Data-driven Time Series Based Prediction in Smart Home Appliance Energy Consumption

Md. Taksir Hasan Majumder Ahsanullah University of Science and Technology 141 142 Love Road, Tejgaon Industrial Area Dhaka, Bangladesh Sharmin Aktar Brac University 66 Mohakhali Dhaka, Bangladesh

ABSTRACT

The rapid development of human population has seen rapid expansion of electric consumptions in buildings and technological application based devices. The necessity of efficient energy management and forecasting energy consumption for devices and buildings know no bounds. Proper development and decision-making follows suite when such criteria are met. Building electrical energy forecasting method using artificial intelligence (AI) methods such as support vector machine (SVM) and artificial neural networks (ANN) is a potential approach for such purpose. In this paper, the possibility of a time-series based approach to an energy consumption prediction problem using state-of-the-art technologies such as LSTM and RNN is explored and hence proved that it indeed works.

General Terms

LSTM, Time Series, Energy Consumption

Keywords

LSTM, RNN, Time Series, Predictive Machine Learning, Smart Home, Energy Consumption

1. INTRODUCTION

The development rate of countries around the globe is increasing in a rate never seen before in human history. The world population growth rate in 2012 is estimated to be about 1.096%. The pattern of population rate around the globe is confirmed to have a growing rate each year, which is concerning. The growing pattern of human population indicates that the demand for accommodation, nation development and others will also keep increasing. Complementary to all these developments, superfluous energy is needed to drive the global demand; and at the same time, the environment needs to be kept safe. Furthermore, the rapid expansion of residential and commercial areas also contributes to the increase of building energy consumption, especially by the industrial growth. At the same time, environmental issues must be taken into account in the development bustle, in order to reduce pollution, carbon footprint and greenhouse effect.

In Europe, 40% of the total energy use and 36% of total carbon dioxide emissions come from buildings [4]. Energy uses in office building in China had been reported in 2008, which consumed

about 70300 kWh/m2 per year. This is 10 to 20 times higher than residential consumption. Besides that, the electricity consumption in Taiwan has increased to 229.20 billion kWh in 2009 from 82.65 billion kWh in 1990. These show that the forecast of energy consumption has become a crucial need in estimating energy usage; hence, the best plan for the development and environment is utterly required.

Besides the use of electrical appliances and devices in buildings, the geographical locations of a building also influence the electric energy consumption, and indirectly affects the forecasting analysis. Geographical differences of the building will make a difference in the use of electrical equipment. This differentiation is related to several factors associated to the particular country, such as the weather condition, surrounding temperature where the country is located, as well as the seasons. The electrical energy consumption generally increases about 2.6% during summer due to the rise of daily temperature. The use of air conditioning indeed consumes the most of electrical energy to provide comfort to the working space of the building. Cooling space will use a lot of energy if the building environment is quite hot. In contrast to temperate regions, the cooling space takes more time and uses less electricity. The main contributions of this paper are:

- —Smart Home Model Smart home systems can automatically take certain decisions. A smart home design with energy reading from appliances is proposed.
- —Long Short Term Memory Application in Time Series Data: Long Short Term Memory (LSTM) is a specific Artificial Neural Network (ANN) architecture, that aims at prediction in sequence-based data, such as speech, text, video, wave signals and so on. LSTM is used as an auto-regressor (which is traditionally implemented by Support Vector Regressiors (SVR) in this paper.
- —Auto-feature Extraction and Noise Removal: The model does not require manual feature extraction. Only data engineering is applied for feeding the data into the base code.
- —**Supervised Regression:** The problem is converted from purely time-series based to supervised regression, where previous moment's data is used as features for the moment for which the energy consumption of an appliance is to be predicted.
- —Multivariate Time Series Forecasting: The problem is also treated as a multivariate time series forecasting, where in order to predict the energy consumption of a problem at time t, energy

Methods	Model complexity	Easy to use	Running speed	Inputs needed	Accuracy
Elaborate engineering	Fairly high	No	Low	Detailed	Fairly high
Simplified engineering	High	Yes	High	Simple	High
Statistical	Fair	Yes	Fairly High	Historical data	Fair
ANNs	High	No	High	Historical data	High
SVMs	Fairly high	No	Low	Historical data	Fairly

Table 1.: Comparison of different approaches which have been adopted so far is shown here. SVM and ANN have most success among all these approaches.

consumption data from time 0 to t-1 is used. Hence, it can be considered a multi-variate time series forecasting adoption.

