Semester 7 Mini-Project
Report

Assessment of Forecasting Strategies on Univariate Time Series Data

Submitted in partial fulfillment of
the requirements for the award of the degree of

Bachelor of Technology
in
Information Technology

Submitted by

Roll No	Names of Students
IIT2016039	Sai Charan Teja
IIT2016040	Anmol Singh Sethi
IIT2016042	Manavdeep Singh

Under the guidance of
Prof. Dr. O.P. Vyas

Department of Information Technology
Indian Institute of Information Technology, Allahabad
Allahabad, Uttar Pradesh, India – 211012

Nov 20, 2019

Department of Information Technology

Indian Institute of Information Technology Allahabad

Certificate

This is to certify that this is a bonafide record of the project presented by the students whose names are given below in partial fulfilment of the requirements of the degree of Bachelor of Technology in Information Technology.

Roll No	Names of Students
IIT2016039	Sai Charan Teja
IIT2016040	Anmol Singh Sethi
IIT2016042	Manavdeep Singh

Prof. Dr. O.P. Vyas
(Project Guide)

Date:Nov, 20 2019

Abstract

Electrical Load Forecasting for demand has always been an integral part of efficient power system planning and operation. Through this report, we aim to solve this Time Series data analysis problem through Long Short Term Memory(LSTM) model which is currently a non-conventional approach for the problem statement. To analyse our results, we compare the results obtained on various(5) data-sets with traditional approaches like Auto Regression Integrated Moving Averages(ARIMA), Auto Regression(AR) and Moving Averages(MA). We discuss and analyse the results obtained.

Contents

Chapter 1

Introduction

The expansion planning of Power systems begins with a forecast of anticipated future load requirement. Estimates of the demand and required energy are integral to an effectively planned system. The generation, transmission and distribution system addition capacities are determined by demand forecasts and the kind of facilities required are determined by the energy forecasts. Energy forecasts are further used to monitor and manage the fuel consumption and procurement rates according the current market prices to maintain an adequate rate of return.

Load forecasting is usually made by the construction of models on relative information such as weather and previous load demand data. Such forecast is majorly reliant for short term load forecasting as long term, such as a month or a year forward would lead to error propagation. Various techniques for power system load forecasting have been proposed in the last few decades, such as ARIMA, AR, MA, ARMA, etc. The idea behind time series approach is based on the assumption that a load pattern is nothing more than a time series signal with known seasonal, weekly and daily variations. These variations help us to present a rough prediction of the current load requirement at the given season, day of the week and time of the day.

There are broadly four categories of load forecasting vis-a-vis

- Very Short term load forecasting
- Short term load forecasting
- Medium term load forecasting
- Long term load forecasting

Very short term load forecasting is done for a few minutes upto an hour and is majorly used for real time evaluation and security.

Short term load forecasting is done for a few hours to a few weeks and is aimed at regulating fuel procurement, short term maintenance scheduling,economic scheduling of required generating capacity.
Medium term forecast of a few months to 5 years ahead are needed for transmission and sub-transmission system planning, maintenance scheduling, setting of prices, so that demand can be met appropriately.
Long term forecasting of upto 20 years is required for regulation policies.

Chapter 2

Problem Statement

2.1 Formal Problem

We are given the problem of forecasting the future demand requirements of electrical load on the basis of analysis done on some given past data. The requirement to approach the problem is through a non-conventional algorithm namely, Long Short Term Memory(LSTM). Results obtained from some established methods like ARIMA, AR are to be compared with the results obtained from LSTM to give a better view about our designated approach.

2.2 Literature Survey

Nataraja.C et al. did a comparative study on Short Term Load Forecasting Models using the load data for the year 2011 and 2012 for the state of Karnataka. Various models like Auto-regressive Model, Auto-regressive Moving Average, Auto-regressive Integrated Moving Average were compared against each other. The pipeline involved a multi-phase process which included steps for the initial development, the tuning and modification which was followed by the prediction and result gathering phase. The tested models had an error range from 13.03 percent to 6.15 percent.[1]

Gao Gao et al. at the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK did a comparative study of two different models namely ARIMA and ANN (Artificial Neural Network) on UK electricity market data. The training was done for a data spread over a span of eight weeks in time. The comparison between these models were made on the basis of root mean squared error. The results of this paper show that Auto-regressive Integrated Moving averages gives better results in comparison to Artificial Neural Network. The ANN model consisted of twenty neurons.[2]

Chikobvu et al. developed a seasonal auto-regressive integrated moving average and predicted daily peak demand of electricity in South Africa from a period of 1996 to 2009. They concluded that SARIMA model produced better results.[3]

Nor Hamizah Miswan et al. compared ARIMA with regression model against the benchmark standard models namely ARIMA and regression models to forecast electricity load demand in a Malaysian city. The parameters used for the comparison of the above listed models included root mean square error along with mean squared error. The results were inclined towards the combined method and thus it is a better model.[4]

Another research conducted by I. A. Iwok et al. drew out the comparisons between uni-variate and multivariate time series models. A data set from Nigeria's Gross Domestic Products were used to compare both approaches. The parameter to evaluate the performance of these two was mean squared error. The conclusion reported that uni-variate time series analysis which is stationary outperformed the other models.[5]

