IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SHORT-TERM ELECTRICITY LOAD FORECASTING USING LSTM

Kushaagra Shukla, Dr. M. Vimala Devi,

School of Computing, SRMIST, Kattankulathur, 603203, Tamil Nadu, India

Abstract: Load forecasts are very important for the smooth working in the field of electricity. It has various uses such as purchasing of energy and production, switching the load, base development, and for evaluating contacts. Energy planning and generation of power play an important role in this scenario. Their knowledge can be used in the development of smart grids. It is of three types: long-term, short-term, and medium-term. A study concerning the already existing approaches of short-term electricity forecasting shows that an example of an improved technique is still in need. In this article, the use of a neural network called long-short term memory has been explored considering tackling of load projecting accuracy issues. The result is analyzed and the steps that can be taken to get better results along with the future scope of the project are discussed. Factors affecting load forecasting accuracy are also discussed. Forecasting permits using energy storage systems to decrease the cost of energy for the consumer. Knowing future electricity consumption along with future electricity prices makes it possible to decide when to engage a battery storage system as opposed to drawing power from the grid. It creates a path for future research and development of high-efficiency algorithms for load forecasting. A comparison of the merits and demerits of various load prediction methods is also performed.

***Index Terms* - LSTM, Load Predictions, Neural Networks, Short-Term Load Forecasting.**

I. INTRODUCTION

Load forecasts have equal importance to that of generation and transmission of energy. Keeping an eye on the system load is the main requirement for any power system. This monitoring can be done on an hourly basis or over the years. Appropriate monitoring of an unconventional power source must be undertaken as these power sources do not show much flexibility. For example, human intelligence cannot control the generation of solar power. The load forecast also impacts the availability of fuel so excess fuel needs to be available for the projected power generation.

At the time of deployment, the load forecast is determined to be performed in three different ways based on the period of prediction. The first is Long-Term Load Forecasting popularly known as LTLF is basically intended for system planning. A period of a decade or two can be covered through this. The second way is Medium-Term Load Forecasting popularly known as MTLF, which is used for fuel supply and maintenance planning. A period of a week or two or a month can be covered. Finally, the third way is Short-Term Load Forecasting popularly known as STLF, which is exercised for the ongoing process and for creating a plan for the power load system. A small period of few hours or a day can be covered.

Why is Load Forecasting (LF) needed? With increment and general growth of energy infrastructure, the demand for electricity is rising apace. To handle this increasing demand expeditiously, so-called smart grids are used. The key feature of demand-side management in intelligent networks is load prediction. Good grid operators will create economical and effective selections. This is why load forecasting is needed.

Factors affecting STLF accuracy:

1.Quality of data: A bad input data may lead to a bad model thus giving a false result
2.Model: Forecast should be good if the model can capture the important features and filter the noise.
3.Load size and composition: Different errors in housing load forecasts and industrial load forecasts.
4.Time of the day: The prediction errors during rest periods at night are usually smaller than errors during the day.
5.Ongoing season: The projection errors in winters and summers are usually greater than the errors in autumn and spring.
6.Exceptional days: The projection errors on particular days, like public holidays are usually usually higher than the errors on normal days.

II. LITERATURE REVIEW

Thouvenot et al. [2] (2016) came up with the idea to use multilevel estimators for non-linear additive models. An approach which was semi-parametric was introduced. It was built on additive models. It was a self-executing method for explaining variable selection in an additive model. It showed in what manner MTLF errors were corrected for STLF.

Mosbah and El-Hawary (2016) [5], IEEE and Life Fellow worked on the multi-layered neural network. A new computational approach for the use of back-propagation multi-layered neural networks was practised to forecast electricity price for the next month on an hour basis, based on important factors pertaining to hours like earlier load, and weather conditions in the Australian market on January 2006 electricity load forecast . It included three different networks to achieve the best performance on the basis of simulated goals.

Another approach, proposed by Yu et al. (2017), was a low-cost coding model for electricity demand in the households in smart grids [3]. Sparse coding refers to the modeling of data signals as the sum of some basic elements. It found its use in signal and image processing. Many natural signals are sparse in appropriately selected bases, e.g. B. Fourier base in speech, curvelets and wavelets in natural images.

In 2017, Wang et al. [4] proposed a Stacked Denoising Autoencoder (SDA). For the hourly prediction of the price of electricity, a class of neural networks along with their extended versions were used. The data collected in the US hubs of Arkansas, Nebraska, Louisiana, Indiana, and Texas were used. Two types of predictions were examined, the hourly online forecast and the hourly forecast for the predictive day. On calculating the results, it was observed that the SDA models were able to accurately predict the electricity prices. Also, an enhanced SDA model can achieve some improvements in the forecasting performance.

