Deep Learning Methods to Predict Electricity Load

Group 3

Nima Rafizadeh, Amirhossein Hajigholam-Saryazdi, Behdad Ehsani

Abstract

The prediction of electricity load is one of the most important routines for an electricity market. Accurate prediction of electricity load is extremely beneficial to market participants in risk management and decision making. In this study, we use data from hourly real-time electricity load in the New York State, which has one of the world's largest deregulated electricity markets, from January 1, 2019, 00:00 to December 31, 2019, 23:00 to assess the performance of advanced deep learning methods in electricity load prediction. We accomplish this first by employing the most common statistical and machine learning methods such as linear regression, random forest, SVR, and neural network, and then progressing to the advanced deep learning methods including RNN, LSTM, GRUs, and the associated stacked models. The results show that stacked RNN is the best deep learning model due to its lowest MSE and MAE, followed by RNN, and stacked LSTM. All the deep learning models, however, outperform the used statistical and machine learning methods.

1 Introduction

The use of data mining, analysis, and prediction has been critical in decision making [32]. Accurate prediction models may aid in risk mitigation and management return on investment [22]. Electricity, as a clean and promising energy source, is crucial in our everyday life due to its environmental friendliness in comparison to conventional energy sources, which are necessary in the economic sector [16]. Numerous nations have deregulated and competitive energy markets in recent decades, compelling electrical power firms to generate electricity at competitive prices [27]. Thus, electricity load, which is controlled by a complex web of elements [26], is a critical indication for the electrical power system. However, developing a high-quality and effective electricity load model is a concern and a challenge for all power market participants, including economists and risk managers, due to electricity's widespread use in our society and inherent characteristics such as nonlinearity, high volatility, and high frequency [15].

Prediction models, in particular, are frequently employed in management and have evolved into a critical component of management science models [1]. Additionally, demand/supply prediction is a critical area of management prediction since the most worrisome and difficult problem for all players is the future usages change. If future usage changes are reliably predicted, decision makers may make the optimal decision and design a realistic strategy to mitigate market risk and then optimize management's economic and social advantages.

Numerous electricity load prediction models have garnered global attention and developed become a critical management strategy for all players [33]. For example, reliable prediction findings may be used to direct consumers and producers' production schedules, assisting participants in maximizing their benefits and managers in developing an ideal power market operating plan. Additionally, prediction is critical in investment management. However, developing a high-quality and effective electricity prediction model is a concern and a challenge for all power market participants, including economists and risk managers, due to the widespread use of electricity in our society and its inherent characteristics, which include nonlinearity, high volatility, and high frequency [15].

Broadly speaking, the models used for electricity load prediction fall into one of two categories: statistical methods and machine learning (ML) methods [18]. Statistical methods such as autoregressive moving average (ARMA) [6], autoregressive integrated moving average (ARIMA) [23], generalized autoregressive conditional heteroskedasticity (GARCH) [10], vector auto-regression (VAR) [21], and Kalman filters (KF) [29] performs better in relatively stable power markets [20] but fail to capture the nonlinear features and rapid changes in electricity load because of their inherent weaknesses [4]. In contrast, ML methods are superior to statistical methods because of their strength of capturing the abovementioned nonlinear features and rapid changes [12]. In this study, we aim to address the electricity load prediction issue using advanced deep learning (DL) methods including LSTM, GRUs, transformers, and bi-directional LSTM. We conclude this study by comparing their performance with the most common statistical and machine learning methods including linear regression, random forest, SVR, and neural network to identify the best model.

2 Literature Review

The literature on electricity load and price prediction began to grow in the early 2000s [28] and [2]. We can categorize the major approaches to energy load prediction in detail into five categories based on [31]'s review: multi-agent, fundamental, reduced-form, statistical, and computational intelligence models. Multi-agent models replicate the system's activity and construct the process by matching demand and supply. The publications of [25] and [35] are excellent examples of this sort of work from the recent works. [25] models wind energy producers, plug-in electric car owners, and demand response program participants as independent agents in a small Spanish market.