A research by N.A. Abd Jalil et al. focuses on Electrical Load Demand Forecasting using exponential smoothing methods. The dataset used was a half hourly demand of Malaysia for one complete year. For the comparison of forecasting accuracy a parameter called Mean Absolute Percentage Error was used. Various smoothening methodologies were applied namely Holt-Winters Taylor, traditional Holt-Winters etc. [6]

V.Venkatesh et al. studied and developed Short Term Load Forecasting Models Using Stochastic Time Series Analysis. They had a dataset for one year which was split into two parts. The first six months were used to train the models and the latter six months were used to test the trained models. Approaches such as AR, ARMA, and ARIMA were developed,[7]

Nima Amjady developed a model for Short-Term Hourly Load Forecasting Using Time-Series Modeling with Peak Load Estimation Capability. The main conclusion of this paper was the fact that better results were attained using this particular technique as compared to the traditional Artificial Neural Network Models or the Box-Jenkins Model.[8]

Hippert et al. in a paper reviewed and evaluated traditional method using neural networks for Short-Term Load Forecasting.[9] Variuos other approaches like using support vector machine for load forecasting for a EUNITE competition[10], using fuzzy neural networks[11], using knowledge-based expert systems[12] have been developed or the same analysis. Another approach develops iterative reweighed least squares algorithm for short term power system load forecasting.[13]

Another research conducted by Mohamed A. Abu-El-Magd et al. drew a comparison between online and offline methods for short-term electric load forecasting. The load demand was also modeled using multivariate time series analysis [14]. G. T. Heinemann and team studied temperature sensitive and non-temoerature sensitive load and did a regression analysis for the same [15].

2.3 Data Set Description

Five data-sets have been used for this comparative study.

2.3.1 Fred Economic Data

This data-set is for Industrial Production: Electric and Gas Utilities and is provided by the Board of Governors of the Federal Reserve System US. It has monthly frequency. The data-set contains the date and the demand column.[16]

2.3.2 Open Power System Data

This is a time series data-set giving load, wind and solar, prices in hourly manner. The data is available for 37 European countries. This data set is provided by Open Knowledge Foundation. Various time series include electricity time consumption or the load, power system modelling etc. Among the many columns, the one used is GB_EAW_load_actual_tso which is the total load in England and Wales in MW as published by National Grid. [17]

2.3.3 SMARD - Strommarktdsten

A data set provided by Federal Network Agency, Germany, contains the electricity market data. This data set was created with the aim of improving

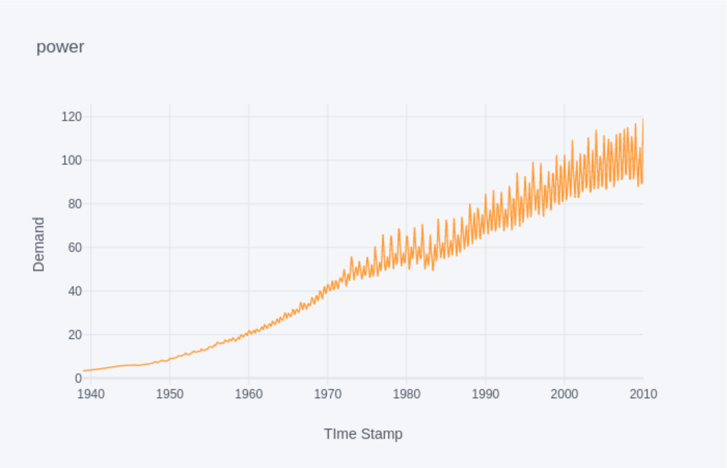

Figure 2.1: FRED Data-set

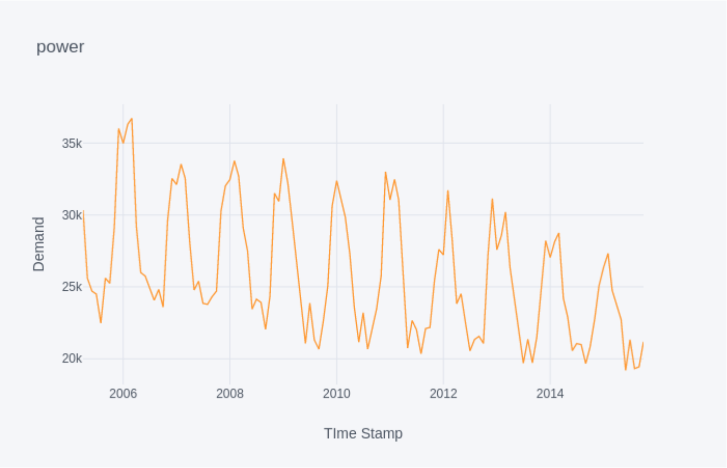

Figure 2.2: Open Power System Data Data-set

transparency. The frequency of the data is daily. The data set contains the value of load for every 15 minutes per day. [18]

