A Survey of Machine Learning's Electricity Consumption Models

Umar Farouk Ibn
Abdulrahman
Christ Apostolic University
College, Kwadaso, Ghana
Kwame Nkrumah University of
Science and Technology, Kumasi,
Ghana

Michael Asante
Kwame Nkrumah University of
Science and Technology, Kumasi,
Ghana

James Ben Hayfron-Acquah
Kwame Nkrumah University of
Science and Technology, Kumasi,
Ghana

ABSTRACT
Electricity is a very important commodity used for both domestic and industrial purposes. It is generated from many sources which include the thermal, coal, nuclear and hydro. Its demand is increasing on regular basis as result of the ever increasing world population coupled with other socio-economic factors. This therefore requires effective predictions of the future needed electricity to sustain it demand. However, predicting the exact amount of electricity for all times is a challenge. Over predictions can lead to wasteful investment whiles under predictions can lead to inadequate electricity supply with eventual blackouts, social unrest and low economic growth. The aim of this paper is to present the various electricity consumption predictions models indicating the machine learning algorithm and the variables used in the modeling

General Terms
Machine learning, electricity prediction

Keywords
Machine learning algorithm, electricity, consumption models, variables

1. INTRODUCTION
Electricity is a very important commodity needed for human life. It is used in houses for heating, cooking, ironing and other important aspects of living. Industries use it for powering plants and production. Offices also use it for their daily operation like powering the computers, use of air condition and so on. It is generated from many sources which include, the thermal, coal, nuclear and hydro. The existence of sufficient supply of electricity helps to propel economic growth and stability of a country. This in turn gives rise to a lot of economic activities and increase in Gross Domestic Product.

Lately, there is a sharp increase in electricity demand as a result of the ever increasing use of various gadgets, tools and increase in population the world over. Inaccurate estimation of the required electricity needs would either lead to power shortages or wasteful investments which may have damming consequence to an economy [1]. Therefore an efficient predictive electricity consumption model is very important tool to every utility provider and a nation at large. It would serves as tool for policy anddecision making for the utility provider. Literature has presented different electricity consumption models, notably on machine learning and classical models. This paper attempts to shed-light on the various works, indicating the machine learning algorithm and variables used in the model building

1.1 Machine Learning
Machine learning is the use of computers to intelligently discover knowledge for the purpose of making decisions [2]. It is grouped into two main types, namely supervised learning and unsupervised learning. In supervised learning all data are labelled such that the algorithm learns to predict the output based on the input data. Where as in unsupervised learning, the data is unlabelled but the algorithm learns the structure from the input data [3].

According to [4] there are nine most frequently used machine learning algorithms, which includes k-means, linear support vector machines (LSVM), logistic regression (LR), locally weighted linear regression (LWLR), Gaussian discriminant analysis (GDA), back-propagation neural network (BPNN), expectation maximization (EM), naive Bayes (NB), and the independent variable analysis (IVA). However in Electricity demand models, other ML's like optimised gene programming [5] and Support Vector Regression[6] have been reported which are not part of[4] assertion.

The most applied machine learning algorithm in electricity predictions are the Artificial Neural Network (ANN) and Support Vector Regression (SVR).[6], [7]. ANN is designed to mimic the working of the human brain whose architecture consists of the input, sometimes with hidden and the output layers respectively which are connected to each other. Each input is assigned a weight which is multiplied to the activation function to obtain the output
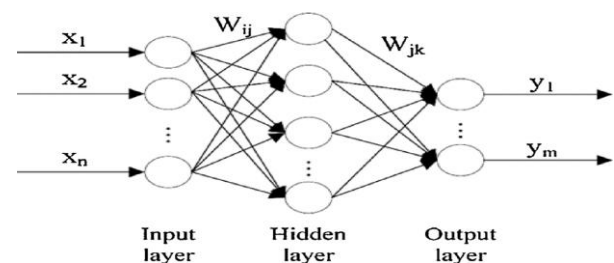

Fig-1 Neural Network Architecture: (Source: [9])

1.2 Variable
Variable plays the most important role in model development and forecasting of electricity consumption. [6] that, building a model requires the identification of variables and parameters that contribute to electricity consumption of a country. Wrong choices of variables can either lead to over projected values or

under projected values.