Fundamental or structural approaches are used to examine the influence of physical and economic variables on electricity loads. In this section of the literature, variables are separately modelled and predicted, often using reduced-form, statistical, or machine learning approaches. [14], for example, create a model for utilizing stochastic processes for the independent variables. Additionally, their strategy takes into consideration the bid stack function of the price drivers and the price of power.

Reduced-form models are based primarily on two techniques: Markov regime switching and jump diffusion. These models are substantially better at handling spikes than structural and statistical models. [11] employs a model known as mean-reverting jump diffusion. Their technique accounts for both the trajectory and statistical characteristics of electricity load and its prices. Additionally, [9] suggests a semi-parametric Markov regime switching model. Robust statistical approaches are used to incorporate model parameters in their methodology. Additionally, it is simpler to estimate, requires less computing effort, and requires fewer distributional assumptions.

Statistical models come in a wide range of flavors, ranging from simple naive procedures [20] to highly evolved systems [36]. As [36] highlighted, in the field of electricity market forecasting, there are univariate and multivariate frameworks. In day-ahead electricity price prediction, participants bid on prices and quantities for the following day's 24 hours. In this respect, the first method is to predict all prices in an univariate framework as a 24-step-ahead projection using a single price series. Another technique, referred to as multivariate framework, is to estimate prices from 24 separate time series as one-step-ahead projections. [20], [8], and [7] illustrate the auto-regressive model in detail. As indicated by [8], [20], and [7], poorly calibrated forecasting algorithms cannot outperform the naive approach. While [7] concludes that the Auto-regressive Integrated Moving Average (ARIMA) model is inferior to the model with exogenous variables in the American PJM market, [8] noted that adding an exogenous variable does not always improve prediction accuracy.

Numerous kinds of computational intelligence models are used in the literature on electricity load and price predictions. [19], [3], and [34] present several early stage articles. [19] predicts electricity loads and prices in the Australian market for the next 16 hours using an Artificial Neural Network (ANN) model. Although their works differ in terms of predicting number of steps ahead, and number of layers and neurons in the ANN model, the results demonstrate that the ANN model outperforms the conventional techniques such as ARIMA. [17] proposes a long short-term memory (LSTM) recurrent neural network-based framework for short-term electricity load forecasting. The framework is tested on a publicly available set of real data, of which the performance is compared to various benchmarks techniques and outperform them. More recently, other variant of LSTM, such as bi-directional LSTM and gated recurrent unit (GRU), are also proposed and combined with other technique to improve the prediction result. [30] presents an intelligent hybrid technique that combines a Convolutional Neural Network (CNN) with a Multi-layer Bi-directional Long-short Term Memory (M-BDLSTM) method to forecast energy consumption. The proposed method achieves better prediction results and smallest error than existing techniques on the same data, thus demonstrating its effectiveness. [24] develops a hybrid sequential learning-based energy forecasting model that employs CNN and GRU into a unified framework for accurate energy consumption prediction. The proposed model becomes an effective alternative to the previous hybrid models in terms of computational complexity as well prediction accuracy, due to the representative features' extraction potentials of CNNs and effectual gated structure of multi-layered GRU.

3 Background and Data

Electricity markets around the world are classified into two categories: 1) regulated electricity market in which vertically integrated monopoly utilities cover the entire value chain with oversight from a public regulator. Price variation is low and strictly regulated by authorities, who set prices mostly on the basis of average costs. 2) deregulated electricity market, colloquially referred to as a competitive electricity market, where market parties other than utilities own power plants and transmission networks. In these circumstances, generators (businesses that generate electricity) sell electricity to a wholesale market, where retail energy suppliers purchase it to sell to customers. Owners and operators of the transmission grid are transmission corporations or utility businesses. Under deregulation, prices are determined by the intersection of supply and demand. As a results, we should utilize the data from a deregulated electricity market in this study.

Given the aforementioned remarks, we exploit the data of the hourly real-time electricity load in the state of New York, which has one of the world's largest deregulated electricity markets, from 1^{th} January 2019, 00:00 to 31^{th} December 2019 23:00, resulting in a total of 8759 observations.[1] Figure 1a illustrates the time series of electricity load. Table 1 also presents the summary statistics.