2. MOTIVATION

For a forecasting purpose, there are three categories of forecasting analysis, named as Short Term Load Forecasting (STLF), Medium Term Load Forecasting (MTLF), and Long Term Load Forecasting (LTLF). Each of these terms has different time ranges and different purposes of forecasting. The definition of time range is slightly different regarding to the research study. STLF usually takes time range that lasts for 24 h. STLF time range is also defined as time range between 24 h, but up to only one week. For MTLF time range, the time range covers from one week, up to one year analysis. For utility companies, they will use the MTLF analysis to estimate the load demand in long term, which is useful for maintenance operation. LTLF analysis involves a longer time frame than STLF and MTLF, and is usually conducted for future infrastructure plans. For forecasting purposes, the precision in forecasting is the main issue in conducting the forecasting analysis; however there are difficulties due to the complexity of the system. Now-a-days, there are many forecasting models developed by combining several methods, in order to solve the difficulties in forecasting and to find the best accuracy in forecasting.

2.1 WSN Perspective

Data can be obtained from measurements of temperature and humidity sensors from a wireless network, which can be used for prediction of weather of a local place and recorded energy use of home appliance. In this paper, data obtained from a wireless sensor (temperature and humidity) network can be used for appliance energy prediction.

Data filtering is needed to remove non-predictive parameters since the number of features is quite high. For the same reason, feature ranking plays an important role with this data. After that, different statistical models could be developed over this dataset. Results obtained from the models can be used to automate many appliances in a smart phone, for example, switching off an AC and switching on the room fan because it was predicted that the AC is going to consume too much energy.

3. RELATED WORK

There are several popular methods used for forecasting building energy consumption, which can be categorized into three, which are Engineering Method, Statistical Method and Artificial Intelligence Method. Among these methods, the most widely implemented method in forecasting is the Artificial Intelligence (AI) Method, which includes Artificial Neural Network (ANN) and Support Vector Machine (SVM). The other two methods of Engineering Methods and Statistical Methods are still applied, but some shortcomings have been identified in both these methods, causing the Artificial Intelligence Method to have priority for application in forecasting analysis. Among the deficiencies identified in Engineering Method are its complexity which contributes to the difficulties to apply it practically and its lack of input information [1]. Meanwhile, the Statistical Methods has been found to have lack of accuracy and not flexible [1].

The application of Artificial Intelligence Method in building energy system had been discussed by Krarti and Dounis in 2003 and 2010 [2, 3], which included forecasting, system modeling and control. For building energy consumption analysis, the ANNs are most widely implemented due to their best accuracy result and the ability of analyzing nonlinear problems. The ANNs also can learn from the historical pattern during analyzing the data. Besides that, it has capabilities in pattern recognition and pattern classification [29]. Some results from the Artificial Intelligence Methods and other methods from previous works are included for performance comparison in the next chapter. As referred from Table 1, as similarly reviewed by Zhao [1] in 2012, the accuracy of ANN and SVM seem to have better promising result for building energy forecasting, based on the historical data analysis. To evaluate the potential Artificial Intelligence capability, we had applied and tested the LSTM in buildings electrical energy consumption forecasting.

Gonzalez and Zamarreno [4], in 2005, used feedback neural network to predict the short term electricity load, which produced an excellent result for load forecasting in buildings. Furthermore, in their research, they suggested that there are three values that need to be studied, which are the number of neurons in hidden layers, the best size of data windows and the algorithm parameters.

Nonetheless, in order to get satisfactory result, too many neurons is not necessarily needed, as explained by Eric [5] in 1996. Yang et al. [6] used adaptive artificial neural networks in building energy prediction. There were two adaptive ANN models used in this study, which were accumulative training and sliding window training. In real measurements, the prediction results were found to be more accurate when using sliding window technique.

Karatasou et al. [7] discussed about the application of ANNs in predicting building energy consumption. In their study, the application of ANNs, together with statistical analysis, was found able to improve the applicability of ANNs in modeling and predicting building energy consumption.

