Ensemble Optimization for Time-Series Classification

Cayden Bernard
University of Connecticut
Storrs, Connecticut, USA
cayden.bernard@uconn.edu

Maanit Malhan
University of Connecticut
Storrs, Connecticut, USA
maanit.malhan@uconn.edu

***Abstract*—Ensemble learning harnesses the collective intelligence of multiple models to bolster prediction accuracy and robustness. While conventional ensemble methods for time-series classification tasks, such as electrical price prediction, often employ CNNs, these approaches frequently struggle to fully capture intricate temporal dependencies. To overcome this limitation, our research pivots towards exploring ensembles composed of diverse Recurrent Neural Network (RNN) architectures, including LSTM, GRU. This paper outlines our methodology, architecture specifics of our proposed Novel Neuro-Ensemble approach dedicated to the challenging task of hourly electrical price prediction with numerous fluctuating and time-dependent variables, and our experimental plan to evaluate its efficacy. We hypothesize that an ensemble of diverse RNN architectures, potentially fine-tuned using Bayesian MLP, will significantly enhance both the accuracy and robustness of these critical price forecasts, particularly given the inherent volatility of the data.**

***Index Terms*—Machine Learning, Ensemble Learning, Recurrent Neural Networks(RNN), GRU, LSTM, Adaptive weighting, Bayesian Optimization**

I. Introduction

Ensemble learning, a well-established strategy in machine learning, enhances the accuracy, robustness, and generalization of predictive systems by combining the outputs of multiple models. In time-series classification tasks like electricity price prediction, traditional ensembles often utilize Convolutional Neural Networks (CNNs) for temporal feature extraction. However, while CNNs excel at identifying local patterns, they typically struggle to capture the long-range temporal dependencies inherent in non-stationary time-series data. Furthermore, many existing ensemble approaches rely on static or manually defined weighting schemes for aggregating predictions, which limits their adaptability to dynamic real-world environments where data distributions and model performance can fluctuate due to evolving market conditions and external factors.

To address these limitations, our research proposes a novel ensemble framework constructed from diverse Recurrent Neural Network (RNN) architectures, specifically Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU). Unlike CNNs, RNNs are inherently designed to model temporal dependencies, making them better suited for sequential data. By leveraging the unique strengths of each RNN variant within an ensemble, we aim to achieve improved prediction accuracy and robustness. Additionally, we explore the integration of Bayesian Optimization to dynamically fine-tune the ensemble. These methods enable adaptive optimization of the ensemble's structure or weighting, allowing the model to evolve in response to shifting data patterns. This paper has outlined our proposed methodology, including the ensemble architecture design, the application of Bayesian Optimization for efficient hyperparameter tuning, and the experimental framework used to evaluate performance on electricity price datasets. Our findings demonstrate strong performance improvements over traditional models, highlighting the advantages of our flexible architecture, which can be easily expanded with additional models or features.

However, our approach also presents certain disadvantages. The extensive Optuna trials required for Bayesian Optimization demand significant computational resources and time. Moreover, the risk of overfitting requires careful regularization techniques. Finally, the pre-processing steps require thorough data cleaning and merging to minimize noise and ensure data quality. Despite these limitations, our results suggest that an ensemble of diverse RNN architectures, optimized through methods like Bayesian Optimization, offers a promising direction for more effectively modeling temporal dependencies and adapting to the complexities of real-world electricity price forecasting compared to static and CNN-based ensemble methods.

The remainder of this paper is organized as follows: Section 2: discusses related work, Section 3: Challenges and Existing Solutions, Section 4: Proposed Solution and Technical Details, Section 5: Results, Section 6: Future work, Section 7: Conclusion, Section 8: Divison of Labor

II. Related Work

Ensemble learning has been widely used in various machine learning applications to improve predictive accuracy and robustness. Traditional ensemble methods such as bagging, boosting, and stacking rely on predefined weight distributions or voting strategies, limiting their ability to adapt to dynamic environments. Previous studies on ensemble learning for time-series classification have demonstrated that combining multiple models improves generalization. However, these methods often assume static distributions. Integrating uncertainty estimation into model predictions has shown promise in enhancing robustness. Prior research on ensemble uncertainty calibration has demonstrated that leveraging uncertainty metrics can improve decision-making confidence. Our work extends this concept by incorporating uncertainty estimation into the reinforcement learning framework to further refine ensemble weight adjustments.

