МРНТИ: 44.29.29 https://doi.org/10.65247/3136-8042-2026-3.03

DATA SCIENCE-BASED MODELING FOR ELECTRICITY CONSUMPTION FORECASTING

*[1]A.KANATNUR, [1]Y.B.AYZEREN
[1]Cyprus International University
(Nicosia, Cyprus)
*kanatnur0501@gmail.com, yaseminbay@ciu.edu.tr

Abstract. Accurate forecasting of electricity consumption is central in modern energy systems for both planning and operational decision making. This research proposes a data science driven methodology for predicting electricity demand and directly compares machine learning with deep learning approaches under identical experimental conditions. Specifically, a tree based model, XGBoost, and a recurrent neural network, Long Short-Term Memory (LSTM), have been evaluated on the same dataset with consistent data preprocessing, splits, and metrics. Currently, tree based models use explicit temporal feature engineering in order to find nonlinear relationships in structured time dependent data. In contrast, LSTM networks learned the temporal dependencies from sequential load observations directly. The performance evaluation of these models is made by RMSE and MAPE. Accordingly, it follows that LSTM yields the highest forecasting accuracy with the superior capturing of temporal dynamics, while XGBoost offers competitive performance with lower computational complexity. This implies that there are some important context-dependent tradeoffs that exist among model accuracy, interpretability, and efficiency when coming to electricity consumption forecasting.
Keywords: Electricity consumption forecasting, Data science, Machine learning, Deep learning, XGBoost, LSTM, time series forecasting, load forecasting, feature engineering, baseline model.

Электр энергиясын тұтынуды болжаудың деректер ғылымына негізделген модельдері

[1]А.Қанатнұр, [1]И.Б.Айзерен
[1]Кипр халықаралық университеті (Никосия, Кипр)

Аңдатпа. Электр энергиясын тұтынуды дәл болжау қазіргі заманғы энергетикалық жүйелерде жоспарлау мен операциялық шешімдер қабылдаудың негізгі факторларының бірі болып табылады. Бұл зерттеуде электр энергиясына деген сұранысты болжауға арналған деректер ғылымына негізделген әдістеме ұсынылып, бірдей эксперименттік жағдайларда машиналық оқыту мен терең оқыту тәсілдері тікелей салыстырылады. Атап айтқанда, ағаш құрылымына негізделген XGBoost моделі мен қайталанатын нейрондық желі - Long Short-Term Memory (LSTM) бірдей деректер жиынында, деректерді алдын ала өңдеудің, бөлу тәсілдерінің және бағалау метрикаларының бірізді шарттарында бағаланды. Қазіргі таңда ағашқа негізделген модельдер құрылымдалған уақытқа тәуелді деректердегі сызықтық емес тәуелділіктерді анықтау үшін уақыттық белгілерді (temporal feature engineering) айқын түрде құрастыруды талап етеді. Ал LSTM желілері жүктеме мәндерінің тізбекті бақылаулары негізінде уақыттық тәуелділіктерді тікелей үйренеді. Модельдердің өнімділігі RMSE және MAPE көрсеткіштері арқылы бағаланды. Нәтижелерге сәйкес, LSTM уақыттық динамиканы тиімді қамтуының арқасында ең жоғары болжау дәлдігін көрсетті, ал XGBoost есептеу күрделілігі төмен бола отырып, бәсекеге қабілетті нәтижелер ұсынды. Бұл электр энергиясын тұтынуды болжау барысында модель дәлдігі, интерпретациялануы және есептеу тиімділігі арасында контекстке тәуелді маңызды теңгерімдердің бар екенін көрсетеді.
Түйін сөздер: электр энергиясын тұтынуды болжау, деректер ғылымы, машиналық оқыту, терең оқыту, XGBoost, LSTM, уақыттық қатарларды болжау, жүктемені болжау, белгілерді құрастыру, базалық модель.

Модели на основе науки о данных для прогнозирования потребления электроэнергии

[1]А.Қанатнұр, [1]И.Б.Айзерен
[1]Кипрский международный университет (Никосия, Кипр)

