IE 434 Deep Dive Project

Project Title: Forecasting Illinois Locational Marginal Prices Using Machine Learning and Deep Learning Models

Team Members: Prithvish Taukari (taukari2), Viswanath Vadlamani (sv56), Yongrin Park (yongrin2)

1. Introduction

Locational Marginal Price (LMP) forecasting is essential for power system operators, utilities, and large consumers who must anticipate hourly price fluctuations caused by load, congestion, outages, weather conditions, and market dynamics. Illinois Hub LMP is particularly volatile, with sudden spikes reaching thousands of dollars.

The goal of this project was to predict hourly Illinois Hub LMP using machine learning and deep learning methods, evaluate the predictive performance of different model families, and determine which features most strongly influence price movements. Our focus was not only on model accuracy but also on understanding the limits of deep learning for this noisy, high-volatility time series.

2. Data Summary

The data used in this project consisted of hourly observations from 2017 to 2025, combining multiple operational datasets from MISO such as Actual LMP, Day-Ahead LMP, and various load-related fields. To construct a unified modeling table, we first standardized timestamps across all files to ensure perfect temporal alignment. Hub-level markets such as Illinois, Indiana, Michigan, and Minnesota were reshaped into a wide pivoted format, producing feature columns like *act_ILLINOIS.HUB*, *da_INDIANA.HUB*, and others that allowed the models to learn cross-regional price relationships directly. Missing values within these hub price series were handled using a mixture of interpolation and forward/backward fills, preserving the continuity of the data without artificially smoothing volatility. Additional engineered features—such as one-hour lag values and 24-hour rolling means—were added to embed short-term temporal dependencies. After cleaning, aligning, and feature engineering, the final dataset contained approximately 65,000 hourly rows and 26 input features. This dataset served as the consistent foundation for the baseline models, neural networks, and deep sequence architectures evaluated in this study.

3. Modeling Approach

Our modeling framework consisted of three major classes of predictive models: linear baselines, feed-forward neural networks, and deep sequence models. We began with Ridge Regression, chosen for its robustness and ability to handle highly collinear features such as hub-to-hub LMP prices. Using a pipeline with a robust scaling step, Ridge served as the benchmark model and established a surprisingly strong performance, suggesting that much of the predictive structure in the data was linear or locally linear. We then implemented a Dense Neural Network trained on the same input features to assess whether nonlinear interactions across hubs and load variables could further improve accuracy. While the network successfully captured nonlinear relationships, its performance did not surpass the simpler Ridge model. Finally, we explored a variety of sequence-based architectures, including LSTM and GRU models trained on rolling windows of historical hourly prices. These models aimed to learn longer-term temporal dependencies but struggled to generalize due to the extreme volatility and irregular spikes inherent in the LMP data. Despite tuning hyperparameters, altering sequence lengths, scaling inputs, and experimenting with expanding-window configurations, sequence models consistently underperformed relative to Ridge and the Dense network. Overall, the modeling results revealed that the strongest predictive signals reside in current- and near-current price features rather than in long autoregressive sequences.

4. Feature Importance Findings

To better understand the drivers of price behavior and to evaluate model interpretability, we conducted a detailed feature importance analysis using both coefficient-based inspection from Ridge Regression and permutation importance techniques. The Ridge coefficients revealed that day-ahead prices, particularly *da_ILLINOIS.HUB* and *da_INDIANA.HUB*, had the largest positive influence on forecasting real-time Illinois LMP. Actual hub prices from neighboring regions such as Michigan and Indiana also ranked highly, reinforcing the interconnected nature of regional electricity markets. Load features, including forecasted load and cleared load, contributed moderate but meaningful predictive value, reflecting demand-driven dynamics in short-term pricing. Permutation importance produced a consistent pattern, with *act_MICHIGAN.HUB* and *da_ILLINOIS.HUB* causing the largest degradation in model performance when shuffled. By contrast, weather-based features, when included, showed near-zero influence on predictive accuracy, suggesting that weather impacts are too diffuse or indirect to affect hourly LMP at this forecasting scale. Overall, these findings demonstrate that

Illinois LMP is shaped predominantly by synchronous price signals from nearby hubs, with load variables providing secondary adjustment effects and historical temporal structure playing a relatively minor role.

5. Technical Challenges & Lessons Learned

Throughout the project, several technical challenges provided insight into the limitations of deep learning for real-world energy forecasting. First, sequence models such as LSTMs and GRUs were unable to learn stable temporal patterns due to the highly volatile and spike-heavy nature of hourly LMP data. Even after experimenting with input scaling, window size adaptations, and different optimization strategies, these models frequently failed to converge or produced unstable test-set results with low explanatory power. Second, integrating weather data—temperatures, wind speed, humidity, and precipitation—did not improve predictive performance for any model class. While logically connected to electricity consumption, weather fluctuations exhibited weak short-term correlation with hourly Illinois LMP, reducing their usefulness in this modeling context. Dimensionality reduction via PCA also proved ineffective, as it did not enhance model accuracy and instead sacrificed interpretability by blending hub-level signals into mixed components. Finally, we observed that price spikes, sometimes exceeding $2,000, played an important role in maintaining realistic model behavior. Attempts to clip or remove these spikes initially improved training stability but resulted in poorer test-set accuracy, indicating that the volatility itself contains valuable information. From these challenges, we learned that sophisticated architectures do not guarantee better performance, and that effective forecasting often relies more heavily on domain-aware feature engineering and model simplicity than on algorithmic complexity.

6. Conclusions

The results of this project reveal a clear and consistent pattern across all modeling approaches: simpler, well-regularized machine learning models outperform more complex deep learning architectures when forecasting highly volatile hourly LMP values in the Illinois Hub. Ridge Regression delivered the strongest performance with an R^2 of approximately 0.897, demonstrating that the relationship between current-hour hub prices and Illinois LMP is largely linear and immediate. Although the Dense Neural Network was able to capture nonlinear interactions within the engineered feature set, it did not surpass Ridge, suggesting that the added model complexity provided limited marginal benefit for this dataset. In contrast, sequence models such as LSTMs and GRUs consistently struggled to model the unstable, spike-heavy temporal structure of LMP data, resulting in significantly lower predictive power. From a feature perspective, the models consistently indicated that cross-hub price signals, particularly day-ahead and real-time prices from neighboring hubs, were the dominant drivers of Illinois LMP. Load features played a secondary role, while long temporal histories and external variables like weather contributed little to predictive accuracy. These findings suggest that market participants and system operators may achieve the most reliable short-term LMP forecasts by focusing on contemporaneous regional price information and by employing simpler models that emphasize interpretability, stability, and strong regularization rather than deep sequence learning.

7. Summary

In summary, this project demonstrates that effective forecasting in energy markets depends more heavily on the characteristics of the data and the structure of the underlying domain than on model complexity alone. Our experiments show that Illinois LMP, despite being a time-dependent series, is driven predominantly by concurrent regional market conditions rather than long-term temporal patterns. As a result, models that emphasize engineered features and cross-sectional relationships, such as Ridge Regression, offer significantly better performance than deep recurrent architectures. The deep learning models struggled not because they were under-parameterized, but because the volatility and irregularity of hourly price spikes fundamentally limit the learnability of long-term patterns. This reinforces an important lesson for applied machine learning: the most powerful model is not always the most complex one, but the one that best aligns with the structure and signal-to-noise characteristics of the problem. By integrating domain knowledge, thoughtful feature engineering, and careful evaluation, we were able to produce interpretable, robust, and highly accurate forecasting results, while also gaining valuable insight into the strengths and limitations of different model families in real-world energy analytics.