IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Smart Electricity Demand Forecasting by using Hybrid ARIMA with LSTM Algorithm

C.BRINDHA
Information Technology
National Engineering College
Kovilpatti, India

S.SIVA KUMAR
Information Technology
National Engineering College
Kovilpatti, India

K.RAJAKALI
Information Technology
National Engineering College
Kovilpatti, India

G.PANDIYA RAJAN
Information Technology
National Engineering College
Kovilpatti, India

Abstract- **For the operation and administration of power systems, demand forecasting, or the projection of future energy demand, is required. Power supply and demand conflicts can be alleviated with effective electricity demand forecasting. Furthermore, accurate load forecasting may help power plants operate more efficiently while also ensuring grid safety. A drop of a few percentage points in forecast accuracy, it is believed, would have a considerable cost impact on enterprises competing in highly competitive power markets. We will use a Deep learning algorithm named Improved in our project to estimate power consumption. We can obtain accurate future predicted output by utilising improved hybrid ARIMA with LSTM.**

Keywords-LSTM, ARIMA, Hybrid ARIMA, Deep Learning, forecast, energy demand

I- INTRODUCTION

Forecasting is the art of forecasting what will happen in the future by taking into account past and present occurrences.

Budgeting vs. Forecasting

Budgeting and forecasting are both instruments that help firms prepare for the future, which is undeniably true. In many respects, though, the two are diametrically opposed. Take a look at the following points:

- Budgeting is preparing a statement that includes a company's many financial operations for a certain time, such as predicted income, costs, cash flow, and investments. It is seldom carried out by a single department, such as the finance department, because it needs input from other departments in order to provide a comprehensive and complete report. As a result, the budgeting process takes some time
- Forecasts are frequently revised monthly or quarterly, although budgets are normally created for the full year. Forecasting allows a corporation to change its budget and distribute more funding to a department as needed, based on what is anticipated. In conclusion, budgets are based on forecasts.

Forecasting Demand for Electricity

A power system's effective planning necessitates the forecasting of numerous critical factors. Variables such as customer numbers, energy consumption, and yearly summer or winter maximum demand are required for identifying future requirements in the following decade, whether planning for network or generating assets. While maximum demand is critical in capacity planning, other factors like customer count and energy usage are necessary for economic evaluation and financial planning. As a result, most utilities periodically disclose their customer count, energy usage, and maximum demand estimates to aid planning and regulatory decisions. . Many utilities, on the other hand, produce customer, energy, and demand estimates separately, which means they may not be internally consistent. The goal of this project is to develop a novel approach for forecasting long-term maximum demand that has two unique characteristics:

1) it assures internal consistency by fitting as a function of other known factors and current predictions, and 2) it is resistant to changing network topologies (i.e., substation transfers).Variation generated by trends (i.e., economic and customer growth) is generally modelled using numerous scenarios, whereas variance driven by consumer response to weather is treated as a probability distribution in long-term forecasting of maximum demand. To explore long-term maximum demand, several Australian utilities and standards have established various degrees of likelihood (10 percent, 50 percent, and 90 percent) of exceedance (POE) [2]. These likelihood of exceedance values are used in network design to find a good balance between capacity investment and the risk of customer interruptions. Since they are the key inputs of production planning at different horizons [1] and storage capacity are currently restricted in terms of consumption demands, accurate energy load forecasting is of fundamental significance for the balancing of the electrical system.

