2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT)
The Maldives National University, Male City (Maldives). Mar 11-12, 2023

2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT) | 979-8-3503-3211-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/GLOBCONHT56829.2023.10087707

Short-term Electricity Price Forecasting Using Interpretable Hybrid Machine Learning Models

Hamza Mubarak
School of Engineering and Built Environment
Griffith University
QLD, 4222, Australia
hamza.mubarak@hotmail.com

Shameem Ahmad
Department of Electrical and Elctronic Engineering
American International University-Bangladesh
Dhaka 1229, Bangladesh
ahmad05shameem@gmail.com

Al Amin Hossain
Military Institute of Science and Technology
Dhaka 1216, Bangladesh
Reverie Power and Automation Engineering Limited
Dhaka 1215, Bangladesh
hossain.amin899@gmail.com

Ben Horan
School of Engineering
Deakin University
Geelong, VIC 3216, Australia
ben.horan@deakin.edu.au

Abdallah Abdellatif
Department of Electrical Engineering
Universiti Malaya
Kuala Lumpur 50603, Malaysia
abdallahh950@hotmail.com

Saad Mekhilef
Mehdi Seyedmahmoudian
Alex Stojcevski
School of Science, Computing and Engineering Technologies
Swinburne University of Technology
Hawthorn, VIC 3122, Australia
smekhilef@swin.edu.au
mseyedmahmoudian@swin.edu.au
astojcevski@swin.edu.au

Hazlie Mokhlis
Jeevan Kanesan
Department of Electrical Engineering
Universiti Malaya
Kuala Lumpur 50603, Malaysia
hazli@um.edu.my
jievan@um.edu.my

Mohamed Becherif
FEMTO-ST Institute
Univ. Bourgogne Franche-Comté, UTBM
CNRS Rue Ernest Thierry Mieg, 90000 Belfort, France
mohamed.becherif@femto-st.fr

***Abstract*—In this paper, a combination of single and hybrid Machine learning (ML) models were proposed to forecast the electricity price one day ahead for the Nord Pool spot electricity market. The proposed models were evaluated based on performance metrics, such as Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Error (MAE). Further, a model interpretation by employing SHapley Additive exPlanations (SHAP) framework to show the impact of each feature in the forecasting output. Based on the SHAP, the lag electricity price EP(t-1) impacts the forecast result most, followed by EP(t-2) and time stamp, respectively. Finally, the results show that hybrid models performed better than single ones, where the LR-CatBoost model surpassed other models and attained 7.94 and 10.49, which are the lowest values of MAE and RMSE respectively. Moreover, the kNN and SVM models performed poorly, achieving the highest RMSE values of 12.88 and 12.39, respectively.**

Keywords— Machine learning, electricity price forecasting, XGBoost, CatBoost, time series, hybrid models.

I. Introduction

Recently, power system planning has incorporated a variety of resources in order to meet the growing demand while adhering to many techno-economic and environmental considerations [1]. All areas of power system operation place a premium on electricity price forecasting (EPF), as seen by the huge number of researchers focusing on operation-related topics [2]. Solving the EPF issue is now becoming increasingly difficult. The applications made available by a price forecasting model are essential to attaining the power transition [3]. Further, the authorities enable owners of renewable power-generating means to earn money on the market by forecasting price changes [4], and they encourage intelligent purposes like self-consumption [5] or battery optimization for automobiles [6]. Additionally, several aspects must be considered to comprehend electricity prices. For instance, energy transition strategies enhance the share of renewable energy in overall output and implement new market regulations like carbon dioxide emission taxes. Furthermore, the pricing methods employed to regulate generation and consumption can potentially produce pleasant and unpleasant price spikes. These spikes can cause enormous losses for unprepared business owners and are hard to predict with conventional prediction methods [7]. Specifically, the present time is characterized by frequent lockdowns that have resulted in significant alterations to the global market.

Many methods were used to forecast the EPF. For instance, the traditional methods, such as the autoregressive integrated moving average (ARMA) [8] and autoregressive integrated moving average exogenous variable (ARMAX) [9] models, were employed to forecast the electricity price and peak load, respectively. Still, these models are predicated mostly on linear relationships. Besides, they are limited in their ability to demonstrate nonlinear correlations. Hence, to cope with the limitation of the traditional method, machine learning (ML) is introduced to assist in forecasting multiple objectives in different areas. For example, ML was used in many fields, such as forecasting photovoltaic power generation [10, 11], heart disease diagnoses [12, 13], and fault diagnosis [14]. ML models are easy to implement. In contrast to deep learning (DL) approaches, which require huge amounts of data and high computational time [15].