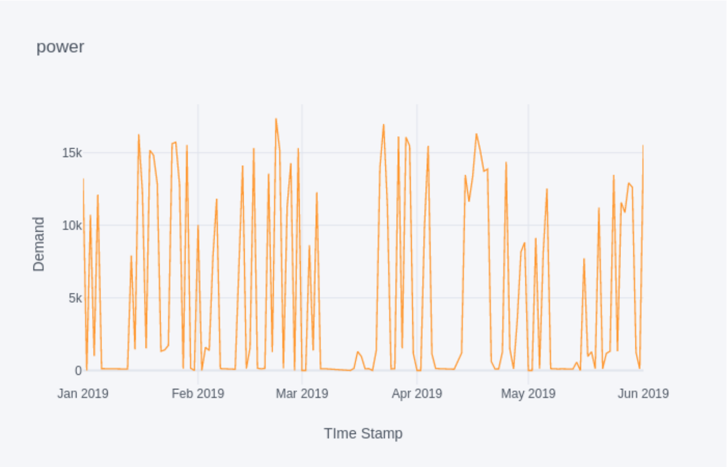

Figure 2.3: SMARD Data-set

2.3.4 Individual household electric power consumption Data

This data set is provided by UC Irvine Machine Learning Repository. The data-set contains 2075259 instances for one household with one minute sampling located in France for a period of 47 months. Nearly 1.25 percent of the values are missing in the data-set. The data-set contains 9 columns out of which two namely date and global_reactive_power have been taken in consideration. [19]

2.3.5 Uttar Pradesh State Load Dispatch Data

This is a real time data provided by Uttar State Load Dispatch Centre. The website provides real time data for schedule, draw, demand, total SSGS, UP Thermal Generation, deviation rate, etc. The data has been web-scrapped for a period of 1 day. [20]

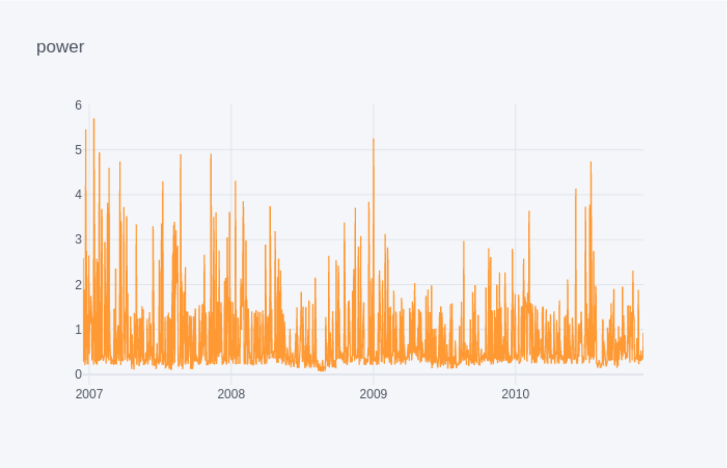

Figure 2.4: Individual household electric power consumptionData Set

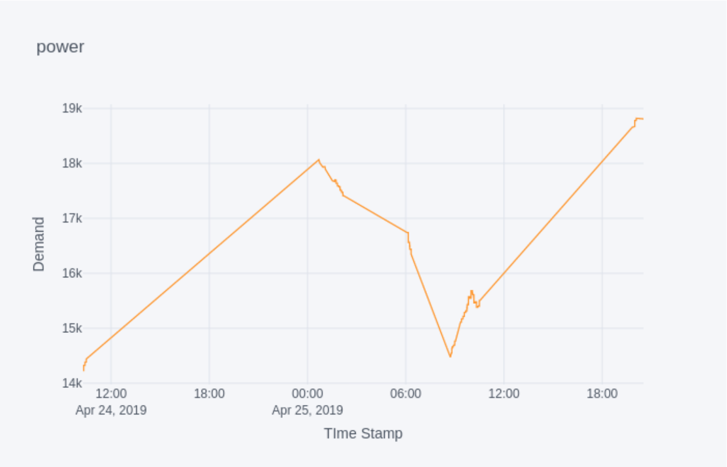

Figure 2.5: Individual household electric power consumptionData Set

Chapter 3

Methodology

The methodology can be represented as the following pipeline:-

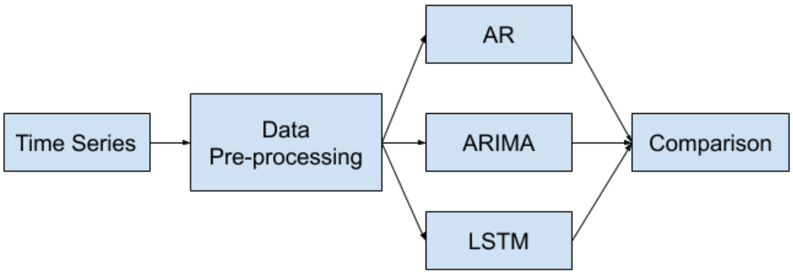

Figure 3.1: Block Diagram of the pipeline

3.0.1 Auto Regression

Regression is a process in which we try to predict one variable with help of one or more other variables. Auto regression is a special type of regression in which prediction of a variable is done with help of the past or previous values of the same variable. For example we can predict the price of gold annually or we can predict the load of electricity over a period of time based on their previous values. Since there are other factors that need to be taken into account for example weather in case of electricity load prediction the values might not be absolutely correct but still it makes more sense to find a pattern over a period of time which gives a much stronger prediction.

For example if we say that prices of gold depend on last two time stamps the equation for the prediction would look something like this:-

$$Price_{t+1} = \phi_0 + \phi_1 \times Price_{t-1} + \phi_2 \times Price_{t-2} \tag{3.1}$$

Once we find out the coefficients a0 and a1 we can predict the values of prices of gold over any time stamp if we have the values of the previous two time stamps. Auto Regression is a very powerful tool when it comes to prediction of the values. This is one of the traditional and benchmark models of forecasting with given time series.

