Hour-ahead electrcity price prediction using LSTM

Daniel Sunil Kumar Ravi*, Md Faisal Kabir, Ph.D.†, Md Habib Ullah, Ph.D.‡
*MS CS, Penn State Harrisburg
†Department of Computer Science, Penn State Harrisburg
‡Department of Electrical Engineering, Penn State Harrisburg

***Abstract*—Forecasting hour-ahead electricity prices is critical for optimizing energy market operations, improving operational efficiency for energy companies, and enabling cost-effective energy decisions for consumers. However, the inherent volatility of electricity prices—driven by demand, supply, weather, and market dynamics—makes precise prediction challenging. We employ a Long Short-Term Memory (LSTM) network to address these challenges. The dataset from the New York Independent System Operator (NYISO), includes 75 input variables such as historical prices and load demand. Preprocessing steps include outlier handling, feature reduction using Principal Component Analysis (PCA), and lagged feature selection to streamline the data. The LSTM model effectively captures temporal patterns, achieving a Mean Absolute Error (MAE) of 1.356, Root Mean Squared Error (RMSE) of 4.065, and Mean Absolute Percentage Error (MAPE) of 6.46%. These results demonstrate the model's ability to predict trends and short-term fluctuations accurately.**

***Index Terms*—Electricity Price Forecasting, Deep Learning, Long Short-Term Memory (LSTM), Wavelet Transformation, Hybrid Models, Time-Series Prediction, Feature Decomposition**

I. Introduction

Accurately forecasting electricity prices has become increasingly important in modern energy markets. It plays a critical role in improving generation scheduling, optimizing bidding strategies for energy companies, and helping consumers manage energy costs more effectively. However, electricity prices are highly volatile, influenced by factors such as fluctuating demand, varying supply, weather conditions, and dynamic market trends. This inherent complexity makes short-term price prediction a challenging yet essential task.

Traditional methods, such as statistical models and simple machine learning approaches, often struggle to handle high-dimensional datasets efficiently. While these methods require less computational power, they fail to capture intricate temporal dependencies and non-linear patterns in electricity price data. On the other hand, modern deep learning techniques, though powerful, often demand significant computational resources and longer training times, especially under high-dimensional inputs.

This research aims to strike a balance by achieving accurate predictions with a simpler and computationally efficient LSTM-based model. By carefully preprocessing the data—reducing dimensionality, handling outliers, and selecting significant features—we ensure that the model remains lightweight while still effectively capturing essential patterns in the data.

This is achieved through a structured pipeline that streamlines the data and model complexity, consisting of:

Outlier Detection and Handling: Outliers are identified using the 3-sigma rule and handled using local window-based averaging to maintain data integrity while accounting for seasonality. Feature Reduction: Principal Component Analysis (PCA) is applied to reduce dimensionality while retaining 95% of the variance, minimizing complexity without sacrificing performance. Lagged Feature Selection: Only significant lagged features are retained to simplify the input space and improve model efficiency.

The LSTM model architecture integrates layers for capturing temporal dependencies, flattening outputs, and applying regularization through dropout layers to prevent overfitting.

This streamlined approach allows the LSTM model to focus on the most relevant patterns in the electricity price data, improving both accuracy and interpretability. The model's performance is evaluated using industry-standard metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), to demonstrate its effectiveness compared to existing methods.

The remainder of this paper is organized as follows: Section 2 reviews related works and highlights their limitations. Section 3 describes the dataset and preprocessing methodology. Section 4 details the proposed LSTM model. Section 5 presents experimental results, followed by conclusions and future work in Section 6.

II. Related Works

Electricity price prediction is a vastly researched field. Several methods, from traditional statistical approaches to advanced machine learning techniques, have been employed in the field, each addressing different aspects of price volatility, seasonality, and the non-linear nature of the data.

A. *Traditional and Statistical Methods*

Statistical methods, such as AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA with Exogenous Inputs (SARIMAX), have been widely used in forecasting electricity prices. These methods work well for stationary data but rely on linear assumptions that often struggle with the multi-dimensional, non-linear nature of electricity prices. McHugh et al. [1] used the NARMAX model to include energy-related factors like historical price and demand. While the model provides accurate hourly predictions, it fails to handle sudden fluctuations and outliers effectively.

Castelli et al. [2] proposed a genetic programming approach to improve forecasting accuracy, incorporating variables like weather and oil prices. Their model outperformed existing methods, though it lacks scalability.

