Harnessing AutoML for Electricity Price Forecasting in Germany: A Comparative Analysis of Models

1st Demetrios N. Papadopoulos
INATECH
University of Freiburg
Freiburg im Breisgau, Germany
dimitrios.papadopoulos@inatech.uni-freiburg.de

2nd Ramiz Qussous
INATECH
University of Freiburg
Freiburg im Breisgau, Germany
ramiz.qussous@inatech.uni-freiburg.de

3nd Anke Weidlich
INATECH
University of Freiburg
Freiburg im Breisgau, Germany
anke.weidlich@inatech.uni-freiburg.de

***Abstract*—Electricity price forecasting is crucial for energy market stakeholders, facilitating informed decisions in trading, investment, and operational planning to optimize portfolios and ensure grid stability. Machine learning (ML) models have proven effective for this task, but their deployment often demands substantial computational resources and expertise in both energy markets and ML techniques. Automated Machine Learning (AutoML) addresses these challenges by automating key steps like feature engineering, preprocessing, model selection, and hyperparameter tuning. This study evaluates the performance of three leading AutoML frameworks—Auto-sklearn, TPOT, and FLAML—for day-ahead electricity price forecasting in Germany, focusing on the year 2024. The evaluation includes testing during two distinct periods characterized by high and low price volatility. By incorporating domain-specific features, these frameworks are benchmarked against a Multi-Layer Perceptron (MLP) to assess their viability in comparison to traditional neural networks. The findings aim to provide insights into the scalability and effectiveness of AutoML for complex, data-intensive forecasting tasks, offering practical solutions for navigating the challenges of Germany's dynamic electricity market.**

***Index Terms*—Electricity Market, AutoML, Price Forecasting**

I. Introduction

Electricity price forecasting has always been crucial for energy system planning, market operations, and investment decision-making. In recent years, there has been a growing need for precise short-term price forecasts. This increased demand is largely because accurate short-term predictions enable consumers to adjust their electricity usage in response to changing prices. By doing so, they help create demand-side flexibility—a key factor in developing a more adaptive and resilient energy system during the ongoing energy transition.

In literature, electricity price forecasting methods can be categorized into three main categories: statistical approaches and simulations models, and game theoretical models [1]. Statistical approaches uses historical data to make predictions, while simulations models depends on either optimization or agent-based modeling techniques to simulate market dynamics and participant behavior. The choice of modeling approach depends on the goal of the forecast, as different methods offer varying advantages in terms of accuracy, interpretability, and computational complexity [2]. In this paper, we focus on comparing machine learning (ML) methods, specifically, rather than exploring other modelling approaches.

ML models have gained significant traction in electricity price forecasting due to their ability to model complex relationships and leverage large datasets. For example, the authors in [3] compared statistical models with prominent deep learning models for forecasting spot electricity prices, finding that deep learning models significantly outperformed statistical models in terms of sMAPE. Conversely, a study by [4] demonstrated that non-deep learning models, such as Lasso, Random Forest, and Support Vector Regression (SVR), performed better than deep learning models (e.g., LSTM, GRU, CNN) in day-ahead electricity price forecasting, particularly in terms of both accuracy and computational efficiency.

These findings suggest that the performance of forecasting models may vary based on the characteristics of the time series being predicted. Consequently, the successful deployment of ML models often requires a combination of domain-specific knowledge and expertise in computer science. To address this challenge, Automated Machine Learning (AutoML) has emerged as a viable solution. AutoML frameworks, such as TPOT, Auto-Sklearn, and FLAML, automate key aspects of the ML development process, including model selection, feature engineering, and hyperparameter tuning, thereby reducing the reliance on extensive manual intervention [5]–[7]. Notably, AutoML has demonstrated even the ability to outperform human experts in various tasks [8].

Despite the increasing interest in AutoML for time-series forecasting, few studies have specifically focused on its application to electricity price forecasting. While some research has evaluated AutoML frameworks across diverse time-series datasets [9], [10], the study by [11] remains one of the few to compare AutoML frameworks in the context of day-ahead electricity prices in the Irish market.

This study aims to bridge this gap by:

979-8-3315-1278-1/25/$31.00 ©2025 IEEE

- **Evaluating** which AutoML framework achieves the best performance within a fixed computational budget,
- **Comparing** AutoML frameworks against the commonly used MLP, and
- **Assessing** the performance of AutoML frameworks under both stable and volatile market conditions.

