ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2020.0300

FORECASTING ELECTRICITY CONSUMPTION OF RESIDENTIAL USERS BASED ON LIFESTYLE DATA USING ARTIFICIAL NEURAL NETWORKS

Isaac Kofi Nti[1], Moses Teimeh[2], Adebayo Felix Adekoya[3] and Owusu Nyarko-Boateng[4]
[1]*Department of Computer Science, Sunyani Technical University, Ghana*
[2]*Department of Electrical and Electronic Engineering, Sunyani Technical University, Ghana*
[3,4]*Department of Computer Science and Informatics, University of Energy and Natural Resources, Ghana*

Abstract

Electricity is the lifeline of almost everything in this 21st century. Residential electricity consumption has seen an increase both locally and globally. Therefore, it has become a global concern of significant importance to promote electrical energy consumption reduction (energy conservation) within the household for a viable development of a nation in the case of resource limitations. The current study seeks to identify the social psychology (lifestyle) factors that significantly influence the residential electricity consumption, and predict future electricity consumption using an artificial neural network (ANN) based on lifestyle data collected from three hundred and fifty (350) households in the Sunyani Municipality. The performance metrics RMSE, MSE, MAPE, and MAE, were used to estimate the performance of the proposed model. The RMSE (0.000726) and MAE (0.000976) of the proposed model compared to (RMSE = 0.0657 and MAE = 0.05714) for Decision Trees (DT) and (RMSE = 0.08816 and MAE = 0.06911) for Support Vector Regression (SVR) shows a better fit of the proposed model. Furthermore, it was observed that the type of vehicle (saloon or sport utility vehicle) used by the head of a household was the most significant lifestyle feature in forecasting residential electricity consumption. Future studies would focus on developing a vigorous model using a combination of weather parameters and several socio-economic factors based on hybrid machine-learning algorithms to increase forecasting accuracy.

Keywords:

Load Forecasting, Lifestyle Data, Hybrid Machine-Learning, Artificial Neural Network

1. INTRODUCTION

Electricity, with the passage of time, has become one of the essential forms of energy to humankind in light of the explicit fact that almost everything depends on electricity [1] [2]. The quantity of electrical energy consumed has become one of the critical concerns of the electricity industry for strategic planning and expansion [3]. Hence, the local and global energy producers consider that the efficient use of electrical energy is a key aspect to be addressed as a priority. Consequently, energy efficiency is a crucial challenge for building sustainable societies. However, the world's primary energy consumption is estimated to increase by 1.6% yearly as a result of increasing incomes, growing populations and the industrialisation of developing countries [4]. This scenario raises issues associated with the increasing paucity of natural resources, the increase in environmental pollution, and the looming menace of global climate change. This warrants a call for the efficient management of electrical energy by both domestic and industrial consumers.

Studies show that 30% of the electricity generated globally is consumed by residential users [5]. In 2015, the global average yearly electricity consumption by residential users reached 10,812kWh. Again, a report in 2017 revealed that 25,019kWh/year of electricity was used in Norway, 11,974kWh/year in the US, 6,400kWh/year in France, 4,656.52kWh/year in the UK, 26kWh/year in Sierra Leone, and an average of 21 372TWh globally, which is 2.6% higher than in 2016 [6] [7]. Similarly, residential electricity consumption in Ghana for the past nine years kept increasing in value, thus 1,996GWh in 2007, 2,168GWh in 2008, 2,275GWh in 2009, 3,060GWh in 2013 to 3,932GWh in 2016 [8].

Notwithstanding, the increase in residential electricity consumption, Nishida et al. [9] suggest that residential (domestic) energy consumption differs depending on the lifestyle of the family. A family lifestyle, according to [10], is a set of factors such as family composition, house type, age, home appliances possessed and their usage, family income, cultural background, social life and lifestyle habits which include how long to stay at home and how to spend holidays. For this reason, it is difficult to grasp the factors that have significant influences on electricity consumption by a residential consumer.