II. RELATED WORKS

G.Q. ZHANG and J.F. GUO's paper, A Novel Method for Hourly Electricity Demand Forecasting, offers a new method for forecasting hourly electricity demand. In electric power networks, short-term load forecasting is becoming increasingly crucial. However, due to the complex impacts of a range of factors on load, accurate forecasting of future power demand is challenging. To solve this issue, a hybrid technique based on support vector regression (SVR) with meteorological parameters and power price is presented. To examine and construct a load forecasting model, SVR was employed. Third, to properly create a forecasting model, an improved adaptive genetic algorithm (IAGA) was used to optimise the particular ratio value combinations of each characteristic parameter, penalty factor C, and Gaussian kernel function. SVR is a good choice for predicting. Merit: It only forecasts in the short future. An Ensemble Approach for Multi-Step Ahead Energy Forecasting of Household Communities is a paper that describes an ensemble approach for multi-step ahead energy forecasting of household communities. This paper considers several prediction horizons for estimating residential communities' overall energy demand and solar energy output. Forecasting community electricity demand and energy output can assist energy grid operators better manage their short-term supply by enhancing the information available to them. Furthermore, in order to make better judgments regarding appliance usage and energy-trading programmes, families will need to know more about their usage and generation trends. The key challenges to be addressed here are the volatility of load consumption caused by consumer behaviour, as well as fluctuation in solar output caused by solar cell specifications, multiple climatic variables, and contextual elements such as time and calendar information. To solve these challenges, we offer a prediction technique that takes into account the most important aspects first and then uses an ensemble learning method in which one Gradient Boosted Regression Tree algorithm is paired with many Sequence to Sequence LSTM networks. The style is unique. Merit: The ensemble technique assists the model in forecasting with high accuracy. Demerit: If the difference between the LSTM and the gradient selected regression tree is large, the accuracy suffers greatly. Roope Tervo, Joonas Karjalainen, and Alexander Jung Roope Tervo, Joonas Karjalainen, and Alexander Jung Roope Tervo, Joonas Karjalainen, and Alexander Jung Roope Tervo, Joonas Karjalainen, and Alexander Jung Roope Tervo, Jo Power grid operators must be able to predict power outages induced by convective storms, which are extremely localised in space and time. The use of machine learning algorithms to detect and track storm cells using weather radar imagery is central to this strategy. The overall prediction technique entails contouring storm cells using CAPPI weather radar pictures with a firm 35-dBZ threshold to identify storm cells, anticipating the path of storm cells and categorising them according to their likelihood for causing harm to power grid operators Storm cells that have been tracked are then categorised using data from weather radar, ground weather observations, and lightning detectors. The fundamental problem is that the training data is substantially skewed due to the rarity of catastrophic weather occurrences. The design has merit since it is accurate and stable. Merit: They only made short-term.

III. SYSTEM IMPLEMENTATION

For demand forecasting, the old system employed a support vector regressing technique, which is only adequate for short-term forecasting. The SVR algorithm's output has a poor level of stability. It's only good for linear forecasting. As a result, the anticipated values deviate greatly from the actual values. We apply a deep learning model called Hybrid ARIMA with Long Short Term Memory in the suggested system, which is ideal for long-term forecasting. We anticipate the future 2 month power demand figures with great accuracy using LSTM.

Data collection

The electricity demand dataset was obtained from the kaggle website. Between 1 January 2015 and 6 October 2020, the dataset spans 2016 days. RRP was negative during several intraday intervals, implying that energy producers were paying consumers rather than the other way around. A brief explanation of the data is provided below.

- demand: float, a total daily power demand in MWh • RRP: float, a recommended retail price in AUD$ / MWh • date: datetime, the date of the recording
- Demand pos RRP: float, a total daily demand in MWh with a positive RRP.
- RRP positive: float, an averaged positive RRP, weighted by intraday demand in AUD$ / MWh
- RRP negative: float, an average negative RRP, weighted by the matching intraday demand in AUD$ / MWh
- demand neg RRP: float, a total daily demand at negative RRP in MWh
- float, a part of the day when demand was transacted at negative RRP
- min temperature: float, minimum temperature in Celsius for the day
- max temperature: float, maximum temperature in Celsius for the day
- solar exposure: float, total daily sunshine energy in megajoules per square metre
- school day: boolean, if pupils were in school on that day • holiday: boolean, if the day was a state or national holiday • rainfall: float, daily rainfall in mm

Data pre-processing

The data is then saved in a pandas data frame after being retrieved using the pandas library. Because our Deep learning model cannot interpret null values, this dataset has a huge number of null values, which we replace with 0.