B. Machine Learning Approaches

In contrast to traditional statistical models, machine learning methods have gained popularity due to their ability to capture complex, non-linear relationships within the data. Among the most commonly used methods are Support Vector Machines (SVM) and Random Forests (RF), which offer robust performance but are often limited in scalability and interpretability for real-time forecasting.

Artificial Neural Networks (ANNs), particularly Deep Learning models, have been applied to model temporal dependencies in electricity prices. Gareta et al. [3] used Neural Networks (NNs) for short-term price forecasting but struggled to capture sharp price peaks. Similarly, Bai et al. [4] incorporated Attention Mechanisms into their GA-LSTM model for residential electricity consumption, improving accuracy. However, the model's complexity challenges real-time applications.

Namani et al. [5] proposed a BiLSTM-Autoencoder hybrid model for real-time price forecasting, which efficiently captures both forward and backward temporal dependencies but can be computationally intensive. Shi et al. [6] introduced a two-stage price forecasting model that utilizes deep neural networks and spike prediction techniques to enhance forecasting accuracy.

C. Hybrid Deep Learning Frameworks

Recent studies have introduced hybrid deep learning models to address the limitations of traditional methods. These models combine different techniques to enhance prediction accuracy. Chang et al. [7] introduced a WT-Adam-LSTM model, which uses Wavelet Transform for feature decomposition and LSTM networks optimized with Adam to improve accuracy. However, this model requires complex preprocessing and tuning, increasing computational overhead.

Other studies, such as Liu et al. [8] and Huo et al. [9], have combined LSTM with Extreme Learning Machines (ELM) and chaotic Artificial Bee Colony (ABC) algorithms, respectively, for better price forecasting. Bai et al. [4] further optimized LSTM using Genetic Algorithms (GA) to improve prediction accuracy. However, the increased complexity and longer training times make real-time applications challenging.

Incorporating Grey Relational Analysis (GRA) into models like PSO-GRA-BP [8] has proven effective for feature selection and performance enhancement. Additionally, Zhang et al. [10] combined BiLSTM with CatBoost for short-term price prediction, further improving forecast accuracy.

D. Emerging Techniques: Attention Mechanisms and Optimization

Recent research has also explored the integration of attention mechanisms in deep learning models to address the challenge of identifying relevant features in large datasets. These mechanisms allow the model to focus on the most influential features and capture long-range dependencies. For example, in attention-based LSTM models, attention layers are added to the LSTM network to weight the importance of different time steps in the input data [9]. This approach enhances the model's ability to handle temporal dependencies more effectively and improves prediction accuracy.

Additionally, optimization techniques like chaotic Artificial Bee Colony (ABC) algorithms [9] have been used to improve convergence and hyperparameter selection in deep learning models. Huo et al. [9] demonstrated the use of chaotic ABC in the optimization of LSTM networks for futures price prediction, showing that the method effectively addresses the slow convergence problems of traditional ABC algorithms. However, these models require significant computational resources, which may limit their use in real-time market forecasting.

E. Challenges and Limitations

Despite significant advancements in machine learning and hybrid models, several challenges remain. First, many of the complex hybrid models, such as AE-BiLSTM and WT-Adam-LSTM, offer high accuracy but require extensive computational power, making them impractical for large-scale, real-time applications. Second, the integration of multiple techniques, such as feature decomposition and hyperparameter optimization, increases model complexity, which can lead to overfitting and reduce model generalizability. Finally, many models still struggle with the trade-off between prediction accuracy and computational efficiency, particularly in dynamic and high-dimensional market environments.

F. Key Motivation

The primary motivation for this work is to develop an electricity price forecasting model that balances high prediction accuracy with computational efficiency. While hybrid models like AE-BiLSTM and WT-Adam-LSTM demonstrate strong performance, their computational complexity makes them unsuitable for real-time applications. On the other hand, simpler models like ARIMA and NARMAX are computationally efficient but fail to capture the intricate non-linear dynamics of electricity prices. This study aims to fill this gap by providing a solution that delivers both high predictive accuracy and operational efficiency.

G. Contribution

The key contributions of this work are as follows:

- A simpler LSTM-based framework is proposed for hour-ahead electricity price forecasting, which incorporates PCA-based dimensionality reduction, outlier detection, and lagged feature selection to reduce complexity while maintaining high prediction accuracy.
- The proposed model achieves a balance between accuracy and computational efficiency, making it suitable for real-time applications in energy markets.
- This research provides a practical and scalable solution for hour-ahead price forecasting, addressing the need for real-time decision-making in dynamic energy markets.