II. Datasets

In this study multiple datasets are used to forecast the day-ahead prices in Germany for the year 2024. Below is a detailed description of the datasets used in this study:

- **German Day-Ahead Electricity Prices**
  - **Source:** SMARD [12]
  - **Description:** Hourly electricity prices for the German-Luxembourg (DE-LU) bidding zone in 2024, serving as the primary target variable for forecasting models.
- **CO_2 Emission Allowance Prices**
  - **Source:** European Energy Exchange AG (EEX) [13]
  - **Description:** Hourly CO_2 emission allowance prices for Germany in €/tCO_2, influencing electricity market dynamics.
- **Electricity Demand (Grid Load)**
  - **Source:** SMARD [12]
  - **Description:** Hourly electricity consumption (grid load) in MWh across Germany, including residual load and hydro-pumped storage consumption data.
- **Scheduled commercial electricity exchanges**
  - **Source:** SMARD [12]
  - **Description:** Hourly scheduled cross-border electricity trading data, specifying exchange volumes (MWh) between Germany and neighboring countries.
- **Electricity generation by technology**
  - **Source:** SMARD [12]
  - **Description:** Hourly electricity generation categorized by:
    * **Renewable Sources:** Biomass, Hydropower, Wind (onshore & offshore), Photovoltaics, Other Renewables.
    * **Conventional Sources:** Nuclear, Lignite, Hard Coal, Fossil Gas, Hydro-Pumped Storage, Other Conventional Energy.

III. Methodology

The primary objective of this study is to assess the accuracy of three leading AutoML tools in day-ahead electricity market (DAM) forecasting under equal computational time constraints. To tackle this issue we forecast the prices within different periods: one period in which the prices do not oscillate much (period A) and one in which the prices fluctuate a lot (period B). To accuratively check if the model can generalize well in unseen data we are splitting the dataset into training, validation and test set. For period A, the training period included data from January 1 up to July 15, 2024, while the validation period spanned from July 15 to August 1, 2024. The testing period covered data from August 1 to August 18, 2024. For period B, the training period ended on November 15, 2024, with the validation period running from November 14 to December 1, 2024. The testing period followed, covering data from December 1 to December 18, 2024. For all pipelines, we have set a day-ahead prediction horizon (24 steps horizon on hourly data) as we reason that is the minimum time required to perform adequate planning whether this is a TSO or a power plant operator. Additionally, to ensure that there is a fair comparison among the AutoML tools, we set the runtime to 60 minutes (wall-clock time) and all pipelines are running in an intel core i5-7600 cpu at 3.50 GHz utilizing all of its cores.

After the data has been collected, we implement a feature engineering step, including feature construction and feature selection which reduces the dimensionality of the model, reducing the computational effort and in parallel improves the model's performance. This step is performed considering the training and validation set and cross-validated using a five-fold time-series cross-validation (for more information see subsection III-B). After performing feature selection on the training/validation set, the chosen features are utilized by the AutoML tools to construct a pipeline, which is then evaluated on the test set.

A. Evaluation metrics

To assess the accuracy of the implemented forecasting pipelines, this study employs evaluation metrics widely recognized within the forecasting community, including the Mean Absolute Error (MAE, Eq. 1) and the Mean Squared Error (MSE, Eq. 2). In these equations, n denotes the number of forecasted time steps, while $\hat{y}^{(i)}$ and $y^{(i)}$ represent the forecasted and actual values, respectively, for the i^{th} observation.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y^{(i)} - \hat{y}^{(i)}\right| \tag{1}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y^{(i)} - \hat{y}^{(i)}\right)^2 \tag{2}$$

Given the MSE's propensity to penalize larger errors more heavily, this metric will be employed as the loss function in all the AutoML tools utilized in this study.

B. Feature engineering

Feature engineering is a critical step in the machine learning pipeline, as it dictates the amount and quality of information available to the model. The process involves both the selection of original features and the creation of derived features to enhance predictive performance. Below, we present the original features passed to the model, followed by the description of ML-derived features and their corresponding transformations.

