Applied Computer Science, vol. 16, no. 1, pp. 5–17
doi:10.23743/acs-2020-01

Submitted: 2019-10-12
Revised: 2019-11-25
Accepted: 2020-03-02

Electricity Demand Forecasting, STLF,
Deep Learning Techniques, LSTM, CNN, MLP

Saheed A. ADEWUYI[*], *Segun AINA* [0000-0001-5080-1760][**],
Adeniran I. OLUWARANTI [0000-0003-4920-6053][**]

A DEEP LEARNING MODEL FOR ELECTRICITY DEMAND FORECASTING BASED ON A TROPICAL DATA

Abstract

Electricity demand forecasting is a term used for prediction of users' consumption on the grid ahead of actual demand. It is very important to all power stakeholders across levels. The power players employ electricity demand forecasting for sundry purposes. Moreover, the government's policy on its market deregulation has greatly amplified its essence. Despite numerous studies on the subject using certain classical approaches, there exists an opportunity for exploration of more sophisticated methods such as the deep learning (DL) techniques. Successful researches about DL applications to computer vision, speech recognition, and acoustic computing problems are motivation. However, such researches are not sufficiently exploited for electricity demand forecasting using DL methods. In this paper, we considered specific DL techniques (LSTM, CNN, and MLP) to short-term load forecasting problems, using tropical institutional data obtained from a Transmission Company. We also test how accurate are predictions across the techniques. Our results relatively revealed models appropriateness for the problem.

[*] Osun State University, Department of Information and Communication Technology, Osogbo, Osun State, Nigeria, saheed.adewuyi@uniosun.edu.ng
[**] Obafemi Awolowo University, Department of Computer Science and Engineering, Ile-Ife, Osun State, Nigeria, s.aina@oauife.edu.ng

1. INTRODUCTION

Electricity demand forecasting is a concept in the power system that is used to describe prediction of the load on the power grid ahead of actual consumption. The power system grid must always bear the demand and thus satisfactorily service its customers. Electricity is an ever demanded commodity due to geometric rise in population (IAE, 2018). In Nigeria, a typical tropical climate, population is a key factor hindering the expected equilibrium of supply and demand of electricity as a commodity, among other factors. Therefore, there exists a gap between the generated power and consequential distribution of the commodity in Nigeria. Although, it was learnt that there is also a problem of carrying capacity of the transmission lines which contributes to supply shortage of electricity to users. As a result, the power stakeholders subject its users to harsh situations of inadequate or epileptic power supply or even blackout sometimes. While keeping this scourge under control, the distribution company will nonetheless need to always carry out demand or load forecasts of electricity consumption ahead of time.

Electricity demand forecasting or load forecasting is categorised into three main groups, generally. This includes Short-Term Load Forecasting (STLF), Medium-Term Load Forecasting (MTLF) and Long-Term Load Forecasting (LTLF) (Hernandez et al., 2012, 2013, 2014). However, in the smart grid system there is also Very Short-Term Load Forecasting (VSTLF) (Hernandez et al., 2012). In any cases, load forecasting is very crucial to all power system operators. In fact, a motivating influence is the market deregulation in most countries. Also, the unbundling of the power sector is a factor. For instance in Nigeria, the formal Power Holding Company of Nigeria (PCHN) is unbundled into three interrelated companies namely Power Generation Company (GenCo), Power Transmission Company (TransCo) and Power Distribution Company (DisCo). Electricity demand forecasting will, therefore, help the power players across classes to manage the power system's load effectively and efficiently. With load forecasting, the Utilities will especially make essential decisions critical to its operation and planning. The decision can be purchasing and power generational. Others are load switching, infrastructure development, capacity planning, maintenance schedules, energy demand, production adjustment, and contract evaluation (Gamboa, 2017; Kuo & Huang, 2018; Sarabjit & Rupinderjit, 2013).