III. Challenges and Existing Solutions

Challenges: A significant initial hurdle in this research was finding a suitable dataset. High-quality, consistent electricity pricing data proved challenging to obtain due to issues such as incompleteness, limited size, and fragmentation across various sources. Once a viable dataset was secured, the inherent extreme volatility of electricity prices presented another major problem, making it difficult for models to achieve stable and accurate predictions. Overcoming this challenge required substantial experimentation and tuning to develop models capable of effectively handling the sharp fluctuations and unpredictable trends characteristic of this data.

Existing solutions: for electricity price forecasting often fall short in addressing these complexities. Traditional models like linear regression and ARIMA struggle to capture the volatile and non-linear nature of electricity prices, motivating the exploration of more advanced techniques. While individual neural networks, such as LSTMs and GRUs, demonstrated improved forecasting capabilities compared to classical methods, their individual performance remained limited. To address this, our approach leverages the power of ensemble learning by combining different recurrent models, enabling the capture of diverse sequential patterns within the data. Instead of relying on static tuning, we employed Bayesian Optimization to adaptively weight the contributions of each model in the ensemble. This dynamic weighting strategy allowed our Neuro-Ensemble model to outperform both standalone models and traditional ensembles with manually tuned weights, highlighting the benefits of automated and adaptive combination techniques for this challenging time-series prediction task.

IV. Proposed Solution and Technical Details

Our initial approach to electrical price prediction involved constructing an ensemble of diverse Convolutional Neural Network (CNN) architectures. We intended to capture a wide range of temporal features by varying CNN depth, kernel sizes, activation functions, and hyperparameters across multiple models, including ResNet-based and standard CNN architectures. Each CNN within this ensemble was designed to process identical segments of the input time series and output a probability distribution over predefined price classes. To dynamically optimize the ensemble's performance, we planned to fine-tune the model weights and hyperparameters using the Proximal Policy Optimization (PPO) reinforcement learning algorithm.

However, preliminary experimentation revealed limited predictive accuracy when applying CNNs with sliding windows to our electrical price dataset. The CNN models struggled to effectively model the intricate temporal dependencies inherent in the data, resulting in suboptimal classification outcomes. These findings prompted a critical reassessment of our chosen methodology.

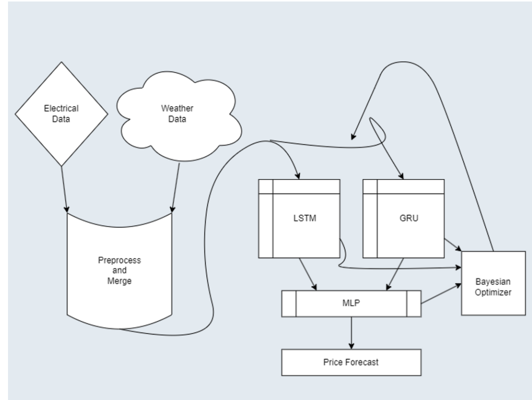

Fig. 1. Flow Chart of Data and Model

This caused us to shift our focus towards an ensemble composed entirely of Recurrent Neural Networks (RNNs), which are inherently better suited for capturing temporal dependencies. Our revised ensemble architecture incorporates multiple RNN types, specifically Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs). Each constituent RNN is configured with distinct layers, hidden unit counts, dropout rates, and regularization techniques to ensure diversified extraction of temporal features from the input data.