System Architecture that has been proposed

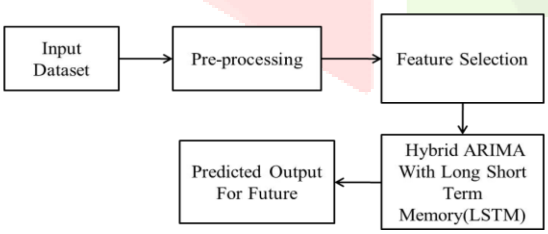


Fig 1 Proposed system Work Flow

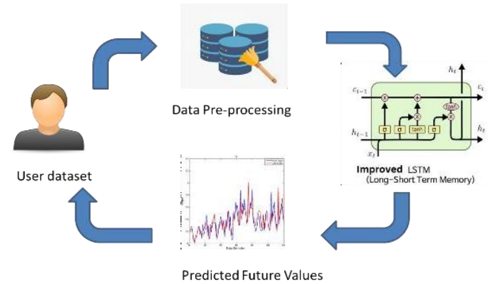


Fig 2 Flow Diagram

Module Description

The data is initially extracted from the input dataset using the pandas package. We then pre-process the data by replacing null values with 0, pick input features for feeding into the hybrid module, construct an upgraded ARIMA with LSTM model with the incremented layer model, insert the extracted features into the hybrid model, and train the machine.

Deep learning

Deep learning gets its name from the fact that it employs numerous levels in the network. A network with a non-polynomial activation function and one unbounded width hidden layer might be utilized as a universal classifier, according to early research. Deep learning is a type of machine learning that uses an unlimited number of bounded-size layers to permit practical application and optimization while preserving theoretical universality under moderate conditions. Deep learning layers are permitted to be diverse and depart substantially from physiologically informed connectionist models for the purposes of efficiency, trainability, and understandability, hence the "structured" component.

ARIMA Algorithm

Time series forecasting has many applications (particularly in the corporate world), and ARIMA is a great place to start. This article's objective is to highlight the model's building pieces (AR, I, and MA) as well as the underlying ideas (auto-correlation and partial autocorrelation).

Lag Factor and Correlation

Consider the case of estimating the price of petrol on a specific day, such as Sunday. It's self-evident and apparent that Sunday's pricing will be determined by Saturday's price. Let's take a look at the impact of Friday's pricing on Sunday's price. Because Sunday's price is determined by Saturday's (2), which is determined by Friday's (1), there may be an indirect transitive relationship between Friday's price and Sunday's price.

The graphic below depicts the aforementioned relationships.

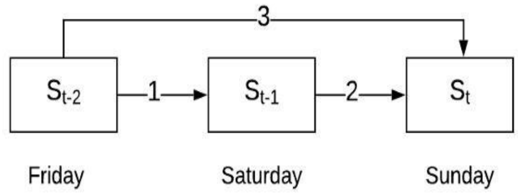


Fig 3 Relationships

Auto-Correlation

Apart from this, there is another relationship that may be difficult to comprehend at first, namely the direct influence of Friday's price on Sunday's pricing (3).Correlation is the term used to describe the reciprocal relationship between day prices. A positive correlation exists when the values of two variables grow (or fall) in lockstep. A negative correlation exists when the value of one variable rises while the value of the other falls (or vice versa). Correlation is a key concept in time series analysis, and it comes in two varieties: auto-correlation, which examines both direct and indirect effects (as shown in the picture above), and partial auto-correlation, which only considers direct effects (as shown in the diagram below).