Certain factors directly influence change in electricity demand of a region. The factors are classified accordingly as economic, demographic and technological. Others are those influenced by policy change and environment (Momani, 2013). For STLF which is the focus of this study, some factors atimes considered are time factors, weather influence and class of the users (Wan, 2014). Accordingly, the time influence can be time of the year, day of the week, and hour of the day. Similarly, there has been observed difference in the demand between weekdays and weekends (Feinberg & Genethliou, 2005; Wan, 2014). More so, there is noticeable change in the demand on holidays, this is lower load than the non-holidays (Wan, 2014). Periodicity of demand occurrence is another factor. STLF usually show noteworthy periodicity. As a matter fact, consumption is a function of daily work and rest period. Thus, it is relevant to be well noted when studying STLF problems. Hourly demands in related days also demonstrate similar patterns (Wan, 2014). Weather factor also greatly influence demand of electricity on the grid. In the tropics where weather situation is average throughout the year, the consumption pattern is obviously different from that of the temperate zones. During the harmattan season and rain season demand patterns exhibit different shapes for obvious reasons. Various weather variables are considered for STLF. Temperature and humidity are the most used for load prediction. Other parameters such as rainfall, wind speed, wind direction, solar irradiations etc. are also considered (Feinberg & Genethliou, 2005; Wan, 2014).

A number of methods have been developed for the prediction of electricity demand from Utilities. A good number of approaches namely, the similar day method, regression analysis, time series, neural networks, expert systems, fuzzy logic, and statistical learning algorithms, are used for STLF (Feinberg & Genethliou, 2005). However, advances in research have led to realisation of more reliable and precise forecasting methods. Some of these methods are classified as classical approach for electricity demand prediction. Some of these methods suffer from imprecise estimation of loads on the power grid. This would, no doubts, cause imbalance in the demand and supply of generated electricity which could lead to supply shortage and wastage (Kuo & Huang, 2018; Wan, 2014). As a result, an accurate method of demand forecasting in the short-term is germane. Machine learning techniques which are capable of learning from data are a more advanced approach to STLF. But, these techniques also have their own problems especially when applied data are a lot plenty while fitting a model for the realisation of its inherent skills on a task. However, the deep learning methods or techniques come with needed rescue. The deep learning techniques are some machine learning algorithms and models that have capabilities to learn tasks and features directly from applied data. Deep learning techniques precisely extract most hidden features underlying and undermining the precision of a model performance on a demand forecasting problem such as the electricity instances. Precise or accurate demand forecasting has advantage of reducing generated costs as well as assists in the reliability of the power sector's responsibilities (Wan, 2014). These, among others, are the promise of deep learning methods. Deep learning is a concept found in the machine learning computing knowledge domain which presents with a lot of research achievements in many areas of computing such as computer vision, signal processing and natural language processing etc. But the techniques have recently been exploited for its applicability to research problems in the power system, especially the electricity demand forecasting.

We therefore investigated some of the deep learning techniques as relevant to the problem area (Adewuyi, Aina, Uzunuigbe, Lawal & Oluwaranti, 2019); and consequently, we exploited only a few models that recent studies have demonstrated to embody positive influence on the STLF research. However, we observed that most of these studies in the problem area were mostly done by applying the temperate climate datasets (Bouktif, Ali, Ali & Mohamed, 2018; Ghullam & Angelos, 2017; Hussein, 2018; Kuo & Huang, 2018; Stuart & Norvig, 2013; Yi, Jie, Yanhua & Caihong, 2013). Most of the approaches implemented were not applicable to nor capable of exploiting feature abstractions characterising a typical tropics dataset needed for attaining a precise STLF model.