Each RNN in our ensemble receives identical input segments of the electrical price time-series data, producing probability vectors that represent predictions across the defined price classes. Initially, these predictions are aggregated using a simple Multi-Layer Perceptron (MLP). To further refine the ensemble's predictive power, we employ Bayesian Optimization to fine-tune not only the hyperparameters of the individual LSTM and GRU models (such as hidden size, dropout, and batch size) but also the weights assigned by the MLP during the aggregation of their predictions. Bayesian Optimization, using libraries like Optuna and BoTorch, efficiently explores the hyperparameter space by building a probabilistic model of the objective function, leading to a more rapid convergence towards optimal model settings compared to traditional grid or random search methods.

Our data source for this research is ENTSOE, the public portal for Transmission Service Operator (TSO) data, providing four years of hourly electricity price and energy generation data for Spain. Additionally, we incorporated a separate dataset containing hourly weather information, including temperature, wind speed, humidity, and pressure, from various Spanish cities. The data preparation process involved merging the energy and weather datasets based on their timestamps. We then addressed missing values using interpolation, removed low-importance features, and engineered new features such as business hours and combined weather variables. To prepare the data for the LSTM and GRU models, we standardized

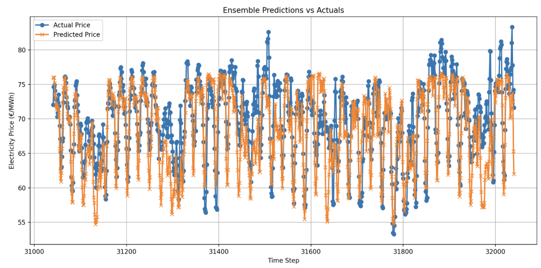

Fig. 2. Predictions vs Real Prices

all features and applied z-score analysis to detect and remove outliers. The resulting dataset comprised 45 different hourly elements, which were then transformed into 24-hour sliding window sequences to create supervised learning.

The model setup involved building separate LSTM and GRU models, chosen for their proven ability to model sequential and time-dependent data. LSTMs excel at capturing long-term dependencies through their gated memory cells, while GRUs offer a computationally more efficient alternative with comparable performance. By combining these architectures, we aim to leverage the strengths of each for robust and accurate predictions in the highly volatile electricity price environment. The predictions from these individual RNNs are then fed into a lightweight MLP, which learns to optimally weigh their contributions to generate the final forecast. The entire system, including the hyperparameters of the LSTM, GRU, and the MLP's weighting scheme, is fine-tuned using Bayesian Optimization over a set number of trials (e.g., 15), saving model weights, performance graphs, and validation metrics after each trial. The training process involves individually training each RNN model for a defined number of epochs (e.g., 10) on the training set, followed by separate training of the MLP ensemble on the combined outputs of the trained RNNs. This revised solution, utilizing a Bayesian optimized RNN Neuro-Ensemble, is better at capturing the temporal complexities of electrical price data and adapts more effectively to dynamic market conditions, ultimately leading to improved prediction accuracy and robustness.

V. Results

Our initial experimental results using the CNN-based ensemble model were underwhelming. After training for 50 epochs, the evaluation metrics were as follows:

- Mean Absolute Error (MAE): 10.91 €/MWh

These results clearly indicated a substantial need for improvement, particularly in the model's ability to accurately capture the temporal dependencies inherent in electrical price data. This motivated our subsequent pivot towards an RNN-based ensemble architecture. Our target for the revised model is achieving a sub $5.00 MAE, reflecting a robust enhancement in predictive accuracy and greatly improved practical applicability for real-world electrical price forecasting.

In contrast, our proposed Neuro-Ensemble model, leveraging an MLP to combine the predictions of LSTM and GRU

