for testing purpose. Then the data will be processed appropriately to train the DNN model in FL framework, during this process we perform model evaluation and when convergence for FL model is reached, we stop the training process and use XAI to explain the model's output prediction result.

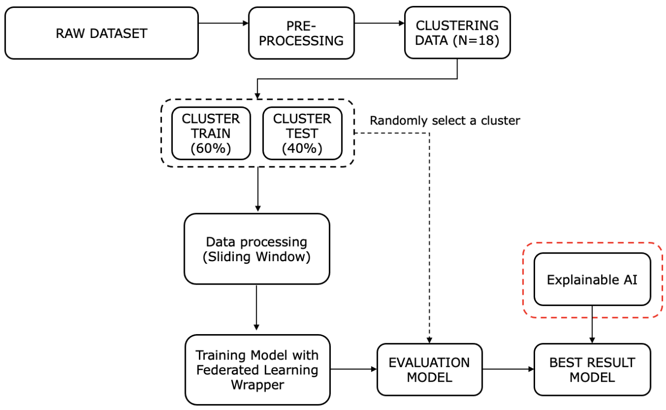

Figure 2: Integrate XAI into Federated learning

3.3 Dataset

This study uses the publicly available SmartMeter Energy Consumption dataset from 5,567 London households[12] for load forecasting in smart buildings. Due to limited data on smart building consumption, we use household data as an alternative method, assuming households represent smart buildings without affecting the scope. The dataset covers energy consumption from November 2011 to February 2014, with half-hourly data including building ID, tariff type, timestamp, and half-hourly electricity consumption (kWh/hh).

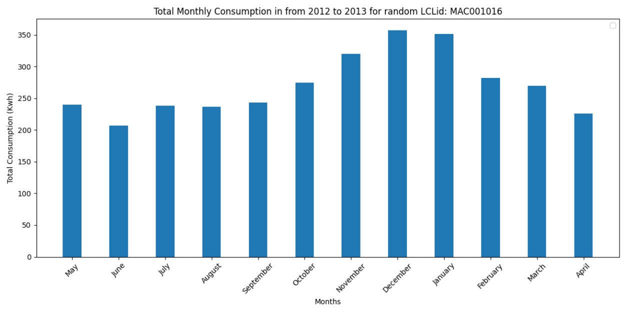

Figure 3: Annual energy consumption trend of a random building.

In the first step of analysis, we review total monthly consumption data to examine variability in energy use and observe yearly fluctuations in consumer patterns. Figure 2 shows the energy consumption of building MAC001016 from May 2012 to April 2013, revealing higher consumption in December, January, and February, similar to building MAC000025[5].

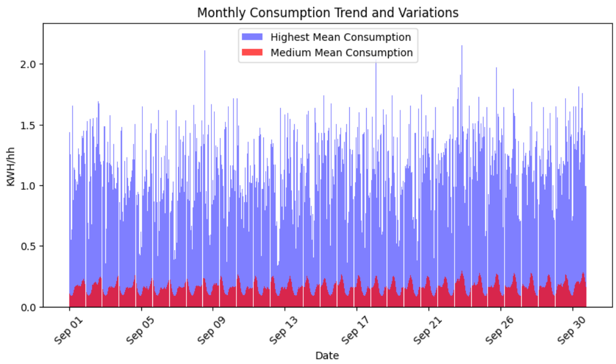

Figure 4: Comparison of daily consumption over a month: high consumption building vs average consumption building[5].

Figure 3 displays the energy consumption of all buildings for September 2012. Buildings with average half-hourly consumption below 0.09 kWh/hh or above 1.35 kWh/hh are considered outliers and removed, reducing the dataset from 5,547 to 4,672 buildings and eliminating potential bias.