3.0.2 Auto Regression Integrated Moving Averages - ARIMA

ARIMA stands for Auto Regressive Integrated Moving Average. It contains two different models AR model and MA model and the I terms is number of times we differentiate the data to make it stationary. AR model is explained before the Model takes previous error residuals into account while predicting. Following is an example equation for MA model. Here the w terms are the previous errors and they are multiplied with theta

$$x_t = \mu + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \ldots + \theta_q\epsilon_{t-q} \tag{3.2}$$

ARMA is combination of AR and MA model. Normally we apply ARMA on stationary data. i.e when the mean and variance is constant over a period of time. There are basically two different ARIMA models seasonal and Non-seasonal ARIMA models that can be used for forecasting. The combined formula for ARIMA can be given by

$$X_t = \mu + \epsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

Normal ARIMA model has three different terms p, d, q.
p — the number of auto-regressive
d — degree of differencing q — the number of moving average terms
In seasonal ARIMA, we just have to add in a few parameters to account for the seasons. We write SARIMA as ARIMA(p,d,q)(P, D, Q)m, m — refers to the number of periods in each season (P, D, Q)— represents the (p,d,q) for the seasonal part of the time series Non seasonal ARIMA doesn't contain P, D, Q values.

For our model, we used the following values for p, d, q, P, D, Q, m respectively for various data-sets:

Data-Set	p	d	q	P	D	Q	m
FRED	2	1	2	1	0	0	12
OPSD	1	0	0	2	0	0	12
SMARD	0	0	0	0	0	0	1
UCI	0	1	1	0	0	1	24
UPSLDC	0	1	0	0	0	0	24

These values were obtained through experimentation and hit and trial. Notice that for SMARD data-set we have all values equal to zero and m equal to 1 because no seasonality or trend was observed in SMARD data-set and thus every next value was given out to be equivalent to previous one.

3.0.3 Long Short Term Memory - LSTM

Long Short Term Memory(LSTM) Networks are a special type of recurrent neural networks, which have been used for various tasks like music completion, handwriting generation among others. They are much more effective to those specific problem areas, than the standard version.
RNNs appeal to computer scientists due to their ability to use recent past data to predict the near future characteristics of the system. But, the problem area is when data older than just the immediately previous plays some part in determining the future output. Here the gap between relevant information and the point where it is needed becomes sufficiently large to be out of the scope of standard RNNs. Theoretically, RNNs can handle such dependencies, but, in practice, they fail to do so.
LSTMs, on the other hand, are capable of learning long term dependencies. They were introduced by Hochreiter Schmidhuber (1997)[21]. LSTMs were designed to solve the long term dependency problem. By their default behavior, they remember long term information. LSTMs have a similar chain like structure like RNNs but they differ in the repeating module such that instead of a single neural network layer, they have four, as shown in the figure below.

The key to LSTMs is the cell state which is the straight line running

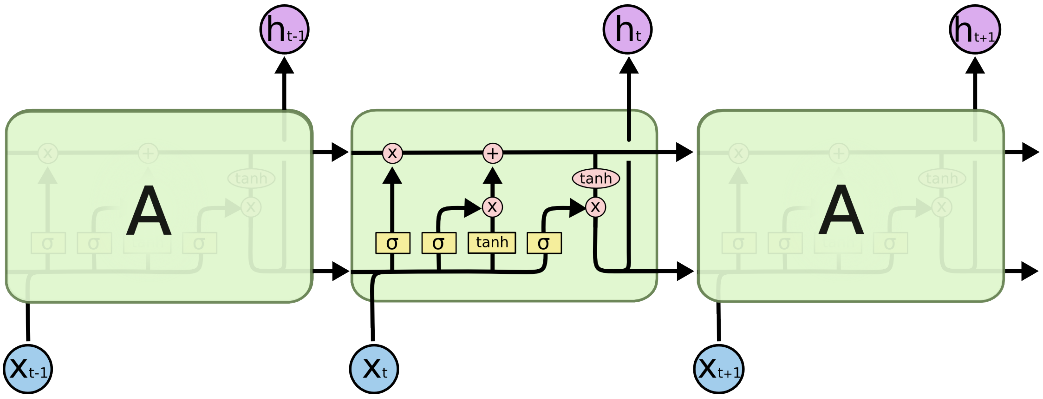

Figure 3.2: four interacting layers in LSTM repeating module

through the repeating module. It is like a conveyor belt which runs through the module with minor interactions which are controlled by gates. Gates are composed of sigmoid neural net layers and multiplication pointwise functions. They output numbers between 0 and 1 where 0 means do not let anything through and 1 means let everything through.

The first step is the forget gate layer, which looks at the current cell state, and from h(t-1) and x(t), decides how much part to keep and what to forget.

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

The next step is the input gate layer in which new information is added to the cell state through a sigmoid neural net layer and a tanh function. In the next step, these two layers are combined to give an update to the state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$C_t^{'} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

The old cell state is now updated to the new cell state by first multiplying with f_t and then adding i_t*C_t.

$$C_t = f_t \times C_{t-1} + i_t \times C_t^{'}$$

Finally, we decide what to output by filtering the previously generated new cell state. This is again done by using a sigmoid layer and a tanh layer.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \times \tanh(C_t)$$

In our model, we have used the following architecture: LSTM layer, dropout layer, LSTM layer, dropout layer, dense layer. The results achieved are discussed further in the paper.