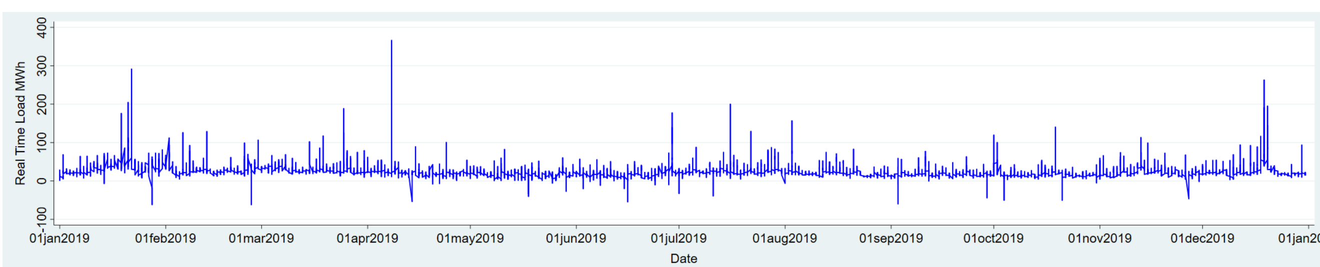
Table 1: Descriptive Statistics for the Hourly Electricity Load

Descriptive Statistics	
Observations	8,759
Mean	18.11
Median	17.59
Maximum	43.91
Minimum	8.251
Standard Deviation	3.613
Skewness	1.206
Kurtosis	5.768

Real time electricity load is the target variable in our prediction model. The prediction model's main features are the real time and day ahead prices. Additionally, some features, including year, month, day, day of week, and hour are decomposed from the Time feature to improve the model's efficiency. An overview of dataset is illustrated in Figure 1b.

Before diving deep into the model section, the single-output and single-time step prediction procedure are presented in Figure 2. The proposed structure aims to predict just the electricity load for the next time step (i.e., next hour).