In this paper, we developed a STLF model that can reliably and accurately predict day-ahead electricity consumptions. We investigated the process underlying the problems and formulated a precise deep learning model for the process. We, therefore, limited the study to three deep learning techniques well established in the literature (Bengio, 2009; Bouktif et al., 2018; Brownlee, 2018; Chengdong, Zixiang, Dongbin, Jianqiang & Guiqing, 2017; Deng, 2013; Deng & Yu, 2013; Ghullam & Angelos, 2017; Hamedmoghadam, Joorabloo & Jalili, 2018; Hosein & Hosein, 2017; Kuo & Huang, 2018; Schmidhuber & Sepp, 1997; Stuart & Norvig, 2013; Wan, 2014) for modelling electricity demand forecasting. The techniques employed are the Long Short Term Memory (LSTM) network, Convolutional Neural Network (CNN) and Multilayer perceptron (MLP). The data is based on three-year historic electricity data collected from the Transmission Company of Nigeria (TCN) and one year weather data collected from the Nigerian Meteorological Agency (NiMet) for an institutional customers. The datasets are preprocessed for various anomalies, such as inputting the missing values and the models are applied to the datasets with the results being analysed for their training, validation and prediction scores. The results show that upon comparing the three techniques, the LSTM model has an average performance across the training, validation and testing metrics.

2. RELATED WORK

Research studies that are adopting deep learning method have increased in the last decade since Hinton, Osindero & Teh (2006) presented a novel work on deep belief network. A vast majority of works in this domain focused computing areas of acoustic, image, natural language, and signal processing. But, the last few years also witness application of deep neural network approach on power system's datasets especially for load prediction purposes. In the study by (Hossein & Hossein, 2017), a STLF system was developed using Deep Neural Network (DNN) techniques. The studies compared DNN with some traditional methods including moving averages, regression trees and support vector regression.

The DNN methods used are DNN without pretraining (DNN-W), DNN with pre-training using Stacked Autoencoders (DNN-SA), standard RNN and RRN-LSTM. The authors conclude that, in all the DNNs and the baseline techniques, the DNN pretrained with SAE had most stable characteristic when run through both 200 and 400 epochs as it outperformed other methods. The research also used the eventual DNN results to demonstrate dynamic pricing possibility especially on peak load reduction application. Ghullam & Angelos (2017) developed a Feed-Forward DNN and Recurrent DNN models to predict short term electricity load and exploit their applicability. The study introduced utilisation of time/frequency feature extraction procedure initiated by the two models which reveals hidden dominant factors responsible for electricity consumption, as it considers prediction accuracy. Temperate climate datasets were applied. These datasets were analysed on time/frequency domains self-reliantly and the frequency domain components are subsequently transformed back to the time domain which results in capturing of latent features useful for accurate day-ahead electricity demand load measurement. The result of combined features of time/frequency analysis with DNNs enables attainment of higher accuracy. In the work done in (Seunghyoung, Hongseok & Jaekoo, 2017), the researchers identified the need to investigate important aspect of Demand-Side Management (DMS), that is, individual customer loads as a forecasting scheme. As a result, the study adopts DNN as forecasting model for various DSM's loads of a region. The work proposed a STLF framework which is DNN-based. The DNNs are trained in two ways namely: DNN with pretraining scheme using RBM and DNN without pretraining scheme, using ReLu as an activating function, in order to forecast daily loads a day-ahead. The DNNs were compared with two shallow network techniques in its evaluation. The problem of training DNN with customer load was investigated and this revealed that overfitting may occur if the training set is small. The results obtained shown that DNNs performances were better than that of the shallow neural networks. This study also utilises temperate climate datasets. Wan (2014) developed a DNN based load forecaster. He also analyses the critical features underlying load forecasting problems as well as discusses demand forecasting factors to include periodicity, time dependency, holiday effects and weather influences. The researcher presented RBM pretraining and discriminative as pretraining technologies for the problem. He thus utilised 3 years' temperate region datasets, following the algorithms, and compared different neural network models. The result showed that DNNs with pretraining compared favourably and are superior to both DNN without pretraining and single layer ANN. This, thus, supports consideration of a deep learning method to forecasting electricity demand. On the application of a 'community' temperate datasets, the study in (Kuo & Huang, 2018) introduced a precise DNN model for energy load forecasting on short-term horizon using deep learning, a Convolutional Neural Network (CNN) approach. The study compared forecasting results of some AI algorithms which are usually used in load forecasting problem with the CNNs-based DNN, exhibiting great accuracy. Furthermore, the study in (Bouktif et al., 2018) developed an LSTM model for a European electricity consumption data (a typical temperate climate study) using various configurations of the LSTM network for forecast of short to medium term aggregate load forecasting. The study also trained some machine learning algorithms and adopts the best as baseline for comparison with the LSTM. The result shows that with LSTM model, accuracy performance is a lot enhanced unlike an optimized machine learning baseline model. Similarly, on long-term forecasting of electricity demand based on LSTM deep architecture, the study in (Agrawal, Muchahary & Tripathi, 2018) notes that the standard methodology for the LSTM is mainly restricted to electricity demand data characterised by the granularity of a month or a year, leading to very low accuracy load prediction. It therefore developed a method for LTLF having hourly resolution. The model is centred on Recurrent Neural Network (RNN) consisting of the LSTM cells. It also took into consideration the long-term relations in sequence electricity demand data.