1) Original Features: These features were selected based on their relevance to the day-ahead price formation in the german energy market and are grouped into the following categories:

- **Renewable Energy Sources:** Biomass [MWh], Hydropower [MWh], Wind offshore [MWh], Wind onshore [MWh], Photovoltaics [MWh], Other renewable [MWh]
- **Conventional Energy Sources:** Lignite [MWh], Hard coal [MWh], Fossil gas [MWh], Hydro pumped storage [MWh], Other conventional [MWh]
- **Market Data:** Total consumption (grid load) [MWh], Net export-import (scheduled commercial exchanges) [MWh]
- **Environmental/Economic Data:** CO_2 Allowances, DE [€/tCO_2] (Carbon dioxide allowances price for Germany), Germany/Luxembourg [€/MWh] (Electricity price for Germany and Luxembourg)

2) ML-Derived Features: To further enrich the model's inputs, additional features need to be constructed and the original features must be handled in a way that enhances the model's ability to understand temporal patterns, periodic behaviours, and autocorrelations in the energy market. The following ML-derived features were incorporated:

- **Behavioral Context Features:** Binary flags that capture human activity patterns such as holidays and work hours:

$$\text{Holiday Flag} = \begin{cases} 1, & \text{if holiday} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

$$\text{Work Flag} = \begin{cases} 1, & 08:00 \leq h \leq 17:00 \text{ and weekday} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

These features differentiate energy usage patterns between working and non-working periods.

- **Time-Related Features:** Temporal components such as minutes, hours, and days are encoded cyclically to capture periodic behaviors:

$$\text{T\_sin} = \sin\left(\frac{2\pi \cdot \text{T}}{\max(\text{T})}\right) \tag{5}$$

$$\text{T\_cos} = \cos\left(\frac{2\pi \cdot \text{T}}{\max(\text{T})}\right) \tag{6}$$

Here, T represents time parameters such as hour or day of the week.

- **Fourier Features:** Seasonal patterns are modeled using Fourier transformations:

$$\text{Fourier\_sin} = \sin\left(2\pi k \frac{t}{sp}\right) \tag{7}$$

$$\text{Fourier\_cos} = \cos\left(2\pi k \frac{t}{sp}\right) \tag{8}$$

where t is the elapsed time, sp is the seasonal period (e.g., hourly, daily, weekly, monthly), and k is the Fourier term, set to $k = 2$ in this study.

- **Lagged Features:** Autocorrelation and temporal dependencies are captured through lagged values of the target:

$$y_{t-i} = B^i y_t, \quad i = 1, \ldots, 96 \tag{9}$$

where y_t is the target variable at time t, and B^i is the backshift operator applied i steps backward.

- **Window Features:** Rolling mean and standard deviation were calculated for two predefined window lengths: one day ($w_d = 24$) and one week ($w_w = 168$). These are expressed as:

$$\text{Win\_Mean}_t = \frac{1}{w} \sum_{i=0}^{w} y_{t-i} \tag{10}$$

$$\text{Win\_Std}_t = \sqrt{\frac{1}{w} \sum_{i=0}^{w} (y_{t-i} - \text{Win\_Mean}_t)^2} \tag{11}$$

where w represents the window length (w_d or w_w) and y_t is the variable at time t.
This feature engineering was specifically applied to carbon dioxide allowance prices. Due to the relatively static nature of these prices on many days, it is crucial to capture significant deviations and trends when they occur. The rolling mean helps identify longer-term trends, while the standard deviation highlights periods of increased volatility.

While this extensive feature set forms a strong foundation for accurate price forecasting, including all features in the model can significantly increase computational time and reduce accuracy. To address these challenges, a wrapper-based feature selection (FS) method using Random Forest is applied, following a greedy forward feature addition strategy.

First, the Spearman correlation with the target variable (day-ahead price) is computed. As an initial filtering step, features with a correlation below 0.1 are removed to reduce computational overhead. In the second step, all features are incrementally introduced into the model, starting with the strongest correlated feature with the target and the model performance (MSE error) is constantly monitored across progressively larger feature subsets. The process continues until all features have been sequentially evaluated, identifying the optimal feature set. This refined subset is then used as input to the AutoML models.

The Spearman correlation coefficient was selected for its ability to capture non-linear relationships. Represented as ρ, it is a non-parametric measure that quantifies the statistical dependence between two variables by evaluating how well their association follows a monotonic function. As expressed in Equation 12, d_i denotes the difference in ranks between corresponding values of the variables x and y, while n represents the total number of data points. The coefficient ranges from -1, indicating a perfect negative correlation, to 1, signifying a perfect positive correlation, with 0 implying no correlation.