Figure 1: Data Visualisation



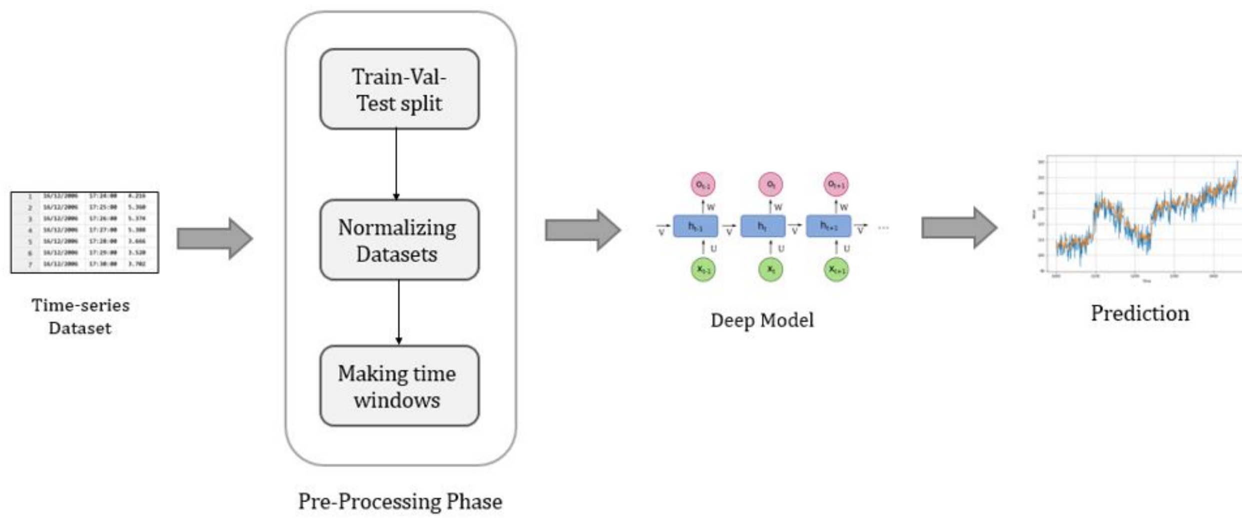
(a) Time Series of Electricity Load

Time	Year	Month	Day	DW	Hour	RealTimePrice	DayAheadPrice	TotalLoad
2019-01-01 00:00	2019	1	1	2	0	19.298	17.954	17.666
2019-01-01 01:00	2019	1	1	2	1	15.858	16.309	14.553
2019-01-01 02:00	2019	1	1	2	2	17.564	14.624	13.988
2019-01-01 03:00	2019	1	1	2	3	10.214	14.532	13.601
2019-01-01 04:00	2019	1	1	2	4	7.262	14.638	13.358
2019-01-01 05:00	2019	1	1	2	5	19.093	14.709	13.432

(b) An Overview of the New York Electricity Load Dataset and Its Proposed Features

[1]The data is publicly available online at `https://www.nyiso.com/`.

Figure 2: Prediction procedure in Deep Learning models



4 Method and Results

Given that we have a labeled data, we employ supervised ML models, as well as DL methods. To be more specific about the DL methods, recurrent methods include three variants, vanilla RNN, long short-term memory (LSTM) [13], a more sophisticated version of RNN called gated recurrent units (GRUs) [5]. Finally, the comparison of ML methods with the DL methods is performed using the New York electricity load dataset.

We briefly explain these advanced methods in the following:

- **RNN:** A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time series data. It is the first algorithm that remembers its input, due to an internal memory.
- **LSTM:** Long Short-Term Memory (LSTM) networks are a type of recurrent neural network being capable of learning long-term dependence in sequence prediction problems.
- **GRU:** Gated recurrent units (GRUs) is a gating mechanism in recurrent neural networks which is like LSTM with a forget gate, but has fewer parameters than LSTM, as it lacks an output gate.
- **Stacked deep models:** Stacked models are obtained by combining two Deep Learning models sequentially.

In order to process time-series using DL models, it is necessary to make a time windows in the pre-processing phase. As it is shown in Figure 3, there are two ways for building time sequences/windows. In the first, the layer returns only the final time step's output, allowing the model time to warm up its internal state before making a single prediction. The layer then returns an output for each input in the second method. This is advantageous when it comes to Stacking RNN layers. In this study, the time window includes data corresponding to the previous 24 time periods for predicting the next hour. In other words, the model uses data from time zero to time 23 to indicate the electricity load for time 24. It should be mentioned that in ML models, the time windowing mechanism is not utilized because ML models use the features of each time step to predict for this time, which is illustrated in Figure 4.