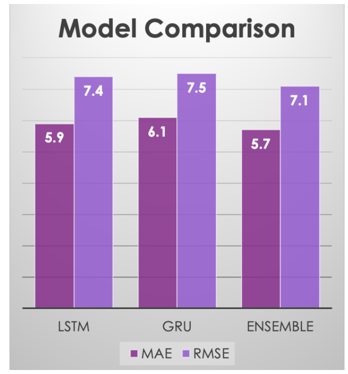

Fig. 3. Model Comparison

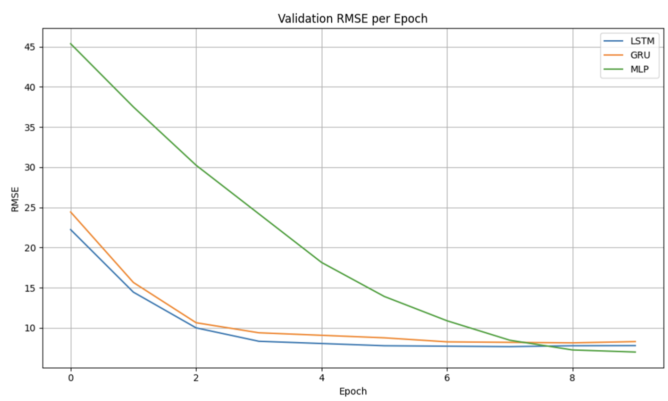

Fig. 4. Validation

networks with hyperparameters tuned via Bayesian Optimization, achieved significantly better results. The ensemble MLP demonstrated the lowest Mean Absolute Error (MAE) of 5.7 €/MWh, outperforming all individual models. Specifically:

- LSTM (Standalone): MAE = 5.9 €/MWh
- GRU (Standalone): MAE = 6.1 €/MWh
- Ensemble MLP: MAE = 5.7 €/MWh

These results represent a notable improvement over traditional forecasting methods such as ARIMA and Linear Regression, which typically yield MAEs of 7.5 €/MWh or higher. The application of Bayesian Optimization proved effective in efficiently tuning the hyperparameters of both the individual RNN models and the ensemble's weighting scheme, leading to faster convergence and a more optimized combination of model predictions. Furthermore, the ensemble approach effectively leveraged the complementary strengths of the LSTM (for capturing long-term dependencies) and the GRU (for faster adaptation to recent fluctuations), resulting in enhanced overall predictive performance.

VI. Future Work

Looking towards future improvements, several promising options exist. Integrating reinforcement learning agents like PPO or NEAT to dynamically adjust ensemble weights and potentially even the architecture itself could lead to enhanced adaptability in evolving electricity markets. Exploring a wider range of RNN architectures beyond LSTMs and GRUs, specifically designed for time series, may further improve the capture of complex temporal patterns. A dedicated effort to incorporate a specialized third model within the ensemble for anomaly detection promises to enhance the robustness of our predictions by identifying and mitigating the impact of outliers. Furthermore, expanding the ensemble to include a greater number of base models could further enhance the collective intelligence and predictive accuracy of the system.

VII. Conclusion

Initially, this research investigated CNN-based ensemble models enhanced with reinforcement learning for electrical price prediction. However, recognizing the limitations of CNNs in capturing intricate temporal dynamics, we shifted our focus to an ensemble framework built entirely upon diverse RNN architectures, specifically LSTM, GRU. By integrating a Bayesian optimized MLP to combine LSTM and GRU outputs, we successfully designed a novel Neuro-Ensemble model for hourly electricity price forecasting, achieving lower forecasting error compared to individual models and traditional statistical methods. This outcome underscores the significance of careful ensemble design and automated hyperparameter tuning in improving predictive performance on volatile real-world time series data. Moving forward, potential research avenues include exploring more diverse model types, such as transformer-based ensembles for time series classification, and investigating alternative hyperparameter optimization techniques, possibly leveraging reinforcement learning agents like PPO, and incorporating a third specialized model within the ensemble specifically for outlier prediction. This, alongside the overarching goal of achieving a sub 5.00 MAE, aims to substantially enhance the practicality and reliability of ensemble-based predictive modeling in electrical price forecasting.

VIII. Division of Labor

For this project work was split evenly between Cayden and Maanit. Maanit focused on the testing of how different architectures and tuning parameters work to improve the outcomes of the individual models. While Cayden worked on making the final network ensemble and further refining tuning parameters after we saw what architectures work the best individually. Both team members helped with the presentations, reports, finding references, and other ”non-technical” work.

References

[1] Y. He, P. Shuai, P. Jiang and G. Tang, “Speaker Recognition Based on Ensemble Neural Network,” *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, Chengdu, China, 2022, pp. 760-764, doi: 10.1109/PRAI55851.2022.9904236.