Single ML models were widely used for forecasting the electricity price. For example, a support vector machine (SVM) is used to predict the electricity price by considering Germany's historical electricity and gas prices [16]. The extreme learning machine (ELM) model was proposed in [17] for EPF, where the results showed that the proposed model attained accurate results with less computation time compared with the previous work. Finally, an extreme gradient boosting (XGBoost) model was presented in [18] for Ontario EPF. Though the obtained results ensured decent values of mean absolute error (MAE) and mean squared error (MSE) however, single models have some drawbacks, such as the incapability of managing complicated problems and producing reliable forecasts. Therefore, hybrid ML was proposed to address these types of issues. A hybrid model from kernel function and principle component analysis along with SVM was presented in [19] for EPF utilizing formerly underused predictive characteristics, like the price records of nearby countries.

CatBoost and XGBoost models could not forecast accurately because quick spike, high volatility and seasonality of the time-series price dataset. With substantial forecasting residuals between real and predicted values, these methods may produce poor outcomes. Furthermore, they are ineffective at recognizing nonlinear time-series behavior and have poor predictive capabilities. Moreover, a number of hybrid methods have been employed to predict power prices in previous publications.

Nevertheless, the majority of researches concentrate on combining linear and DL algorithms. DL methods has complex architecture and demands enormous computer resources. Therefore, this study to improve EPF performance combined ensemble tree-based models with linear regression.

The remaining paper is structured as follows. Section II described the methodology of the proposed models and shows the model interpretation by employing the SHAP framework. Section III demonstrates the forecasting results which are obtained using the proposed hybrid ML models. Finally, to sum up, a conclusion will be presented in Section IV at the end of the paper.

II. METHODOLOGY

This part clarifies the preparation and partitioning of the used Nord Pool spot market dataset and the evaluation metrics which have been used to measure the performance of the proposed models. Further, the working principle of the proposed models have been presented. Lastly, the model interpretation by using the SHAP framework will be illustrated to show the impact of each feature in the forecasting results.

A. Data preparation and partitioning.

The proposed forecasting models were implemented in this work using an actual energy cost dataset gathered from the electricity market of Nord Pool spot [20]. The electricity price for Nord Pool spot market is collected 24 hours daily for an interval of one hour. Further, for the year 2021, the time series data is considered which have been accumulated into a single CSV file. The whole chosen dataset is divided into two parts; 80% of the data is used in the training phase to develop the forecasting model, whereas the remaining 20% of the data is employed in the testing phase for evaluation [21]. Throughout the training stage to fine-tune the model, time-series forecasts have been performed using sliding validation and the models are continually trained K times until the convergence is reached. In this case, a value of 24 is chosen for K.

Figure 1 and Figure 2 represent the autocorrelation and the partial one, where the lag days can be seen from these figures, which impact the forecasting results. The first and the second lag have the highest positive influence on the model output, followed by the third lag (k-3), but it impacts it negatively. The rest followed the same pattern but less impacted the final forecasting results.

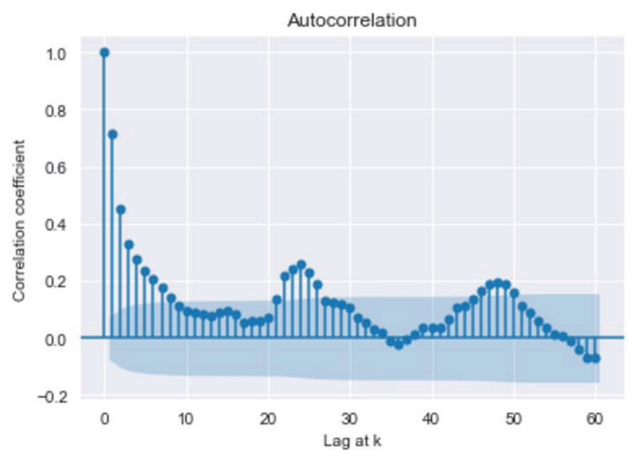


Fig. 1. The autocorrelation for lag days in EPF.