Аннотация. Точное прогнозирование потребления электроэнергии является ключевым элементом современных энергетических систем как для планирования, так и для принятия оперативных решений. В данном исследовании предлагается методология прогнозирования спроса на электроэнергию, основанная на науке о данных, а также проводится прямое сравнение методов машинного обучения и глубокого обучения в идентичных экспериментальных условиях. В частности, древовидная модель XGBoost и рекуррентная нейронная сеть Long Short-Term Memory (LSTM) были оценены на одном и том же наборе данных при одинаковых процедурах предварительной обработки данных, разбиения выборок и метриках оценки. В настоящее время модели на основе деревьев используют явное конструирование временных признаков (temporal feature engineering) для выявления нелинейных зависимостей в структурированных временных данных. В отличие от них, сети LSTM непосредственно обучаются временным зависимостям на основе последовательных наблюдений нагрузки. Оценка качества моделей проводилась с использованием показателей RMSE и MAPE. Полученные результаты показывают, что модель LSTM обеспечивает наивысшую точность прогнозирования благодаря более эффективному захвату временной динамики, тогда как XGBoost демонстрирует конкурентоспособную производительность при меньшей вычислительной сложности. Это свидетельствует о наличии контекстно-зависимых компромиссов между точностью модели, ее интерпретируемостью и вычислительной эффективностью при прогнозировании потребления электроэнергии.
Ключевые слова: прогнозирование потребления электроэнергии, наука о данных, машинное обучение, глубокое обучение, XGBoost, LSTM, прогнозирование временных рядов, прогнозирование нагрузки, конструирование признаков, базовая модель.

Introduction

The forecasting of the quantity of electricity which can be consumed has emerged as a significant issue given that the power sector is becoming more and more complex and dynamic. Increasing urbanization, increasing industrialization, and the overall electrification of homes and industries have significantly changed the way electricity is consumed globally. Since electricity generation and consumption need to be in sync in real-time, even small errors in forecasting can result in inefficient use of resources, increasing costs, and instability in the grid [GEA, 2024; Bouktif1 S. et al., 2018].

The overall energy consumption worldwide has increased by nearly 90% since 2000, and in consequence, there is a higher requirement to improve power infrastructure to be effective and consistent [GEA, 2024)]. As the fluctuations in demand increase, so does the need to forecast accurately to allocate loads, plan infrastructure, and maintain them. Even though conventional mathematical systems have frequently been used to accomplish these functions, their reliance on linearity and assumptions of linearity is limiting in their potential to accommodate irregular and multi-seasonal characteristics of present-day energy consumption patterns. Recent studies have increasingly focused on information-based systems that are more likely to accommodate complex time structures [Misiurek K. et al., 2025].

With recent advancements in data science techniques, energy load forecasting techniques have also been significantly influenced by capturing nonlinear relationships and temporal patterns that were difficult to represent with statistical models. Traditional models like autoregressive and econometric models require high linearity and stationarity properties to fit into the model, but with recent electric energy consumption data that is highly variable and has multiple seasonality and structural changes, these models fail to deliver [Misiurek K. et al., 2024]. However, with the application of tree-based models like ML models, they have shown promising results by using temporal patterns like lagged variables and calendar variables [Roy M. et al., 2025].

On the other hand, the application of deep learning methods, especially recurrent neural networks and Long Short Term Memory (LSTM) networks, has also received increased attention due to their capability to learn the temporal dependencies directly from the data without the need for extensive feature engineering. Recent review studies have particularly stressed that models based on LSTMs are highly effective for learning the long-term seasonal patterns in electricity load forecasting

problems [Kasprzyk M. et al., 2024]. Nevertheless, such studies have also indicated that the relative performance of various machine learning and deep techniques is largely dependent on the context, which again emphasizes the importance of fair comparison.

There are many different forecasting models available; however, most studies are still focused on individual models rather than comparing them. In addition, different studies use a wide range of data sets, forecasting periods, preprocessing techniques, and evaluation criteria. This makes it hard to compare the results or find out which model is actually performing the best. There is still a need to compare machine learning and deep learning models head to head. This is still a concern, as stated in the latest review articles. To correctly measure the real progress regarding the power usage forecast, we need benchmarks of consistent evaluation and fair comparison under equal conditions for all models [Misiurek K. et al., 2024].

Apart from the methodology employed, there is practical significance of accurate electricity consumption prediction that can have a huge impact on the performance of utilities in grid management and energy planning. Precise short-term predictions facilitate better decisions about power generation, reserve capacity, and electricity tariffs, minimizing the losses from either an undersupply or oversupply of electricity. Practical significance of the issue is especially evident in the context of developing countries and countries in transition, where decision making on infrastructure investments has to be conducted in terms of limitations on computational and financial resources.

Kazakhstan serves as a concrete example. With the ongoing rapid modernization of the country’s power sector and increasing integration of renewable energy sources and smart grid technology [Zhakiyev N. et al., 2022], there have been several recent works in which various individual methods of electricity load forecasting were used on Kazakhstani load datasets: autoregression and neural networks for load in Astana [Sagadatova N. et al., 2023]; deep recurrent neural networks for national load consumption [Kabdygali S. et al., 2023]; hybrid regression methods for Astana's demand characteristics [Satan A. et al., 2025]; machine learning algorithms used on 20 years of national consumption records [Madin V. et al., 2025]. None of these works use machine learning and deep learning methods in direct comparison under the same pre-processing, splitting, and evaluation conditions, thus limiting the applicability of their findings for model choice in real-world settings. It is this gap in the literature that the current research aims to address: only such a comparison will be able to provide electricity consumers in Kazakhstan and elsewhere with an answer as to which method provides the optimal combination of predictive ability and computational cost for them.