Figure 3: Making Time Windows

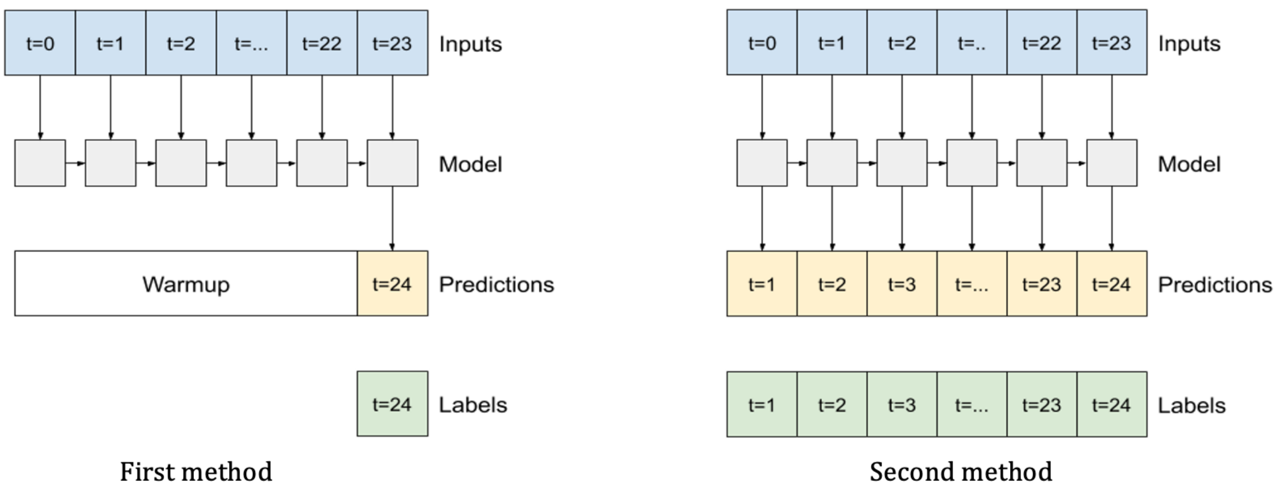
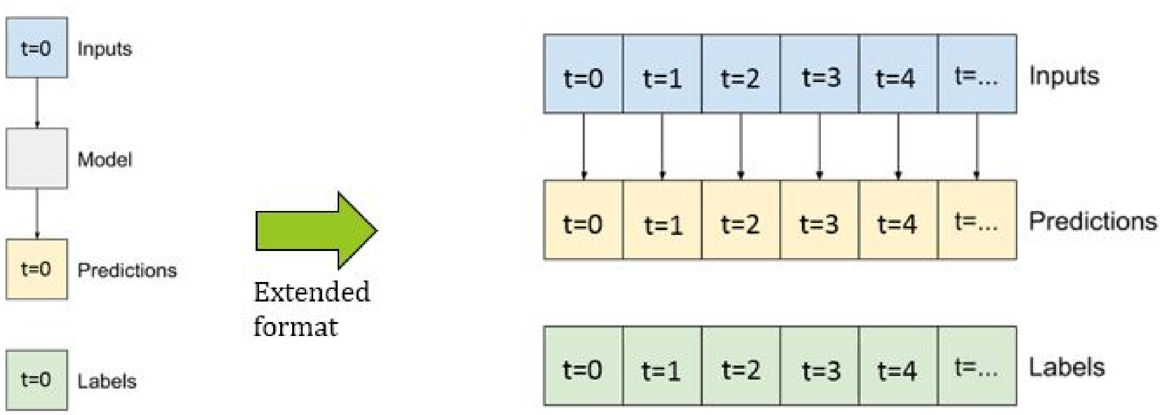


Figure 4: Prediction structure in ML models



After making the time sequences, it is time to train our Deep Learning models. Hence, the comprehensive hyper-parameter tuning was conducted on all DL models; the results are presented in Table 2. It should be mentioned that the GridSearchCV function, along with the set of hyper-parameters, are utilized for all ML models.

Table 2: Hyperparameters for Each Model

Model	**Hidden Dimension**	**Optimizer**	**Learning Rate**
RNN	64	Adam	1e-4
LSTM	64	Adam	1e-5
GRU	64	AdamW	1e-4
Stacked RNN	64	Adam	1e-5
Stacked LSTM	64	AdamW	1e-5
Stacked GRU	64	AdamW	1e-5

Figures 5a, 5b, and 5c show the prediction results from the ML models, DL models, and stacked DL models, respectively. Based on the ML results, the prediction goes from bad to worst as time elapses, and they can not capture ups and downs perfectly. In contrast, DL models do it somewhat well because DL models utilize the historical time sequence as an input instead of just features of one time step used in ML models.

Table 3 indicates error comparison of all models in this study. It is obvious that all DL models outperform all ML models. Also, by comparing DL and stacked DL methods, it can be concluded that stacking improves the results for each DL model. A further finding is that GRUs' error is lower than LSTM in the test set. However, RNN outperforms both GRUs and LSTM based on test set results.