Despite the difficulties associated with electrical load forecasting, knowing the expected residential load demand is desirable for electrical energy generators and distributors to make the right decision ahead and for consumers to know how future energy consumption (demand) will change in line with their future lifestyle [1] [11]. Further, it helps to accurately determine the right time for buying and selling electrical energy, which contributes to costs savings or even earning an income. Moreover, Kong et al. [12] argue that as the power system is facing an evolution toward a new flexible, intelligent, and collaborating system with sophisticated infiltration of renewable energy generation, a short-term electric load forecasting for residential (individual) electricity customers, plays a progressively more essential role in the future grid planning and operation [12].

In [2], the authors argue that the basic unit of electrical energy consumption is the home. Hence, the reduction in residential electricity consumption will reduce the consumption of electrical energy in society. This raises the need to find the features/factors that have significant effects on household electricity consumption, and to address such features/factors to minimize electricity consumption. Simultaneously, the consumers could also be informed of their energy consumption pattern in the future, so that they could plan in time and make an appropriate decision. Therefore, in affirmation to [13] opinion, energy management systems (EMS) are required to monitor the generation, distribution, consumption, and storage and also to make the paramount decisions according to input signals and the user's requirements and preferences.

Based on the issues detailed above, the current study seeks to propose a predictive model for useful and accurate prediction of future electricity demand by domestic electrical energy users based on an artificial neural network (ANN) using comprehensive the lifestyle data about the household through a questionnaire. The general goal of the current study is broken into three sub-objectives as follows:

- To identify which lifestyle data is a high predictor of residential electricity consumption.
- To propose a machine-learning framework to predict future residential electricity consumption based on household lifestyle data.
- Measure the performance of the proposed model against other state-of-the-art models based on accuracy and error metrics.

To reduce the amount of energy consumption by domestic electricity users, one needs to understand the energy needs and consumption patterns of the consumers and predict the expected consumption of customers in advance. By doing so, the following benefits could be in order:

- An accurate prediction of energy demands could provide useful information to make decisions on energy generation and purchase,
- An accurate prediction would have a significant impact on preventing overloading and allowing efficient energy storage by supply authorities,
- Further, if consumers know their tendency for energy consumption in the future, it would help them to make informed decisions on their purchase of household appliances, daily usage of household appliances, and financially be prepared for the same.

In this paper, we aim to answer the following questions:

1. To what extent could the lifestyle data of the consumers of electrical load be predicted?
2. How possible is it to predict the electrical load reliably?

The remaining sections of the current paper are organised as follows: section 2 presents a review of the pertinent literature on electricity demand predictions, the social-psychological factors that influence residential electricity consumption, methods for modelling energy consumption, and machine learning and its application in different fields. In section 3, we discuss the materials and methods adopted for the current study. Section 4 presents the experimental setup, the outcome of the study, and discussions. Section 5 concludes the study and outlines the direction for future studies.

2. LITERATURE STUDY

This section discusses the pattern of electrical energy consumption globally and locally, the social-psychological factors the affect residential electrical energy consumption, methods of modelling energy consumption, machine learning, and its application in energy modelling and finally related studies on electricity prediction.

2.1 ELECTRICITY CONSUMPTION LOCALLY AND GLOBALLY

In 1990, 71% of the people in the world had access to electricity. This figure has risen over the years that in 2016 it became 87% and 88.87% in 2017. The primary reason for this drastic increase is the constantly growing urbanization [14] - [16].

Electrical energy consumption may be classified as residential (domestic), commercial (non-residential) or industrial. Residential or domestic refers to the home or a dwelling where people globally live from day-to-day. Also, domestic electricity consumption happens on a considerably small scale compared to commercial users (businesses) who deal with heavy-duty machinery and lighting or appliances [17]. Fig.1 shows the electricity consumption per capita (kWh per person) in 2017. The individual consumption (IC) of electricity in every country using Eq.(1) shows that there is a higher disparity in electricity consumption among developed countries.

$$IC = \frac{\text{Yearly consumption of country}}{\text{Population of country}} \quad (1)$$