3.4 Data Preparation for Federated Learning in load forecasting

Firstly, in the preprocessing data for preparing the data to train the DNN model using the FL framework, we clustering the data into 18 clusters using K-Means clustering to provide better generalization for deep learning models when used for load forecasting[4]. This also improves the aggregation phase in the distribution station, where local models from smart buildings with similar performance are better able to achieve unbiased aggregation to arrive at a new global model [5]. The electricity consumption data from October 1, 2011, to February 28, 2013, are used for training purposes. The entire training and testing data are divided into a 60:40 ratio for training and testing purposes, respectively as this result [5].

3.5 Using TensorFlow Federated framework and Model architecture

In this study, we use the lightweight DNN model architecture in the FL framework that has been shown to perform well on this dataset in the previous study [5]. We build training and inference (simulation) flows using the Python-based Tensorflow-Federated (TFF) library and related tools to simulate the model training process on each building and the weighting process for the global model. Each building has a unique ID used to access its specific data streams and clusters throughout the training and evaluation phases.

The model architecture described in [5] is reused. The model consists of 4 layers, including 2 hidden layers, and the overall model follows the shape $[16, 8, 4, 1]$. The activation function is ReLU. The author uses RMSE and MAE to evaluate the loss for the training process, and all three metrics (RMSE, MAE, and MAPE) for the experimental task. The learning rate is set to 0.01 for all local models. In our study, we use input data consisting of 336 half-hourly electricity consumption records, equivalent to a week's electricity consumption data. The output of the model will be the predicted continuous electricity consumption for one day, equivalent to 48 half-hour intervals.

3.6 Performance metrics

To evaluate the error of the forecast results, this study uses metrics including average building consumption including MAE, RMSE, and mean absolute percentage error (MAPE). These metrics allow a comprehensive evaluation and comparison of the prediction accuracy independent of the average building consumption [5].

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |X(i) - Y(i)|$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (X(i) - Y(i))^2}$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \left|\frac{X(i) - Y(i)}{X(i)}\right| \times 100$$

Where X represents the actual value and Y represents the predicted value, with N being the number of data points.

3.7 Explainable Artificial Intelligence (XAI)

Building ML or AI systems as shown in Figure 5 with human-understandable explanations is central to XAI research, addressing the limited interpretability of black-box models. XAI enhances understanding through interpretable models (e.g., decision trees) or post-hoc explanations, which provide interpretability after training. Local interpretability focuses on individual predictions, while global interpretability examines aggregated input contributions across the model.

XAI supports human-centered design, enabling users to evaluate AI trustworthiness and transparency by assessing explanation quality. This helps energy operators make informed decisions, detect anomalies, and enhance grid efficiency and stability.

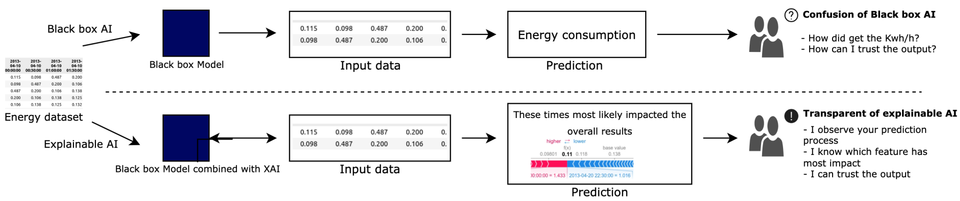

Figure 5: XAI in the context of load forecasting analysis

In this study, we apply SHAP (SHapley Additive exPlanations) to the model for the prediction results to explain the important features that influence the output. The Shapley value, derived from cooperative game theory, ensures a fair distribution of payoffs among players. In the context of XAI, the Shapley value attributes the contribution of each feature to the overall prediction. The importance of feature i is defined by the Shapley value[9]:

$$\phi_i = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|! \left(|N| - |S| - \mathbf{1} \right)! \left[f\left(S \cup \{i\} \right) - f\left(S \right) \right]$$

Where N is the set of all features, S S as a subset of N to consider the permutation of features excluding i, and f is considered as the objective function of the prediction model.

Based on the Shapley value, it provides relevance values for various features in the input tabular data, indicating the contribution of each feature to the model's output. SHAP offers a unifying framework for feature attribution, helping clients understand the relative importance of different Time-of-day consumption attributes in forecasting future outcomes. Furthermore, there are visualization techniques that create visual representations of the features learned by the model.