This paper proposes a methodology for electricity consumption forecasting using data science techniques that enables a fair comparison between machine and deep learning techniques, with the intent to fill the aforementioned research gap. In particular, the performances of a machine learning model, XGBoost, are compared with those of a deep learning model, represented by Long ShortTerm Memory. These two models have been compared on the same dataset, with the same data splits, and the same evaluation metrics. In particular, the accuracy of each model has been evaluated through Root Mean Squared Error and Mean Absolute Percentage Error. This research aims at suggesting practical insights to researchers and practitioners about how to choose an appropriate model for electricity consumption forecasting, hence providing more efficient and reliable strategies for managing electricity consumption. The results of this research will be used to provide practical insights to researchers and practitioners regarding how to choose an appropriate model for electricity consumption forecasting and develop more efficient and reliable strategies in managing electricity consumption [Kasprzyk M. et al., 2024]. In particular, the methodological framework adopted in this research is presented in the next section.

Main body

Electricity consumption forecasting is, by its very nature, a time series model with strong temporal dependencies, seasonal patterns, and nonlinear effects. More recently, however, with advancements

made in data science, it has been shown that it is possible to use machine learning and deep learning models to enhance the accuracy of the forecasts made compared to traditional statistical models, particularly for high-resolution data sets [Kasprzyk M. et al., 2024; Roy M. et al., 2025]. However, it is highly dependent on how well it is able to model or learn the temporal effects.

Tree-based machine learning models, such as XGBoost, rely on feature engineering to incorporate temporal features such as lag effects and moving averages. These models can effectively approximate non-linear relationships given proper feature engineering due to their computationally efficient nature [Roy M. et al., 2025]. Deep learning models like Long Short Term Memory (LSTM) approach time series forecasting differently. They can learn patterns and relationships directly from data without feature engineering. These models can be effective when long-term trends need to be learned and predicted because they do not require feature engineering [Kasprzyk M. et al., 2024].

In this study, a unified framework is presented to enable a direct comparison between baseline, machine learning, and deep learning models. For a comparative study, the data used for each model is the same, and the train-test splits are identical, as are the evaluation metrics. In this research, lag-based forecasting, a feature-engineered XGBoost, and a sequence-based LSTM are implemented, and the performance of each model is tested for electricity consumption forecasting. An interesting finding is that increasing model complexity is not always associated with increasing accuracy, and simpler models can actually do the job with the same accuracy [Li K., 2023].

Methods

The data set used for the analysis of the problem is the Electricity Load Diagrams Electricity Load Diagrams 2011-2014, 2015. As provided by the repository, the dataset contains electricity consumption records collected exclusively between January 2011 and December 2014. No data outside this time period were available or included in the analysis. Consequently, all experiments, model training, and evaluations conducted in this study are strictly based on electricity load data corresponding to the 2011–2014 period. The data set includes records of electricity demand from 370 consumers and is recorded at 15-minute intervals [Trindade A., 2015]. Each column represents the data from a particular consumer, and the timestamp represents the time at which the data was recorded.

There were missing and zero-valued readings recorded at the early stages of the data set, and data cleaning was done at the beginning of the analysis. The target variable includes the energy load data. The forecast includes the prediction of the data based on the recorded data.

For the machine learning model, the data was subjected to feature engineering based on the time of the data set recorded. Pre-aggregation at an hourly level was done as a pragmatic approach to reduce noise in the data set while still holding the characteristics of the data set for XGBoost modeling.

In the case of deep learning, an LSTM was trained by using sequential input windows so that it could learn dependencies from time series without much feature extraction. The two models were trained on the same splits to ensure a proper comparison.

The performance of the two models was compared based on MAPE and RMSE.

$$RMSE = \sqrt{\left(\frac{1}{N}\right) * \sum_{i=1}^{N} \left(\left(y_i - y_{hat_i}\right)^2\right)}$$

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

The overall research design and procedural steps of this study are illustrated in Figure 1, which outlines the transition from data acquisition to the final comparative analysis of the forecasting models.