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{12}$$

C. AutoML Modeling

In this study, we evaluate the performance of three prominent AutoML tools—TPOT, Auto-sklearn, and FLAML—for electricity price forecasting. Their results are benchmarked against a baseline model: a multilayer perceptron (MLP) featuring three hidden layers with 300 neurons each, using the tanh activation function and trained with the LBFGS solver for 300 epochs. The MLP architecture was selected because it is widely employed in the price forecasting literature, either as a standalone model or as a building block of more complex frameworks [14]–[16].

TPOT (Tree-Based Pipeline Optimization Tool): TPOT is a genetic programming-based AutoML framework that automates the design of machine learning pipelines. It searches through a space of possible pipelines by evolving them over successive generations, selecting the most promising ones based on their performance. TPOT optimizes pipeline components such as preprocessing, feature selection, model selection and hyperparameters optimization [5].

Auto-sklearn: Auto-sklearn extends the scikit-learn library with AutoML capabilities. It employs a Bayesian optimization approach to search for the best combination of algorithms and hyperparameters. Auto-sklearn leverages meta-learning which uses knowledge from previously-made predictions in similar datasets. Additionally, it includes ensemble construction techniques to combine multiple pipelines for improved robustness and accuracy [7].

FLAML (Fast and Lightweight AutoML): FLAML is a lightweight AutoML framework designed for efficient and cost-effective optimization of machine learning pipelines. Unlike other AutoML tools, FLAML focuses on minimizing resource usage by employing a cost-aware search strategy. It adaptively selects models and hyperparameters based on computational cost and performance, making it particularly suitable for scenarios with limited computational resources or time constraints [6].

Each of these AutoML tools are based on different optimization and ML techniques, ranging from the genetic programming of TPOT to the meta-learning capabilities of Auto-sklearn and the cost-effective approach of FLAML. By applying these tools to the price forecasting problem, we aim to identify the strengths and limitations of each framework in terms of prediction accuracy, computational efficiency, and adaptability to the dataset.

IV. Results and Discussion

A. Results on FS

In this section, we present the results of the feature selection (FS) process for both scenarios.

First, we note that a substantial reduction in dimensionality was achieved in both pipelines. Initially, 380 features were generated for both periods. After selection, 162 features remained for period A, corresponding to a 57.37% reduction, while 169 features were retained for period B, representing a 55.53% reduction. This reduction implies that AutoML tools require less computational effort when optimizing hyperparameters and selecting techniques to enhance model performance.

TABLE I
Selected Features for Period A

Feature	Lags (hours)
Germany/Luxembourg [€/MWh]	0-4, 12, 17, 21–24
Lignite [MWh]	0–5, 21–24
Biomass [MWh]	0, 1, 4–8, 14–20, 23, 24
Fossil gas [MWh]	0–3, 16–18, 22–24
Hard coal [MWh]	0, 1, 16–19
Hydropower [MWh]	0–4, 11–13, 21–24
Hydro pumped storage [MWh]	0, 1, 2, 5–8, 11, 12, 15–19, 22–24
Photovoltaics [MWh]	0–2, 6–14, 18–24
Wind offshore [MWh]	4–17
Wind onshore [MWh]	0–3, 6–14, 22
Net export-import [MWh]	0–3, 7–9, 22–24
Total (grid load) [MWh]	8–12, 16–22
CO_2 Allowances,DE [EUR/tCO_2]	0
CO_2_wind_mean_1_day	0
CO_2_wind_std_1_day	0
CO_2_wind_mean_7_day	0
Hour	0
Hour_cos	0
WorkingHour_flag	0
Weekend	0
Cos_2880_2	0
Day_of_week	0
Day_of_week_sin	0

Examining the selected features in Tables I and II, we observe that features closely related to the target variable (which is shifted one day ahead) were prioritized. In both pipelines, lagged features from a few hours before the forecasting event were selected, followed by mid-day values, and then again, lagged values from the previous day for DAM prices, hydro pump storage, total grid load, and net export-import. In contrast, renewable energy sources—such as wind, photovoltaics, and biomass did not exhibit a clear selection pattern, as features from various lag periods were chosen.