4 Illustrative example

In this chapter, we present the details of the experimental results from our implementation of federated learning in this study [5], as well as our proposed XAI approach.

4.1 Implementation Federated learning results

We conduct model training using two methods: the centralized method and the FL framework. For the centralized method, the model goes through 21 epochs, while for the FL framework, the model converges after 21 rounds of global weight updates. The model hyperparameters are similar in both DNN models. Our experimental results are consistent with those in [5], where we found that the lightweight DNN method using the FL framework achieved comparable or better accuracy than the traditional method on the test clusters. Figure 7 shows the RMSE of the FL framework model for a random client, as presented in Table 1. It is lower than that of the traditional centralized model. This suggests that the FL model can predict more accurately and has less error. The studies using the DNN method have shown impressive results in the load forecasting task. .

LCLid	Cluster	Date
MAC000123	8	01/03/2023 - 28/02/2024

Table 1: Data for comparison between centralized and FL model

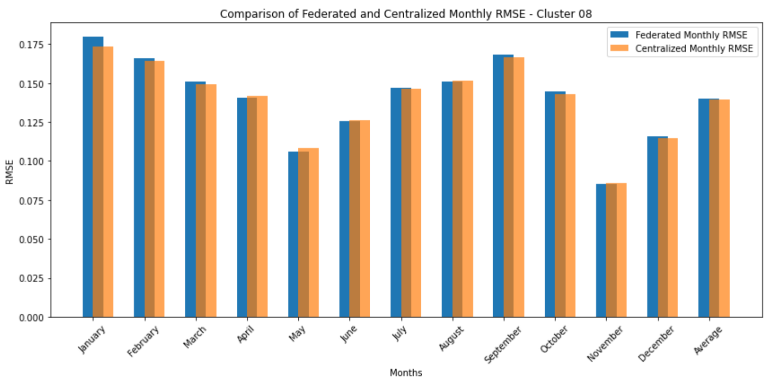

Figure 6: Energy profile yearly trend of a random smart building.

After completing the training, we conducted an evaluation of the electricity consumption over the course of a day for a few random buildings. Figure 8 shows the prediction results for three random smart buildings, as presented in Table 2, for a randomly selected period of the year. It can be observed that electricity consumption varies throughout the year. The results in the chart indicate that our model tends to predict higher consumption than the actual values during peak periods. However, the error between the actual and predicted values is negligible.

LCLid	Cluster	Date
MAC003770	8	17/04/2013 - 24/04./2013
MAC000946	8	22/09/2013 - 29/09/2013
MAC000201	8	20/02/2013 - 27/02/2014