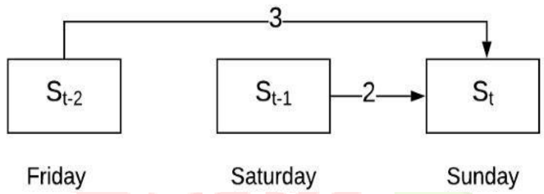


Fig 4 In partial auto-correlation, only direct effects are taken into account.

Partial auto-correlation
Factor of lag

Because we calculated the influence on Sunday's price using the price two days previously.

Correlation(Sunday,Saturday) => lag factor of 1
Correlation(Sunday, Friday) => lag factor of 2
Correlation(Sunday, Thursday) => lag factor of 3

Lag factor concept

Before going on to ARIMA, it's necessary to grasp auto-correlation and partial auto-correlation since they're crucial for choosing the proper parameters for your model.

Auto Regressive (AR)

The Auto Regressive (AR) model is a form of regression model in which the dependent variable is reliant on its own prior values. Partially auto-correlation is the sort of correlation that exists here. The AR model's equation is provided below.

$$Y_t = \beta_1 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p}$$

-----equ. (1)

The correlation between... that lagged observation and the present observation determines the relative weights (1, 2...p) of the associated lagged observations.

Take note of the (p) in the equation...

The lag order is defined as (p). It denotes the number of past lag observations we incorporate in the model, that is, the number of lags that have a meaningful association with the current observation.

Moving Average (MA)

To create a better estimate for the present time-period, the Moving Average (MA) model analyses how inaccurate you were in projecting values for prior time-periods. Basically, the errors from the lagged data are included into this model. The auto-correlation between these previous (lagged) observation mistakes and the present observation determines their consequences. In some ways, this is comparable to the AR model, which takes partial auto-correlations into account.

$$Y_t = \beta_2 + \omega_1 \varepsilon_{t-1} + \omega_2 \varepsilon_{t-2} + \ldots + \omega_q \varepsilon_{t-q} + \varepsilon_t$$

--- equ(2)

The words denote the mistakes detected at various delays, and the weights (1, 2...q) are derived statistically based on the correlations. Take note of the (q) in the equation...The size of the moving window is represented by (q), which is the number of lag observation mistakes that have a substantial influence on the current observation. It's similar to lag order (p), except it takes into account mistakes rather than observations. The MA model complements the AR model by taking past time-period mistakes into account, resulting in a more accurate estimate. When the AR and MA equations are combined, we obtain

$$\begin{aligned} Y_t = {} & (\beta_1 + \beta_2) \\ & + (\Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p}) \\ & + (\omega_1 \varepsilon_{t-1} + \ldots + \omega_q \varepsilon_{t-q} + \varepsilon_t) \end{aligned}$$

Combined equation for ARMA ..... equ(3)

Stationary

So far, the models we've looked at (AR and MA) have assumed that the series is stationary. This also implies that "stationary" is a pre-requisite for using these models with any time series.

But…what is Stationary?

In order for a time series to be stationary, it must meet the following three criteria...

1. The mean () remains constant.
2. The standard deviation () is always the same.
3. There is no such thing as seasonality.

The plot against time may usually be used to visually analyse these situations.

Integrated (I)

Let's imagine you come across a series with a "non-constant" mean, for example. The mean is plainly increasing with time, indicating that the series is not stagnant.

We'll be OK if we can only get rid of this increasing tendency.

Consider the disparities between consecutive time steps as one method.

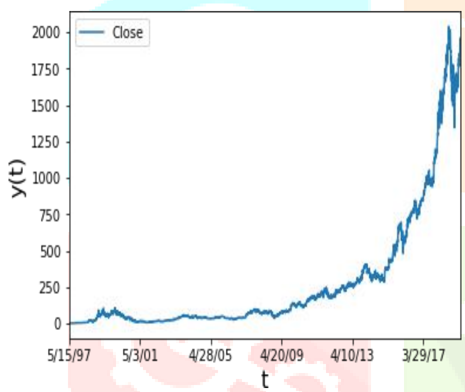


Fig 5. Amazon stock value over time

This is the same as executing a form change.

$$Z_t = Y_{t+1} - Y_t \qquad \text{---equ(4)}$$

Transformation

The standard deviation is also constant, and there is no seasonality, therefore the series is now stationary.