Literature has reported divergent and the justification for inclusion of various variables in building electricity consumption models. However these vary from country to country and electricity needs. Most variables used for projection are socio-economic factors like Foreign Direct Investment (FDI), Gross Domestic Product (GDP), Import, Export, population and so on. For example, it is thought that, an increase in population would invariably lead to increase in electricity consumption and a small population would translate to small electricity consumption. In the case of exports, a country would need reliable and sufficient electricity to produce goods for exports. Therefore a country that exports more goods would need a lot of electricity all things being equal. Also, many countries import goods that are electro-centric like mobile phones, television sets, computers, electric-ovens, refrigerators and so on. For a country whose large population uses those gadgets would push for high increase in electricity consumption. Alternatively if the country does not depend on those gadgets, the country is likely to record low electricity consumption. GDP measures the wealth of people living in a country at a particular time. It is argued that, as people's wealth increases, there is a high probability of life style change. The people tend to demand certain goods or service that would include many electrical appliances which directly increase the country's electricity consumption. Therefore an increase in GDP would also lead to increase in electricity demand, all things being equal. However the inclusion of these variables in any predictive modeling requires empirical evidence.

2. ELECTRICITY CONSUMPTION MODELS

Electricity predictions models thrive on the choice of appropriate machine learning algorithms and the use of appropriate input variables. Electricity consumption models have looked at short term, events, households and some long period consumption/demand. For example, [10] used linear regression models to forecast electricity consumption of Italy using a dataset of 1970-2009. In building the model, historical consumption, GDP, Gross Domestic product per capital and population were used as input variables. They opined that, the model was compared with national forecast which showed a deviation of +-1 in the best case and +-1 in the worst case respectively which was attributed to the time span.

[11] used ANN to predict long term energy consumption for Greek. The multi layer perception a type of feed forward network together with Yearly ambient temperature, installed power capacity, yearly per resident electricity Consumption, G.D. P were used to provide accurate prediction for the years 2005-2008, 2010, 2012 and 2015 respectively. The work was compared with similar implementation using regression model and support vector machines, the results of the ANN out-performed the Regression and Support vector

[12] used two heuristic algorithms to estimate the electricity demand for Iran. The estimation was done using Gross Domestic Product, population, number of Customers and average price of electricity. Data for the research was based on 21 years (1980-2000) actual data which was obtained from the ministry of power which provided average errors of 1.36% for the linear model and 1.56% for the non-linear models respectively.

[6] developed a model to predict Turkey's electricity

consumption using Support vector regression. Four socio-economic factors were employed in building the model. These are population, GDP, import and export respectively.

[7] used Elmart Recurrent Network to predict the consumption of household electricity for the City of Palermo, Italy. The research achieved a percentage prediction error of 1.5% for the mean and 4.6% for the maximum error.

[13] used seven (7) different M.L approaches to model a predictor for hourly residential electrical consumption. The aim was to identify which is the best for predicting the next hour residential building consumption. The prediction was based on data obtained from residential homes called Campbell Greek data set. Result showed that Least Squares Support Vector Machine (LS-SVM) is the best approach for predicting every house future electrical consumption and NN works better for commercial building. The accuracy of the models was verified using C.V., MAPE, and Mean Bias Error but did not indicate the accuracy level.

[14] employed particle swarm optimisation and artificial bee colony to estimate long term electricity demand/consumption for Turkey. Four inputs namely Gross Domestic Product, population, import and exports were used in building the forecasting models. The two models were compared and the Colony optimisation showed better performance.

[1] used both regression and neural network to predict the energy demand for the residential sector of the U.S.A. Seven independent variables namely resident population, GDP, household size, median household income, cost of residential electricity, cost of residential natural gas and cost of residential heating oil which all came from different sources. The research concluded that the two models show different trends but their performance where similar at the test level. They did not indicate any accuracy testing metrics used.

[15] present architecture to forecast accurate electricity demand andgeneration. The proposed architecture consisted of stream and predictive analytics. The stream engine would be responsible for collecting data from the smart metres, analyzing and segmenting the categories of customer. Also the predictive analytics made use of linear Multivariate regression which was used for the prediction. Finally, an experimental result was conducted using five days data which was obtained from Red Electrical Espanol. A comparison was made with Multivariate linear regression with Adaptive Windowing but the Multivariate regression performed better

[16] developed a long term electrical energy consumption model using optimised regression and ANN based on different historical data types for developed and developing countries (USA and IRAN). The research concluded that, forecasted electricity consumption was somehow different and improved particle swam (ipso)-ANN is most accurate for prediction regardless of the type of economy. A mean absolute percentage error of 1.94% and 1.51% were achieved for Iran and the USA respectively. However they employed GDP, Energy imports, Energy exports, and population as the input variables

[17] predicted annual electricity consumption for Iran using - Simple recursive Neural Network with time delay (NARX ANN), Multi-Layer Perceptron (MLP) and Auto Regressive Integrated Moving Average (ARIMA). Data from 1982-2010 together with population, Gross National Product(GRIP), import and export were used for the predictive model. A comparison was made of the three (3) models and the NARX

ANN performed better than the MLP and ARIMA time series models.

