# Electricity Price Forecasting using Convolution and LSTM Models

Dhruv Aditya Mittal*, Shaowu Liu* and Guandong Xu*
*Advanced Analytics Institute, University of Technology Sydney.
{dhruv.mittal, shaowu.liu, guandong.xu}@uts.edu.au

***Abstract*—Electricity Market uses Demand and Supply chain strategy. Also, it is prone to random fluctuations that directly impact profit. Therefore forecasting demand becomes very important to mitigate the consequences of price dynamics. This paper proposes a Deep Learning model using Long Short Term Memory (LSTM) and Convolution Neural Network to forecast future electricity prices on the Australian electricity market and compares them with other state of the art models. We have selected evaluation metrics to prove that our model outperforms the other existing models for electricity price prediction.**

***Index Terms*—Electricity Price Forecasting, LSTM, Convolution, Neural Networks**

## I. Introduction

Electricity is one of the most vital part of our everyday's life. With the emergence of Renewable Energy sources, distributed energy sources, and deregulation, the electricity market is becoming more sophisticated and more unpredictable every day. Many countries now use Demand and Supply (DS) strategy to supply electricity to the consumers through a deregulated market instead of using conventional power systems where the production of electricity is provided by enormous centralized power plants [7].

For efficiently implementing the DS strategy, predicting future price values is extremely important for the electricity suppliers in the market. Accurate predictions play an important role in power system planning and operation, risk assessment, and another decision-making process. According to [13], even a 1% improvement in Mean Absolute Percentage Error (MAPE) can help reduce the cost to about 0.1–0.35% for short term price forecasting, which is approximately 1.5 million dollars/year for a medium-size utility with a 5GW peak load [11]. The principle target of an electricity market is to minimize the cost of electricity through competition and maximize the efficient generation and utilization of electricity [8]. Real-Time electricity price values show random patterns and fluctuations where the price suddenly peaks or drops from its normal average value. For example, the price value can sometimes become 10 times the average value and sometimes even go negative or zero. Also, electricity price differs from other assets and commodities because of its unique features like non-storability, oligopolistic generation, and the obligation of having a constant balance between the supply and demand sides [5]. Therefore, forecasting the price of electricity accurately has become a major challenge to this industry and has gained the attention of many researchers worldwide. Figure 1 shows the factors affecting electricity prices in the real-world scenario.