Recurrent neural networks were used in the load prediction of Bianchila et al. (2018) [6]. This class of mathematical models is gaining new interest in researchers today, replacing many of the practical implementations of predictive methods earlier based on static systems. Regardless of the inarguable expressiveness of some architectures, their recurring nature made them difficult to understand and tough to train. Recently, many new significant families of recurring architectures have come up, that are not fully explored in the context of load forecasting.

Bedi and Toshniwal [1] (2018) proposed an Empirical Mode Decomposition (EMD) type model for power gauging. This strategy was a mix of EMD technique and the long-haul stockpiling system model for evaluating the power necessities for each season, multi-day, and the time interim of that one day. For accomplishing this, the EMD's calculation was a separated arrangement flag which depended on the heap time into a few Intrinsic Mode Functions (IMFs) and their residuals. Following that, a model dependent on LSTM was given preparing exclusively for the majority of the removed IMFs and residuals. At the end, to learn the general power request, the estimated output of all IMFs was assembled.

Mi et al. (2018) proposed an improved Exponential Smoothing-Gap model [7]. For improving the accuracy of prediction, this article came up with a STLF technique on the basis of an improved or enhanced exponential gray type model. He first determined the main factor that affected the electricity load, by making use of the gray correlation analysis. Afterwards, the STLF was performed by using an enhanced grayscale type multi-variable model. The improved predictive model first pro-cessed the original power load data by making use of the first exponential smoothing system. Second, the gray prediction method was set with the help of a background value that gets optimized using one of the smoothed sequences that matching the exponential movement. Lastly, for restoring the predicted values he used the inverse exponential smoothening.

III. EXISTING LOAD PREDICTION METHODS/APPROACHES

Here are some existing and widely applied load prediction techniques with their merits and demerits mentioned in a very brief manner:

1.Similar Day Approach: It is easier to implement and has flexibility of application. But here, the accuracy is less and considers very few parameters.

2.Multiple Linear Regression: The accuracy is good here, also easy implementation, improvisation and automation. There is a problem in function selection, it requires explanatory variables, and long history is required.

3.Time Series Analysis: Simple trend Formulation and easy to use. It is not a very flexible method and uses certain boundaries for prediction.

4.Artificial Neural Network (ANN): Knowledge required in statistics and the related domain is minimal; during normal days this method has good accuracy. It contains heavy computation with over-parameterization. It is strenuous to explain and shows low accuracy in acute weather conditions.

5.Expert System: It is an easy method, having human interface making it practically usable. But, has disadvantages like making it less accurate due to the effect of less data which means little knowledge.

6.Support Vector Machines: It uses a simple linear equation making it accurate and flexible. Problems occur while choosing a suitable kernel and is complex to use.

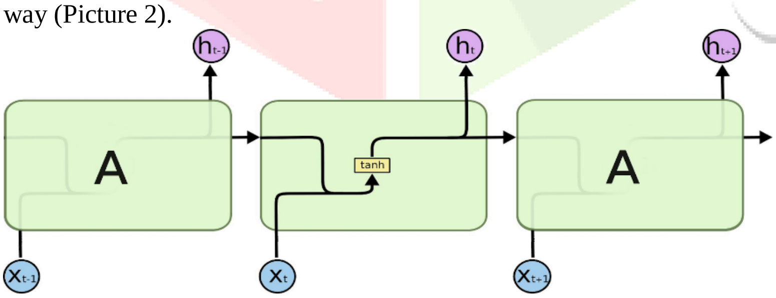
IV. PROPOSED SYSTEM

Basically, conventional neural systems each time manufacture an independent information vector and have no memory idea to assist them with undertakings that require memory. One prior endeavor to deliver this was to utilize a basic neuron-arranged criticism approach during which the yield was sustained once more into the contribution to offer the setting of the last observed info. These were alluded to as "Recurrent Neural Networks" (RNNs). While these RNNs worked somewhat, that they had a genuinely vast decrease, so any noteworthy utilization of them caused a problem called the Vanishing Gradient Problem. For this issue, RNNs are found ill-suited to most true issues in light of this issue. Accordingly, another route must be found to affect the setting memory. This is where “Long Short -Term Memory” (LSTM) was instituted as a retort. Similar to RNN neurons, the neurons in LSTM keep a chunk of data inside their system to require under consideration handling successive and worldly issues without the difficulty of the disappearing slope influencing their execution.