Additionally, CO_2 prices were only relevant at timestamps close to the start of the forecasting timing, which aligns with the author's expectations since they remain unchanged throughout extended periods. Regarding time-related features, only a few were selected, with Fourier-transformed features appearing exclusively in the first scenario.

Overall, we find that the selected features for both periods have a significant degree of overlap. This suggests that once an ML model is established, frequent re-evaluation of feature extraction may not be critical.

B. Results on Forecasting Periods

This section presents the day-ahead price prediction horizon, with the numerical results for periods A and B shown in Tables III and IV. As previously mentioned, the validation (val) metric represents the cross-validated score obtained from the validation set, while the test set evaluates model performance

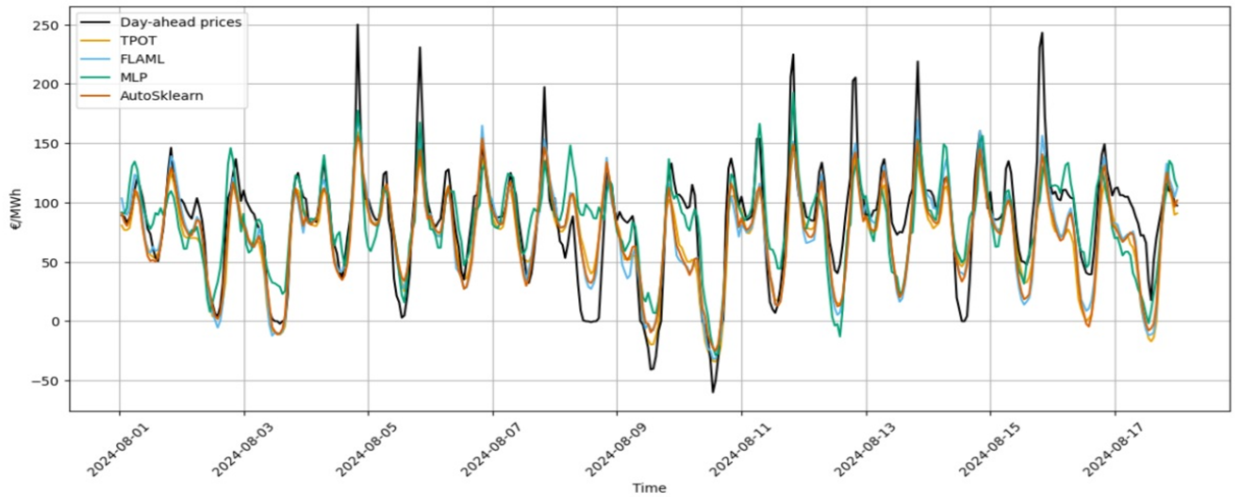

Fig. 1. Visual representation of the forecasting results for period A[1]

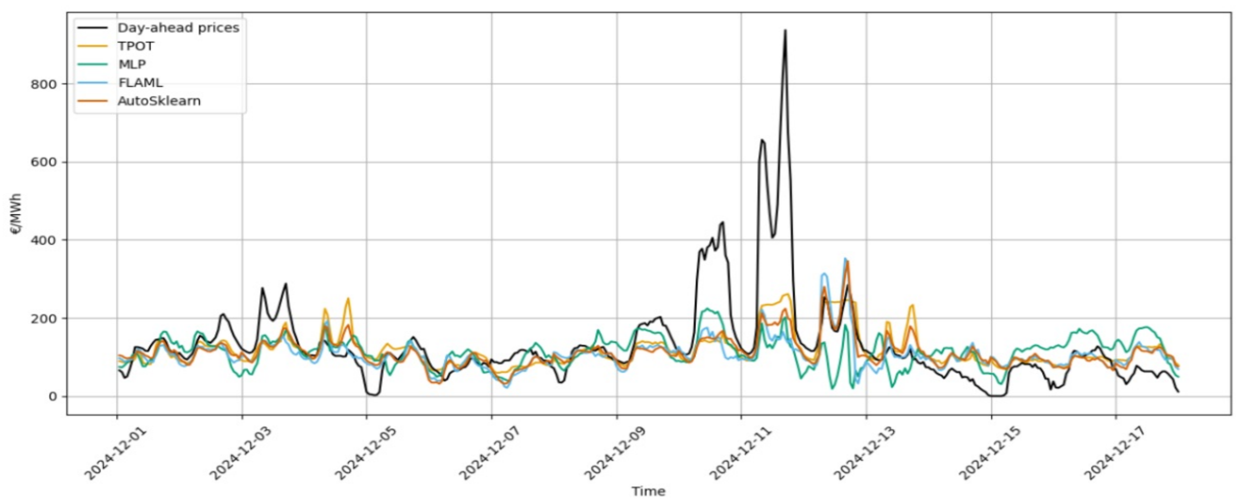

Fig. 2. Visual representation of the forecasting results for period B[1]

on unseen data. Additionally, Figures 1 and 2 provide a visual comparison of the models.