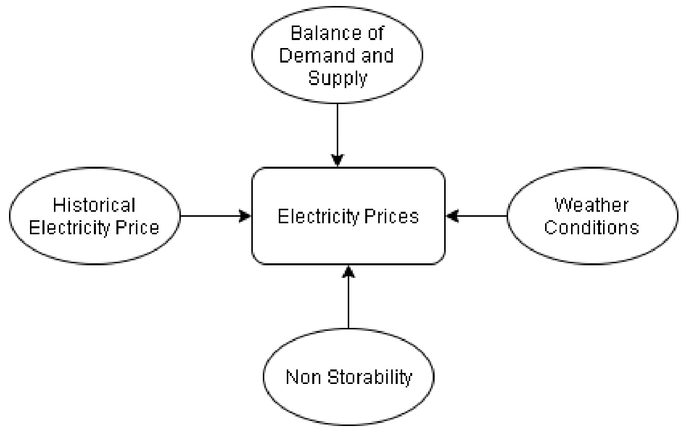


Fig. 1. Factors affecting the electricity prices

Jakaša et al. [4] has used the famous Autoregressive Integrated Moving Average (ARIMA) model for day ahead forecasting using 10 years of European Energy Exchange data. The ARIMA model cannot be used for every time series prediction problem because of its statistical linear properties. It requires the data to be stationary i.e. to have constant mean and variance throughout its range, which is difficult to obtain in the real-world scenario, because random fluctuations occur anytime and there are rapid variations and high-frequency changes in the price. Even if the data is made stationery by using some pre-processing techniques, the random peaks could not be easily forecasted. Also, the selection of the perfect ARIMA parameters is difficult and requires high computation costs [2]. The same limitation applies to other statistical models like Moving Average, AR, ARMA etc, so we only took ARIMA model for our baseline comparison. Dudek & G. [2] used a random forest regressor for the short-term load forecasting of power generation. Chikkakrishna et al. [1] used SARIMA (Seasonal ARIMA) and PropehtFb for short-term traffic predictions on hourly traffic data in Tamil Nadu (India). Ranjbar et al. [10] used four-layered Artificial Neural Networks (ANN) for Electricity Price Forecasting and evaluated the model by changing the size of hidden layers in the Neural Network Architecture and then picking up the best one by using MAPE (Mean Absolute Percentage Error) as the evaluation metrics. Ranjbar et al. [10] used ANN because of its ability to learn complex and non-linear features using previous historical data. Deep learning methods, which have been used in many learning tasks (e.g., misinformation detection [3], linguistics [9], and portfolio optimization [12]), are also used for electricity price forecasting. Jiang et al. [6] used Long Short-Term Memory (LSTM) for the day ahead forecasting of Electricity Prices in the Australian market at Victoria (VIC) region and Singapore market. LSTMs can retain previous information for making predictions on current information, which makes them a suitable candidate for the time series prediction task. Khan et al. [11] uses convolution network and compares the performance with the multilayered perceptron model.

The key contributions mentioned in this paper are as follows: In this paper, we have tried to purpose a hybrid model using convolution, LSTM, and multilayered perceptron layers that could be used to forecast electricity prices in the Australian electricity market. The accuracy and efficiency of the model are compared with other popular deep learning and statistical models using three evaluation parameters. The evaluation metrics used in our analysis are RMSE (Root Mean Squared Error) (Equation (3)), which is used to calculate the standard deviation of the predicted errors, MSE (Mean Squared Error) (Equation (2)) is used to calculate squared deviation and MAPE (Mean Absolute Percentage Error) (Equation (4)) is used to calculate mean deviation percentage.

The organization the paper is as follows: Introduction is given in Section I. Methodology is discussed in Section II followed by experimental results and analysis in Section III. Section IV summarizes the paper with concluding remarks.

## II. Methodology

### A. Data collection and processing

The data set used for our task contains price data from 1st January 2000 to 22nd April 2020 (Approximately 20 years) with a half-hourly frequency rate in the NSW (New South Wales) region of Australia. The flow diagram of the proposed model is shown in Figure 4. We used 20 years of previous data for training rather than 1-2 years, so that the model can make predictions based on the learning from various different scenarios and generalize well on testing dataset. The data input used for Deep Learning Methods is the Rolling window data set in which past values are used as input and the current value is considered as an output. $WindowSize$ is the number of past values in the input for one output. We have also removed the outliers from the data. The outliers are removed to get rid of unnecessary noise and random peaks in the dataset, which helps the model to learn more efficiently. The data points with a price value of less than or equal to zero are only $0.013\%$ of the entire dataset. We have removed data points having a price value greater than 200 as there are only $0.1\%$ data points in that range. The mean value of the price is $48.81$ AUD (Australian Dollars). Figure 2 shows the daily re-sampled and Figure 3 shows monthly re-sampled price data after removing the outliers. Data distribution for training, validation, and testing is taken to be $60\%$, $20\%$, and $20\%$ respectively.