In summary, the related work on STLF problem using deep learning approaches have one way or the other demonstrated and shown success on the application of any of the feed-forward or the recurrent deep learning techniques to electricity demand forecasting. However, the data utilised for these studies differs in terms of the climate of the study area. Most of the reviewed works were done for the temperate region. In any situations, there is need to estimate electricity load forecasting problem with interest on the climatic condition of the study area. Weather, being an essential aspect of the climate, appears a big influence on electricity demand, considering how change in weather condition of a place could shape its load profile (Feinberg & Genethliou, 2005; Hernandez et al., 2012; Momani, 2013). Therefore, this study is a typical tropical climate study for an institutional customer type. It is also a typical study for modelling load forecasting problems in the tropics using deep architectures.

3. DEEP LEARNING BASED LOAD FORECAST

3.1. Convolutional neural network

Convolutional Neural Network (CNN) is one of the most popular techniques for deep learning with images and videos. However, its applicability to quantitative data such as the electricity demands is sketchy. The internal representation and architecture of CNN is as in Adewuyi et al., (2019). Furthermore, in order to obtain a richer representation of the applied data, the hidden layer was stacked, so as to obtain multiple feature maps (Hosein & Hosein, 2017). Therefore, in this case, 3 layers of CNN were stacked, with each comprising 128 neurons and 3 kernel size.

3.2. Long Short-Term Memory (LSTM)

LSTM is another deep learning architecture utilised in this study. The introduction to its concepts is documented in our previous work (Adewuyi et al., 2019). However, in this paper, the LSTM deep method was defined as a sequential model characterised by 128 neurons of 3 stacked architecture.

Furthermore, in order to have effective model training, we utilised Truncated backpropagation Through Time (TBPTT) algorithm defined in (Sutskever, 2013) and adjudged as the most practical method for training RNNs models (Brownlee, 2018; Ronald & Jing, 1990)

3.3. Multilayer Perceptron

The Multilayer Perceptron (MLP) is characterised by a system of input, hidden and output layers of neurons. However, unlike an MLP, a perceptron will only have all its input directly connected to its output (Stuart & Norvig, 2013). We, therefore, discuss a single hidden layer MLP for easy digest of its concept.

Assume a situation of n samples of data inputs x_1, x_2,x_3,..., x_{n-2}, x_{n-1}, x_n and corresponding outputs y_1,y_2,y_3,..., y_{n-2}, y_{n-1}, y_n. Therefore, we evaluate the hidden layer input $\hat{y}_h$, expressed as:

$$\hat{y}_h = \rho(\sum_{k=1}^{n} x_k \times w_{kl} - \omega_l) \tag{1}$$

where: w_{kl} – is the weight on data input k and hidden neuron l,
ω_l – is the lth neuron bias, while
$\rho(x)$ – is the sigmoid activation function (Hosein & Hosein, 2017; Stuart & Norvig, 2013) defines as:

$$\rho(x) = \frac{1}{1+e^{-x}} \tag{2}$$

Following this is the output estimation of the MLP. We defined this as:

$$\hat{y}_j = \rho\left(\sum_{l=1}^{m} x_{lj} \times w_{lj} - \omega_j\right) \tag{3}$$

where: m – is the total inputs neurons, j, at the output layer.