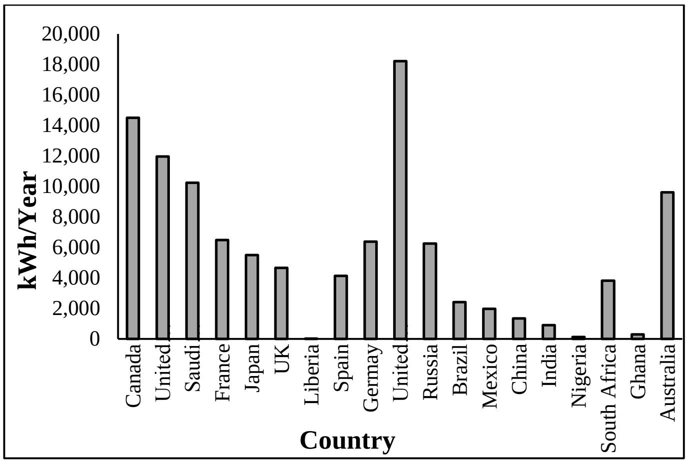

Fig.1. Electricity consumption per capita [6]

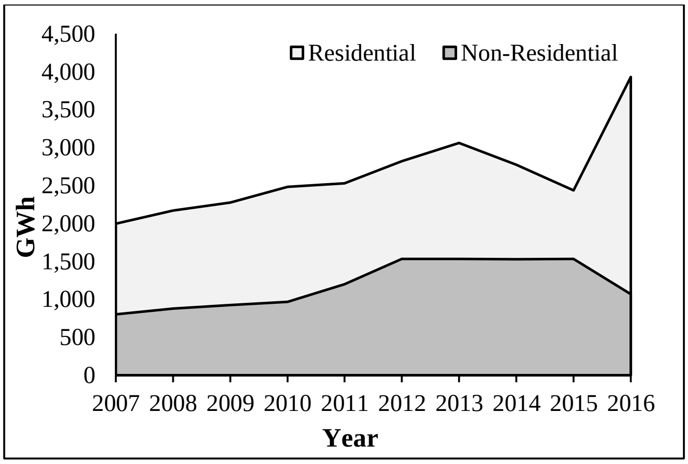

Fig.2. Residential and commercial electricity demand in Ghana from 2007 - 2016

The situation in Ghana is of no difference; a report by the Energy Commission of Ghana shows that residential electricity demand has increased over the past nine-year as compared with commercial/non-residential demand for electricity (Fig.2). The average electricity consumption in 2016 in Ghana was 2,647GWh for residential and 1,196GWh for commercial [8].

2.2 SOCIAL PSYCHOLOGICAL FACTORS INFLUENCING RESIDENTIAL ELECTRICITY CONSUMPTION BEHAVIOUR

As per [2], the home is the basic unit of electrical energy consumption, and hence, there is a need to study and identify the factors that contribute to energy consumption in the home. The literature argues that diverse families have diverse structures, ideological concepts, and cultural backgrounds. Every family is anticipated to have different electricity load profiles based on the influence of categories of factors and their interactions. Subsequently, a different load profile also mirrors diverse family types and consumption patterns [2] [5] [9] [13] [18]. The section briefs some of the household characteristics that are reported to influence domestic electricity consumption.

The Family Size (FS), study reveals that there is a direct positive relationship between the FS and electricity consumption (EC), thus FS α EC. The family size includes the nuclear family (children and parents), extended family, and domestic staff in the house [2] [9].

Size of the House (SH): the SH (the type of house) and the number of rooms in a house is also believed to partially affect the electricity consumption pattern in a home [2] [9].

The Age Composition of Family Members (ACFM), the age of the family member significantly influences the electricity consumption in a home [9] in their studies. Reports showed that electricity consumption in a home is minimal (lesser) when the age of the family members is below fifty (50) years and above sixty-five (65) and higher when the age of the family members is within fifty to sixty-five (50-65) years [2].

Family Economic Situation (FES), the family FES thus the disposable family income, and the family income is believed to influence household electricity demand [2] [9].

Time Staying in the House (TSH), the daily average time spent in the house, is reported to influence the electricity consumption in the home [9].

Educational Level (EL): the EL of the family member is believed to influence the consumption of electricity home Gram-Hanssen et al. (2005) cited in [2]. However, some researchers disagree with its effects on electricity consumption [19].

Time for Watching TV (TWTV): the time spent on watching television is argued to impact electricity consumption in the house [9].