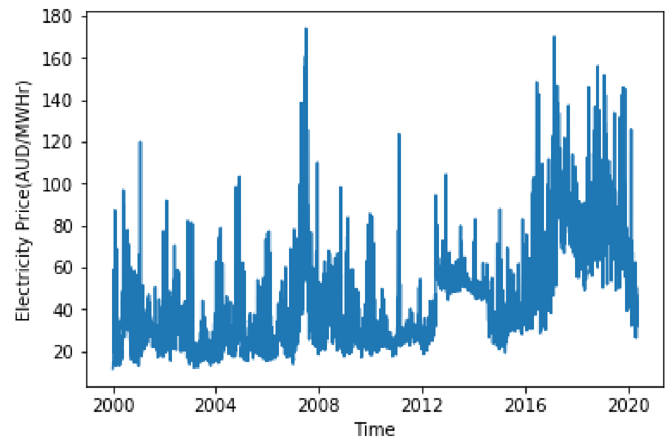


Fig. 2. Daily Resampled Data After Pre-Processing

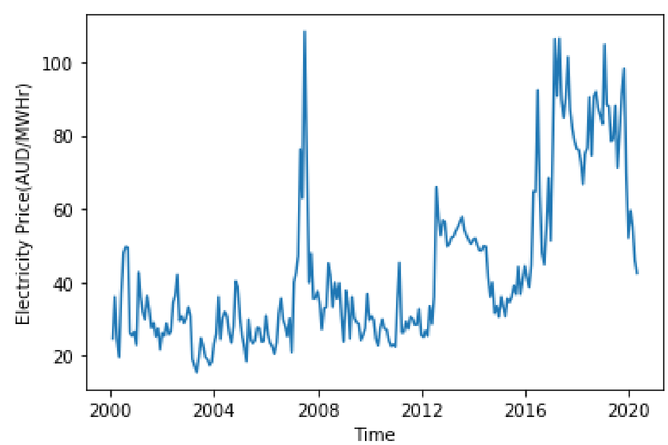


Fig. 3. Monthly Resampled Data After Pre-Processing

### B. Proposed Model

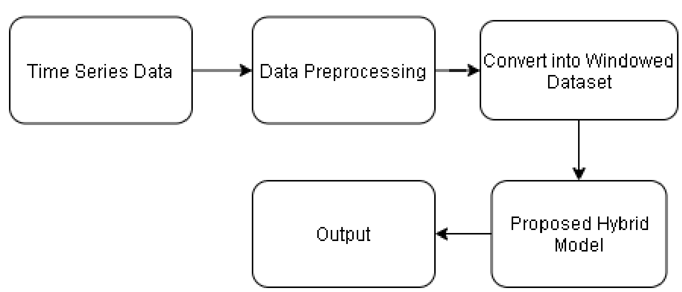


Fig. 4. Flow Chart

In this section, the implementation of the proposed model is discussed. Figure 5 shows the hybrid model architecture used for our task. The best hyperparameters are selected by executing the model on different sets of hyperparameters and then choosing the one with least MSE (Equation (2)). The $WindowSize$ is taken as 10 and the total number of learn-able parameters in the proposed model is 5173. The first layer of the proposed network is the convolution1D layer. Convolution with a 1-D Dimension is selected because the time series data is also of the same dimension and allows us to use larger filter sizes. The number of filters and kernel size of the convolution layer are both taken as 3. Convolution neural network consists of multiple hidden layers and applies filters to data for better performance and data mapping. Convolutions can extract deep features that are independent of time. They are highly noise-resistant models and are able to extract features and create informative representations of time series automatically. To apply the kernel filters to the matrices and assign them weights, ReLu (Rectified Linear Unit) activation function is used in the first layer.

The output of the first convolution layer is then passed into the LSTM layer. LSTM's have the capability to learn information from historical data and use it for later predictions. The Number of LSTM units used in our model are: $WindowSize + 1$, we took the $WindowSize$ to be 10 so number of units become 11. A single LSTM cell consists of 3 gates: input gate, output gate, and forget gate. The basic description of these gates is as follows.

1) Input Gate: It controls the input flow entering into the LSTM Cell.
2) Output Gate: It controls how the value in the cell is used to compute the output activation of the LSTM cell.
3) Forget Gate: It decides whether the specific information should remain in the LSTM cell.

The output from the LSTM layer is then flattened so that the information can pass through the multi perceptron layers. We have added two perceptron layers having 50 nodes and 1 node respectively. The activation used for the perceptron layers is 'Tanh' activation function (Equation (1)) as it provides more steeper derivatives and more range for the output $(-1, 1)$ as compared to Sigmoid $(0, 1)$ activation function. The output of the model is then scaled up from the range $(-1, 1)$ to its original range and is then evaluated using evaluation metrics provided in Section III.

$$a = \frac{e^z}{e^z + e^{-z}} \tag{1}$$

## III. Experimental Results and Analysis

The experiment was conducted on the Australian Electricity Market Price data set using Deep Learning and Statistical Models. The Deep Learning models selected for the comparison are simple convolution neural network model with perceptrons, simple LSTM model with perceptrons, simple convolution neural Network with simple LSTM model, and perceptrons (Proposed Model). The statistical models selected