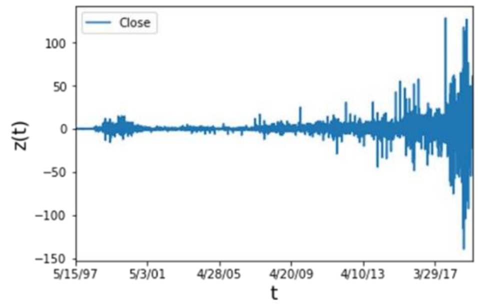


Fig. 6 Transformation applied

The letter I stands for Integrated (though it has nothing to do with integration). It simply implies that rather than forecasting the time series itself, we will predict the changes between one time step and the next. It's worth noting that we've used the first order difference, which is a single phase of term differencing. This can be repeated several times to keep the series steady.

$$Z_t = Y_{t+1} - Y_t \qquad \ldots d = 1$$
$$Q_t = Z_{t+1} - Z_t \qquad \ldots d = 2 \qquad \text{---equ(5)}$$

This order of differencing (d) is an essential ARIMA parameter that impacts the model's performance. So, to revise, the final ARIMA model will look like this.

ARIMA(p, d, q) where,
p => lag order
d => order of differencing
q => size of moving average window

LONG SHORT TERM MEMORY (LSTM)

LSTM features feedback connections, unlike normal feed forward neural networks. It can handle not just individual data points (such as photos), but also complete data streams (such as speech or video). LSTM may be used for tasks like future prediction, un-segmented, linked handwriting identification, speech recognition, and anomaly detection in network traffic or IDSs, for example (intrusion detection systems).A cell, an input gate, an output gate, and a forget gate make up a typical LSTM unit. Because there might be delays of undetermined duration between critical occurrences in a time series, LSTM networks are well-suited to categorising, processing, and generating predictions based on time series data. LSTMs were created to solve the problem of vanishing gradients that can occur while training standard RNNs. In many cases, LSTM has an advantage over RNNs, hidden Markov models, and other sequence learning approaches due to its relative insensitivity to gap length.

LSTM STRUCTURE

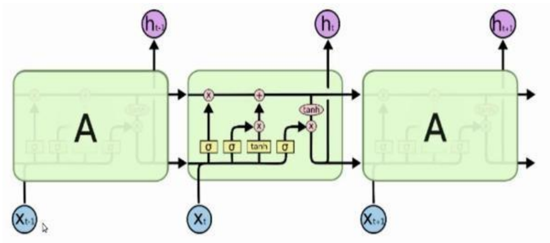
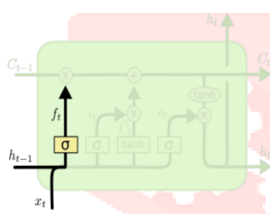


Fig 6. LSTM STRUCTURE

STEPS IN LSTM

STEP-1 The first stage in our LSTM is to select which information from the cell state will be discarded. The "forget gate layer," a sigmoid layer, makes this judgement. It examines the files ht1ht1 and xtxt and returns a number between 00 and 11 for each number in the Ct1Ct1 cell state. A 11 indicates "totally keep this," whereas a 00 indicates "absolutely discard this." Let's return to our earlier example of a language model attempting to anticipate the next word based on the preceding ones. In this case, the cell state may include the gender of the current subject, allowing the appropriate pronouns to be used. When we come upon a new topic, we get excited.in order to employ the proper pronouns. When we observe a new topic, we desire to forget about the previous subject's gender.