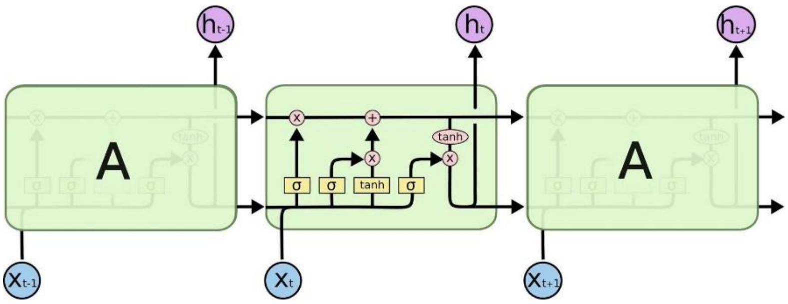
LSTM Networks:

They became a better version of RNN since they showed long-term dependency. Hochreiter & Schmidhuber launched them in 1997. LSTMs work on a variety of problems and are widely used these days. LSTMs are specially designed to deal with long-term dependence issue. Long-term memory is their speciality and not something which is difficult to deal with. All RNNs have the ability to take the shape having a sequence of neural networks’ repeated module. For normal RNNs, the replicating module has a very simple structure, like one tanh layer (Picture 1).

In LSTMs, the repeating module has a distinct structure. In spite of one layer, it has four layers that interconnect using a very peculiar way (Picture 2).



Picture 1: A regular RNN with single layer



Picture 2: A regular LSTM with multiple layers.

V. MODULES

Data Loading: Dataset can be created using surveys or by collecting a dataset from different sources using APIs or public sources. Loading data in the proper format to be useful to process it is also a big task. For this project, the data has been extracted from a public source API and processed as a CSV file.
Data Pre-processing: The images provided in a dataset need to be pre-processed before loading for the training. Proper reshaping and resizing needs to be done in order to train the dataset.
Training of Dataset: The model can be trained on a particular architecture. In this case, the architecture used is Bi- LSTM. The output of previous layer would be fed to the next layer which is the forward layer and the output of the forward layer would then be fed backward.

VI. ALGORITHMS/METHODOLOGIES USED

Bi- LSTM
It has two LSTMs involving inputs in both forward and backward directions. It is a sequence processing model. Bi-LSTMs primarily help in augmenting the amount of information available in the network thereby leading to the improvement of the content for the algorithm (e.g. knowing the order of succeeding and preceding words in a sentence).
TensorFlow:
It facilitates the path in multiple ways like securing information, model preparation, satisfying expectations, and improvising future results made by the Google Brain group. For mathematical calculation and massive scope AI, this open-source library is used. It combines together with an enormous magnitude of AI and detailed learning (also referred to as neural systems administration) calculations. It makes them helpful because of a typical representation. It employs Python to provide a supportive front-end API for developing applications with the structure while executing those applications in superior C++. It is often used across a variety of tasks but specially features specialization in training and inference of deep neural networks. It is used in systems using sound recognition, audio/video recognition, and self-driven cars.

Keras:
It is a TensorFlow running , Python written , deep learning API. It was developed for enabling fast experimentation. A good research requires the ability to travel from idea to result as fast as possible. It is a higher-level API of Tensorflow. It is an interface for solving machine learning problems. It provides necessary ideas and acts as a building block for developing and shipping ML solutions. It is highly scalable and provides cross-platform qualities to its users: you'll run Keras on TPU or large clusters of GPUs, and you'll export your Keras models to run within the browser or on a mobile device.

Google Colab:
Collaboratory, or “Colab” for a brief, maybe a product from Google Research and it is particularly compatible with Machine Learning and Data Analysis as well. Users can program in python through their google browsers. It is a hosted Jupyter notebook service that does not require any setup to run or execute its files. Many computing resources like GPUs are provided free access through this. By using Colab, you'll use Keras, TensorFlow, and OpenCV just by installing and importing.

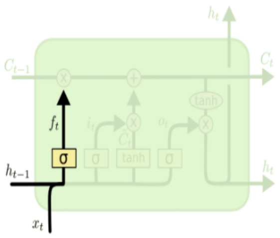
VII. DATASET DESCRIPTION

A single household’s electrical power consumption was calculated for about a period of four years with a single -minute rate. Some necessary parameters and values are mentioned .
Source:
Georges Hebrail and Alice Berard.
Data Set details:
A house in Sceaux, Paris was used as a specimen to take 2075259 measurements to be used as input. The duration was December 2006 to November 2010.
Notes:
1.The active energy utilised per-minute within the household in watt-hour via electrical devices, excluding the values of the three sub-meterings can be represented as: (1000/60)*global-active-power – [sub-metering[(3) +(2) +(1)]].
2.This data also has few empty values within the measurements in approximately 1.25% of the rows. For example, the value on 28th April 2007 is missing.