Consequently, an update of the weight follows. This is necessary in order to reduce error rate (Brownlee, 2018; Hosein & Hosein, 2017; Stuart & Norvig, 2013) estimated as:

$$e = \hat{y}_j - y_j \tag{4}$$

So, this study defined MLP also as sequential model also with 128 neurons and 3 stacked layers.

4. ANALYSIS

4.1. Data Description

The applied datasets primarily consists of 25,751 samples of 21 features representing three years institutional electricity consumptions from the national grid and some weather data. The power data were recorded at hourly intervals throughout the years and it represents a typical tropics electricity demand. The power data was collected at the Transmission Company of Nigeria (TCN) 132/33kV, Ile-Ife while the weather parameters were taken at the Nigerian Meteorological Agency (NiMet) situates at Ido-Osun Aerodrome, Osogbo, Nigeria. These datasets were initially split into two parts, 80% training set and 20% test set. The test set was consequently broken into 60% validation set and 40% test set. As can be inferred, the datasets are characterised by both electrical and non-electrical features. The electrical features include the load taken at each hour of the day and other lagged loads such as the previous one hour, previous two hours, previous day same hour, previous day previous hour, previous day previous two hours, previous two days same hour, previous two days previous hour, previous two days previous two hours, previous week same hour, average of past twenty-four hours, average of past seven days, day of the week, weekend-day and holiday. Understanding what type of data our problem represents, suggests that we carefully culled out 'holiday' consumptions from our dataset as stated herein. The non-electrical data features include weather parameters such as relative humidity, dry bulb and wet bulb. We also included calendar features such as the actual date load was recorded, time load was recorded and date/time load was taken in a particular day.

4.2. Modelling Method

As a data preprocessing strategy, our data is cleansed by removing noisy data such as an instance of incorrect date value. Missing data was handled by *imputing* values using the mean strategy (Swalin, 2018). Furthermore, we prepare the data by scaling the numeric data and transforming the categories. This is with a view to improving the stability of the network and modelling performance. We, therefore, standardised our data to have zero mean and unit variance. This will enable the model to be more robust to new data. For categorical data, we adopt a one-hot encoding approach as in (Swalin, 2018). This prevents ordering, leading to model poor performances.

As a compilation strategy, we carefully specified some parameters for the network training purposes. We simply set Adam optimiser to the following configuration: exponential decay rate = $1e-3$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$ which is in line with the algorithm in (Swalin, 2018) and specified below:

Require: α: *Stepwise*
Require: $\beta_1\beta_2 \in (0,1)$: *Exponential decay rate for the moment estimates*
Require: $f(\theta)$: *Stochastic objective function with parameter*
Require: θ_0: *Initial parameter vector* $m_0 \leftarrow 0$: *Initialises first moment vector*
$v_1 \leftarrow 0$: *Initialises second moment vector*
$t \leftarrow 0$: *Initialises timestep*
while θ_0 unconverges ***repeat***
$t \leftarrow t + 1$
$g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$: *Gets gradient w.r.t stochastic objective at time-step* t.
$m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$: *Updates biased first moment estimate.*
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$: *Update biased second raw moment estimate.*
$\widehat{m_t} \leftarrow m_t/(1 - \beta_1^t)$: *Computes the bias-corrected first raw moment estimate.*
$\widehat{v_t} \leftarrow v_t/(1 - \beta_2^t)$: *Computes the bias-corrected second raw moment estimate.*
$\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m_t}/(\sqrt{\widehat{v_t}}+\epsilon$: *Updates parameters.*
end while
return θ_t: *Resulting parameters,*
where $\leftarrow$ *means equal to.*

Furthermore, in order to configure the models loss function, we considered the Mean Squared Error (MSE), which is defined as below:

$$MSE = \epsilon[(y_m - \bar{y}_m)^2] \tag{5}$$

As a model fitting approach, we specified a matrix of input features and corresponding output patterns. Our model is, therefore, fit by means of the Truncated Backpropagation Through Time (TBPTT) algorithm as in (Sutskever, 2013).