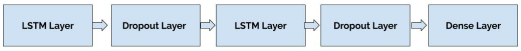

Figure 3.3: our LSTM pipeline

3.0.4 Implementation Plan and Timeline

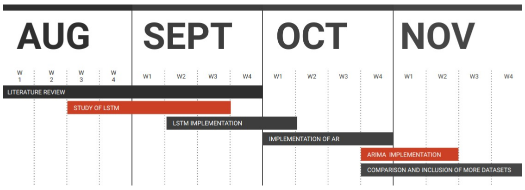

Figure 3.4: our implementation plan and timeline

As can be seen in the image above, we went on with the implementation of the solution to the problem statement by starting with literature review of similar papers written in the past which talk about some of the ways in which the problem has been tackled. Then we went on to implement the algorithm of choice for our approach, i.e. LSTM and then followed it by other benchmark algorithms like Auto Regression and ARIMA. Further, after successful implementation of each, we compared and compiled the results obtained through two parameters on five different data-sets, namely r2 error and root-mean-square error. We present a tabulated presentation of the same further.

Chapter 4

Results Obtained

4.0.1 Outcomes

The outcomes of the experiments performed are as follows:-

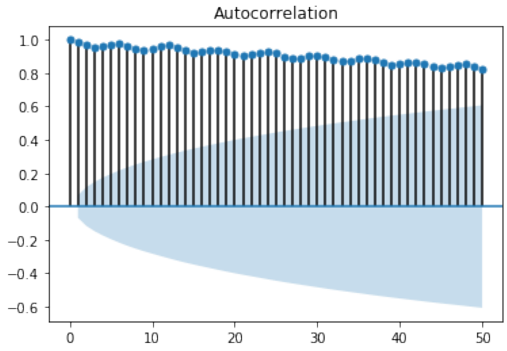

Figure 4.1: Auto Correlation Plot for FRED Data-set

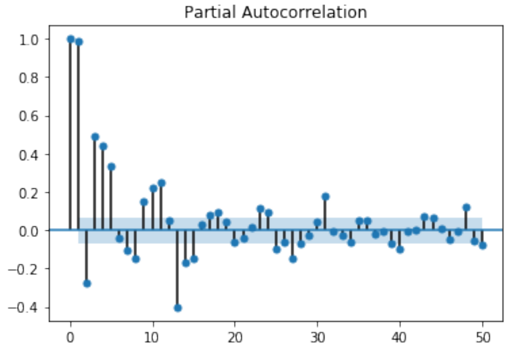

Figure 4.2: Partial Auto Correlation Plot for FRED Data-set

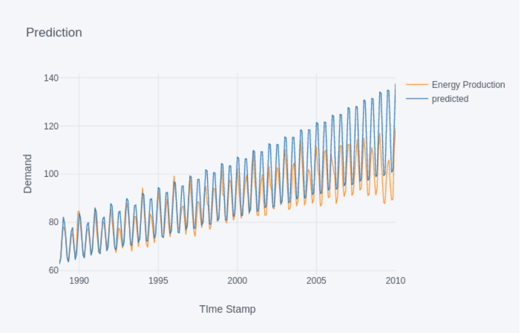