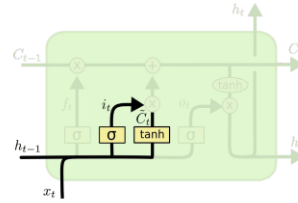


$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \ + \ b_f\right)$$

Fig 7. LSTM STRUCTURE step 1

STEP-2

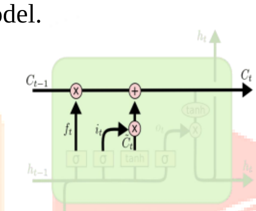
The next stage is to determine what additional data will be stored in the cell state. There are two components to this. The "input gate layer," a sigmoid layer, decides which variables to update first. A tanh layer produces a CtCt vector of new candidate values, which may then be added to the state. We'll merge these two in the following phase to make a state update. We'd want to replace the old subject's gender in the cell state with the new subject's gender in our language model.



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

Fig 8. LSTM STRUCTURE step 2

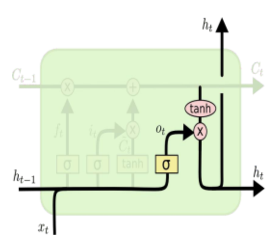
STEP-3 It's time to switch from the old Ct1Ct1 to the new CtCt cell state. We already know what to do as a result of the previous phases; all we have to do now is put it into action. We multiply the previous state by ftft, disregarding the parts we previously selected to ignore. We then put it in a CtitCt. This is the most recent set of candidate values, scaled by the percentage change in each state value. This is where we would erase the old subject's gender information and replace it with the new information we agreed on in the previous phases in the case of the language model.



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

Fig 9 LSTM STRUCTURE step 3

STEP-4 Finally, we must determine what we will produce. This output will be dependent on the condition of our cells, but it will be filtered. First, we run a sigmoid layer to determine which aspects of the cell state will be output. The cell state is then passed through xanthan (to force the values to be between 11 and 11) and multiplied by the output of the sigmoid gate, resulting in just the parts we choose to output. Because it just observed a subject, the language model might wish to output information relevant to a verb in case that's what comes next.. The language model may want to output information related to a verb since it just saw a subject.



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

Fig 9 LSTM STRUCTURE step 4

IV- RESULTS& DISCUSSION

RESULTS:

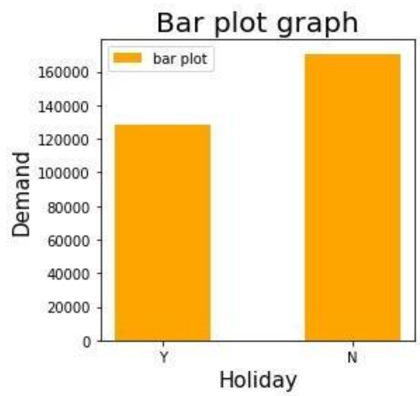


Fig 10 Histogram of dataset taken

```
Epoch 37/60
63/63 [==============================] - 12s 198ms/step - loss: 0.0194
Epoch 38/60
63/63 [==============================] - 12s 192ms/step - loss: 0.0198
Epoch 39/60
63/63 [==============================] - 12s 192ms/step - loss: 0.0194
Epoch 40/60
63/63 [==============================] - 12s 193ms/step - loss: 0.0192
Epoch 41/60
63/63 [==============================] - 12s 192ms/step - loss: 0.0195
Epoch 42/60
63/63 [==============================] - 12s 194ms/step - loss: 0.0193
Epoch 43/60
63/63 [==============================] - 12s 195ms/step - loss: 0.0196
Epoch 44/60
63/63 [==============================] - 12s 192ms/step - loss: 0.0191
Epoch 45/60
63/63 [==============================] - 12s 193ms/step - loss: 0.0193
Epoch 46/60
63/63 [==============================] - 12s 191ms/step - loss: 0.0191
Epoch 47/60
```