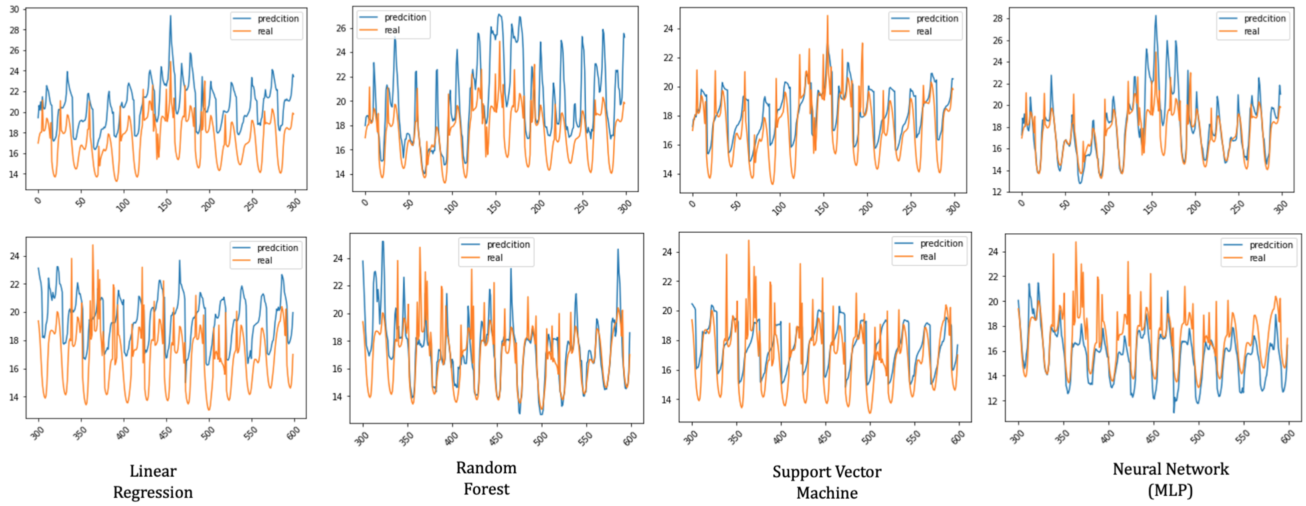
Table 3: Error Comparison

Error	**LR**	**RF**	**SVR**	**MLP**	**RNN**	**LSTM**	**GRU**	**Stacked RNN**	**Stacked LSTM**	**Stacked GRU**
MAE	2.82	1.65	1.18	1.87	0.98	1.15	1.07	0.93	1.09	1.08
MSE	9.73	5.35	2.53	5.64	1.79	2.28	2.12	1.74	2.17	2.07