Figure 4.3: AR Predictions for FRED Data-set

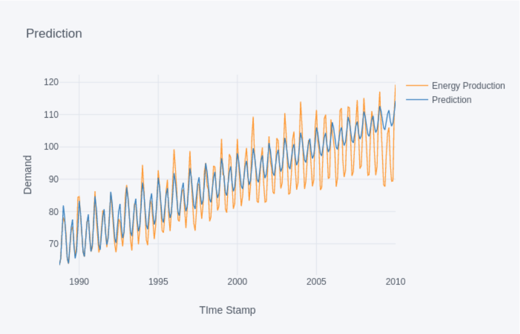

Figure 4.4: ARIMA Predictions for FRED Data-set

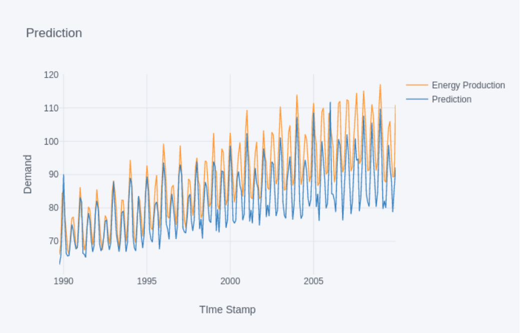

Figure 4.5: LSTM Predictions for FRED Data-set

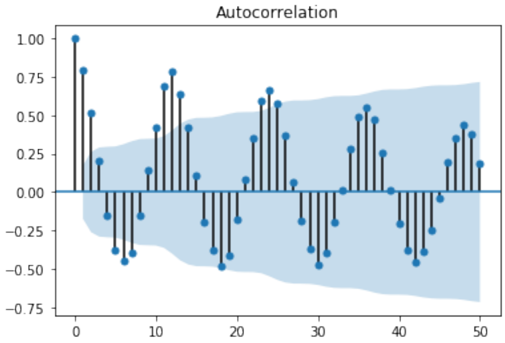

Figure 4.6: Auto Correlation Plot for OPSD Data-set

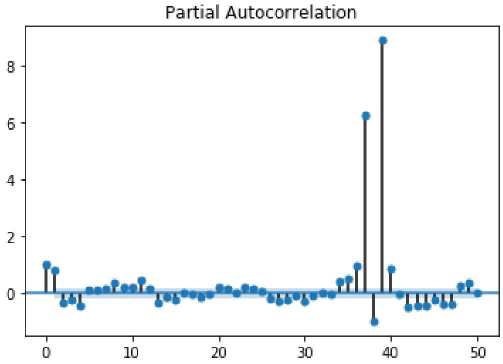

Figure 4.7: Partial Auto Correlation Plot for OPSD Data-set

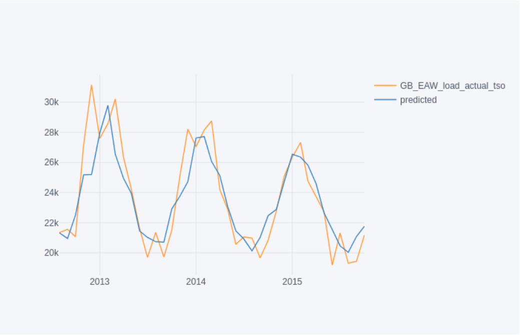

Figure 4.8: AR Predictions for OPSD Data-set

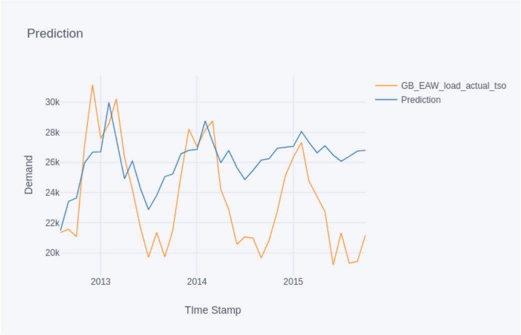

Figure 4.9: ARIMA Predictions for OPSD Data-set

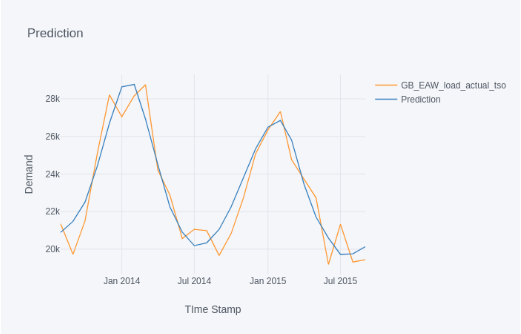

Figure 4.10: LSTM Predictions for OPSD Data-set

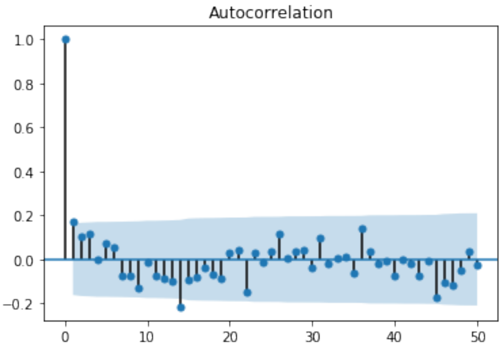

Figure 4.11: Auto Correlation Plot for SMARD Data-set

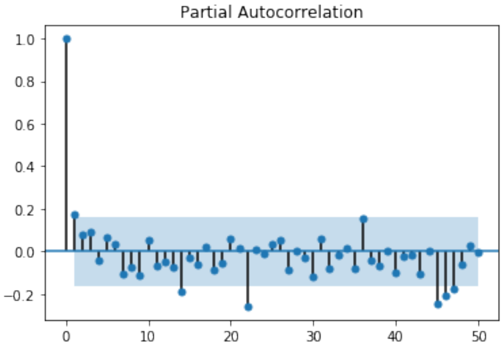

Figure 4.12: Partial Auto Correlation Plot for SMARD Data-set

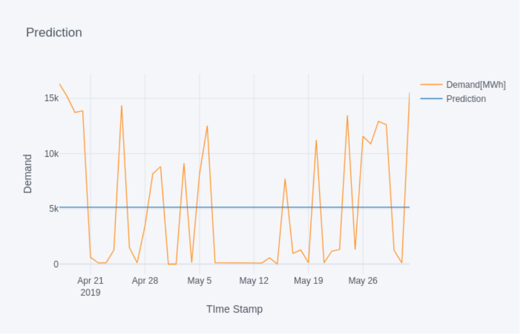

Figure 4.13: ARIMA Predictions for SMARD Data-set

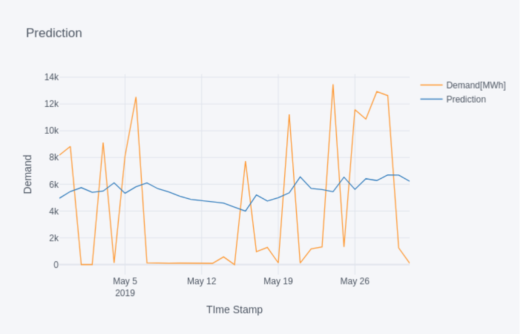

Figure 4.14: LSTM Predictions for SMARD Data-set

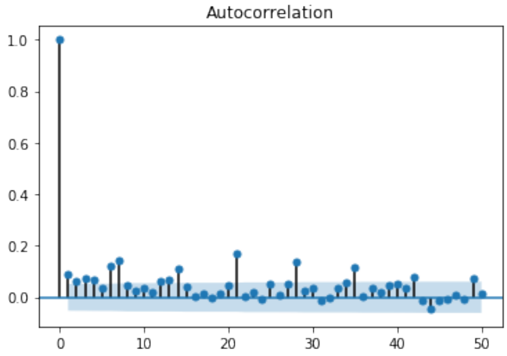

Figure 4.15: Auto Correlation Plot for UCI Data-set