The Artificial Neural Network (ANN) was used by Azadeh et al. [8] in 2008 to forecast long term electricity consumption in energyintensive manufacturing industries. The comparison between Table 2. : The following features are some of the most important features. Since we adopt a time-series based solution, time-stamps of the recorded values are necessary.

Feature	Description
date	time year-month-day hour:minute:second
Appliances,	energy use in Wh lights
T1	energy use of light fixtures in the house in Wh
RH_7	Temperature in ironing room , in Celsius
T7	Humidity in laundry room area, in %
Wind speed	(from Chievres weather station), in m/s

the ANN and nonlinear regression model using error analysis showed that the ANN has good forecasting value for electricity usage.

Using ANN is not something new in forecasting energy consumption. However, earlier attempts were centered around using simple Multilayer Perceptrons (MLP) to predict energy consumption. The works in [9, 10, 11] reflect this, as the authors use MLP to build a predictive algorithm. The problem with using MLP is that it can not take the flow of time into account. The sequence information is very crucial in such problems, and hence MLP can not reliably utilize these time-driven features. The authors in [12] attempts to mitigate this by using hourly feedback data, which somewhat improves the performance. The authors in [13] demonstrates comparative performance of various types of approaches, all of which do not utilize time driven data. The works in [14, 15, 16, 17] also suffer from the same problem.

4. PROPOSED METHOD

A smart home system where appliances are connected with sensors is assumed. The sensors monitor power consumption of the appliances and periodically send them to the base station. Figure 1 demonstrates the proposed setup for such a system.



Fig. 1: An IOT-based smart home model for prediction of appliance energy consumption

4.1 Recurrent Neural Networks (RNNs)

RNNs are heavily used in NLP tasks. The big advantage of the RNN is that it is able to effectively use data from previous time steps. Figure 2 is what a small piece of an RNN looks like.

At the bottom there are word vectors (xt, xt-1, xt+1). Each of the vectors has a hidden state vector (ht, ht-1, ht+1), which also



Fig. 2: Recurrent Neural Nets for Sequence

correspond to the same time stamp. The hidden state in each module of the RNN is a function of both the current vector and the hidden state vector at the previous time step. The state at time t follows the equation shown at Figure 3.

$$h_t = \sigma \left(W^{(hh)} h_{t-1} + W^{(hx)} x_{[t]} \right)$$

Fig. 3: State at time, t

There is a weight matrix W^{hx} which will be multiplied with the input, and there is a recurrent weight matrix W^{hh} which is multiplied with the hidden state vector at the previous time step. These recurrent weight matrices maintain the same value across all time steps. This is the key point of RNNs. It is very different from a traditional 2 layer NN for example. In that case, there is a distinct W matrix for each layer (W1 and W2). Here, the recurrent weight matrix is the same through the network. To get the output (Y) of a particular module, this will be h times W^S , which is another weight matrix.

The most distinct difference from a traditional NN is that an RNN takes in a sequence of inputs (words in this case). This can be contrasted to a typical CNN where just a singular image works as an input. With an RNN, however, the input can be of any length, given it represents a sequence. Additionally, the order of inputs in this sequence can largely affect how the weight matrices and hidden state vectors are finalized after training is finished, unlike ANNs. The idea is that, the hidden states, after training, will exploit the values of previous steps to capture the information from the past (the previous time steps). Figure 4 shows the state loop.



Fig. 4: State loop

4.2 LSTM

Long Short Term Memory networks, usually just called LSTMs, are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hochreiter & Schmidhuber [18]. They work tremendously well on a large variety of problems, and are now widely used.

LSTMs are explicitly designed to avoid the long-term dependency problem. All recurrent neural networks have the form of a chain of repeating modules of neural network. In standard RNNs, this repeating module (as seen in Figure 5) will have a very simple structure, such as a single tanh layer. However, more complexities can be added into the structure of an LSTM (as seen in Figure 6). The equations of the gates after such variations are included can be seen in Figure 7.



The repeating module in a standard RNN contains a single layer.

Fig. 5: Repeating modules



The repeating module in an LSTM contains four interacting layers.