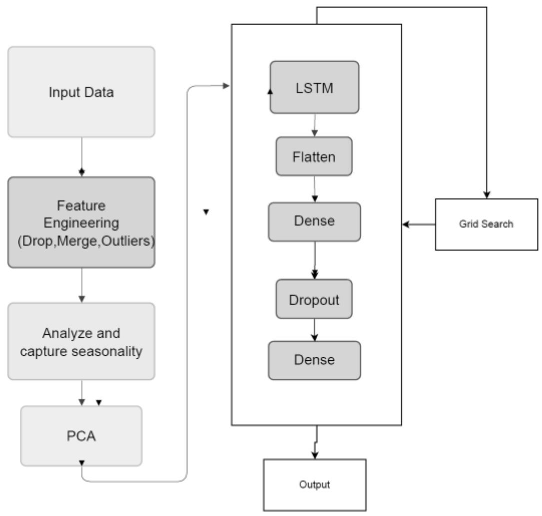

Fig. 1. Model workflow

This work offers a promising solution to the challenges in electricity price forecasting by providing accurate, efficient, and scalable models for real-time market applications.

III. Methodology

This section presents the framework used for electricity price forecasting. The methodology involves several critical steps, including dataset understanding, feature engineering, preprocessing, model development, and hyperparameter tuning.

A. Dataset Understanding

The dataset used for this study comes from the New York Independent System Operator (NYISO), containing features such as historical electricity prices, demand, and generation data. The dataset is analyzed to understand the data distribution, relationships between features, and to identify potential issues such as missing values, outliers, and redundant features.

B. Our Approach

The proposed methodology begins with feature understanding and preprocessing steps, followed by the development of a deep learning model for forecasting electricity prices. The approach is designed to efficiently handle temporal dependencies and non-linear relationships in the data, ensuring that the model captures both short- and long-term patterns.

1) Adding Time Features: To better capture temporal patterns in the dataset, we introduce time-related features such as the hour of the day, day of the week, and month of the year. These features help the model understand seasonal patterns in electricity price fluctuations. This addition enhances the model's ability to predict prices based on periodic trends.

2) Features Understanding, Relevance, and Lagged Features: Feature relevance is determined by examining the relationship between the input features and the target variable (electricity price). Additionally, we explore lagged features, where past price values are used to predict future prices, recognizing their impact on price forecasting.

3) Feature Merging Based on Correlation: Highly correlated features are merged to reduce redundancy and improve the model's performance. For instance, features like demand and system generation are often correlated and are combined into a single composite feature, simplifying the model and reducing overfitting.

4) Outlier Impact, Detection, and Replacement Strategy: Outliers in the dataset can significantly affect model performance, so a robust outlier detection mechanism is applied. We use statistical methods such as the Z-score to detect and replace outliers. For example, any data point that lies beyond three standard deviations from the mean is considered an outlier and is replaced with the mean of the surrounding data points.

5) Seasonality: Seasonal patterns in electricity prices, such as daily or yearly fluctuations, are crucial for accurate forecasting. We ensure that these seasonal effects are captured through feature engineering (e.g., adding time-based features like hour, day, and month). This helps the model identify periodic trends in electricity prices, enhancing the model's predictive accuracy.

6) PCA and Train/Test Split: Principal Component Analysis (PCA) is used to reduce the dimensionality of the dataset while retaining most of the variance. This step helps in improving the computational efficiency and reducing the risk of overfitting. The data is split into training (75

C. LSTM Model Development

The LSTM network, a deep learning model designed for time-series forecasting, is used to predict electricity prices. The LSTM model is capable of learning long-term dependencies in sequential data, making it an ideal choice for this problem.

1) LSTM Architecture: The architecture consists of:

- **Input Layer**: Takes the preprocessed features as input.
- **LSTM Layer**: Captures temporal dependencies and patterns.
- **Dropout Layer**: Prevents overfitting by randomly setting a fraction of input units to 0 during training.
- **Dense Layer**: Outputs the final prediction for electricity prices.

