Intelligent Demand Forecasting Using Deep Learning

Lithicka Anandavel[1], Ansh Sharma[1], Naveenkumar S.[2], Suresh Sankaranarayanan[2*], Anis Salwa Mohd Khairuddin[3*]

[1]Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

[2]Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

[3]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Malaysia

[*]Corresponding Author
doi: https://doi.org/10.21467/proceedings.141.7

ABSTRACT

One type of energy demand is the electricity demand, which measures the electricity consumption Wh (watt-hour). Forecasting this electricity demand is very crucial and plays a fundamental role in the electrical industry, as it provides the basis for making decisions in the operation and planning procedures of power systems. Forecasting is important for development experts and are of great interest to energy authorities, power utilities, and private investors. Inaccurate projections can have disastrous social and economic implications, whether they over-or under-predict demand. Supply shortages and forced power outages occur from underestimating demand, wreaking havoc on productivity and economic growth. Overestimating demand can result in overinvestment in generation capacity, financial hardship, and, eventually, higher power costs. This paper has validated several methodologies such as ARIMA, XGBOOST, LSTM and Bi-LSTM towards forecasting the energy demand for different regions of Australia during different season. The models were validated towards energy demand forecasting in terms of error and accuracy resulting in LSTM with 2 layers outperforming the other models.

Keywords: Electricity demand, Short-term forecast, Machine Learning.

1 Introduction

Forecasting the demand for energy is essential for power generation and planning. Accurate power demand forecasts provide a better knowledge of the electricity network development and generation required to meet future demand in a sustainable manner. It is crucial to forecast energy consumption and peak electricity demand. For the reliable operation of power systems, accurate information on electricity demand is necessary. There has been good amount of work done in forecasting the energy demand using standard machine learning and deep learning technique which is evident from the literature.

Researchers in [1] have forecasted demand using the Box–Jenkins time series technique, particularly the ARIMA. So, towards achieving their objective, historical demand data from January 2010 to December 2015 were used. Multiple "ARIMA" models were built and evaluated using four performance criteria: the "Akaike criterion (AIC)", the "Schwarz Bayesian criterion (SBC)", the "maximum likelihood criterion", and the standard error criterion. ARIMA is the chosen as the best model that minimises the four preceding criteria (1, 0, 1).

There has also been work [2] which incorporates all of the most recent statistical techniques employed in the fields of demand forecasting. In the findings, it has been shown that the ensemble of the results of the

Proceedings DOI: 10.21467/proceedings.141; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-957605-4-1

timeseries model and the regression-based model produces a superior result by eliminating over- and under-forecasting and bringing the forecast values closer to the actual values.

Research in [3] worked towards predicting the peak electricity demand by developing a time series model suitably. This study considered the "Naïve", "Seasonal Holt-Winters", "ARMA", "ARAR" algorithm, and Regression with "ARMA" Errors time series models. Using the forecasting accuracy criteria of "Mean Absolute Error (MAE)", "Root Mean Square Error (RMSE)", and "Mean Absolute Relative Percent Error", the performance of these various models was examined "(MARPE)".

There has been work reported in [4] which have worked on developing three demand forecasting for electricity based on parallel implementation of the "adaptive Holt Exponential smoothing" approach. The paper has taken French half-hourly electricity demand data for validating their effectiveness with an average MAPE of 0.491% proving the efficacy of the third model.

Electrical power Load forecasting have been worked in [5] by analysing the pattern of electrical load and predicting the future of electrical load demand for short, medium, and long term. The authors here have applied time series methodology and forecasting was conducted for daily electrical power load based on Kuwait electricity network for three years.

Research has been done in [6] on evaluating the "recurrent neural networks (RNN)" based on three recurrent units which are (1) a traditional "tanh" unit (2) a "long short-term memory (LSTM)" and (3) "Gated Recurrent Unit (GRU)". The evaluation was performed on dataset pertaining to polyphonic music data and raw speech signal data. From the validation it has been shown that both "GRU" and "LSTM" outperformed the "tanh" unit.

There has been work reported in [7] towards behavioural analysis of "Bi-LSTM" and "LSTM" model towards time series forecasting. From the analysis it was proved that the Bi-LSTM outperformed ARIMA and LSTM in terms of prediction. In addition, Bi-LSTM models reached equilibrium much slower than LSTM model.