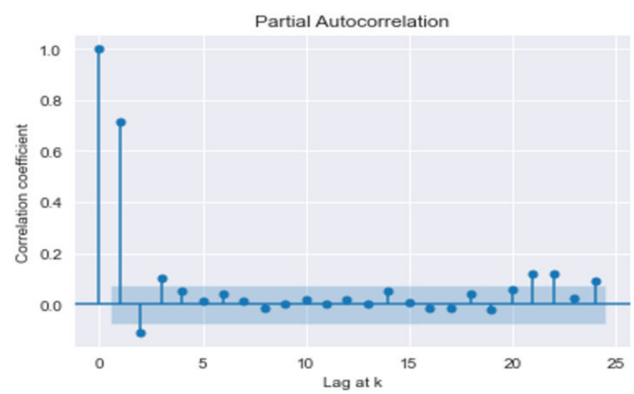


Fig. 2. The partial autocorrelation for lag days in EPF.

The normalization of the data was achieved by applying the standard deviation. The following equations (1-4) outline this procedure [22]. The mean is indicated by μ, whereas the referred σ is used to define standard deviation. In addition, from dataset $data_i$ is referred to the value of the data at i and M represents the size of the dataset. In Eq.(3), how before training the data is standardized is depicted and in Eq.(4), to evaluate the efficacy of the testing data compared to the trained data, how the real data forecasted ($EPF_{Fore.-Real}$) is forecasted is elaborated [23]. Using Python 3.8, all the experiments of this study are performed on a desktop PC which consisted of 16 GB memory, microprocessors with six Core i7-9750H, and Max-Q NVIDIA GeForce GTX 1660 Ti.

$$\mu = \frac{1}{M}\sum_{m=1}^{M} Data_{set} \quad (1)$$

$$\sigma = std(Data_{set}) = \sqrt{\frac{\sum_{m=1}^{M}(data_i - \mu)^2}{M}} \quad (2)$$

$$Data_{set}{}^{std.} = \frac{(Data_{set} - \mu)}{\sigma} \quad (3)$$

$$EPF_{Fore.-Real} = \sigma \cdot O_{Fore.-stand.} + \mu \quad (4)$$

B. Performance Metrics

In Eqs. (5-7), the explanation for all models' performance indices namely Root Mean Square Error (RMSE), MSE and MAE respectively are presented. The forecasted and the real values are represented A and B respectively.

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(A - B)^2} \quad (5)$$

$$MSE = \frac{1}{M}\sum_{i=1}^{M}(A - B)^2 \quad (6)$$

$$MAE = \frac{1}{M}\sum_{i=1}^{M}|(A - B)| \quad (7)$$

C. Machine Learning Models

In this study, ML models, including hybrid ML models (LR-CatBoost, LR-XGBoost, LASSO-CatBoost, and LASSO-XGBoost), as well as single models (LASSO, LR, CatBoost, XGBoost, RFR, kNN, and SVR) were used for EPF utilizing the Nord Pool spot electricity market and to determine the best predictive model which attains the lowest error according to performance matrices. The basic models will be elaborated on in the following subsections.

1) Categorical boosting (CatBoost)

CatBoost is a developed gradient boosting method introduced by Prokhorenkova [24] and Dorogush [25], and it effectively functions with categorical characteristics with the lowest information loss. CatBoost is unique in comparison to other gradient boosting methods. Firstly, it employs ordered boosting, an efficient variation of gradient boosting methods, to combat target leakage [25]. Secondly, this approach can be applied to tiny datasets. Finally, CatBoost can manage category characteristics. This is often performed during the pre-processing stage which comprises mostly of substituting the initial categorical variables with one or more numerical values. Calculating statistics based on the label values of the samples is an alternative method for dealing with categorical attributes. For example, suppose that D is a provided dataset of observations D ={(X_i, Y_i)} $i = 1 \dots, n$, where $X_i = (x_{i,1}, \dots, x_{i,m})$ is a vector of m features, some numerical, some categorical, and $Y_i \in R$ is a label value. The easiest method is to replace the category with the mean label value across the entire train dataset. CatBoost employs a more effective approach that decreases overfitting and permits the utilization of the entire dataset for training. By performing a random permutation of the dataset and calculating the average label value for the instance with the identical classification value placed prior to the supplied one in the variation. Assume $\sigma = (\sigma_1, \dots, \sigma_n)$ be the variation, hence $x_{\sigma_p,k}$ is replaced with the following equation:

$$\frac{\sum_{j-1}^{p-1}[x_{\sigma_j,k} = x_{\sigma_p,k}]\, Y_{\sigma_j} + a.P}{\sum_{j-1}^{p-1}[x_{\sigma_j,k} = x_{\sigma_p,k}] + a} \quad (8)$$

Further, the addition of the previous values includes the previous value P and the previous weight parameter a > 0. It is worth mentioning that integrating previous values is a frequent approach that aids in the reduction of low-frequency category noise.