[9] worked on energy forecasting for event venues. They used two machine learning algorithms namely the Feed Forward Neural Network and Support vector Regression which was applied to three different time granularities namely Daily, *Hourly and 15 minutes*. Data for the research was obtained from the London Hydro for the period 1 January 2013 to 31 March, 2014 with a total of 43,680 points which includes the reading date, time, consumption and demand. The authors concluded that ANN works better for the daily predictions than SVR but no superiority over the other in terms of hourly and 15 minutes data.

[18]used ensemble of Neural Network to predict the consumption of heating energy for the Norwegian University of Science and Technology campuses. Data for the research was obtained from heating meters, electricity consumption and the main meter installed by the district supplier. They further used outside mean daily temperature (tm), heating consumption of the previous day (HCp),day of the week (WD), maximum daily temperature (tmax), relative humidity (φ),total daily solar radiation (SR) and month of the year(MY) as the input variables. To determine the accuracy of the model, Root Mean Square (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient of determination were used.

[19]modeled the annual electricity demand of Turkey for the years 2014-2018 using multi-layer Artificial Neural Network and six variables namely population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature were used. The significance of each was first tested statistically which proved that inflation percentage and average winter are not too significant according their criterion and therefore were dropped in the final modelling.

[20]Birim and Tumturk (2016) designed a model to forecast electricity consumption for Turkey using ANN with four (4) independent input variables. The variables are population, import, employment and natural gas. However, the variables were determined after an initial identification of six variables which included export and GDP but the two (2) were not found statistically relevant after running multiple regressions on all the variables. The predicted results differed slightly from the official results but were attributed to values of the year 2009.

[21] used Time Delay Artificial Neural Network to predict short term and photo voltaic smart meter profiles and also investigated different spatial aggregation levels. Data for the research were obtained from 1WB a public utility of the city of Basel for a period between April 2014 and September, 2015 which consisted of 40,000 small consumers and 15,000 large consumers. Mean Absolute Percentage Error (MAPE) and Normalised Root Mean Square Error (NRMSE) were employed to evaluate the performance of the forecasting algorithm which achieved 3.36% MAPE and 5.26% for the sum of all available load profiles for small household and a MAPE of 12.6%. For commercial and industrial loads, a MAPE of 1.96 % was achieved showing a great significance

[5] employed optimised gene expression programming with GDP , Population, Import, and Export of goods and services to forecast long term (2017-2030) electricity consumption for five Asian countries. Two different historical data were used. The results outperformed ANN, SVR, RULE BASED and optimised quadratic rule.

[22]used deep recurrent neural network to predict electricity consumption for both commercial and residential building for medium to long term horizon. They reported that the model Recurrent Neural Network model *A* and *B* performed better than the 3 layered perceptron for commercial buildings but the model performs less to Multi Layer Perceptron for 1 year horizon as the number of buildings increases.

Table -1: Summary of Electricity Consumption Prediction Models

Author(s)	M.L Algorithm	Variables Used	Number of variables
Bianco et al (2009)	Linear Regression	Historical, consumption, GDP, GDP per capital and population	4
Ekornomou (2010)	Artificial Neural Network	Yearly ambient temperature, Installed power capacity, Yearly per resident electricity Consumption, GDP	4
Amjadi et al (2010)	two heuristic algorithms	Gross Domestic Product, population, number of Customers and average price of electricity	4
Kavaglioglu (2011)	Support Vector Regression	Population, GDP, import and export	4
Marvuglia and Messineo (2012)	Elmart recurrent Artificial Neural Network	weather data, Historical Time series, HC Index(Estimated No of AC in Dwellings)Humidex (estimate discomfort)Radiation , atmospheric pressure	6
Edwards et al (2012)	Least Squares Support Vector Machine	Sensor Data	1
Kiran et al(2012)	Particle Swarm & Artificial Bee Colony	GDP, Population, import and export	4

Kialashaki and Reisel(2013)	Regression and Neural network	Resident population, GDP, Household size, Median household income, Cost of residential electricity, Cost of residential natural gas, Cost of residential heating oil	7
Couceiro et al (2014)	Multivariate Regression	Meteorological variables (basically temperature, humidity and wind speed)	
Ardakani and Ardehali (2014)	Optimised Regression and ANN	GDP, Energy imports, Energy exports and population	4
Kargar and Charsoghi (2014)	NARX Neural Network, Perceptron NN and ARIMA	Population, Gross National Product (GRIP), import and export.	4
Gunay (2016)	Artificial Neural Network	Population, GDP per capita, inflation percentage, unemployment percentage, average summer temperature and average winter temperature	6
Birim and Tumturk (2016)	Multi Linear Regression and Artificial Neural Network	population, import, employment and natural gas	4