Research has been done in [8] towards electricity demand prediction reliably using "Adaptive Neuro Fuzzy Inference System (ANFIS)". The focus of the work was considered towards electrical load forecasting for medium term in a "Greater Accra" Region of "Mallam" town which is densely populated. From the validation, it was proved that "ANFIS" prediction model is dependent on quality of data and tuning parameters.

Lastly work has been done [9] on yearly electricity demand forecast of Bangladesh using "multivariate time series". The authors in this work have introduced two exogenous variables which are Population and "GDP". "Autoregressive Integrated Moving Average" with Exogenous which is "ARIMAX (0,1,1)" that show comparatively better performance than the state of art model with the lowest "Alkaline Information Criterion (AIC)" values.

Based on the literature reviewed, it is clear that there have been several works done pertaining to time series models like ARIMA and deep learning like LSTM, Bi-LSTM for forecasting of energy demand. This work focuses on forecasting the demand of energy of different states of Australia for four different seasons using machine learning and deep learning methodologies. This work will provide a clear indication on how much energy demand is for different regions for four seasons to assist the utility to make wise decision in producing energy supply accordingly. The rest of paper is organised as follows. Section 2 gives the details on time series, machine learning and deep learning models used. Section 3 gives details on results and analysis of energy demand forecasting using different models. Section 4 is the concluding remarks of paper with future work.

2 Materials and Methods

We in this work have deployed different models pertaining to ARIMA, XGboost, Deep learning models like LSTM, Bi-LSTM. The details of model implemented are explained in the forthcoming sections

2.1 Time Series Method

Time series are comprised of three components: trend, seasonality, and irregular components, often known as residuals [10]. Traditionally, ARIMA models are generalised by Box and Jenkins which are used for forecasting time series [11]. ARIMA has been the standard technique for forecasting time series for many years. The success of ARIMA models has spawned significant study in time series analysis. Even though the ARIMA has been a success, it has some significant shortcomings [12]. Nonlinear interactions between variables are challenging to model in simple ARIMA models, for instance. An “Autoregressive Integrated Moving Average” model is a regression analysis method where the strength of one dependent variable in relation to another variable is measured. The parameters used in “ARIMA” are standard notation which are integer values as follows:

“p”: lag difference of model also called as lag order.

“d”: difference between the raw observation number of times; referred as degree of differencing

“q”: the moving average window size and also called as moving average order.

2.2 XGboost

“XGBoost” is a “gradient boosting” approach that is based on decision-tree-based ensemble Machine Learning algorithm. “XGBoost” stands for “Extreme Gradient Boosting” where parallel tree boosting is enabled and is one of the top machine learning models for regression, classification, and ranking problem. The main intention behind gradient boosting is towards increasing the performance of machine learning model and computation speed. “XGBoost” is responsible for building trees in parallel manner as opposed to sequential manner like “GBDT (Gradient-boosted decision trees)”. Level wise technique is employed in “XGBoost” also possess lot of methodologies for model optimisation, computer environment changes and algorithm improvisation. The algorithm also employs lot of base learners towards constructing an additive model.

2.3 Long Short-Term Memory (LSTM)

The “Long Short-Term Memory Network (LSTMN)” is an augmented “RNN (sequential network)” for storing the information indefinitely. Classical “RNNs” are responsible for tracking long term dependencies input sequences arbitrarily. When training RNN using back-propagation, the long-term gradients are “vanishing” due to computation involved in training process. In terms of Vanilla-RNN, the backpropagation employed for training can vanish or explode. These problems are solved with “LSTM”. In “LSTM”, gradients can also flow freely. LSTMs still suffer from exploding gradient issue.

2.4 Bi-LSTM

The Bidirectional LSTM or biLSTM is an improved or enhanced form of LSTM. Information travels from backward to forward in unidirectional LSTMs, but bi-directional LSTMs use hidden states to transmit

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

information from backward to forward and forward to backward. It is a sequence-processing model composed of two sequences: one that takes input in the forward direction, and the other that takes it in the backward direction. To determine the output y at time t, both forward and backward activations would be evaluated.

3 Results and Discussion

The data for this study was obtained from the "Australian Energy Market Operator (AEMO). AEMO" is an Australian independent organization that was established by the "Council of Australian Governments (COAG)" and whose ownership is shared between government and industry. Figure 1 shows the dataset of the New South Wales region. The authors extracted historical 30-minute Energy Demand time-series data from "AEMO (Australian Energy Market Operator)" for the state of NSW from January 2015 to September 2021. This data was split into training and test, the last three months were used as test data while the rest of the data was used for training. The data from 2015 through September 2021 was used to train the model, and the remaining three months of 2021 were forecasted.