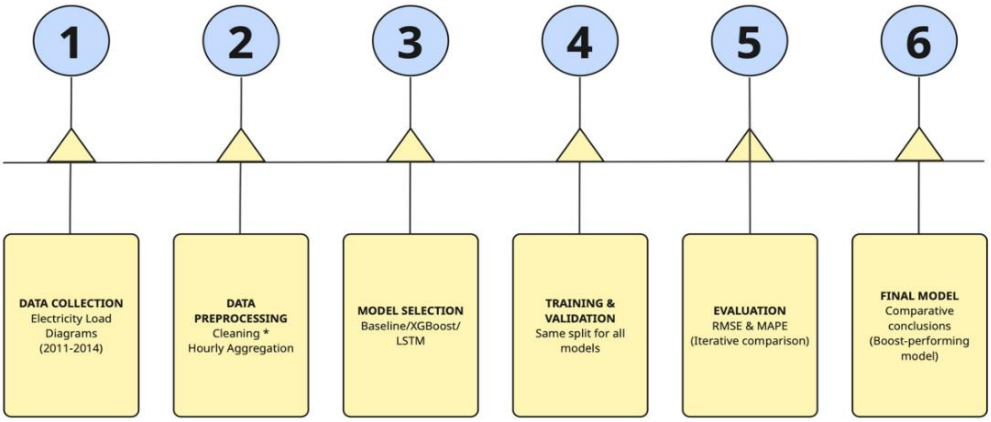


Figure 1. Workflow Diagram of the Electricity Load Analysis

LSTM model

Long Short Term Memory (LSTM) neural networks have been utilized for including sequential dependencies in the electricity consumption data. It should be noted that the primary rationale behind developing the LSTM neural networks was to avoid the 'vanishing gradient problem.' By using this property, a long-term dependence can be included in the data, and this has made the use of the LSTM neural network very popular in forecasting, particularly in electricity energy demand forecasting, as the data shows strong non-linearity and seasonality in the patterns [Kasprzyk M. et al., 2024; Alexiadis V. et al., 2025].

In this study, the Electricity Load time series was aggregated into hourly intervals and then into a supervised learning problem with a sliding window approach: the sampling of time series data on fixed-size window size. Each input sequence is defined by a number of previous hourly intervals, and the target variable is defined by the electricity usage at the next point in time. In this way, an LSTM neural network can easily provide temporal dependencies of the data without intensive human intervention, which would normally be needed when using other tree-based machine learning models. Likewise, the same split strategy used for the baseline and XGBoost models was applied in this study.

In the case of the LSTM model, the loss function was considered as the minimization of Mean Squared Error, and the minimization was done via a gradient-based optimization approach. In that process, the LSTM model was able to update its memory states in such a way that it could learn short-term changes along with capturing seasonal variation, which is present in the electricity consumption data. Later, the LSTM model forecasting performance was verified on the test data through the Root Mean Squared Error and Mean Absolute Percentage Error, which also allowed us to compare the LSTM model performance with other baseline and machine learning models under similar experimental conditions. Similar kinds of LSTM-based formulations have seen to perform well in electricity load forecasting, where the dependencies are more significant [Roy M., 2025].

Results

Table 1 shows the performance of the baseline lag models and the LSTM model, as measured by RMSE and MAPE. Baseline benchmarks for Naive-24 and Naive-168 lag forecasting models were considered for the analysis. The performance of the Naive-168 model had lower error values for forecasting in comparison to the Naive-24 model (565.60 for RMSE and 4.47% for MAPE in contrast to 745.68 for RMSE and 5.78% for MAPE in the Naive-24 model). This confirmed the presence of strong weekly seasonality within the energy consumption data set. The 15-minute-based original series was used for aggregation purposes on an hourly scale using 60-minute time intervals

for ease of modeling along with the presence of daily or weekly seasonality within the series. Figure 2 – Actual vs. Predicted Electricity Consumption Values Using XGBoost.

Table 1

Baseline forecasting performance

Model	RMSE	MAPE (%)
Naive-24 (t–24)	745.68	5.78
Naive-168 (t–168)	565.60	4.47
LSTM	148.70	1.28

Comparing the performance of the XGBoost model with respect to the baseline methods, it can be seen that the XGBoost model performed significantly better. Compared to the Naive-168 model, the XGBoost model has significantly lower forecasting errors compared to the baseline. This shows that the time-based feature engineering has been able to account for a large portion of the seasonality.