2) Training Procedure: The training procedure involves splitting the data into training, validation, and testing sets. The LSTM model is trained for 50 epochs to capture the temporal patterns. The model's performance is evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

D. Hyperparameter Tuning

Hyperparameter tuning is performed to optimize the LSTM model's performance. The key hyperparameters, such as the

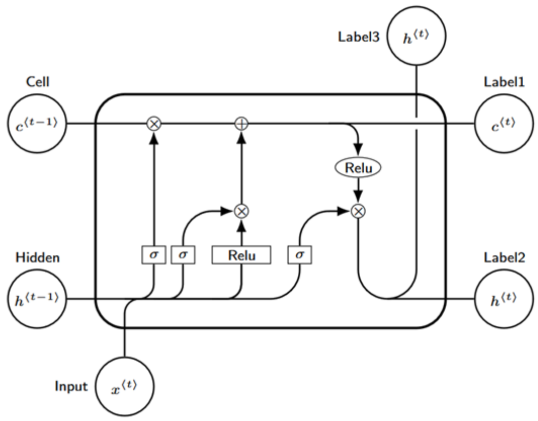

Fig. 2. LSTM internal structure

number of LSTM units, learning rate, and batch size, are tuned using grid search or random search techniques. This process ensures that the model is well-fitted and capable of providing accurate predictions.

IV. Results

This section presents the performance evaluation of the models used for electricity price forecasting. The models evaluated include CNN, CNN-LSTM, XGBoost, and LSTM. The evaluation is based on three key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The models were trained and tested using a dataset of electricity prices, and their performance was measured across these metrics.

A. Performance Metrics

- **Mean Absolute Error (MAE):** This metric measures the average magnitude of errors in the predictions, without considering their direction. A lower MAE indicates better predictive accuracy.
- **Root Mean Squared Error (RMSE):** This metric measures the square root of the average squared differences between predicted and actual values. RMSE gives a higher weight to large errors, making it sensitive to outliers.
- **Mean Absolute Percentage Error (MAPE):** This metric expresses prediction accuracy as a percentage, indicating how much the forecasted values deviate from the actual values.

Model	MAE	RMSE	MAPE
CNN	1.591	4.939	7.211
CNN-LSTM	1.501	4.379	6.844
XGBoost	1.476	4.203	6.798
LSTM	1.356	4.065	6.46

TABLE I
MODEL PERFORMANCE COMPARISON

B. Model Comparison

Table I shows the performance of all models based on MAE, RMSE, and MAPE.

C. Observations

From the results, we can make the following observations:

- **LSTM** performed the best across all three metrics. It achieved the lowest MAE (1.356), RMSE (4.065), and MAPE (6.46), indicating that it provides the most accurate and reliable forecast compared to other models.
- **XGBoost** was the second-best performer, with an MAE of 1.476, RMSE of 4.203, and MAPE of 6.798. While it performed better than CNN and CNN-LSTM, its accuracy was slightly lower than LSTM, indicating that XGBoost may not capture the temporal dependencies in the data as well as LSTM.
- **CNN** and **CNN-LSTM** performed relatively poorly compared to LSTM and XGBoost. CNN had the highest MAE (1.591) and RMSE (4.939), and the highest MAPE (7.211), which suggests that it struggles to capture the temporal patterns in the electricity price data effectively. CNN-LSTM, despite incorporating LSTM layers, did not significantly outperform CNN, likely due to its lower ability to handle the temporal dependencies.

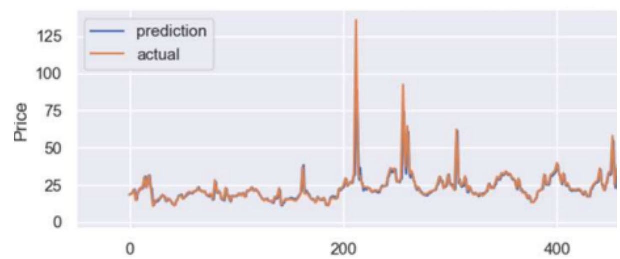

Fig. 3. Actual vs Predicted data visualization

V. Conclusion

In this study, we proposed a framework for forecasting electricity prices using machine learning models, with a particular focus on Long Short-Term Memory (LSTM) networks. The models evaluated included CNN, CNN-LSTM, XGBoost, and LSTM, with performance measured based on three key metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

The results demonstrated that the LSTM model outperformed all other models, achieving the lowest values for MAE, RMSE, and MAPE. This indicates that the LSTM model effectively captured the temporal dependencies in the electricity price data, making it the most accurate and reliable model for short-term forecasting in this context. XGBoost showed promising results, though it lagged behind LSTM in terms of predictive accuracy. On the other hand, CNN and CNN-LSTM performed less effectively, likely due to their limited capacity to capture temporal patterns in time-series data.