Another factor that is believed to affect the consumption of electricity in a home is the social status of the family (SSF). However, social status is seen to have a different influence on electricity consumption. According to [20], a positive correlation exists between the socio-economic status of a family and household electricity consumption. On the other hand, [21] argues that there is no correlation between the family members' economic status and household electricity demand.

Additional influences such as the location of the family house, the type of car used by the family, the occupation of the family members, and the gender composition of the family are believed to influence the consumption pattern of households.

2.3 METHODS FOR MODELLING ENERGY CONSUMPTION

Energy modelling is seen to be a difficult task for engineers, academicians, and professionals [1]. Due to this, numerous methods have been adopted for energy modelling of which electrical energy is of no exception. According to [5], the techniques for modelling residential energy can be clustered into two categories, namely bottom-up and top-down. These categorisation terminologies are concerning the hierarchal position of data inputs as compared to the housing segment. These two main techniques are further divided into 8 [5], as presented in Fig.3.

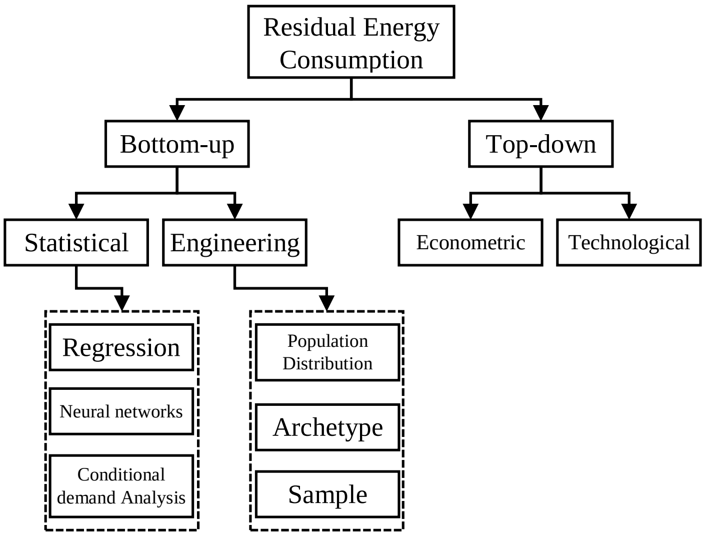

Fig.3. Top-down and bottom-up modelling techniques for predicting residential energy consumption

In the top-down techniques, the models utilize the assessment of the entire residential sector energy consumption and other relevant variables to attribute the energy consumption to the characteristics of the housing sector in totality. On the other hand, "bottom-up" models compute the energy consumption of individual or groups of households and then generalize these results to symbolize the region or nation.

2.4 MACHINE LEARNING

Machine-Learning (ML) is referred to as the science of making computer programs and robots to study and conduct themselves like humans do whiles having an autonomous-enhancement in their learning over-time by nursing them data and data in the form of observations and real-world interactions [22]. Various ML algorithms have been applied in diverse real-world situations, which include (i) Education for prediction of students' academic performance [23]. (ii) Finance for predicting future stock price movement [1] [24]. Energy for predicting monthly energy demand based on historical data [1] (iii) agriculture for predicting annual crop yields [25] [26] (iv) Health healthcare [27] [28]. However, ML application is not limited to only the area above the economy.

2.5 RELATED WORKS

Energy management is a global concern. As a result, several studies to promote energy efficiency through energy management have been carried out locally and globally. This section discusses, in brief, some of the studies and the techniques adopted.

The study by [1] proposed a soft-computing technique predictive model for predicting the monthly electricity demand in the Bono region of Ghana based on historical weather and electricity demand data. Using a multi-layer perceptron (MLP), decision tree (DT), and support vector machine (SVM), the author achieved an accuracy of 80.57% for DT, 95% for MLP, and 67.2% for SVM. However, the study concluded that consumer's lifestyles could not be left out when predicting electricity demand. Hence, future studies should consider this feature for high prediction accuracy. Likewise, an enhanced convolutional neural network (ECNN) and enhanced support vector regression (ESVR) predictive model for forecasting short-term electricity price and load in smart grids based on historical load demand and electricity price was proposed in [29]. XG-Boost (XGB), decision tree (DT), Recursive Feature Elimination (RFE), and random forest (RF) were used for feature selection and extraction. The authors, in their conclusion, reported that their proposed ESVR achieved almost 1% improved accuracy than the existing Support Vector Regression (SVR).