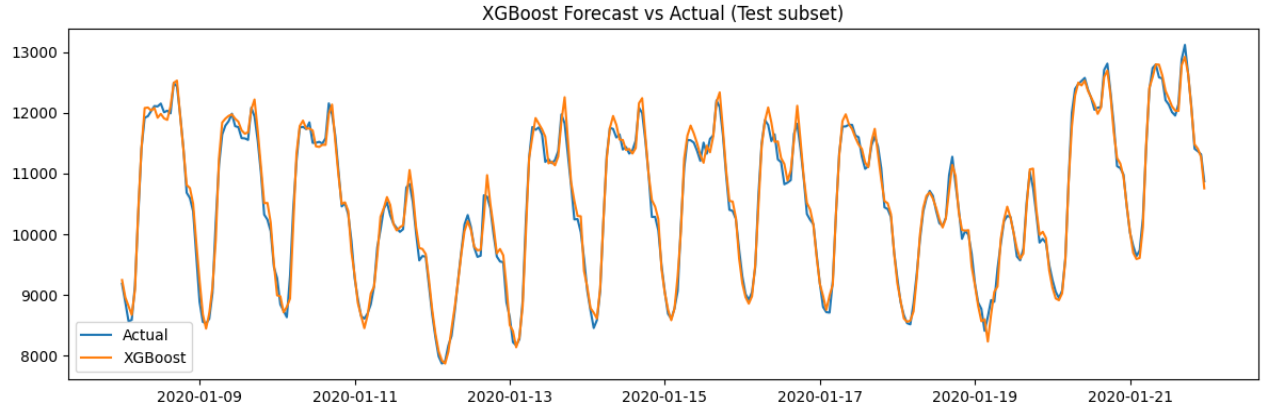


Figure 2. Comparison of actual and XGBoost predicted electricity consumption values over a subset of the test period (hourly resolution)

Model comparison results

The forecasting performance of the three models on the same hourly test split is summarized in Figure 3. Both the learning-based models outperformed the Naive-168 model by a significant margin. XGBoost model obtained an RMSE = 155.80 and MAPE = 1.32%, whereas the LSTM model obtained the lowest errors with an RMSE = 148.70 and MAPE = 1.28%. This shows that both the engineered tree-based learning model and the sequence-based deep learning model can effectively capture strong seasonal patterns in the electricity consumption data, though the LSTM model obtained a minor accuracy improvement over the former model. Figure 3 shows a comparison between the actual and predicted values of the electricity consumption data obtained by the LSTM model on a selected subset of the test data at the hourly level.

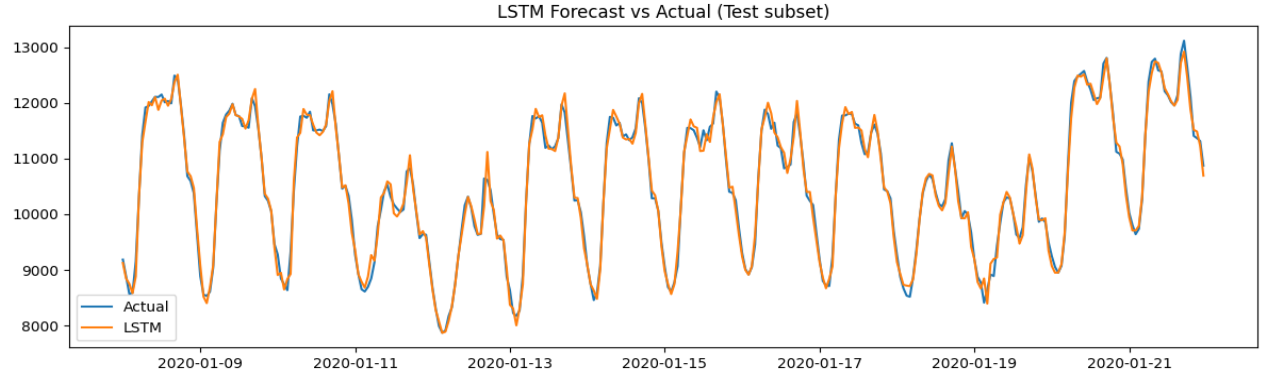


Figure 3. Comparison of actual and LSTM predicted electricity consumption values over a subset of the test period (hourly resolution)

Figure 4 presents the residual analysis of the LSTM model on the test dataset, including residuals over time and the distribution of residuals at hourly resolution.