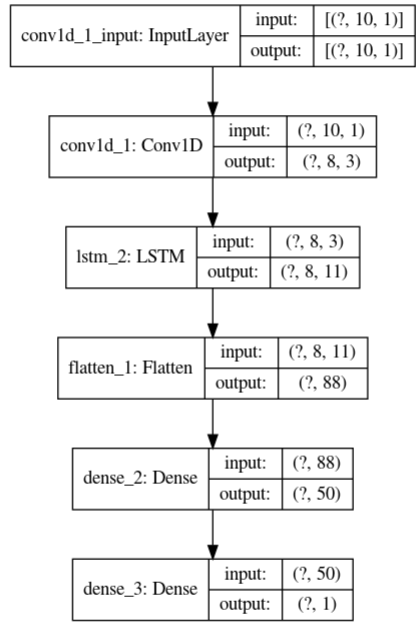


Fig. 5. Model Architecture

are ARIMA and Prophet models. The data set used for the experimental analysis contains approximately 35K data points from the year 2000 to 2020. Data pre-processing and data distribution are mentioned in Section II. Deep Learning Models are trained by using Adam optimizer and it uses MSE (Equation (2)) as the loss function. The training was done for 100 epochs for each of the models. Figure 6 shows the graphical analysis of forecasts made by our proposed model. It could be observed that the model is able to map the random behavior of the time series and gives reliable forecasts.

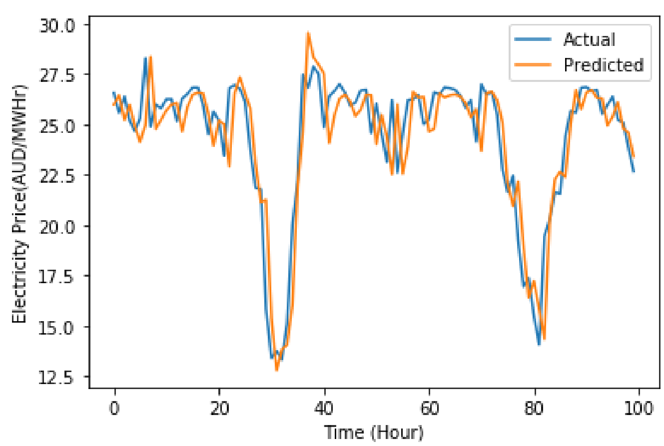


Fig. 6. Forecast vs. Actual Price using Proposed Model

ARIMA and Prophet models were fitted using the pre-processed time series data without converting it into a windowed dataset. From the experiments and analysis of the results, it was observed that these models fail to model the randomness in the time series data. These models work well when there is a certain observable pattern in the time series data without any random fluctuations. Figure 7 shows the forecasts made by the prophet model. It could be observed that the model is not able to forecast the random fluctuations happening in the time-series data. Simple Convolution and LSTM models were able to forecast the random fluctuation but their accuracy is less than the proposed model.

For the performance evaluation we have selected three evaluation metrics: MSE (Equation (2)), RMSE (Equation (3)) and MAPE (Equation (4)).

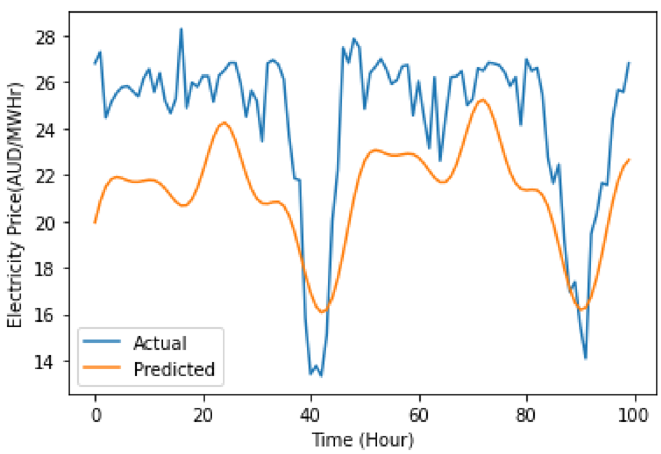


Fig. 7. Forecast vs Actual Price using Prophet Model

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{3}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right| \tag{4}$$

| Model | MSE | RMSE | MAPE |
|---|---|---|---|
| ARIMA | 290.96 | 17.05 | 12.89 |
| Conv1D with Perceptron | 264.85 | 16.274 | 9.76 |
| LSTM with Perceptron | 234.48 | 15.313 | 10.01 |
| Conv1D with LSTM and Perceptron | 191.85 | 13.881 | 8.9 |
| Prophet | 270.86 | 16.45 | 10.90 |