Table 2: Evaluation process data

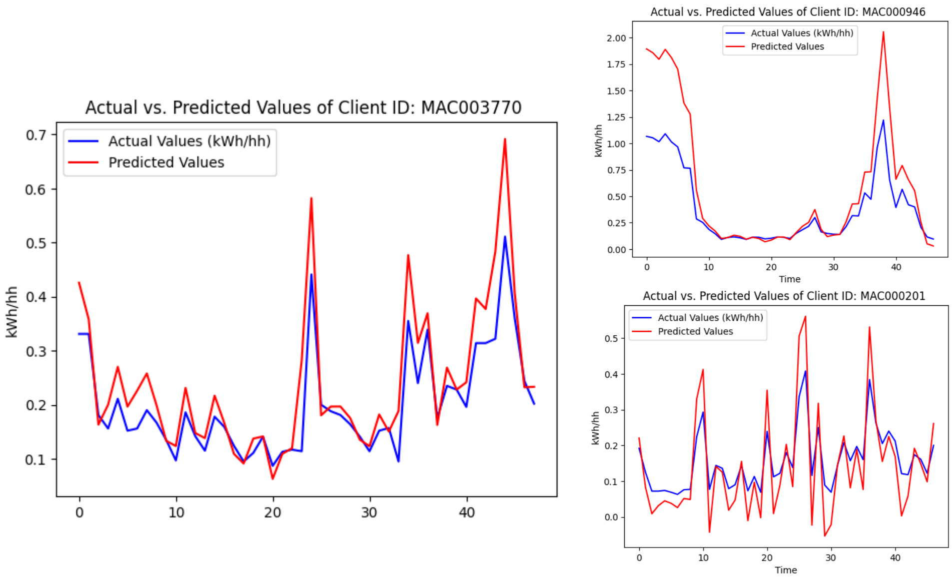

Figure 7: Load forecast predicting on random smart buildings

4.2 SHapley Additive exPlanation

First, consider an overview of the SHAP values of the model to identify which features have the most significant impact on the model's output.

This experiment is conducted on a random ID from the testing dataset, with results for MAC000123 as shown in Table 3

LCLid	Cluster	Date
MAC000123	8	10/04/2013 - 17/04/2013

Table 3: A random smart building for the SHAP Violin plot

To observe which periods within the time frame have the most impact on the model's output, we can use a violin plot. This is a summary plot that provides an overview of the distribution and variability of SHAP values for each feature. Individual violin plots are stacked according to the importance of each feature in the model output, with the importance measured by the total absolute SHAP values for each feature[1].

Therefore, we conduct experiments on all samples over the energy consumption of one week to predict consumption for a day.

According to the plot as shown in Figure 9, the waveform illustrates the SHAP value density of each feature and indicates the range over which each feature's impact is positive or negative. This allows us to understand which ranges of each feature have more impact on the model output.

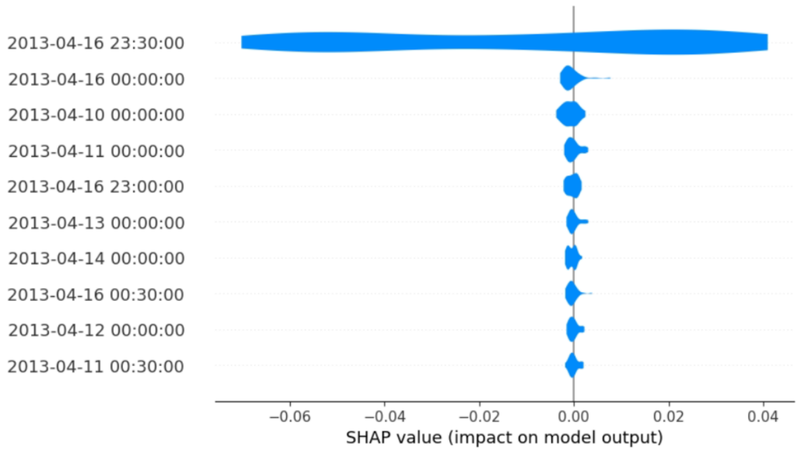

Figure 8: SHAP Violin plot for the global model

Then, the time at 23:30:00 contributes the most to the model's consumption prediction.

Subsequently, this experiment is tested on a specific sample taken from client MAC000123, as shown in Table 3. is analyzed using a waterfall plot[1] as depicted in Figure 10, which provides a local interpretation by focusing on a single sample. This approach is useful for examining all of its features, enabling the identification of model outliers and SHAP outliers.