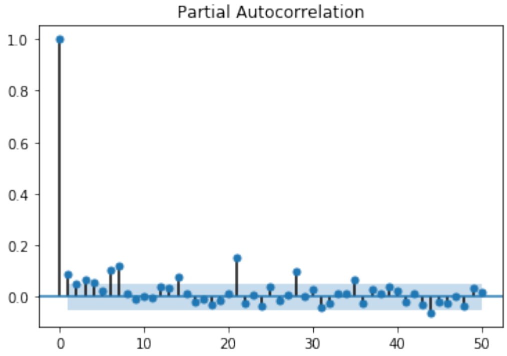

Figure 4.16: Partial Auto Correlation Plot for UCI Data-set

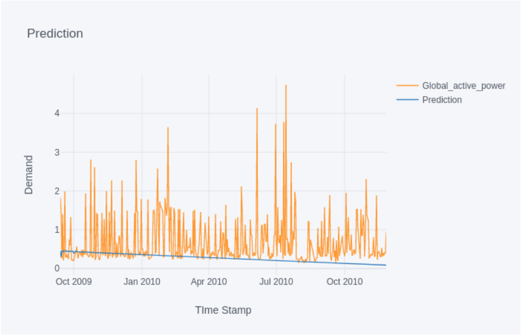

Figure 4.17: ARIMA Predictions for UCI Data-set

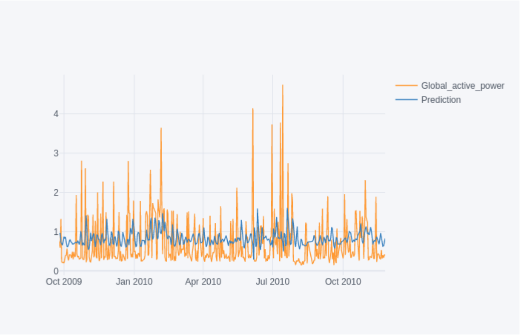

Figure 4.18: LSTM Predictions for UCI Data-set

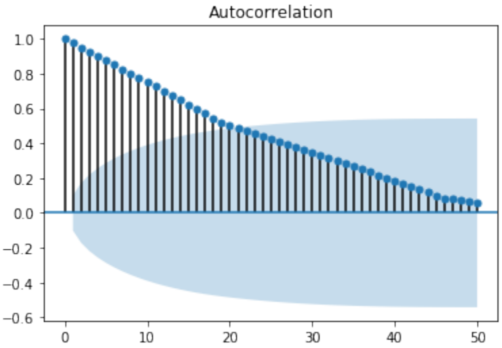

Figure 4.19: Auto Correlation Plot for UPSLDC Data-set

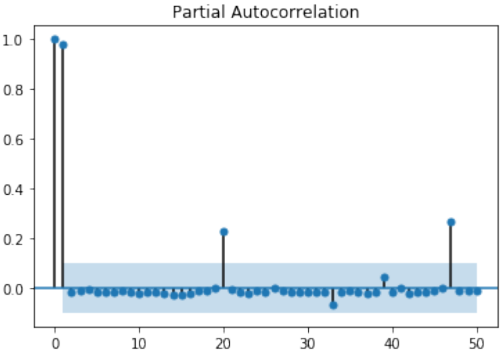

Figure 4.20: Partial Auto Correlation Plot for UPSLDC Data-set

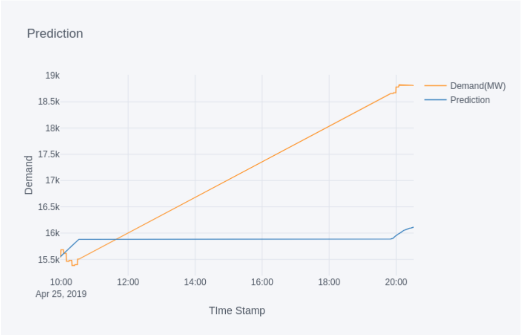

Figure 4.21: ARIMA Predictions for UPSLDC Data-set

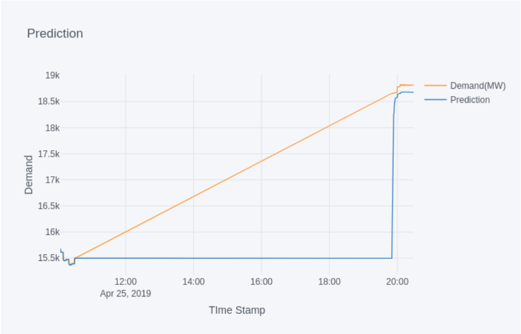

Figure 4.22: LSTM Predictions for UPSLDC Data-set

Comparison (R2 Score)			
Data-set	AR	ARIMA	LSTM
FRED	0.558	0.825	0.360
OPSD	0.761	-0.212	0.8647
SMARD	-	-1.347	-56.4874
UCI	-	-0.481	-13.052
UPSLDC	-	-0.2292	0.954

Comparison (Root Mean Square Error)			
Data-set	AR	ARIMA	LSTM
FRED	8.247	5.193	7.992
OPSD	1646.63	3710.969	1035.05
SMARD	-	5864.92	5184.04
UCI	-	0.780	0.6526
UPSLDC	-	1780.71	334.540

4.0.2 Disposal of comments given by respected Board Members during first evaluation

Comment 1: Three different data-sets should be taken

We have used five different datasets for this project.