Also, a hybrid of a feature selection technique and a multivariate linear regression model for forecasting predicts seasonal power demand of households Tokyo based on consumers' lifestyle data was proposed by [9]. The study showed that lifestyle data is highly efficient for electricity demand forecasting. Household factors, which included age, family composition, house type, and others, served as the independent variables. However, the study concentrated on classifying the future load demand for residential electricity users as base-power consumption, summer-power consumption, and winter-power consumption, but did not predict the actual energy value. Similarly, in [12], a short-term domestic load forecasting based on LSTM recurrent neural network using a publicly available set of real residential smart meter data was proposed. The proposed model demonstrated comparable results with other state-of-the-art models. However, the effectiveness of the proposed model against other models was only based on the MAPE metric, which makes it hard to compare to other works.

Similarly, the association between domestic electricity consumption and income was analysed for Algeria within 1970–2013, by valuing a domestic electricity consumption per capita demand function based on the Gross Domestic Product (GDP) per capita. Though the study did not predict future residential electricity demand, it concluded that stimulating growth among Algerians could be useful to reduce the residential electricity consumption, as a higher income level might use more energy-efficient electrical machines. In [18], authors combined household-reported electrical gadget data and online sales of a residential electrical appliance to forecast the present and future residential electricity demand in Nigeria. The study concluded that the economic conditions of consumers should be a key feature for future studies. Furthermore, the household energy use and appliance ownership in Ireland were examined using Logit Regression (LR) analyses on a massive micro-dataset [21]. The study results revealed that the methods of space and water-heating used by a household are even more critical than electrical machines in clarifying residential energy usage.

The above literature reveals that there are numerous related studies on forecasting electricity demand, but only a few studies are based on forecasting residential electricity demand based on lifestyle information of a household, which also includes the life stage. Again, as the lifestyle of a household differs from region to region as argued in [2] [9], it makes a little difficult to a generalized an electricity demand predictive model based on lifestyle data from one or two regions. Moreover, the current study becomes the first residential electricity demand prediction in Ghana based on users' lifestyle to the best of our knowledge.

3. MATERIALS AND METHODS

This section presents the materials and methods adopted for the implementation of the proposed residential electricity demand predictive model based on household lifestyle data using machine artificial neural networks.

3.1 SAMPLE SELECTION AND LOCAL AREA DATA

A total of 350 households were randomly selected from 6 communities within the Sunyani Municipality. A semi-structured questionnaire designed approach was adopted to design 14 sets of questions to obtain the lifestyle of participants. The questionnaire was pretested with five households in Sunyani area 4 to examine how easy it was to understand and provide adequate answers needed from the participants. Comments from these five households were added to fine-tune the questionnaires. The questionnaires were then administered within 35 days with the help of three trained research assistants. The Table.1 shows the household distribution of the current study.

Table.1. Household Distribution

Community	No. of household	Percentage (%) (n = 350)
Area 4	33	9.43%
New Dorma	67	19.14%
Yawhema	39	11.14%
Kotokrom	55	15.71%
Abesim	65	18.57%
Baakoniaba	68	19.43%
Estate	23	6.57%

3.2 RESEARCH DESIGN

The cross-industry process for data mining (CRISP-DM) modelling process [30] was adopted for this study (Fig.4). The CRISP-DM method offers an organised approach for planning data mining and forecasting studies. Its iterative process offers a continuously learning and improves communication of insight and forecasting power. Most of the tasks involved in this model can be executed in a different order, and it will often be necessary to do a volte-face to previous tasks and echo certain activities. The CRISP-DM model has six phases, namely: study objectives, data understanding, data preparation, modelling, evaluation, deployment/publish results. The first phase study objective covers ascertaining the aim and objectives, which, when well-understood, leads to the gathering of the correct data.