Fig 10 Hybrid ARIMA LSTM training

```
Performing stepwise search to minimize aic
 ARIMA(1,1,1)(0,0,0)[0] intercept   : AIC=45051.367, Time=2.27 sec
 ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=45157.766, Time=0.13 sec
 ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=45146.712, Time=0.24 sec
 ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=45121.745, Time=0.49 sec
 ARIMA(0,1,0)(0,0,0)[0]             : AIC=45155.768, Time=0.09 sec
 ARIMA(2,1,1)(0,0,0)[0] intercept   : AIC=44435.078, Time=1.22 sec
 ARIMA(2,1,0)(0,0,0)[0] intercept   : AIC=44837.602, Time=0.53 sec
 ARIMA(3,1,1)(0,0,0)[0] intercept   : AIC=44437.023, Time=2.15 sec
 ARIMA(2,1,2)(0,0,0)[0] intercept   : AIC=44430.390, Time=2.42 sec
 ARIMA(1,1,2)(0,0,0)[0] intercept   : AIC=44493.571, Time=0.91 sec
 ARIMA(3,1,2)(0,0,0)[0] intercept   : AIC=44424.102, Time=2.92 sec
 ARIMA(3,1,3)(0,0,0)[0] intercept   : AIC=44382.345, Time=2.59 sec
 ARIMA(2,1,3)(0,0,0)[0] intercept   : AIC=44387.441, Time=2.30 sec
 ARIMA(3,1,3)(0,0,0)[0]             : AIC=44379.782, Time=2.39 sec
 ARIMA(2,1,3)(0,0,0)[0]             : AIC=44384.380, Time=2.12 sec
 ARIMA(3,1,2)(0,0,0)[0]             : AIC=44421.082, Time=2.41 sec
 ARIMA(2,1,2)(0,0,0)[0]             : AIC=44425.165, Time=1.41 sec

Best model:  ARIMA(3,1,3)(0,0,0)[0]
Total fit time: 26.632 seconds
```

Fig.11 ARIMA Training

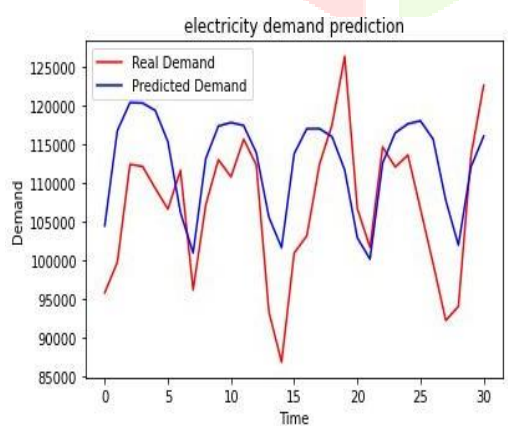


Fig.12. Existing LSTM result

	Electricity Demand
0	121711.3
1	121426.2
2	119327.5
3	112961.9
4	103653.8
5	102209.4
6	114006.1
7	117318.7
8	117748.2
9	116489.3
10	111426.8
11	103211
12	101499.1

Fig 13. Electricity Demand

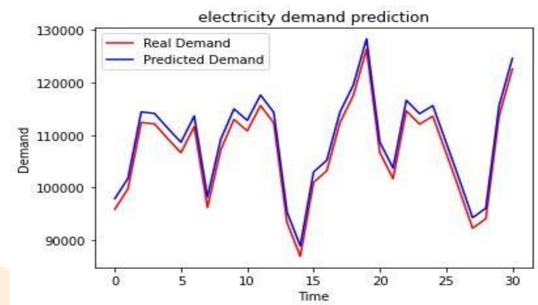


Fig 14. Hybrid ARIMA LSTM output

V-CONCLUSION

For estimating future power demand data, we employed a Hybrid ARIMA with LSTM model in our research. It is a deep learning approach capable of training and predicting non-linear data for long-term prediction. For training purposes, we used power demand data that spans over 6 years and is a nonlinear dataset. Using this model, we anticipate future 2 months of power consumption values after training. Hybrid ARIMA with LSTM predicts reliable and consistent outcomes, and its patterns match those of the current dataset. As a result, our model is completely trained and capable of accurately predicting future values.