Comparison of LSTM with ARIMA is required

Comparison using two parameters namely Root Mean Square Error and R2 Score have been used. For all data-sets except FRED, LSTM gives a better result as compared to ARIMA. Fred data-set has a very good pattern because of which better results are shown in ARIMA.

Acknowledgments

We are extremely grateful to our project supervisor Prof. Dr. O.P. Vyas for his guidance and support at every stage of the project. We are also thankful to all the friends who have supported and helped in every phase of the project till now.

Thank you.

November 2019
Indian Institute of Information Technology, Allahabad

References

[1] Nataraja, C. Gorwar, Mahesh Shilpa, G.N. Harsha J, Shri. (2012). Short term load forecasting using time series analysis: A case study for Karnataka, India. International Journal of Engineering Science and Innovative Technology. 1. 45-53.

[2] Gao, Gao Lo, Kwoklun Fan, Fulin. (2017). Comparison of ARIMA and ANN Models Used in Electricity Price Forecasting for Power Market. Energy and Power Engineering. 09. 120-126. 10.4236/epe.2017.94B015.

[3] Chikobvu, Delson Sigauke, Caston. (2012). Regression-SARIMA modelling of daily peak electricity demand in South Africa. Journal of Energy in Southern Africa. 23. 23-30. 10.17159/2413-3051/2012/v23i3a3169.

[4] Miswan, Nor Mohd Said, Rahaini Anuar, S.H.H.. (2016). ARIMA with regression model in modelling electricity load demand. 8. 113-116.

[5] Iwok, Iberedem Okpe, A. (2016). A Comparative Study between Univariate and Multivariate Linear Stationary Time Series Models. American Journal of Mathematics and Statistics. 2016. 203-212. 10.5923/j.ajms.20160605.02.

[6] Jalil, N.A. Ahmad, Maizah Mohamed, N.. (2013). Electricity load demand forecasting using exponential smoothing methods. World Applied Sciences Journal. 22. 1540-1543. 10.5829/idosi.wasj.2013.22.11.2891.

[7] “Load Forecasting Bibliography”, Phase I, IEEE Transactions on Power Apparatus and Systems, Vol.PAS-99, No.1 January/February 1980.

[8] Short term load forecasting using time series modeling with peak load estimation capability”, IEEE Transactions on Power Systems, Vol.16, No.3 August 2001

[9] Hippert, H.s Pedreira, Carlos Souza, Reinaldo. (2001). Neural Networks for Short-Term Load Forecasting: A Review and Evaluation. Power Systems, IEEE Transactions on. 16. 44 - 55. 10.1109/59.910780.

[10] “Load forecasting using support vector machines: A study on EUNITE competition 2001”, IEEE Transactions on Power Systems, Vol.19, No.4, November 2004.

[11] “Short term load forecasting using fuzzy neural networks”, IEEE Transactions on Power Systems, Vol.10, No.3 August 1995.

[12] “Short term load forecasting for fast developing utility using knowledge-based expert systems”, IEEE Transactions on Power Systems, Vol.17, No.4, May 2002.

[13] “Short term power system load forecasting using the iteratively reweighed least squares algorithm”, Electric power system research, 19(1990) pp.11-12

[14] M. A. Abu-El-Magd and N. K. Sinha, ”Short-Term Load Demand Modeling and Forecasting: A Review,” in IEEE Transactions on Systems, Man, and Cybernetics, vol. 12, no. 3, pp. 370-382, May 1982. doi: 10.1109/TSMC.1982.4308827

[15] G. T. Heinemann, D. A. Nordmian and E. C. Plant, ”The Relationship Between Summer Weather and Summer Loads - A Regression Analysis,” in IEEE Transactions on Power Apparatus and Systems, vol. PAS-85, no. 11, pp. 1144-1154, Nov. 1966.

[16] Board of Governors of the Federal Reserve System (US), Industrial Production: Electric and gas utilities [IPG2211A2N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/IPG2211A2N, November 18, 2019.

[17] ”Open Power System Data. 2019. Data Package Time series. Version 2019-06-05. https://doi.org/10.25832/time_series/2019-06-05. (Primary data from various sources, for a complete list see URL).”

[18] ”SMARD Version 2019-19-11. https://smard.de,November 19, 2019”

[19] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[20] Uttar Pradesh State Load Dispatch Centre. 2019. UP Generation Summary. 2019 Version 11-19-2019. http://www.upsldc.org/real-time-data/2019-11-19

[21] Hochreiter, Sepp Schmidhuber, Jürgen. (1997). Long Short-term Memory. Neural computation. 9. 1735-80. 10.1162/neco.1997.9.8.1735.