In period A, all AutoML frameworks outperformed the neural network benchmark across both validation and test sets. However, in period B, which features extreme price fluctuations, TPOT and AutoSklearn failed to surpass the benchmark in validation. This presents a challenge, as poor validation performance may discourage forecasting experts from considering these models for the test set. Such an approach could lead to misinterpretations, particularly in this case, where TPOT and FLAML ultimately outperform the benchmark on unseen data, with TPOT emerging as the best-performing model. This result is expected, as the test set MSE for period B is nearly ten times higher than in validation due to extreme price spikes (above 300 €/MWh), making it difficult to determine the best model in advance. This highlights the need for meta-estimators and ensemble forecasting approaches to address such challenges.

Furthermore, the architecture of the AutoML models significantly influences performance. In period A, TPOT performed best in validation but was outperformed by other models in the test set, indicating overfitting. In contrast, in period B, TPOT ranked the worst in validation but achieved the best performance on unseen data, suggesting underfitting. This behavior stems from TPOT's genetic algorithm, which can evolve highly diverse regressors. Conversely, FLAML performed best in validation for period B, but since it is tree-based, it primarily predicts within the training data range. As a result, FLAML exhibited the weakest test set performance among the AutoML tools in period B.

V. Conclusion

This research assessed the effectiveness of three prominent AutoML frameworks—TPOT, Auto-sklearn, and FLAML—in

[1]The plots were designed using the Okabe-Ito color palette, which is colorblind-friendly.

TABLE II
SELECTED FEATURES FOR PERIOD B

Feature	Lags (hours)
Germany/Luxembourg [€/MWh]	0–4, 10–14, 21–24
Lignite [MWh]	0–7, 9–14, 20–24
Biomass [MWh]	0, 1, 4–8, 15–20, 23, 24
Fossil gas [MWh]	0–3, 22–24
Hard coal [MWh]	0–5, 9–12, 22–24
Other conventional [MWh]	18
Hydropower [MWh]	6, 7, 16–19
Hydro pumped storage [MWh]	0–2, 4–8, 12, 15–20, 22–24
Photovoltaics [MWh]	0– 2, 7–12, 18–24
Wind offshore [MWh]	4–18
Wind onshore [MWh]	0–14
Net export-import [MWh]	0–3, 7, 8, 22–24
Total (grid load) [MWh]	8–12, 16–21
CO_2 Allowances,DE [EUR/tCO_2]	0
CO_2_wind_mean_1_day	0
CO_2_wind_std_1_day	0
CO_2_wind_mean_7_day	0
Hour	0
Hour_cos	0
WorkingHour_flag	0
Cos_2880_2	0
Day_of_week	0
Day_of_week_sin	0

TABLE III
NUMERIC FORECASTING RESULTS FOR PERIOD A

Model	MSE		MAE	
	Val	Test	Val	Test
MLP	775.948	967.198	21.453	23.590
TPOT	456.508	815.548	16.486	21.707
FLAML	500.241	713.855	17.607	20.568
AutoSklearn	542.389	727.531	17.455	20.821

forecasting day-ahead electricity prices in Germany, benchmarking their performance against an MLP model. The evaluation was carried out across two distinct periods: one characterized by stable prices and the other by significant fluctuations.

The findings reveal that during periods of relatively stable electricity prices, all AutoML models outperformed the neural network baseline in both the validation and test phases. However, in highly volatile market conditions, validation results proved to be less reliable, as TPOT and Auto-sklearn underperformed the benchmark in validation but delivered superior accuracy on the test set. When considering test data, FLAML demonstrated the best performance in terms of MAE and MSE during stable price periods, whereas TPOT emerged as the most effective forecasting model under volatile market conditions.