Phase two covers visualisation of the research data, to give a better understanding, and identification of discrepancies, and deviation in the dataset. In the third phase, the obtained dataset is pre-processed by identifying missing values, if any, and treating them appropriately by scaling or by normalisation of the dataset for better/accurate prediction. The dataset for the current study was scaled in the range of [0, 1] and by dividing every value in the dataset by the maximum value to get the new value. The feature extraction and selection techniques are applied in this same phase to the select the significant feature. In this study, the root means square error (RMSE) metric is used to select the significant feature. The next phase is the modelling phase, where the selected features serve as input (independent) variables to our MLP model to predict the expected (dependent variable) future residential electricity demand. The predicted values $(\hat{y})$ are then compared with the actual values (y) to evaluate the performance of the model in the evaluation phase. The error metrics used in this study are discussed in detail in section 3.4 of this paper. Lastly, the deployment or publish results phase; at the stage, the results and findings obtained from the empirical study are communicated to the scientific community.

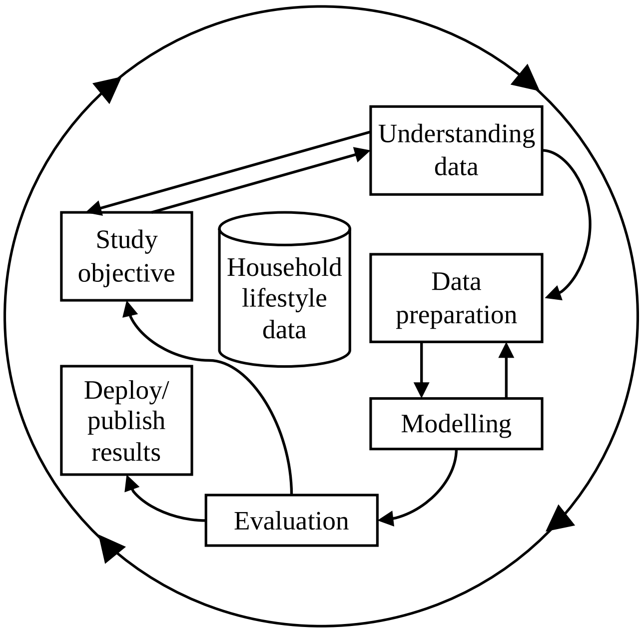

Fig.4. Adopted Study Framework [30]

3.3 MODELLING APPROACH

ANN algorithm was adopted for the current study based on its high accuracy levels in energy prediction, as discussed in section 2.5 of this paper. A supervised ML technique was used in this study, where household lifestyle served as input parameters, while actual monthly electricity consumption of participants from the supply authority served as the output target in the proposed network. Implementation was done with Python on the Anaconda framework.

Data mining and classification problems are usually expedited with optimization problem "feature selection" for choosing the essential features from various input sets while producing the same output. Although several algorithms have been developed, on the other hand, none is universally accepted to be the finest for all conditions. Therefore researchers are still trying to come up with improved solutions [31]. Artificial Neural Networks (ANNs) are a computer-based copy of the human brain composed of many "neurons" that work together to accomplish the desired purpose. They can be used for pattern recognition, classification, image matching, feature extraction, prediction, and noise reduction.

ANNs can learn and generalise, as mentioned earlier [32]. The proposed model comprises of two ANN multi-layer perceptron model. Thus, backpropagation trained ANN (BPANN_1) for feature selection and (BPANN_2) for prediction (Fig.5). The current study adopted [32] ANN algorithm for feature selection. The approach aimed at varying the weights of the different features to minimise the error in the actual value (y) and the predicted value $(\hat{y})$. So that the feature that possess lower weight was considered not important hence, rejected.

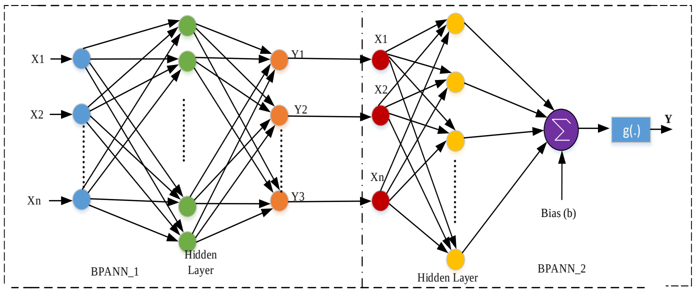

Fig.5. Proposed ANN Framework

For BPANN_1, let N be the number of observations N_{neuron} and K_x the number of features. Two classes produced (-1/+1). Where X is the matrix, and the rows denote the features and, the columns signify the observations, and y indicates the class of an observation. All features were given equal weights at the start. For each feature (X_x) is associated with a weight δ_k.