2) eXtreme Gradient Boost (XGBoost)

The eXtreme Gradient Boost (XGBoost) method is a revised variant of the gradient boosting decision tree algorithm (GBDT) that generates boosted trees simultaneously and effectively [26]. Along these lines, the XGBoost method is a scalable and computationally fast technique. The algorithm's concept is to construct better decision trees (expressed as $\sum_{n=1}^{N}\widehat{f}_n(x_i), f_n \in F$ where F is the set of all decision trees, and f_n is a decision tree in F) to optimize the objective function $F_{obj}(t)$, the regularization goals are reduced as demonstrated in Eq. (9) to learn the set of decision trees employed in the model.

Equations (10-11) demonstrates the formulation of the XGBoost technique.

$$F_{obj}(t) = \mathcal{L}(t) + \Omega(f_t) \quad (9)$$

$$\mathcal{L}(t) = l\big(y_i, \hat{y}_i^{(t-1)} + f_t x_i)\big) \quad (10)$$

$$\Omega(f_t) = \alpha T + \frac{1}{2}\lambda\, \|w\|^2 \quad (11)$$

where l denotes the differentiable convex loss function, y_i is the real value, $\hat{y}_i^{(t-1)}$ is the previous round forecast at $t - 1$, and $f_t(x)$ denotes the next decision tree at t round. Further, the $\Omega(f_t)$ shows the regularization factor, T is the number of constructed trees, α is the learning rate, w is the leaves' weights, and λ is the regularization parameter.

3) Linear Regression (LR)

Linear regression (LR) is a frequently employed technique for the predictive analysis of continuous data that involves defining the set of factors (x) that have a significant impact on forecasting the outcome variable (y). Usually, LR is used to predict the impact of variation in the dependent variable due to any magnitude of variation in the independent variable(s) [42, 43]. The procedure is carried out utilizing the following Equation (12):

$$y = \beta_0 x_0 + \beta_1 x_1 + C \quad (12)$$

Where β and C are parameters trained by employing the least-squares method, the values of x corresponding to the set of electricity prices, time stamp, and the lag values of electricity prices. While the value of y corresponds to the forecasted electricity price. For the purposes of electricity price forecast, LR was implemented in python by employing the sklearn library using the default hyperparameters.

4) Least Absolute Shrinkage and Selection Operator (LASSO)

For regression analysis, LASSO has been a popular technique. LASSO is fundamentally a regularized LR model with variable selection. Reducing the objective function (obj) stated in Eq.(13) yields the coefficients of the LASSO model (β) [27]. LASSO's goal is to add two parts, the first part responsible for the summation of square residues. While the second part is accountable to the regularization component which decreases variation in order to improve long-term forecasts.

$$obj = min_{\beta}\left\{\sum_{i=1}^{M}\left((g_i - \sum_{j=1}^{Z}\beta_j h_{ij})\right)^2 + \gamma\sum_{j}^{Z}|\beta_j|\right\} \quad (13)$$

The response variables are represented by g_i while h_{ij} denoted to the explanatory variables and finally the corresponding coefficients is demonstrated by β_j . The regularization part $\gamma\sum_j^Z|\beta_j|$ reduces the coefficients of the explanatory variables, hence preventing the variables from being overfit to the training data. The reduction can be fine-tuned by a positive value of the regularization parameter γ. Where when γ is equal to zero, the obj approaches the form of LR. In another word, if the value of γ increased, the variance will decrease.

5) The proposed methodology for EPF.

In this work, different ML models were considered to forecast the electricity price one day ahead. Figure 3 illustrates the whole process, starting with selecting the inputs, such as the time index and the electricity prices (from t till t-24). After that, the data preparation and partitioning will be done to split the data into the training and testing phase. The trained data will enter into the proposed single or hybrid models to forecast the electricity price one day ahead EP(t+1) and compare it with the testing data. The following step will evaluate the models forecast based on performance metrices to opt for the model with the lowest error. Finally, model interpretation by using the SHAP framework will be investigated to show the impact of each feature on the forecasting results.