Fig. 6: LSTM variation

5. EXPERIMENTAL RESULTS

5.1 Dataset

An open dataset is used due to the lack of a physical model. The dataset can be found here at, goo.gl/MPPE6o. Some features are

Input gate (current cell matters) $i_t = \sigma \left(W^{(i)} x_t + U^{(i)} h_{t-1} \right)$ Forget (gate 0, forget past) $f_t = \sigma \left(W^{(f)} x_t + U^{(f)} h_{t-1} \right)$ Output (how much cell is exposed) $o_t = \sigma \left(W^{(o)} x_t + U^{(o)} h_{t-1} \right)$ New memory cell $\tilde{c}_t = \tanh \left(W^{(c)} x_t + U^{(c)} h_{t-1} \right)$

Fig. 7: Gates in LSTM

described in Table 2. Some examples of statistical distribution is given in Figure 8.

T5 numeric	Null 0%	Mean 19.6	15.3	25.8
RH_5 numeric	Null 0%	Mean 50.9	29.8	96.3
T6 numeric	Null 0%	Mean 7.91	-6.07	28.3
RH_6 numeric	Null 0%	Mean 54.6		99.9
T7 numeric	Null 0%	Mean 20.3		26.0

Fig. 8: Statistics of some of the features are shown here.

5.2 Data Engineering

Neural networks like Long Short-Term Memory (LSTM) recurrent neural networks are able to almost seamlessly model problems with multiple input variables. Usually classical linear methods can be difficult to adapt to multivariate or multiple input forecasting problems. Hence, time-series forecasting with LSTM models have more potential for the prediction of energy consumption.

- —The data is split into train and test by 80:20 ratio.
- —The data requires cleaning. The first step is to consolidate the date-time information into a single date-time so that it can be used as an index in Pandas Library. The first date column is converted into numeric sequence starting from Integer 1 (as seen from Table 3 and Table 4).
- —Several columns were removed since they are empty columns. —Next, all features are normalized, then the dataset is trans-
- formed into a supervised learning problem. All previous time series data (from t = 0 to t = t-1) are made available to t = t data using LSTM gates. Figure **??** shows the new dataframe after supervised formula.
- -Figure 9 shows the LSTM model used. 100 neurons were used in the LSTM layer.

5.3 Graphs

Figure 10 shows the Mean Squared Error (MSE) after the training of the LSTM in different epochs with only t = t - 1 supervision. Figure 11 shows the Mean Squared Error after the training of the LSTM in different epochs from t = 0 to t = t-1 supervision. Clearly, the more history the model can look into, the better it does.

Table 3. : The data contains exact time-stamp which can not be proper input to an LSTM. However, the consistent time-gap among the time-stamps allows the exploitation of a typical time to integer conversion.

Serial No	Date	Appliance
1	17:00	60
2	17.10	60
3	17:20	50
4	17:30	50
5	17:40	60
6	17:50	50
7	18:00	60
8	18:10	60
9	18:20	60

Table 4. : Here, for each 10-minute progression, the equivalent integer value of the time-stamp is increased by 1. As a result, the time-series data from 3 is now exploitable by the LSTM network.

Serial No	Date	Appliance
1	1	60
2	2	60
3	3	50
4	4	50
5	5	60
6	6	50
7	7	60
8	8	60
9	9	60

Layer (type)	Output	Shape	Param #
lstm_1 (LSTM)	(None,	100)	52000
dense_1 (Dense)	(None,	1)	101
Total params: 52,101			

Trainable params: 52,101 Non-trainable params: 0

Fig. 9: LSTM model architecture



Fig. 10: LSTM model with less time-series based data performs poorly. The x-axis denotes the number of epoch of training, whereas the y-axis denotes the validation loss.



Fig. 11: LSTM model exhibits stronger performance when it is allowed to look into more of past data. The x-axis denotes the number of epoch of training, whereas the y-axis denotes the validation loss.

Table 5. : MSE loss has comparatively less reduction rate after several epochs during earlier training phase. After the loss optimization is stabilized during 70^{th} epoch, the training is stopped.

Time taken	Epoch Number	Training Loss	Validation Loss
5	45	0.0047	0.0044
5	46	0.0047	0.0044
5	47	0.0047	0.0044
5	48	0.0047	0.0044
5	49	0.0047	0.0044
5	50	0.0046	0.0043
5	51	0.0046	0.0043
5	52	0.0046	0.0043
5	53	0.0046	0.0043
5	54	0.0046	0.0043
5	55	0.0046	0.0043

Clearly, the final LSTM model can generalize auto-regression quite well given from the low MSE value of 0.43 only out of over 14000 samples training. Table 5 shows the the training of the LSTM network stabilizes along with the MSE value of validation data.