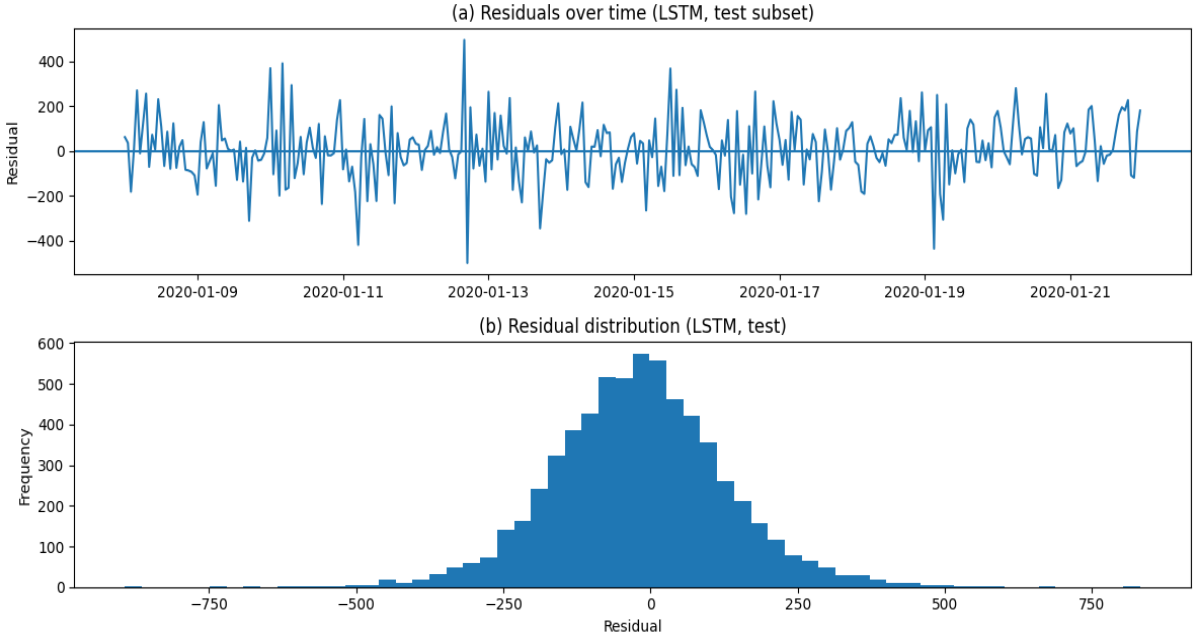


Figure 4. Residual analysis of the LSTM model on the test set: (a) residuals over time and (b) residual distribution (hourly resolution)

Residual analysis indicates that LSTM prediction errors are centered around zero without systematic drift over time. The residual distribution is approximately symmetric, suggesting the absence of significant bias in the model forecasts.

Discussion

One of the important observations based on the experimental outcomes is the high performance of the Naive-168 baseline model, which is based on the prominent weekly seasonality of the electricity consumption patterns. However, an important methodological consideration based on these observations is that, especially for data with prominent periodic patterns, persistence-based models can already deliver high accuracy. Moreover, the relatively moderate outperformance of the learning-based models over the baseline model indicates that it is not always the case that an increase in model complexity will lead to significant increases in model performance. From an operational perspective, it is important to consider whether complex models are always necessary [Alexiadis et al., 2025].

The distinction between these two learning-based approaches serves to emphasize the importance that real-world modeling offers in terms of understanding modern energy usage. In this regard, XGBoost determines that, with effective temporal feature engineering, seasonality and nonlinear trends in electric load data can be effectively modeled. As such, XGBoost is effective in dealing with complex variation in data usage patterns without necessarily involving computationally expensive mechanisms. Moreover, with XGBoost, there exists an understandable mechanism for such effectiveness, and these assertions apply in line with what has been demonstrated in literature, indicating that, with effective temporal feature engineering, gradient boosting-based techniques and models are effective and reliable as long as they are applied in a structured time series forecasting approach [Yang B. et al., 2025].

On the other hand, the LSTM model relies on sequential learning as opposed to the use of feature construction. For instance, even as the results were slightly better, there was a sense of diminishing returns as one transitions from feature-based machine learning models to deep learning models. This result agrees with the recent research indicating that deep learning models do not improve

significantly in terms of accuracy unless the data presents complex dependencies [Yang B. et al., 2025].

From a methodological point of view, the obtained results emphasize the necessity to formulate standard schemes of model evaluation in the field of electricity consumption forecasting research, as in the majority of the existing articles, the training and evaluation data sets used are different, and therefore, the comparison of the forecasting model performance is difficult. By using the same model evaluation scheme, this research gives a better understanding of the actual differences among the evaluated model performances.

The results obtained in electricity consumption forecasting problem solve can be considered satisfactory, despite the need to observe several limitations. First of all, it should be emphasized that the data used for electricity consumption forecasting was unique publicly available data. Therefore, it is not automatically applicable for general purposes since it was recorded in a definite location at a definite time. Secondly, it is considerable that the data of one kind was analyzed in the research. At the same time, there was no weather or socio-economic information. It is evident that such factors influence electricity consumption in general, particularly in cases where demand varies considerably in one direction or another.

The 15-minute data were transformed into hourly average values. This reduced the noise in the data and made the calculations easier, but it smoothed important short-term variations that the operators need to know about. Finally, this study focused on point forecast evaluation using RMSE and MAPE measures. Forecast uncertainty was not addressed, which leaves one important evaluation criterion uncovered for future work. This is consistent with past studies which have found that enhanced training goals and loss optimization can improve performance for deep learning models in energy forecasting techniques [Nussipova F. et al, 2024].