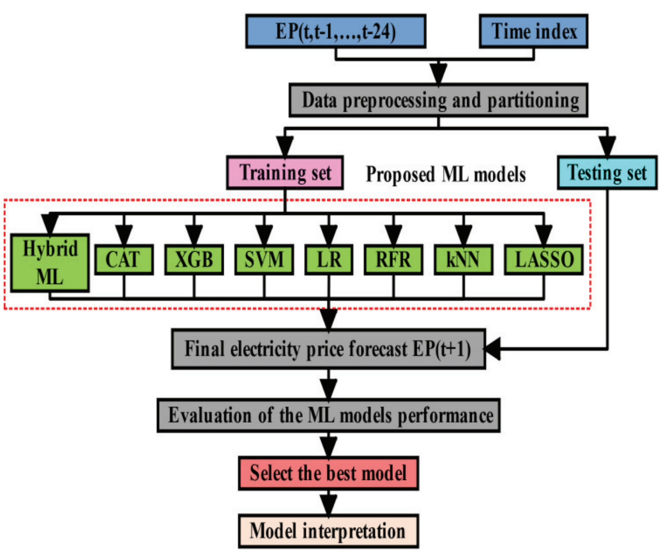


Fig. 3. The flowchart of the proposed ML models for EPF.

D. SHapley Additive exPlanations (SHAP) framework.

To comprehend how a model makes forecasts, it is important to interpret the model. The SHAP framework is applied in this study to understand the behavior of ML models and the interplay of their properties. SHAP was introduced by Lundberg and Lee to explain the different forecasts of game-theory-based ML models [28]. Further, SHAP leverages cooperative game theory by providing a relevance score to each characteristic depending on its impact on the model's forecast. SHAP's central concept is to reorganize the Shapley value issue from analyzing how association participants provide to association value to analyzing how specific model characteristics impact to the model's yields. Figure 4 demonstrates that the lag electricity price one day EP(t-1) impact the forecast result most, followed by EP(t-2) and time stamp, respectively. The other features follow the same pattern but with less influence on the output results. The summary plot for the EPF is illustrated in Figure 5. The figure shows how the EP(t-1) impacts the model output most indicating them by the red dots. Where the high values contribute more to make the forecasting results more accurate, whereas the blue ones impact the model output negatively. Also, it is worth to mention that the SHAP figures confirmed the autocorrelation in Figure 2 shown that the EP(t-24) also have impact on the forecasting output.

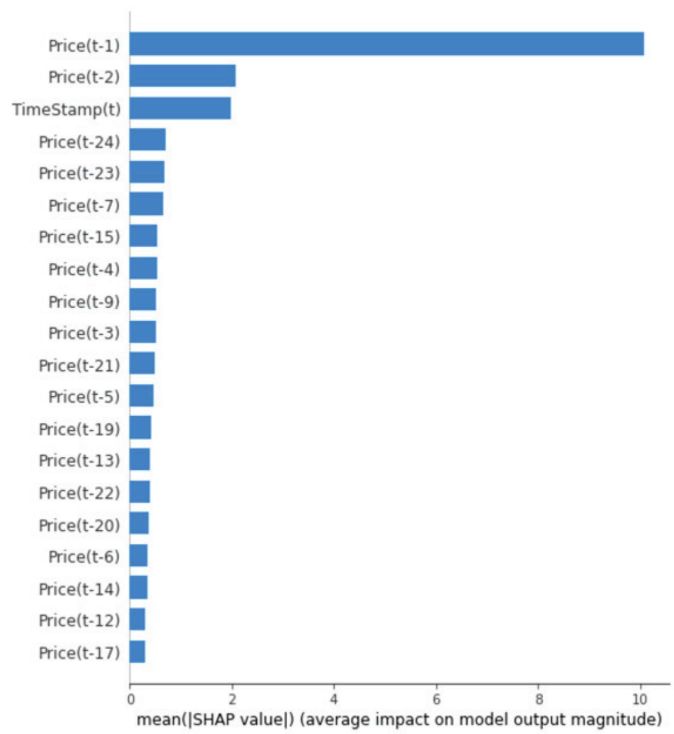


Fig. 4. The impact of each feature in EPF.