SETTLEMENTDATE	TOTALDEMAND
01-01-2015 0:30	6920.94
01-01-2015 1:00	6668.71
01-01-2015 1:30	6327.71
01-01-2015 2:00	6056.02
01-01-2015 2:30	5827.97
01-01-2015 3:00	5711.33
01-01-2015 3:30	5653.85
01-01-2015 4:00	5625.27
01-01-2015 4:30	5618.95
01-01-2015 5:00	5580.51
01-01-2015 5:30	5589.51
01-01-2015 6:00	5646.63
01-01-2015 6:30	5806.57
01-01-2015 7:00	5998.21
01-01-2015 7:30	6221.81
01-01-2015 8:00	6361.85

Figure 1: *Demand Data Set of NSW Region*

A statistical analysis was used to interpret the pattern and trend in the time series data. Auto Regressors, specifically SARIMA, is the statistical model that has been employed. So, towards evaluating the nature, economics and other time varying systems, "Autoregressive" models are used. The reason being that they operate on the assumption that previous values have an effect on current values. The result obtained from these two models are provided in the table 1.

Table 1: *SARIMA performance*

MSE	RMSE	ACCURACY	MAPE
256664.1992	506.6203699	3.478%	6.0409

However, because autoregressive models rely solely on historical data to predict future prices, they implicitly presume that the fundamental forces that drove past prices would remain constant over time. If the underlying dynamics in question are shifting, such as if an industry is undergoing rapid and unprecedented technological transition, this might lead to unexpected and wrong projections.

Consequently, XGBoost is applied towards improving the modelling results and efficiency with its lookback feature. For this case, we used the xgbregressor method so that the output is continuous. The learning rate was set to 0.1, max depth as 5, the number of trees were set to 1000 using the n_estimator parameter and finally n_jobs parameter was used to set the number of cores as 32. Five evaluation metrics, RMSE, MSE, MAE, MAPE and R2 score, were used to analyze the model's performance. Table 2 shows the result obtained from this model.

Table 2: *XGBoost performance*

MSE	RMSE	MAE	MAPE	R2 SCORE
691039.335232	831.287757	24.818078	0.297813	0.210223

In terms of Deep learning, these models (RNN, LSTM, and GRU) including BiLSTM were built and trained using Tensorflow Keras. The parameters that were tuned are the learning rate, number of layers and the number of units in the hidden layer. The dropout value was fixed as 0.5, the batch size was also fixed to 32 and a call back function was used for the epochs used to train the model. Similar to XGBoost, the accuracy of each of these models were examined using five evaluation metrics. Table 3 shows the MAPE values derived by every model for each evaluation metric. Lr denotes learning rate, layer is the number of hidden layers, and the unit represents the number of units in each hidden layer.

Table 3: *MAPE for different Deep learning models*

		MAPE		
	RNN	LSTM	BILSTM	GRU
lr=0.001, layer=1, unit=32	0.009	0.009	0.010	0.008
lr=0.001, layer=1, unit=64	0.009	0.008	0.010	0.007
lr=0.01, layer=1, unit=32	0.078	0.068	0.009	0.036
lr=0.01, layer=1, unit=64	0.278	0.129	0.008	0.095
lr=0.01, layer=1, unit=128	0.132	0.064	0.056	0.016
lr=0.001, layer=2, unit=64,32	0.010	0.008	0.008	0.010
lr=0.01, layer=2, unit=64,32	0.133	0.012	0.017	0.047
lr=0.01, layer=2, unit=128,32	0.140	0.123	0.008	0.049
lr=0.01, layer=2, unit=128,64	0.246	0.011	0.042	0.042

From the above results remarkable performance was observed through deep learning-based approaches to the demand forecasting problem is due to the LSTM model with 2 layers (64 and 32) and learning rate

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

0.001. LSTMs provide a vast array of parameters, including learning rates and input and output biases. Therefore, no precise modifications minimize the difficulty to update each weight to O (1), comparable to Back Propagation Through Time (BPTT), which is advantageous. The proposed model is computed based on these hyperparameters over data divided into four seasons (Summer, Autumn, Winter, Spring). The result is tabulated below in table 4.