Attribute details:
1.Date: It is defined as the date in the format of date in 2 digits, followed by months in 2 digits and years in 4 digits.
2.Time: It is defined as the time in the format of hours in 2 digits , followed by minutes in 2 digits and seconds in 2 digits.
3.Global-Active-Power: The active power mean per minute in KW (kilowatts) for a single brood.
4.Global-Reactive-Power: The reactive power averaged per minute in KW in a single family.
5.Voltage: The average voltage rated per minute in volts.
6.Global-Intensity: Global Current Intensity averaged per minute in amperes for a single household.
7.Sub-Metering-1: It is defined as the first energy sub-metering of active energy in watt-hour. It relates to the kitchen which has a dishwasher, a micro-wave and an oven, but the hot plates are powered by gas and not electricity.
8.Sub-Metering-2: It is defined as the second energy sub-metering of active energy in watt-hour. It relates to the refrigerator and the washing machine with a clothes dryer.
9.Sub-Metering-3: It is defined as the third energy sub-metering of active energy in watt-hour. It relates to the air-conditioner and the electrical water-heater.

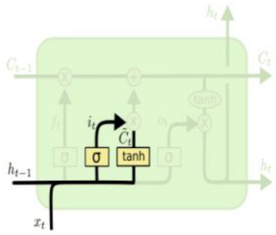
VIII. SYSTEM IMPLEMENTATION/WORKING PRINCIPLE IN THEORY

An LSTM cell is composed of 3 Gates (Forget, Input and Output) and four layers in theory.
1.The forget gate layer determines which information needs to be dismissed. (Picture 3)
2. The input gate layer chooses the values which must be updated in the cell state (Ct). The tanh layer produces a vector of fresh possible values that are added to Ct. (Picture 4)
3. The results from the previous steps to be combined to update each corresponding value in Ct. (Picture 5)
4. The sigmoid(σ) layer needs to be run in order to determine what parts of the Ct will be called output. (Picture 6)
5. The Ct is then put in tanh and multiplied to the outcome of the σ layer in order to get only the decided parts as output (Picture 6).



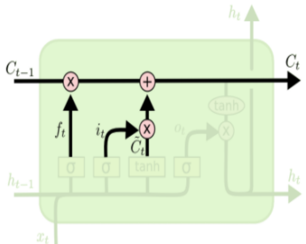
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

Picture 3



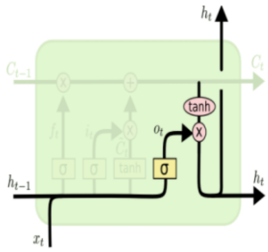
$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

Picture 4.



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Picture 5.



$$o_t = \sigma\left(W_o \; [h_{t-1}, x_t] \; + \; b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

Picture 6.

Where,
xt – input
Ct – cell state (the horizontal line where the information flows).
ht – hidden state (filtered cell state).
W, b – weights and biases which are randomly initialised during the training step , written in matrices and vectors respectively. Here, weights are Wf , Wi and Wo , and biases are bf , bi , bo matching to f, i, o respectively.
σ – It is known as sigmoid layer
tanh – tanh layer
f,i,o – resultant vectors going to the cell state. At a time 't', they become ft, it and ot .
[ht-1,xt]- concatenation of past prediction and current input.
'+' and '*' are addition and multiplication operators respectively.