These findings provide a basis for further research. The involvement of external factors in the calculation of variables, the probabilistic method, the use of different time intervals, prediction ranges, and hybrid models can be useful in enhancing the reliability and practicability of the prediction. The findings of this study were sufficient to provide a direction towards effective ways of enhancing the accuracy of prediction using machine learning models and deep learning models. However, a particular model can only be effective within a particular scenario. It is paramount to compare each model based on the requirements of the scenario and the interpretability of the findings [Li K., 2023].

Conclusion

In this research, we proposed a data science platform for electricity usage prediction, was introdused a structured investigation of the existing baseline, machine learning, and deep learning models on an equal footing, the implementation of the Naive lag-based models, XGBoost, and LSTM models using the Electricity Load Diagrams 2011-2014 data set for hourly usage data. We have used RMSE and MAPE as the metrics for the proposed model. By using the baseline models, the improvement relative to the existing seasonal predictions can be measured.

The achieved results confirm that Naive-168 was able to properly reproduce the rhythm inherent in weekly electricity consumption, thus proving that the data does contain strong temporal patterns. At the same time, both learning-based approaches were able to significantly outperform the baseline method. This, in turn, indicates that currently applied methods of electric load prediction indeed heavily rely on data-centric approaches rather than weaker heuristic-style methods. Furthermore, in addition to this, XGBoost was able to obtain a smaller prediction error on account of properly designed features, namely demand values, averages, and calendar features, without sacrificing efficiency in any way.

Among the tested models, LSTM achieved the best accuracy in forecasting, albeit marginally better than XGBoost, which learned time-dependent patterns directly from input windows. This shows that the class of recurrent neural networks is effective in learning complex time dependencies in

electricity demand. However, the margin between LSTM and XGBoost was not large, which suggests that feature-based models can be competitive in terms of accuracy while having much less computation complexity. Therefore, the choice between LSTM and XGBoost should not be made based on accuracy alone, but also on other criteria.

This study shows that there is no single universally optimal model for electricity consumption forecasting. The best model choice depends on the resolution of the available data, the strength of seasonality within that data, and the specific requirements placed on the model itself. Where computational resources are limited and/or easily interpretable results are required, the tree-based model may be the preferable option.

Of course, beyond the numerical results, this study is important to the field because it shows the importance of proper benchmarking in the forecasting of electric power consumption. By comparing the simpler models to the more complex machine learning and deep learning approaches, it is easy to understand the added value that these approaches can provide. This is important to ensure that expectations are not overstated regarding the ability to use these approaches in the real world [Review of Methods and Models for Forecasting Electricity Consumption, 2024].

The results also suggest that the issue of forecasting accuracy needs to be considered in conjunction with other factors such as the practical constraints. Although the results of the deep learning models such as LSTM are good in learning the temporal relationships, the increased computational cost and lack of interpretability of the results may limit the usage of the models. On the other hand, the feature-based models such as XGBoost have a good balance of efficiency and accuracy.

Lastly, the proposed framework is also useful as a guide for conducting comparative studies for electricity demand forecasting in the future. Using a similar data preprocessing methodology, evaluation measures, and baselines, it is possible to ensure that any reported improvements are valid and reproducible. As the energy landscape continues to change with increasing penetration of renewable energy and demand volatility, accurate and understandable forecasting models will become ever more vital tools for decision-makers. As such, the results of this study are not only useful for providing insights but also for guiding the development of data-driven research for electricity consumption forecasting. In this sense, the study directly addresses the motivating gap identified at the outset: for practitioners in Kazakhstan and similar energy markets, it provides one of the first controlled, head-to-head comparisons needed to choose between accuracy-oriented and resource-efficient forecasting approaches with confidence.

References

GEA (2024) Global electricity demand rises by 90% since 2000. *Global Energy Association,* August, 19 [Electronic resource]: URL: https://globalenergyprize.org/en/2024/01/19/global-electricity-demand-rises-by-90-since-2000/ (date of access: 09.12.2025).

Bouktif1 S., Fiaz1 A., Ouni A., Serhani1 M.A. (2018) Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies*, 2018, *11*(7), 1636; https://doi.org/10.3390/en11071636 [Electronic resource]: URL: https://www.mdpi.com/1996-1073/11/7/1636 (date of access: 29.01.2026).

Misiurek K., Olkuski T., Zyśk J. (2025) Review of Methods and Models for Forecasting Electricity Consumption. *Energies*, 18, 4032. DOI: https://doi.org/10.3390/en18154032 [Electronic resource]: URL: https://www.mdpi.com/1996-1073/18/15/4032 (date of access: 13.12.2025).

Roy M., Pyltsov V., Hu Y. (2025) IISE PG&E Energy Analytics Challenge 2025: Hourly-Binned Regression Models Beat Transformers in Load Forecasting. arXiv preprint, 16 May. DOI: https://doi.org/10.48550/arXiv.2505.11390 [Electronic resource]: URL: https://arxiv.org/abs/2505.11390 (date of access: 19.12.2025).