Table 4: *Seasonal performance*

	SUMMER	AUTUMN	WINTER	SPRING
RMSE	73.251618	111.633766	139.815398	140.483480
MAE	55.906736	76.412985	88.658871	92.090652
MAPE	0.009190	0.011652	0.012672	0.013260
R2 SCORE	0.993688	0.985781	0.982614	0.976602

4 Conclusions

In this study, various forecasting methods has been examined and compared to determine which method is more superior for predicting energy demand. Based on the above-mentioned findings, it was discovered that the LSTM model with two layers (64 and 32) and a learning rate of 0.001 outperformed the other models with a MAPE value of 0.008 and an R2 score of 0.995. In future, the improvement brought about by deep learning by optimising the algorithms and creating hybrid models in an effort to improve accuracy would be carried out.

5 Declarations

5.1 Study Limitations

All other regions were left out of this study; only New South Wales was included. This was done to make it simpler to evaluate the different approaches. It was believed that the remaining regions would behave similarly. Only seven years of data—from 2015 to 2021—were used to minimise the training period.

5.2 Acknowledgements

This research was done in collaboration with the Universiti Malaya, Malaysia as a part of SATU Joint Research Scheme and funded by Research Grant from Universiti Malaya (Project Number ST005-2021). In addition, we like to thank SRMIST for providing AWS Cloud support for finishing the project work.

5.3 Competing Interests

There have been no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

5.4 Publisher's Note

AIJR remains neutral with regard to jurisdiction claims in published maps and institutional affiliations.

References

[1] Fattah J, Ezzine L, Aman Z, El Moussami H, Lachhab A, "Forecasting of demand using ARIMA model", *International Journal of Engineering Business Management*. January 2018. doi: 10.1177/1847979018808673.

[2] Silaparasetti, Teja & Das Adhikari, Nimai & Domakonda, Nishanth & Garg, Rajat & Gupta, Gaurav, “An Intelligent Approach to Demand Forecasting”, *ICICT*. October, 2017.

[3] Razak, Fadhilah & Shitan, Mahendran & Hashim, Amir & Zainal Abidin, Izham, “Load Forecasting Using Time Series Models”, *Jurnal Kejuruteraan*. December, 2009. doi: 21. 10.17576/jkukm-2009-21-06.

[4] Laouafi, Abderrezak & Mordjaoui, Mourad & Dib, Djalel, “Very short-term electricity demand forecasting using adaptive exponential smoothing methods”, *15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*. December, 2014. doi: 10.1109/STA.2014.7086716.

[5] Eisa Almeshaiei, Hassan Soltan, “A methodology for Electric Power Load Forecasting”, *Alexandria Engineering Journal*, Volume 50, Pages 137-144, Issue 2. June, 2011. doi: doi.org/10.1016/j.aej.2011.01.015.

[6] Chung, Junyoung & Gulcehre, Caglar & Cho, KyungHyun & Bengio, Y, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*”, NIPS 2014 Workshop on Deep Learning*. December 2014. doi: doi.org/10.48550/arXiv.1412.3555

[7] S. Siami-Namini, N. Tavakoli and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," *2019 IEEE International Conference on Big Data (Big Data)*, pp. 3285-3292. February, 2020. doi: 10.1109/BigData47090.2019.9005997.

[8] Amevi Acakpovi, Alfred Tettey Ternor, Nana Yaw Asabere, Patrick Adjei, Abdul-Shakud Iddrisu, "Time Series Prediction of Electricity Demand Using Adaptive Neuro-Fuzzy Inference Systems", *Mathematical Problems in Engineering*, vol. 2020, Article ID 4181045, 14 pages. August, 2020. doi: doi.org/10.1155/2020/4181045

[9] F. Rabbi, S. U. Tareq, M. M. Islam, M. A. Chowdhury and M. Abul Kashem, "A Multivariate Time Series Approach for Forecasting of Electricity Demand in Bangladesh Using ARIMAX Model," *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1-5. December, 2020. doi: 10.1109/STI50764.2020.9350326.

[10] Shumway RH, Stoffer DS, “Time series analysis and its applications (with R examples)”, *Springer Texts in Statistics*. January, 2011. doi: doi.org/10.1007/978-3-319-52452-8

[11] G. E. P. Box and G. M. Jenkins, “Time Series Analysis: Forecasting and Control”, 2nd ed, *San Francisco: Holden-Day*, 1976.

[12] Earnest A, Chen MI, Ng D, Sin LY, “Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore*”, BMC Health Serv Res*. May, 2005. doi: 10.1186/1472-6963-5-36.

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922