4.3. Evaluation Method

As an evaluation method, the developed model was based on the three performance evaluation metrics in (Adewuyi et al., 2019).

5. RESULTS

The three models were compared against one another based on the set performance evaluation criteria. To begin with, the learning curves of the developed LSTM model was visualised to see how close or far it is from the regression line. This was done for an experimental cycle of the training set spanning 100 epochs. The same was done for the candidate models CNN and MLP, also with a view to observing how close they are to an ideal solution. This is shown in Fig. 1, Fig. 2 and Fig. 3.

Also, the validation and prediction tests of the developed LSTM model as well as its candidates were carried out. This was with a view to determining how capable they are at forecasting electricity demand from the grid when fed with unseen load data. This validation and test methods follow that a performance analysis of the three techniques were done by comparing the candidate models with the LSTM model using their MSE, RMSE and MAE scores when experimented at different epochs as shown in Tab. 1. Prediction scores of the three applied techniques show that our LSTM model out-performed the CNN and MLP, its candidates, across metrics and epochs. The models were shown our datasets during the training, evaluation and testing phases, each at 40, 60, 80 and 100 epochs to observe their prediction accuracy. Although, we noticed an instance of MLP superiority over LSTM and CNN, but this occurred at an early stage of the experiment, 40 'epoch', precisely. This would amount to nothing significant because a further exposure of our datasets shown supremacy of the LSTM model over others for our problem.

a) **b)**

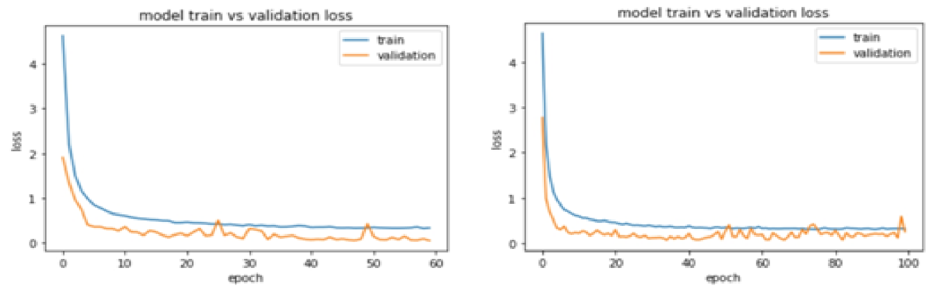

Fig. 1. LSTM model learning losses at 60 and 100 epochs

a) **b)**

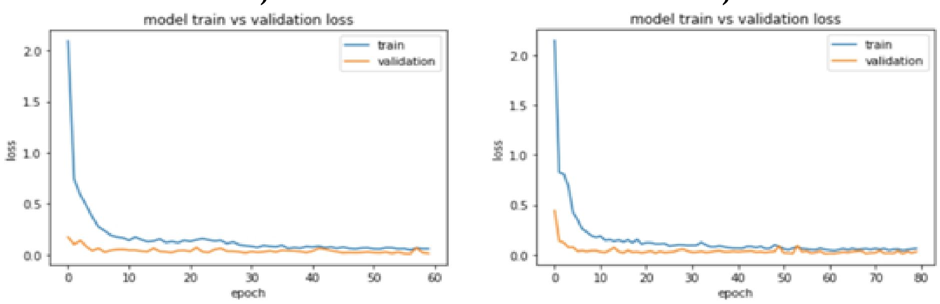

Fig. 2. CNN model learning losses at 60 and 80 epochs

a) b)