Table 1 presents various electricity consumption models indicating the machine learning algorithm and the variables used, which are the determinants of electricity consumption

3. DISCUSSION

The essence of building electricity predictive models is to offer a solution that would optimally forecast the needed electricity by the populace at all time. Albeit models use machine learning predictive algorithms like Neural Network, Support Vector Regression and so on, they thrive much on variables used, which would invariably increase the efficiency or reduce the efficacy of the built model. The choices and inclusion of variables are country or demand specific. Therefore those choices are informed by their relevance which are determined statistically, Gunay (2016).

4. CONCLUSION

In this paper, various machine learning electricity predictive models have been presented. Firstly, the various machine learning algorithms used for electricity forecasting have been indicated. Secondly due to the importance of variables in model building, the choices for the model development have been indicated where necessary.

5. REFERENCES

[1] Kialashaki, A., & Reisel, J. R. (2013). Modeling of the energy demand of the residential sector in the United States using regression models and artificial neural networks. *Applied Energy*, *108*, 271-280

[2] Wang, L. (2016). Machine learning in big data. *International Journal of Advances in Applied Sciences*, *4*(4), 117-123.

[3] ML, 2018, https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-alg2018orithms/ Accessed on 18th February, 2018

[4] Tu, C., He, X., Shuai, Z., & Jiang, F. (2017). Big data issues in smart grid–A review. *Renewable and Sustainable Energy Reviews*, *79*, 1099-1107.

[5] Kaboli, S. H. A., Fallahpour, A., Selvaraj, J., & Rahim, N. A. (2017). Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming. *Energy*, *126*, 144-164.

[6] Kavaklioglu, K. (2011). Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. *Applied Energy*, *88*(1), 368-375.

[7] Marvuglia, A., & Messineo, A. (2012). Using recurrent artificial neural networks to forecast household electricity consumption. *Energy Procedia*, *14*, 45-55.

[8] Ahmad, A. S., Hassan, M. Y., Abdullah, M. P., Rahman, H. A., Hussin, F., Abdullah, H., & Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, *33*, 102-109.

[9] Grolinger, K., L'Heureux, A., Capretz, M. A., & Seewald, L. (2016). Energy forecasting for event venues: big data and prediction accuracy. *Energy and Buildings*, *112*, 222-233.

[10] Bianco, V., Manca, O., & Nardini, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, *34*(9), 1413-1421.

[11] Ekonomou, L. (2010). Greek long-term energy consumption prediction using artificial neural networks. *Energy*, *35*(2), 512-517.

[12] Amjadi, M.H., Nezamabadi-pour, H and Farsangi, M.M, Estimation of electricity demand of Iran using two heuristic algorithms,Energy Conversion and Management 51 (2010) 493–497

[13] Edwards, R. E., New, J., & Parker, L. E. (2012). Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings*, *49*, 591-603

[14] Kıran, M. S., Özceylan, E., Gündüz, M., & Paksoy, T. (2012). A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey. *Energy conversion and management*, *53*(1), 75-83.

[15] Couceiro, M., Ferrando, R., Manzano, D., & Lafuente, L. (2012, May). Stream analytics for utilities. Predicting power supply and demand in a smart grid. In *Cognitive Information Processing (CIP), 2012 3rd International Workshop on* (pp. 1-6). IEEE.

[16] Ardakani, F. J., & Ardehali, M. M. (2014). Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. *Energy*, *65*, 452-461.

[17] Kargar, M. J., & Charsoghi, D. K. (2014). Predicting annual electricity consumption in Iran using artificial neural networks (NARX). *Indian J. Sci. Res*, *5*(1), 231-242.

[18] Jovanović, R. Ž., Sretenović, A. A., & Živković, B. D. (2015). Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, *94*, 189-199.

[19] Günay, M. E. (2016). Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. *Energy Policy*, *90*, 92-101.

[20] Birim, Ş., & Tümtürk, A. (2016). Modeling and Forecasting Turkey's Electricity Consumption by using Artificial Neural Network. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *25*(1), 192-208.

[21] Zufferey, T., Ulbig, A., Koch, S., & Hug, G. (2016, September). Forecasting of smart meter time series based on neural networks. In *International Workshop on Data Analytics for Renewable Energy Integration* (pp. 10-21). Springer, Cham.,

[22] Rahman, A., Srikumar, V., & Smith, A. D. (2018). Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Applied Energy*, *212*, 372-385.

[23] Tso, G. K., & Yau, K. K. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, *32*(9), 1761-1768.