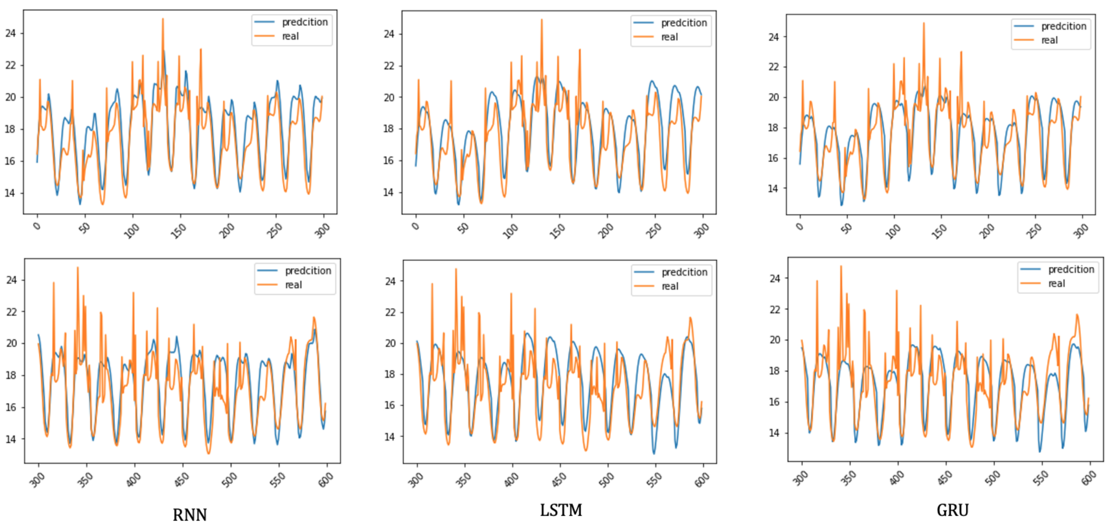
5 Conclusion

An accurate estimate of market demand is a key input into a generating company's decision-making operations for producing energy. As machine learning approaches have shown to be better than statistical methods due to their ability to capture nonlinear features and rapid changes, we aim to address the electricity load prediction issue by utilizing hourly real time electricity prices in the state of New York from January 1, 2019, 00:00 to December 31, 2029, 23:00 to evaluate the effectiveness of advanced deep learning approaches for electricity load prediction. This is accomplished by first employing the most common statistical and machine learning methods such as linear regression, random forest, SVR, and neural network, and then progressing to the advanced deep learning methods including RNN, LSTM, GRUs, and the associated stacked models. The results show that stacked RNN is the best deep learning model due to its lowest MSE and MAE, followed by RNN, and stacked LSTM. All the deep learning models, however, outperform the used statistical and machine learning methods.

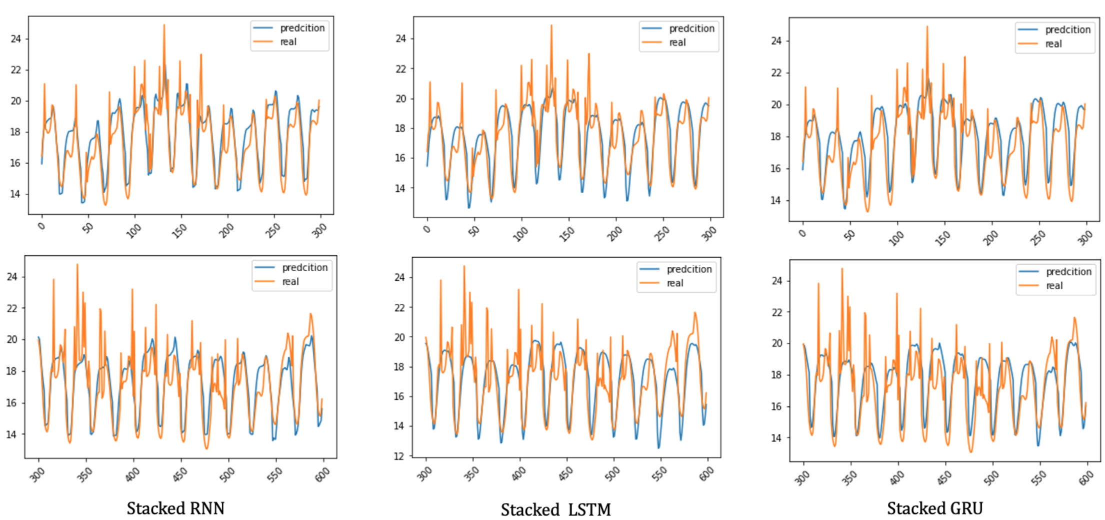
Figure 5: Prediction Results



(a) Prediction Results of ML models



(b) Prediction Results of DL models



(c) Prediction Results of DL stacked models

The vertical axis illustrates the electricity load, and the horizontal axis show hours.

References

[1] J. E. Boylan and A. A. Syntetos. Forecasting in management science. *Omega*, 6(40):681, 2012.

[2] D. W. Bunn. Modelling prices in competitive electricity markets. 2004.

[3] J. P. d. S. Catalão, S. J. P. S. Mariano, V. Mendes, and L. Ferreira. Short-term electricity prices forecasting in a competitive market: A neural network approach. *electric power systems research*, 77(10):1297–1304, 2007.

[4] S.-C. Chan, K. M. Tsui, H. Wu, Y. Hou, Y.-C. Wu, and F. F. Wu. Load/price forecasting and managing demand response for smart grids: Methodologies and challenges. *IEEE signal processing magazine*, 29(5):68–85, 2012.

[5] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*, 2014.

[6] F.-L. Chu. Forecasting tourism demand with ARMA-based methods. *Tourism Management*, 30(5):740–751, 2009.

[7] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE transactions on power systems*, 20(2):1035–1042, 2005.

[8] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3):1014–1020, 2003.

[9] M. Eichler and D. Türk. Fitting semiparametric markov regime-switching models to electricity spot prices. *Energy Economics*, 36:614–624, 2013.

[10] R. C. Garcia, J. Contreras, M. Van Akkeren, and J. B. C. Garcia. A GARCH forecasting model to predict day-ahead electricity prices. *IEEE transactions on power systems*, 20(2):867–874, 2005.