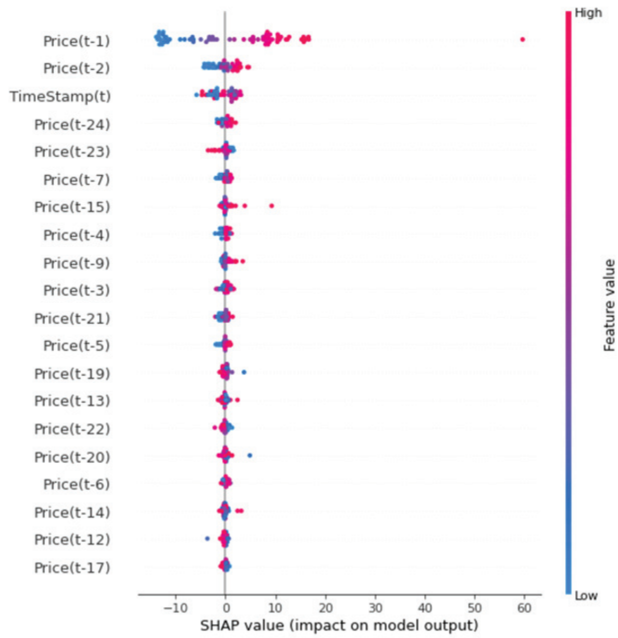


Fig. 5. The summury plot for the each feature in EPF.

III. Results and discussion

In this part, the forecasting of electricity price results from different proposed ML models will be discussed and show the performance of each model compared to the other based on evaluation indexes.

To begin, Table I presented the obtained results for one day ahead for EPF from the ML models according to the performance metrics, such as MSE, RMSE, and MAE. Overall, it can be seen from the table that the hybrid models performed better compared to the single ones. The LR-CatBoost model managed to attain the lowest RMSE and MAE with values of 10.49 and 7.94, respectively, followed by LASSO-CatBoost and LR-XGBoost with a value of 10.5, 8.03 and 10.7, 8.1 for RMSE and MAE, respectively. The kNN and SVM models performed poorly with achieving the highest RMSE value of 12.88 and 12.39, respectively. It is worth to mention that LR-CatBoost attain the lowest MSE value whereas kNN achieved the highest.

The combination between two models will result in taking advantage of the first model to cope with the other shortcomings, and vice versa. In the case of LR-CatBoost, the LR was selected because it is efficient at extrapolating trends, but it is incapable of learning interactions. Hence, it followed by CatBoost since it is has the capability to capture the dependency between variables and estimate the residuals from the linear model. From Fig. 6, it is evident that hybrid models achieved the lowest MAE and RMSE values compared to single models.

Figure 7 shows the forecasted results over one day for the single proposed model (Fig.7 (a)) and the hybrid ones (Fig.7 (b)) compared to the ground truth. The LR-CatBoost, LASSO-CatBoost, and LR-XGBoost forecast the electricity price nearly similar to the actual values, whereas models like kNN and SVM has been found has huge deviations from the centerline which leads to large deviations from the ground fact.

TABLE I. The performance evaluation for the ML models

Models	Performance Evaluation		
	MSE	***RMSE***	***MAE***
RFR	136.93	11.7	8.71
kNN	165.99	12.88	9.88
XGBoost	138.37	11.76	8.77
CatBoost	127.86	11.31	8.33
SVM	153.55	12.39	8.12
LASSO	130.89	11.44	8.81
LR-XGBoost	130.5	11.42	8.78
LASSO-XGBoost	122.44	11.07	8.39
LASSO-CatBoost	110.3	10.5	8.03
LR-XGBoost	114.52	10.7	8.1
LR-CatBoost	110.13	10.49	7.94