REFERENCE

[1] D. Bunn and E. D. Farmer, *Comparative Models for Electrical Load Forecasting. New York, NY, USA: Wiley, 1985.*

[2] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via arma model identification including non-Gaussian process considerations," *IEEE Trans. Power Syst., vol. 18, no. 2, pp. 673–679, May 2003.*

[3] A.Antoniadis,X.Brossat, J. Cugliari, and J.-M. Poggi, "Aprediction interval for a function-valued

forecast model: Application to load forecasting," *Int. J. Forecasting, vol. 32, no. 3, pp. 939–947, 2016.*

[4] H. Cho, Y. Goude, X. Brossat, and Q. Yao, "Modeling and forecasting daily electricity load curves: A hybrid approach," *J. Amer. Stat. Assoc.*, vol. 108, no. 501, pp. 7–21, 2013.

[5] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *Int. J. Forecasting, vol. 30, no. 2, pp. 357–363, 2014.*

[6] J.R. Lloyd, "Gefcom2012 hierarchical load forecasting: Gradient boosting machines and Gaussian processes," *Int. J. Forecasting, vol. 30, no. 2*, pp. 369–374, 2014.

[7] D.C. Park, M. El-Sharkawi, R. Marks, L. Atlas, and M. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. Power Syst., vol. 6, no. 2, pp. 442–449, 1991.*

[8] W. Charytoniuk, M. S. Chen, and P.van Olinda, "Nonparametric regression based short-term load forecasting", IEEE Transactions on Power Systems, vol. 13, no. 3, pp. 725-730, Aug. 1998, doi: 10.1109/59.708572.

[9] H. H. Çevik, M. Çunkaş, "Short-term load forecasting using fuzzy logic and ANFIS," Neural Computing and Applications, vol. 26, no. 6, pp. 1355–1367, Aug. 2015.

[10] K. H. Kim, H. S. Youn, and Y. C. Kang, "Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method," IEEE Transactions on Power Systems, vol. 15, no. 2, pp. 559 - 565, May. 2000, doi: 10.1109/59.867141.

[12] Y. Chen, P. B. Luh, C. Guan, Y. G. Zhao, L. D. Michel, M. A. Coolbeth, P. B. Friedland, and S. J. Rourke, "Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks," IEEE Transactions on Power Systems, vol. 25, no. 1, pp. 322-330, Feb. 2010, doi: 10.1109/TPWRS.2009.2030426.

[13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," IEEE Transactions on Power Systems, vol. 16, no. 1, pp. 44-55, Feb. 2001, doi: 10.1109/59.910780.

[14] R. Zhang, Z. Y. Dong, Y. Xu, K. Meng, and K. P. Wong, "Short-term load forecasting of Australian National Electricity Market by an ensemble model of extreme learning machine," IET Generation, Transmission & Distribution, vol. 7, no. 4, pp. 391-397, Apr. 2013, doi: 10.1049/iet-gtd.2012.0541.

[15] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One-Hour-Ahead Load Forecasting Using Neural Networks," IEEE Transactions on Power Systems, vol. 17, no. 1, pp. 113 - 118, Jan. 2002

[16] V. H. Ferreira and P. A. Silva, "Toward estimating autonomous neural network-based electric load forecasters," IEEE Transactions on Power Systems, vol. 2, no. 4, pp. 1554-1562, Nov. 2007, doi: 10.1109/ TPWRS.2007.908438.

[17] B. J. Chen, M. W. Chang, and C. J. Lin, "Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001," IEEE Transactions on Power Systems, vol. 19, no. 4, pp. 1821-1830, Nov. 2004.

[18] S. Fan, and L. N. Chen, "Short-term load forecasting based on an adaptive hybrid method," IEEE Transactions on Power Systems, vol. 21, no. 1, pp. 392-401, Feb. 2006