[11] H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3): 1225–1261, 2006.

[12] H. Ghoddusi, G. G. Creamer, and N. Rafizadeh. Machine learning in energy economics and finance: A review. *Energy Economics*, 81:709–727, 2019.

[13] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[14] S. Howison and M. Coulon. Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices. *The Journal of Energy Markets*, 2:29–69, 2009.

[15] S. Islyaev and P. Date. Electricity futures price models: Calibration and forecasting. *European Journal of Operational Research*, 247(1):144–154, 2015.

[16] P. Jiang, H. Yang, and X. Ma. Coal production and consumption analysis, and forecasting of related carbon emission: evidence from China. *Carbon Management*, 10(2):189–208, 2019.

[17] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang. Short-term residential load forecasting based on lstm recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1):841–851, 2017.

[18] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan. A review on the forecasting of wind speed and generated power. *Renewable and sustainable energy reviews*, 13(4):915–920, 2009.

[19] P. Mandal, T. Senjyu, and T. Funabashi. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy conversion and management*, 47(15-16):2128–2142, 2006.

[20] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola. Forecasting next-day electricity prices by time series models. *IEEE Transactions on power systems*, 17(2):342–348, 2002.

[21] H. Nyberg and P. Saikkonen. Forecasting with a noncausal VAR model. *Computational statistics & data analysis*, 76: 536–555, 2014.

[22] D. Önkal, K. Z. Sayım, and M. Lawrence. Wisdom of group forecasts: Does role-playing play a role? *Omega*, 40(6): 693–702, 2012.

[23] P. Ramos, N. Santos, and R. Rebelo. Performance of state space and ARIMA models for consumer retail sales forecasting. *Robotics and computer-integrated manufacturing*, 34:151–163, 2015.

[24] M. Sajjad, Z. A. Khan, A. Ullah, T. Hussain, W. Ullah, M. Y. Lee, and S. W. Baik. A novel cnn-gru-based hybrid approach for short-term residential load forecasting. *Ieee Access*, 8:143759–143768, 2020.

[25] M. Shafie-Khah and J. P. Catalão. A stochastic multi-layer agent-based model to study electricity market participants behavior. *IEEE Transactions on Power Systems*, 30(2):867–881, 2014.

[26] N. A. Shrivastava and B. K. Panigrahi. A hybrid wavelet-ELM based short term price forecasting for electricity markets. *International Journal of Electrical Power & Energy Systems*, 55:41–50, 2014.

[27] N. Singh, S. R. Mohanty, and R. D. Shukla. Short term electricity price forecast based on environmentally adapted generalized neuron. *Energy*, 125:127–139, 2017.

[28] B. Szkuta, L. A. Sanabria, and T. S. Dillon. Electricity price short-term forecasting using artificial neural networks. *IEEE transactions on power systems*, 14(3):851–857, 1999.

[29] H. Takeda, Y. Tamura, and S. Sato. Using the ensemble Kalman filter for electricity load forecasting and analysis. *Energy*, 104:184–198, 2016.

[30] F. U. M. Ullah, A. Ullah, I. U. Haq, S. Rho, and S. W. Baik. Short-term prediction of residential power energy consumption via cnn and multi-layer bi-directional lstm networks. *IEEE Access*, 8:123369–123380, 2019.

[31] R. Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1081, 2014.

[32] W. Yang, J. Wang, T. Niu, and P. Du. A novel system for multi-step electricity price forecasting for electricity market management. *Applied Soft Computing*, 88:106029, 2020.

[33] Z. Yang, L. Ce, and L. Lian. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Applied Energy*, 190:291–305, 2017.

[34] J. Zhang and C. Cheng. Day-ahead electricity price forecasting using artificial intelligence. In *2008 IEEE Canada Electric Power Conference*, pages 1–5. IEEE, 2008.

[35] F. Ziel and R. Steinert. Electricity price forecasting using sale and purchase curves: The x-model. *Energy Economics*, 59:435–454, 2016.

[36] F. Ziel and R. Weron. Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, 70:396–420, 2018.