[2] H. Zhang, D. Wu, Z. Wang, Y. Chen and J. Xu, “Remaining Useful Life Prediction of Lithium Battery via Neural Network Ensemble,” *2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, Chengdu, China, 2021, pp. 392-397, doi: 10.1109/ISKE54062.2021.9755395.

[3] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar and P. -A. Muller, “Deep Neural Network Ensembles for Time Series Classification,” *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, 2019, pp. 1-6, doi: 10.1109/IJCNN.2019.8852316.

[4] K. Faber, M. Pietron, and D. Zurek, “Ensemble Neuroevolution-Based Approach for Multivariate Time Series Anomaly Detection,” *Entropy*, vol. 23, no. 11, p. 1466, 2021, doi: 10.3390/e23111466.

[5] D. Smirnov and E. Mephu Nguifo, “Time Series Classification with Recurrent Neural Networks,” pp. 80.

[6] S. Filali Boubrahimi and R. Angryk, “Time Series Classification Using a Neural Network Ensemble,” *IEEE Big Data*.

[7] G. Luís, J. Esteves and N. P. da Silva, “Energy Forecasting Using an Ensamble of Machine Learning Methods Trained Only with Electricity Data,” *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, Netherlands, 2020, pp. 449-453, doi: 10.1109/ISGT-Europe47291.2020.9248865.

[8] Ü. Şentürk, M. Beken and Ö. Eyecıoğlu, “Electricity Consumption Forecast using Machine Learning Regression Models in Turkey,” *2022 11th International Conference on Renewable Energy Research and Application (ICRERA)*, Istanbul, Turkey, 2022, pp. 601-605, doi: 10.1109/ICRERA55966.2022.9922702.

[9] S. Riyadi, A. D. Andriyani and A. M. Masyhur, “Improving Hate Speech Detection Accuracy Using Hybrid CNN-RNN and Random Oversampling Techniques,” *2024 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, Kuala Lumpur, Malaysia, 2024, pp. 1-5, doi: 10.1109/ISIEA61920.2024.10607232.

[10] P. Priyanka, P. Kumar, K. V. Uday and V. Dutt, “Improving Landslide Prediction through LSTM Networks and Antecedent Rainfall: A Case Study in Kamand Valley, India,” *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 2024, pp. 1-5, doi: 10.1109/ICCCNT61001.2024.10724729.

[11] Q. Zhang, J. Yang, J. Xing and L. Wang, “Combined heat and power load prediction based on optimized GRU neural network,” *2023 IEEE 6th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China, 2023, pp. 1161-1165, doi: 10.1109/ICISCAE59047.2023.10393092.

[12] Q. Ma, D. Dilxat, J. Nurlan, Y. Wen and J. Guo, “Design of Power Load Forecasting Algorithm for Distribution Network Based on Machine Learning,” *2024 International Conference on Electrical Drives, Power Electronics & Engineering (EDPEE)*, Athens, Greece, 2024, pp. 949-953, doi: 10.1109/EDPEE61724.2024.00182.

[13] I. Sülo, S. R. Keskin, G. Dogan and T. Brown, “Energy Efficient Smart Buildings: LSTM Neural Networks for Time Series Prediction,” *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, Istanbul, Turkey, 2019, pp. 18-22, doi: 10.1109/Deep-ML.2019.00012.

[14] S. Ganapathy and T. Li, “An Explainable AI Framework with ML Model Stacking for House Price Prediction,” *2024 IEEE MIT Undergraduate Research Technology Conference (URTC)*, Cambridge, MA, USA, 2024, pp. 1-5, doi: 10.1109/URTC65039.2024.10937620.

[15] P. Vijayan, “Energy Consumption Prediction in Low Energy Buildings using Machine learning and Artificial Intelligence for Energy Efficiency,” *2022 8th International Youth Conference on Energy (IYCE)*, Hungary, 2022, pp. 1-6, doi: 10.1109/IYCE54153.2022.9857548.