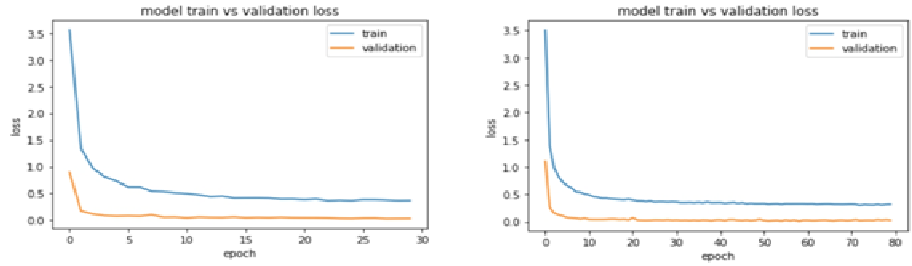

Fig. 3. MLP model learning losses at 30 and 80 epochs

Tab. 1. Comparison analysis of the deep learning techniques performances

Technique	Training Score			Validation Score			Test Score		
	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
				40 Epochs					
LSTM	6.36	2.52	2.40	5.60	2.37	2.34	5.53	2.35	2.34
CNN	6.58	2.56	2.55	6.51	2.55	2.53	6.58	2.57	2.56
MLP	5.85	2.42	2.39	5.47	2.34	2.31	5.32	2.31	2.30
				60 Epochs					
LSTM	16.80	4.10	3.02	5.96	2.44	2.42	6.09	2.47	2.46
CNN	6.71	2.59	2.55	6.30	2.51	2.47	5.80	2.41	2.40
MLP	6.14	2.48	2.46	5.80	2.41	2.40	5.76	2.40	2.40
				80 Epochs					
LSTM	9.04	3.01	2.58	6.03	2.46	2.45	6.05	2.46	2.44
CNN	7.09	2.66	2.61	6.62	2.57	2.52	6.58	2.57	2.55
MLP	6.05	2.46	2.44	5.86	2.42	2.41	5.76	2.40	2.40
				100 Epochs					
LSTM	6.04	2.46	2.36	5.49	2.34	2.19	5.29	6.37	5.78
CNN	6.54	2.56	2.53	6.28	2.51	2.48	2.30	2.52	2.40
MLP	6.08	2.47	2.44	5.77	2.40	2.39	2.28	2.52	2.40

6. CONCLUSIONS

In this paper, we demonstrate an application of deep learning techniques for electricity demand forecasting. We considered three deep architectures (LSTM, CNN and MLP) for our study. The applied datasets, which is a typical University electricity consumption, underpins the concepts of weather influence on loads especially in the tropics. These three techniques were compared to see which deep method would perform best. Thus, we found out that the LSTM model, of the three techniques out-performed others. We, therefore, establish that deep learning approach is a suitable approach to solving the problem of electricity demand forecasting.

REFERENCES

Adewuyi, S., Aina, S., Uzunuigbe, M., Lawal, A., & Oluwaranti, A. (2019). An Overview of Deep Learning Techniques for Short-Term Electricity Load Forecasting. *Applied Computer Science*, *15*(4), 75–92. doi: 10.23743/acs-2019-31

Agrawal, R. K., Muchahary, F., & Tripathi, M. M. (2018). Long term load forecasting with hourly predictions based on long-short-term-memory networks. In *2018 IEEE Texas Power and Energy Conference (TPEC)* (pp. 1–6). College Station, TX.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundation and Trends in Machine Learning, 2*(1), 1–127.

Bouktif, S., Ali, F., Ali, O., & Mohamed, A. S. (2018). Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies*, *11*, 1636–1656.

Brownlee, J. (2018). *Deep learning for time series forecasting: Predicting the future with MLPs, CNNs and LSTMs in Python*. V1.2 ed. M. L. Mastery.

Chengdong, L., Zixiang, D., Dongbin, Z., Jianqiang, Y., & Guiqing, Z. (2017). Building energy Consumption prediction: An extreme deep learning approach. *Energies*, *10*(10), 1525–1545.

Deng, L. (2013). A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Transactions on Signal and Information Processing*, *3*(2). doi:10.1017/ATSIP.