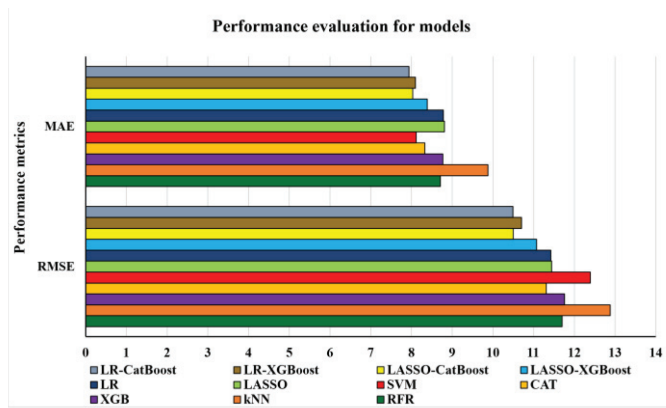
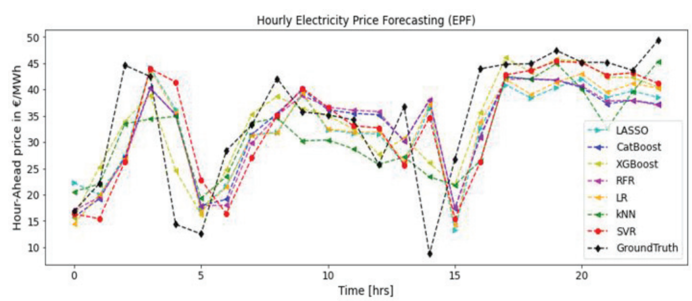
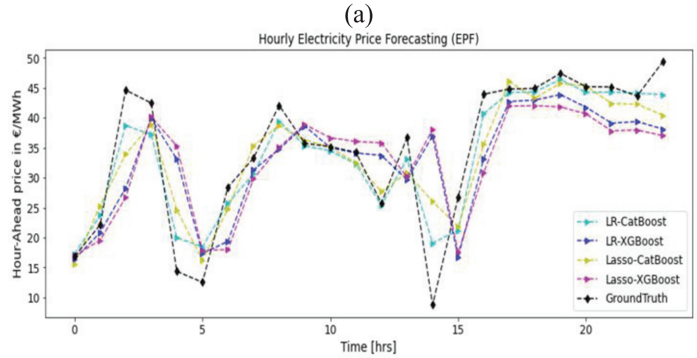


Fig. 6. Comparison between different ML models in terms of RMSE and MAE.



(a)



(b)

Fig. 7. The ML models EPF results compared to the actual data in (a) single models and (b) hybrid models.

IV. CONCLOUSION

The study proposes a combination of single and hybrid ML models to predict Nord Pool spot electricity market price one day. On the basis of performance indicators such as RMSE, MSE, and MAE, the proposed models were evaluated. In addition, an interpretation of the model using the SHAP framework demonstrates the effect of each attribute on the predicted output. According to SHAP, the lag electricity price EP(t-1) impacts the forecast result most, followed by EP(t-2) and time stamp, respectively. Lastly, the results demonstrate that hybrid models performed better than their single counterparts, with the LR-CatBoost model achieving the lowest MAE and RMSE values of 7.83 and 10.49, respectively. In addition, the kNN and SVM models performed badly, with RMSE values of 12.88 and 12.39, respectively. Generally, combining LR model with ensemble tree-based models result in enhance EPF performance.

The extension of the work will mainly focus on proposing a hybrid machine learning model with hyperparameter optimization and implementing different inputs that contribute to changes in the electricity price, such as the price of natural gas, oil, and other factors. Finally, these inputs will be represented using SHAP to show their impact on electricity price forecasting.

REFERENCES

[1] H. Mubarak *et al.*, "Optimum distribution system expansion planning incorporating DG based on N-1 criterion for sustainable system," *Sustainability,* vol. 13, no. 12, p. 6708, 2021.

[2] H. Mubarak *et al.*, "Operational Cost Minimization of Electrical Distribution Network during Switching for Sustainable Operation," *Sustainability,* vol. 14, no. 7, p. 4196, 2022.

[3] S. Ahmad, U. K. Jhuma, H. Mubarak, S. Mekhilef, and H. Mokhlis, "Real Time Power Management for AC Microgrid Prioritizing Renewable Energy Sources, Battery Storage and Critical Load," in *2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET)*, 2022: IEEE, pp. 981-986.

[4] H. Mubarak, H. Mokhlis, N. N. Mansor, M. Mohamad, A. S. M. Khairuddin, and S. Afzal, "Optimal Distribution Networks Expansion Planning with DG for Power Losses Reduction," in *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 2021: IEEE, pp. 1-5.

[5] L. Amabile, D. Bresch-Pietri, G. El Hajje, S. Labbé, and N. Petit, "Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties," *Advances in Applied Energy,* vol. 2, p. 100020, 2021.

[6] J. Mei, Y. Zuo, C. H. Lee, X. Wang, and J. L. Kirtley, "Stochastic optimization of multi-energy system operation considering hydrogen-based vehicle applications," *Advances in Applied Energy,* vol. 2, p. 100031, 2021.