6. CONCLUSION

The building of energy consumption forecasting using LSTM as an autoregressor is presented here. Time series method is currently used in the field of buildings energy forecasting and has received much attention from many researchers, with regard to its advantage to cope with the complexity of buildings system that is influenced by many buildings parameters. Among the AI methods, the Neural Network and Support Vector Machine are the widely used models. The researches conducted using these models mostly aimed to find better forecasting performances, in terms of the forecasting accuracy. However, the experiments in the paper show that LSTM can be fruitful to due to their sequence-based data exploration nature. The low MSE value for LSTM autoregression from multivarite supervision supports such claim. Hence, such a method can be considered for building smart home systems, which may leverage energy consumption prediction of electronic appliances. Although traditional approaches such as SVMs and

ANNs can somewhat predict energy consumption to a certain degree, advanced sequence-based neural networks such as LSTM can be exploited as an autoregressor for the same purpose more effectively. Clearly, the LSTM approach can be utilized in such a real world application.

7. REFERENCES

- Zhao, Hai-xiang, and Frdric Magouls. "A review on the prediction of building energy consumption." Renewable and Sustainable Energy Reviews 16.6 (2012): 3586-3592.
- [2] Krarti M. An overview of artificial intelligence-based methods for building energy systems. J Sol Energy Eng 2003;125(3):12.
- [3] Dounis AI. Artificial intelligence for energy conservation in buildings. Adv Build Energy Res 2010;4(1):267?99.
- [4] Gonzalez Pedro A, Zamarreno JM. Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. Energy Build 2005;37 (6):595?601
- [5] Eric SBG. Prediction and control using feedback neural networks and partial models. Zurich: Swiss Federal Institute of Technology Zurich; 1996.
- [6] Yang Jin, Rivard Hugues, Zmeureanu R. On-line building energy prediction using adaptive artificial neural networks. Energy Build 2005;37(12):1250?9.
- [7] Karatasou S, Santamouris M, Geros V. Modeling and predicting building's energy use with artificial neural networks: methods and results. Energy Build 2006;38(8):949?58.
- [8] Azadeh A, Ghaderi SF, Sohrabkhani S. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. Energy Convers Manage 2008;49(8):2272?8.
- [9] Ekonomou, L. "Greek long-term energy consumption prediction using artificial neural networks." Energy 35.2 (2010): 512-517.
- [10] Kalogirou, Soteris A., and Milorad Bojic. "Artificial neural networks for the prediction of the energy consumption of a passive solar building." Energy 25.5 (2000): 479-491.
- [11] Ekici, Betul Bektas, and U. Teoman Aksoy. "Prediction of building energy consumption by using artificial neural networks." Advances in Engineering Software 40.5 (2009): 356-362.
- [12] Gonzalez, Pedro A., and Jesus M. Zamarreno. "Prediction of hourly energy consumption in buildings based on a feedback artificial neural network." Energy and buildings 37.6 (2005): 595-601.
- [13] Tso, Geoffrey KF, and Kelvin KW Yau. "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks." Energy 32.9 (2007): 1761-1768.
- [14] Hoeynck, Michael, and Burton W. Andrews. "Sensor-Based Occupancy and Behavior Prediction Method for Intelligently Controlling Energy Consumption Within a Building." U.S. Patent Application No. 12/183,361.
- [15] Panizza, M., et al. "Electrochemical treatment of wastewaters containing organic pollutants on borondoped diamond electrodes: prediction of specific energy consumption and required electrode area." Electrochemistry communications 3.7 (2001): 336-339.
- [16] Dong, Bing, Cheng Cao, and Siew Eang Lee. "Applying support vector machines to predict building energy con-

sumption in tropical region." Energy and Buildings 37.5 (2005): 545-553.9.

- [17] Inoue, Koji, Tohru Ishihara, and Kazuaki Murakami. "Way-predicting set-associative cache for high performance and low energy consumption." Proceedings. 1999 International Symposium on Low Power Electronics and Design (Cat. No. 99TH8477). IEEE, 1999.
- [18] Hochreiter, Sepp, and Jrgen Schmidhuber. "Long shortterm memory." Neural computation 9.8 (1997): 1735-1780.