Deng, L., & Yu, D. (2013). Deep learning: Methods and Applications. *Foundations and Trends in Signal Processing, 7*(3-4), 197–387.

Feinberg, E. A., & Genethliou, D. (2005). Load forecasting. In J. H. Chow, F. F. Wu, J. Momoh (Eds.), *Applied Mathematics for Restructured Electric Power Systems*. *Power Electronics and Power Systems, Springer* (pp. 269–285). Boston, MA.

Gamboa, J. (2017). Deep learning for time-series Analysis. *arXiv: 1701.01887[cs. LG]*.

Ghullam, M. U., & Angelos, K. M. (2017). Short-term power load forecasting using deep neural networks. *ICNC*, *10*(1109), 594–598.

Hamedmoghadam, H., Joorabloo, N., & Jalili, M. (2018). Australia's long-term electricity demand forecasting using deep neural networks, *arXiv:1801.02148 [cs.NE]*.

Hernandez, L., Baladron, C., Aquiar, J. M., Calavia, L., Carro, B., Sánchez-Esguevillas, A., Cook, D. J., Chinarro, D., & Gomez, J. (2012). A study of relationship between weather variables and electric power demand inside a smart grid/ smart world. *MDPI Sensors*, *22*(9), 11571–11591. doi:10.3390/s120911571

Hernandez, L., Baladron, C., Aquiar, J. M., Calavia, L., Carro, B., Sánchez-Esguevillas, A., Cook, D. J., Chinarro, D., & Gomez, J. (2013). Short-term load forecasting for micro-grids based on artificial neural networks. *MDPI Sensors*, *6*(3), 1385–1408.

Hernandez, L., Baladron, C., Aquiar, J. M., Calavia, L., Carro, B., Sánchez-Esguevillas, A., Perez, F., Fernández, A., & Lloret, J. (2014). Artificial neural network for short-term load forecasting in distribution systems. *MDPI Energies, 7*(3), 1576–1598.

Hosein, S., & Hosein, P. (2017). Load forecasting using deep neural networks. In *Proceedings of the Power and Energy Society Conference on Innovative Smart Grid Technologies* (pp. 1–5). IEEE.

Hussein, A. (2018). *Deep Learning Based Approaches for Imitation Learning* (Doctoral dissertation). Robert Gordon University, Aberden, Scotland.

International Energy Agency (IEA) Publications and data (n.n.). Retrieved August 12, 2018 from https://www.iea.org

Kuo, P., & Huang, C. (2018). A high-precision artificial neural networks model for short-term energy load management. *Energy*, *11*(1), 213– 226.

Momani, M. A. (2013). Factors Affecting Electricity Demand in Jordan. *Energy and Power Engineering*, *5*, 50–58.

Ronald, J. W., & Jing, P. (1990). An Efficient Gradient-Based Algorithm for On-Line Training of Recurrent Network Trajectories. *Neural Computation, 2*, 490–501.

Sarabjit, S., & Rupinderjit, S. (2013). ARIMA Based Short Term Load Forecasting for Punjab Region. *IJSR*, *4*(6), 1919–1822.

Schmidhuber, J., & Sepp, H. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Seunghyoung, R., Hongseok, K., & Jaekoo, N. (2017). Deep neural network based demand side short term load forecasting. *Energies MDPI*, *10*(1), 3–23.

Stuart, R., & Norvig, P. (2013). *Artificial Intelligence A modern Approach*. Second ed. Prentice Hall.

Sutskever, I. (2013). *Training Rucurrent Neural Net-works* (Doctoral dissertation). Computer Science, University of Toronto, Toronto.

Swalin, A. (2018). *How to Handle Missing Data.* Towards Data Science. https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4 on 18/01/19.

Wan, H. (2014). Deep Neural Network Based Load Forecast. *Computer Modelling and New Technologies*, *18*(3), 258–262.

Yi, Y., Jie, W., Yanhua, C., & Caihong, L. (2013). A New Strategy for Short-Term Load Forecasting. *Hindawi*, *2013*, 208964. doi:10.1155/2013/208964