[7] H. Mubarak, H. Mokhlis, N. N. Mansor, A. Abdellatif, L. Wang, and M. A. Muhammad, "Reduction of Electricity Cost in Distribution Systems Based on Optimal Switching Configuration," in *2022 IEEE International Conference on Power Systems Technology (POWERCON)*, 2022: IEEE, pp. 1-6.

[8] O. A. Karabiber and G. Xydis, "Electricity price forecasting in the Danish day-ahead market using the TBATS, ANN and ARIMA methods," *Energies,* vol. 12, no. 5, p. 928, 2019.

[9] K. Goswami, A. Ganguly, and A. K. Sil, "Day ahead forecasting and peak load management using multivariate auto regression technique," in *2018 IEEE Applied Signal Processing Conference (ASPCON)*, 2018: IEEE, pp. 279-282.

[10] A. Abdellatif *et al.*, "Forecasting Photovoltaic Power Generation with a Stacking Ensemble Model," *Sustainability,* vol. 14, no. 17, p. 11083, 2022.

[11] H. Mubarak *et al.*, "A hybrid machine learning method with explicit time encoding for improved Malaysian photovoltaic power prediction," *Journal of Cleaner Production,* p. 134979, 2022.

[12] A. Abdellatif, H. Abdellatef, J. Kanesan, C.-O. Chow, J. H. Chuah, and H. M. Gheni, "An Effective Heart Disease Detection and Severity Level Classification Model Using Machine Learning and Hyperparameter Optimization Methods," *IEEE Access,* vol. 10, pp. 79974-79985, 2022.

[13] A. Abdellatif, H. Abdellatef, J. Kanesan, C.-O. Chow, J. H. Chuah, and H. M. Gheni, "Improving the Heart Disease Detection and Patients' Survival Using Supervised Infinite Feature Selection and Improved Weighted Random Forest," *IEEE Access,* vol. 10, pp. 67363-67372, 2022.

[14] M. Hakim *et al.*, "Bearing Fault Diagnosis Using Lightweight and Robust One-Dimensional Convolution Neural Network in the Frequency Domain," *Sensors,* vol. 22, no. 15, p. 5793, 2022.

[15] A. Hammoudeh, M. I. Al Saaideh, E. A. Feilat, and H. Mubarak, "Estimation of synchronizing and damping torque coefficients using deep learning," in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, 2019: IEEE, pp. 488-493.

[16] A. Shiri, M. Afshar, A. Rahimi-Kian, and B. Maham, "Electricity price forecasting using Support Vector Machines by considering oil and natural gas price impacts," in *2015 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, 2015: IEEE, pp. 1-5.

[17] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Transactions on Power Systems,* vol. 27, no. 4, pp. 2055-2062, 2012.

[18] S. Albahli, M. Shiraz, and N. Ayub, "Electricity price forecasting for cloud computing using an enhanced machine learning model," *IEEE Access,* vol. 8, pp. 200971-200981, 2020.

[19] K. Wang, C. Xu, Y. Zhang, S. Guo, and A. Y. Zomaya, "Robust big data analytics for electricity price forecasting in the smart grid," *IEEE Transactions on Big Data,* vol. 5, no. 1, pp. 34-45, 2017.

[20] N. Pool, "Historical market data," *Nordpoolspot. com,* 2016.

[21] A. N. Alkawaz, A. Abdellatif, J. Kanesan, A. S. M. Khairuddin, and H. M. Gheni, "Day-Ahead Electricity Price Forecasting based on Hybrid Regression Model," *IEEE Access,* 2022.

[22] T. Ahmed, S. Mekhilef, R. Shah, and N. Mithulananthan, "An assessment of the solar photovoltaic generation yield in Malaysia using satellite derived datasets," *International Energy Journal,* vol. 19, no. 2, 2019.

[23] U. K. Das *et al.*, "SVR-based model to forecast PV power generation under different weather conditions," *Energies,* vol. 10, no. 7, p. 876, 2017.

[24] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," *Advances in neural information processing systems,* vol. 31, 2018.

[25] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: gradient boosting with categorical features support," *arXiv preprint arXiv:1810.11363,* 2018.

[26] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785-794.

[27] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological),* vol. 58, no. 1, pp. 267-288, 1996.

[28] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems,* vol. 30, 2017.