These findings highlight the importance of choosing the right model for time-series forecasting tasks, especially when capturing complex temporal dependencies is crucial. The success of the LSTM model in this study suggests that similar approaches can be applied to other time-series prediction tasks in the energy sector and beyond.

Future work can explore more complex models to increase the performance oriented projects, the real-time application of the model in dynamic electricity markets can provide valuable insights into its practical viability for operational forecasting.

VI. Discussion

The results from the LSTM model demonstrate its strong performance in predicting electricity prices on an hour-ahead basis, surpassing traditional machine learning models like CNN, CNN-LSTM, and XGBoost. This highlights LSTM's ability to effectively capture the temporal dependencies and intricate patterns in time-series data, which is crucial for accurate short-term price forecasting in the energy market.

XGBoost showed competitive results, but it did not outperform LSTM, suggesting that while it can handle complex relationships, it might not fully capture the sequential nature of electricity price data. The CNN and CNN-LSTM models, despite incorporating advanced features, performed poorly, indicating that they might not be well-suited for time-series prediction without more tailored adjustments to handle temporal dependencies.

The proposed methodology, including outlier detection, feature reduction via PCA, and lagged feature selection, contributed to the model's efficiency and accuracy. These preprocessing steps streamlined the data, ensuring that the LSTM model focused on the most relevant patterns, making it both computationally efficient and accurate.

Overall, our findings emphasize the significance of selecting the right model for time-series forecasting and suggest that LSTM is a promising approach for electricity price prediction, with potential applicability in other dynamic forecasting problems in the energy sector.

References

[1] C. McHugh, S. Coleman, and D. Kerr, "Hourly electricity price forecasting with NARMAX," *Machine Learning with Applications*, vol. 9, p. 100 383, Sep. 2022, ISSN: 2666-8270. DOI: 10.1016/j.mlwa.2022.100383.

[2] M. Castelli, A. Groznik, and A. Popovič, "Forecasting Electricity Prices: A Machine Learning Approach," en, *Algorithms*, vol. 13, no. 5, p. 119, May 2020, ISSN: 1999-4893. DOI: 10.3390/a13050119.

[3] R. Gareta, L. M. Romeo, and A. Gil, "Forecasting of electricity prices with neural networks," *Energy Conversion and Management*, vol. 47, no. 13, pp. 1770–1778, Aug. 2006, ISSN: 0196-8904. DOI: 10.1016/j.enconman.2005.10.010.

[4] Z. Bai, "Residential electricity prediction based on GA-LSTM modeling," *Energy Reports*, vol. 11, pp. 6223–6232, Jun. 2024, ISSN: 2352-4847. DOI: 10.1016/j.egyr.2024.06.010.

[5] S. Namani, M. A. Amin, M. O. Miah, A. A. A. Khan, M. F. Kabir, and M. H. Ullah, "Real-time energy price forecasting using bilstm-autoencoder deep learning model," *IEEE Transactions on Power Systems*, pp. 1–10, 2024.

[6] W. Shi, Y. Wang, Y. Chen, and J. Ma, "An effective Two-Stage Electricity Price forecasting scheme," *Electric Power Systems Research*, vol. 199, p. 107 416, Oct. 2021, ISSN: 0378-7796. DOI: 10.1016/j.epsr.2021.107416.

[7] Z. Chang, Y. Zhang, and W. Chen, "Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform," *Energy*, vol. 187, p. 115 804, Nov. 2019, ISSN: 0360-5442. DOI: 10.1016/j.energy.2019.07.134.

[8] E. Liu, J. Li, A. Zheng, H. Liu, and T. Jiang, "Research on the Prediction Model of the Used Car Price in View of the PSO-GRA-BP Neural Network," en, *Sustainability*, vol. 14, no. 15, p. 8993, Jan. 2022, ISSN: 2071-1050. DOI: 10.3390/su14158993.

[9] L. Huo, Y. Xie, and J. Li, "An Innovative Deep Learning Futures Price Prediction Method with Fast and Strong Generalization and High-Accuracy Research," en, *Applied Sciences*, vol. 14, no. 13, p. 5602, Jan. 2024, ISSN: 2076-3417. DOI: 10.3390/app14135602.

[10] L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: Catboost and xgboost approaches," *Expert Systems with Applications*, vol. 241, p. 122 686, 2024, ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2023.122686.