TABLE I
PERFORMANCE EVALUATION

It could be noted from Table I that the proposed model with both Convolution and LSTM layers outperforms other deep learning and statistical models mentioned above in the Section III as its error is lowest in each of the evaluation metrics used. The worst performing models on our data were Prophet and ARIMA models. However, these models perform well when we try to forecast for large forecasting intervals. However, for short-term forecasting (2-3 hours) deep learning methods show more promising results.

## IV. Conclusion

Recent changes in the electricity market have made forecasting future prices very difficult. Getting good prediction accuracy is directly proportional to the profit made. This paper provides comparative analysis of the performance of different models. Data pre-processing techniques are introduced to have better performance. The proposed model outperforms other existing models for short term price forecasting as it can fit in the random behavior of the time series data.

## V. Acknowledgement


The authors would like to thank the Australian Energy Market Operator for proving the dataset for our research.


## References


[1] Naveen Kumar Chikkakrishna, Chitirala Hardik, Kancherla Deepika, and Narendula Sparsha. Short-term traffic prediction using sarima and fbprophet. In *2019 IEEE 16th India Council International Conference (INDICON)*, pages 1–4. IEEE, 2019.

[2] Grzegorz Dudek. Short-term load forecasting using random forests. In *Intelligent Systems' 2014*, pages 821–828. Springer, 2015.

[3] Md Rafiqul Islam, Shaowu Liu, Xianzhi Wang, and Guandong Xu. Deep learning for misinformation detection on online social networks: a survey and new perspectives. *Social Network Analysis and Mining*, 10(1):1–20, 2020.

[4] Tina Jakaša, Ivan Andročec, and Petar Sprčić. Electricity price forecasting—arima model approach. In *2011 8th International Conference on the European Energy Market (EEM)*, pages 222–225. IEEE, 2011.

[5] LianLian Jiang and Guoqiang Hu. Day-ahead price forecasting for electricity market using long-short term memory recurrent neural network. In *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 949–954. IEEE, 2018.

[6] Zahoor Ali Khan, Sahiba Fareed, Mubbashra Anwar, Afrah Naeem, Hira Gul, Arooj Arif, and Nadeem Javaid. Short term electricity price forecasting through convolutional neural network (cnn). In *Workshops of the International Conference on Advanced Information Networking and Applications*, pages 1181–1188. Springer, 2020.

[7] Dimitrije Kotur and Mileta Žarković. Neural network models for electricity prices and loads short and long-term prediction. In *2016 4th International Symposium on Environmental Friendly Energies and Applications (EFEA)*, pages 1–5. IEEE, 2016.

[8] Niranjan Kumar et al. Market clearing price prediction using ann in indian electricity markets. In *2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS)*, pages 454–458. IEEE, 2016.

[9] Qian Li, Ziwei Li, Jin-Mao Wei, Yanhui Gu, Adam Jatowt, and Zhenglu Yang. A multi-attention based neural network with external knowledge for story ending predicting task. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1754–1762, 2018.

[10] M Ranjbar, S Soleymani, N Sadati, and AM Ranjbar. Electricity price forecasting using artificial neural network. In *2006 International Conference on Power Electronic, Drives and Energy Systems*, pages 1–5. IEEE, 2006.

[11] Umut Ugurlu, Oktay Tas, and Umut Gunduz. Performance of electricity price forecasting models: Evidence from turkey. *Emerging Markets Finance and Trade*, 54(8):1720–1739, 2018.

[12] Nhi NY Vo, Xuezhong He, Shaowu Liu, and Guandong Xu. Deep learning for decision making and the optimization of socially responsible investments and portfolio. *Decision Support Systems*, 124:113097, 2019.

[13] Hamidreza Zareipour, Claudio A Canizares, and Kankar Bhattacharya. Economic impact of electricity market price forecasting errors: A demand-side analysis. *IEEE Transactions on Power Systems*, 25(1):254–262, 2009.