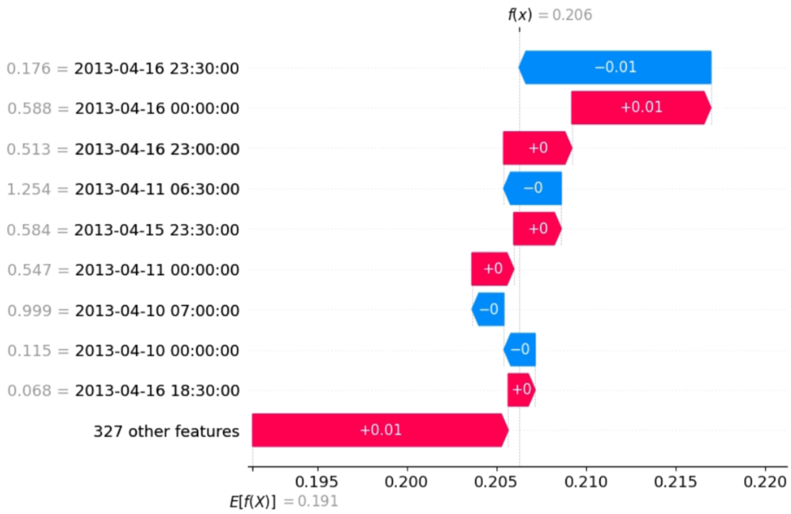

Figure 9: SHAP Violin plot for the global model

The features of the model are listed along the y-axis, with each feature's value for the given sample displayed in gray. The corresponding SHAP value for each feature is shown in the main panel, with an arrow indicating the direction of influence for each feature. If a row is colored red (blue), it indicates that the SHAP value increases (decreases) the prediction compared to the average or expected prediction.

5 Discussion and conclusion

This study re-implemented[5] the lightweight DNN model in the FL framework to model load forecasting using publicly available electricity consumption datasets. Our implementation simulates real-world power grid conditions and applies XAI for model interpretation. Accordingly, the FL method has been proven to maintain higher accuracy than the traditional method while ensuring data security for customers. Using SHAP for XAI, we enhance the explainability of the model by highlighting important features that affect forecast results. However, the effectiveness of XAI also depends on the diversity and richness of the data, too few features in the data can reduce the explainability of the model.

In summary, this study demonstrates the use of XAI to improve transparency and reliability in load forecasting within the FL framework for the smart building scenario. Future work will focus on testing the model on additional datasets and expanding XAI applications in complex data contexts.

References

[1] Shap. URL `https://shap.readthedocs.io`.

[2] Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez, and Francisco Herrera. Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence. *Information fusion*, 99:101805, 2023.

[3] CL Athanasiadis, TA Papadopoulos, GC Kryonidis, and DI Doukas. A review of distribution network applications based on smart meter data analytics. *Renewable and Sustainable Energy Reviews*, 191:114151, 2024.

[4] Kasun Bandara, Christoph Bergmeir, and Slawek Smyl. Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert systems with applications*, 140:112896, 2020.

[5] Abhishek Duttagupta, Jin Zhao, and Shanker Shreejith. Exploring lightweight federated learning for distributed load forecasting. In *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2023.

[6] Nastaran Gholizadeh and Petr Musilek. Federated learning with hyperparameter-based clustering for electrical load forecasting. *Internet of Things*, 17:100470, 2022.

[7] Muhammad Akbar Husnoo, Adnan Anwar, Nasser Hosseinzadeh, Shama Naz Islam, Abdun Naser Mahmood, and Robin Doss. A secure federated learning framework for residential short term load forecasting. *IEEE Transactions on Smart Grid*, 2023.

[8] Joohyun Jang, Woonyoung Jeong, Sangmin Kim, Byeongcheon Lee, Miyoung Lee, and Jihoon Moon. Raid: Robust and interpretable daily peak load forecasting via multiple deep neural networks and shapley values. *Sustainability*, 15(8):6951, 2023.

[9] Tucker AW Kuhn HW. Shapley ls. a value for n-person games. contributions to the theory of games. *Annals of mathematical studies*, 1953.

[10] Aida Mehdipour Pirbazari, Mina Farmanbar, Antorweep Chakravorty, and Chunming Rong. Short-term load forecasting using smart meter data: A generalization analysis. *Processes*, 8(4):484, 2020.

[11] Mohammad Mustafa Taye. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, 12(5):91, 2023.

[12] UKPowerNetworks. Smartmeter energy consumption data in london households, 2014. URL `https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households`. Accessed: 2024-11-15.

[13] Fatma Yaprakdal and Merve Varol Arısoy. A multivariate time series analysis of electrical load forecasting based on a hybrid feature selection approach and explainable deep learning. *Applied Sciences*, 13(23): 12946, 2023.