Kasprzyk M., Pełka P., Oreshkin B.N., Dudek G. (2024) Enhanced N-BEATS for mid-term electricity demand forecasting. *arXiv preprint, 2 December*. DOI:

https://doi.org/10.48550/arXiv.2412.02722 [Electronic resource]: URL: https://arxiv.org/abs/2412.02722 (date of access: 19.12.2025).

Li K., Huang W., Hu G., Li J. (2023) Ultra-short term power load forecasting based on CEEMDAN-SE and LSTM neural network. *Energy and Buildings*, 279, 112666. DOI: https://doi.org/10.1016/j.enbuild.2022.112666 [Electronic resource]: URL: https://www.sciencedirect.com/science/article/abs/pii/S0378778822008374?casa_token=TRUhcOeOrw8AAAAA:azXyuDkf8Omith7wOLO2Zk3fUWJuUvOHyexxSIjYGS7KghrYrNOo9AooKNewc-kVAKbKgvtvaw (date of access: 29.01.2026).

Trindade A. (2015) Electricity Load Diagrams 2011-2014. *UCI Machine Learning Repository,* University of California, Irvine. DOI: https://doi.org/10.24432/C58C86 [Electronic resource]: URL: https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014 (date of access: 25.01.2026).

Alexiadis V., Drakaki M., Tzionas P. (2025) LSTM-based electricity demand forecasting in smart buildings. *Electronics, 14*(22), 4456. DOI: https://doi.org/10.3390/electronics14224456 [Electronic resource]: URL: https://www.mdpi.com/2079-9292/14/22/4456 (date of access: 23.12.2025).

Yang B., Gül M., Chen Y. (2025) Comparative analysis of deep learning and tree-based models in power demand prediction: Accuracy, interpretability, and computational efficiency. *Journal of Building Physics, 49*(1), 127-169. DOI: https://doi.org/10.1177/17442591241234567 [Electronic resource]: URL: https://journals.sagepub.com/doi/10.1177/17442591251333144 (date of access: 25.01.2026).

Nussipova F., Rysbekov S., Abdiakhmetova Z., Kartbayev A. (2024) Optimizing loss functions for improved energy demand prediction in smart power grids. *International Journal of Electrical and Computer Engineering* (IJECE), 14(3), 3415–3426. DOI: https://doi.org/10.11591/ijece.v14i3.pp3415-3426 [Electronic resource]: URL: https://d1wqtxts1xzle7.cloudfront.net/115869009/99_34872_IJECE_DB_F-libre.pdf (date of access: 01.02.2026).

Zhakiyev N., Kalenova A., Khamzina A. (2022) The Energy Sector of the Capital of Kazakhstan: Status Quo and Policy towards Smart City. International Journal of Energy Economics and Policy, 12(4), 414–423. DOI: https://doi.org/10.32479/ijeep.13126 [Electronic resource]: URL: https://doi.org/10.32479/ijeep.13126 (date of access: 20.08.2026).

Sagadatova N., Talas B., Nugumanova A., Zhakiyev N. (2023) Forecasting Electricity Consumption: Case Study in Astana. Scientific Journal of Astana IT University, 14(14), 16–25. DOI: https://doi.org/10.37943/14TRMF1662 [Electronic resource]: URL: https://journal.astanait.edu.kz/index.php/ojs/article/view/285 (date of access: 20.08.2026).

Kabdygali S., Omirgaliyev R., Tursynbayev T., Kayisli K., Zhakiyev N. (2023) Deep Recurrent Neural Networks in Energy Demand Forecasting: A Case Study of Kazakhstan's Electrical Consumption. Scientific Journal of Astana IT University, 16(16). DOI: https://doi.org/10.37943/16YIKA8050 [Electronic resource]: URL: https://doi.org/10.37943/16YIKA8050 (date of access: 20.08.2026).

Satan A., Zhakiyev N., Nugumanova A., Friedrich D. (2025) Hybrid feature-based neural network regression method for load profiles forecasting. Energy Informatics, 8, 19. DOI: https://doi.org/10.1186/s42162-025-00481-0 [Electronic resource]: URL: https://link.springer.com/article/10.1186/s42162-025-00481-0 (date of access: 20.08.2026).

Madin V., Salykova O., Ivanova I., Bizhanova O., Aldasheva D. (2025) Enhancing Electricity Consumption Forecasting in The Republic of Kazakhstan Using Machine Learning. Journal of Applied Engineering and Technological Science (JAETS), 6(2), 1166–1196. [Electronic resource]: URL: https://journal.yrpipku.com/index.php/jaets/article/view/7425 (date of access: 20.08.2026).

Поступила в редакцию: 10.02.2026
Одобрена: 31.03.2026
Первая публикация: 26.06.2026