TABLE IV
NUMERIC FORECASTING RESULTS FOR PERIOD B

Model	MSE		MAE	
	Val	Test	Val	Test
MLP	1218.321	11253.343	26.572	63.194
TPOT	1799.552	8448.086	34.761	50.655
FLAML	1096.534	10644.551	24.682	54.327
AutoSklearn	1194.660	9283.432	27.747	51.344

REFERENCES

[1] N. Singh and S. R. Mohanty, "A review of price forecasting problem and techniques in deregulated electricity markets," *J. Power Energy Eng.*, vol. 03, no. 09, pp. 1–19, 2015.

[2] R. Qussous, N. Harder, M. Schafer, and A. Weidlich, "Increasing the realism of electricity market modeling through market interrelations," in *2022 Open Source Modelling and Simulation of Energy Systems (OSMSES)*. IEEE, Apr. 2022.

[3] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S030626191830196X

[4] C. O'Leary, C. Lynch, R. Bain, G. Smith, and D. Grimes, "A comparison of deep learning vs traditional machine learning for electricity price forecasting," in *2021 4th International Conference on Information and Computer Technologies (ICICT)*, 2021, pp. 6–12.

[5] R. S. Olson, N. Bartley, R. J. Urbanowicz, and J. H. Moore, "Evaluation of a tree-based pipeline optimization tool for automating data science," in *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, ser. GECCO '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 485–492. [Online]. Available: https://doi.org/10.1145/2908812.2908918

[6] C. Wang, Q. Wu, M. Weimer, and E. Zhu, "Flaml: A fast and lightweight automl library," in *MLSys*, 2021.

[7] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," in *Advances in Neural Information Processing Systems 28 (2015)*, 2015, pp. 2962–2970.

[8] M. Hanussek, M. Blohm, and M. Kintz, "Can automl outperform humans? an evaluation on popular openml datasets using automl benchmark," in *2020 2nd International Conference on Artificial Intelligence, Robotics and Control*, ser. AIRC'20. New York, NY, USA: Association for Computing Machinery, 2021, p. 29–32. [Online]. Available: https://doi.org/10.1145/3448326.3448353

[9] P. Pilyugina, S. Medvedeva, K. Mosievich, I. Trofimov, A. Kostromina, D. Simakov, and E. Burnaev, "A large-scale empirical study of aligned time series forecasting," *IEEE Access*, vol. 12, pp. 131 100–131 121, 2024.

[10] G. Westergaard, U. Erden, O. A. Mateo, S. M. Lampo, T. C. Akinci, and O. Topsakal, "Time series forecasting utilizing automated machine learning (automl): A comparative analysis study on diverse datasets," *Information*, vol. 15, no. 1, 2024. [Online]. Available: https://www.mdpi.com/2078-2489/15/1/39

[11] C. O'Leary, C. Lynch, and F. Ghassemi Toosi, "A comparative analysis of automated machine learning libraries for electricity price forecasting," *Applied Computer Systems*, vol. 29, no. 2, pp. 43–52, 2024. [Online]. Available: https://sciendo.com/article/10.2478/acss-2024-0020

[12] SMARD, "Strommarktdaten," 2024, accessed: 10 Feb. 2025. [Online]. Available: https://www.smard.de/en/downloadcenter/download-market-data

[13] European Energy Exchange (EEX), "EU ETS Spot, Futures, and Options Markets," 2024, accessed: 10 Feb. 2025. [Online]. Available: https://www.eex.com/en/markets/environmental-markets/eu-ets-spot-futures-options

[14] M. B. B. Dias, G. R. S. Lira, and V. M. E. Freire, "Methodology for multi-step forecasting of electricity spot prices based on neural networks applied to the brazilian energy market," *Energies*, vol. 17, no. 8, 2024. [Online]. Available: https://www.mdpi.com/1996-1073/17/8/1864

[15] M. Longo, D. Zaninelli, P. Siano, and A. Piccolo, "Evaluating innovative fcn networks for energy prices' forecasting," in *2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, 2016, pp. 315–320.

[16] K. Gao, Y. Jiang, Q. Ma, and C. Zhang, "Electricity price forecasting model based on MLP-mixer," in *Second International Conference on Digital Society and Intelligent Systems (DSInS 2022)*, J. Hu and X. Yang, Eds., vol. 12599, International Society for Optics and Photonics. SPIE, 2023, p. 1259907. [Online]. Available: https://doi.org/10.1117/12.2673665