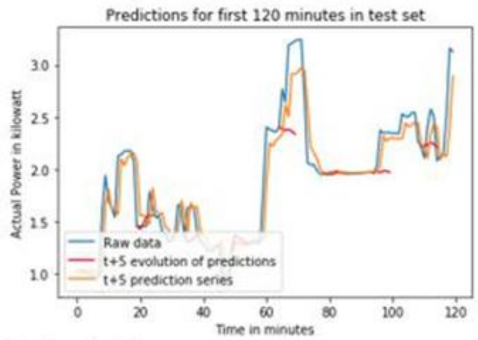
IX. RESULTS AND DISCUSSION

After implementing the LSTM model in Python for STLF using Keras and TensorFlow in Google Colaboratory, the predicted output is mentioned. (Picture 7). A set of real time data from a residential smart meter was used to test the framework(code). In individual residential households, the LSTM approach is way better than any other available rival algorithm.
It was observed that LSTM had following advantages:
1.The fruitful usage of the electricity in the modern times.
2.The intended LSTM based deep learning approach surpasses other regression models.
3.The proposal will provide aid for dynamic learning.
4.It is time efficient.
5.The model has a good accuracy rate.

```
Loading data...
Loaded data from csv.
Center data so mean is zero (subtract each data point by mean of value: 1.091607820345784.)
Data..: (2049270, 10)

Data Loaded. Compiling...

Compilation Time.: 0.027462244033813477
WARNING:tensorflow:From /usr/local/lib/python3.6/dist-
packages/tensorflow/python/ops/math_ops.py:3066: to_int32 (from
tensorflow.python.ops.math_ops) is deprecated and will be removed in a future version.
Instructions for updating:
Use tf.cast instead.
Train on 1752125 samples, validate on 92218 samples
Epoch 1/1
1752125/1752125 [==============================] - 342s 195us/step - loss: 0.1916 -
val_loss: 0.1546
Plotting predictions...
```




```
Duration of training
(s)
: 385.7290394306183]\
```

Picture 7.

X. CONCLUSION

Since the electrical appliances are becoming more and more popular, it is a matter of concern to forecast the electrical demands. The way and means to forecast the energy consumption became more and more complex. Due to the replacement of the conventional techniques by the more advanced techniques, it is possible to invent and adapt several methods of optimization which can be used variously in electrical industry. Accurate monitoring of the energy system requires advanced methods and smart grids. New energy challenges like renewable energy generation forecasts and demand-response forecasts are presented because of smart grid investments and technologies. The advent of the smart grids has enlivened the age-old energy forecasting fields. Many advanced techniques are invented but none of them are without disadvantages. A good forecast requires satisfactory results to be included having different parameters in order to be practically implementable.

XI. FUTURE WORK

In the light of the submitted paper, many load forecasting methods with several advantages and disadvantages are referred. As far as the case of machine learning models is concerned, the scope of improvement is always there. This can be done by including more layers to the LSTM architecture and then testing the accuracy for the models. There is a good scope of future work with the application of other methods and analysis. It is a possibility to compare the analysis of the results obtained from the practical load predictions. Once the analysis is done, a feedback can be given with a recommendation of the best available method based on the study conducted.

REFERENCES

[I] JATIN BEDI AND DURGA TOSHNIWAL, (Member, IEEE), “Empirical Mode Decomposition Based Deep Learning for Electricity Demand Forecasting”, Received July 26, 2018, accepted August 15, 2018, date of publication August 29, 2018, date of current version September 21, 2018.

[II] Vincent Thouvenot, Audrey Pichavant, Yannig Goude, Anestis Antoniadis, and Jean-Michel Poggi, “Electricity Forecasting Using Multi-Stage Estimators of Nonlinear Additive Models”, IEEETRANSACTIONS ON POWER SYSTEMS, VOL. 31, NO. 5, SEPTEMBER 2016.

[III] Chun-Nam Yu, Member, IEEE, Piotr Mirowski, Member, IEEE, and Tin Kam Ho, Fellow, IEEE, “A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids”, IEEE TRANSACTIONS ON SMART GRID, VOL. 8, NO. 2, MARCH 2017.

[IV] Long Wang, Student Member, IEEE, Zijun Zhang, Member, IEEE, and Jieqiu Chen, “Short- Term Electricity Price Forecasting With Stacked Denoising Autoencoders”, IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 32, NO. 4, JULY 2017.

[V] Hossam Mosbah and Mohamed El- Hawary, Life Fellow, IEEE, “Hourly Electricity Price Forecasting for the Next Month Using Multilayer Neural Network”, CANADIAN JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING, VOL. 39, NO. 4, FALL 2016.

[VI] Filippo Maria Bianchi1a, Enrico Maiorinob, Michael C. Kamp_meyera, Antonello Rizzib, Robert Jenssena, “An overview and comparative analysis of Recurrent Neural Networks for Short Term Load Forecasting”, a Machine Learning Group, Dept. of Physics and Technology, UiT The Arctic University of Norway, Troms_, Norway b Dept. of Information Engineering, Electronics and Telecommunications, Sapienza University, Rome, Italy, 20 July, 2018.

[VII] Jianwei Mi , Libin Fan, Xuechao Duan and Yuanying Qiu, “Research Article Short-Term Power Load Forecasting Method Based on Improved Exponential Smoothing Grey Model”, Received 29 August 2017; Accepted 13 February 2018; Published 25 March 2018.