Let D_δ represent a diagonal matrix representing the weights of the features.

$$\delta_k \in [0,1] \quad k = 1, 2, ..., K_x \tag{2}$$

Let $\delta_{K_{x+1}}$ is the weight linked to the fixed input (bias)

$$\delta_{K+1} = 1 \tag{3}$$

Let X_δ be as follows:

$$X_\delta = D_\delta [x:1] \tag{4}$$

The network is then defined as one hidden layer and one linear combination, where S_w of dimension (N,N_{neuron}) represent the output of the hidden layer.

$$S_w = th(W_o X_w)^t \tag{5}$$

where W_0 is a random matrix with dimension (N, $N_{\text{neuron}} + K_{x+1}$). Let Y_w be the output of the network.

$$Y_\delta = S_\delta W \tag{6}$$

where, W is the weight vector of dimension (N, N_{neuron}, 1), δ_k represents the weight of feature and X_k represents a feature degree of importance. The weights of different features were varied to minimise the error, and if the coefficient δ_k is high, then its corresponding feature X_k is of impotence. While if δ_k is low, then feature X_k is of no importance. A feature was considered extra important when its weight is more significant. The aim here was to ensure that the weight of the unimportance feature reduces nearly to zero thereby minimising the errors (E). E is computed as expressed in Eq.(7) given by [32].

$$E = \frac{1}{2}\left(Y_\delta - Y\right)^t \times \left(Y_\delta - Y\right) \tag{7}$$

Finally, features with corresponding significance weight more than 0.5 were picked, while features with corresponding weights less than 0.5 were treated as unimportant features and were rejected.

For BPANN_2, let the output of BPANN_1 be presented by dataset (DS) comprises (x_1,y_1), (x_2,y_2),…,(x_n,y_n) where $x_i \in R^n$. With one hidden layer and fifty hidden neuron, proposed BPANN_2 learns function $f(x)$ is given in Eq.(8). The maximum iteration was set to 5000, optimizer = Limited-memory BFGS (lbfgs).

$$f\left(x\right) = W_{2g} \times \left(W_1^T x + b_1\right) + b_2 \tag{8}$$

where $W_1 \in \Re^m$ and W_1, b_1, $b_2 \in \Re$ are parameters for BPANN_2, and W_1, W_2 represents the input and hidden layer weights, respectively; b_1, b_2 represents biased at the hidden and output layer, respectively. The activation function ($g(.)$:$R \rightarrow R$) adopted for BPANN_2 in the current study was the hyperbolic tan is in Eq.(9). The square Error loss function (Eq.(10)) was adopted. With an initial random weight, the loss function was optimised by repeatedly updating the weights. We compute the loss and propagates backwards pass from output to the previous layer, providing individual weight parameter with a modernise value meant to reduce the loss.

$$g\left(z\right) = \left\{\frac{\left(e^z - e^{-z}\right)}{\left(e^z + e^{-z}\right)}\right\} \tag{9}$$

$$Loss\left(\hat{y}, y, \mathrm{w}\right) = \frac{1}{2} y\|_2^2 + \frac{\alpha}{2}\|w\|_2^2 \tag{10}$$

3.4 EVALUATION METRICS

The following error metrics discussed by [1] are used in measuring the performance of the proposed forecast model:

Root Means Squared Error (RMSE): This index estimates the residual between the actual value and predicted value. A model has better performance if it has a smaller RMSE. An RMSE equal to zero represents a perfect fit.

Mean Absolute Percentage Error (MAPE): This index indicates an average of the absolute percentage errors; the lower the MAPE, the better.

The Correlation Coefficient (R): This criterion reveals the strength of relationships between actual values and predicted values. The correlation coefficient has a range from 0 to 1, where higher R means it has an excellent performance measure.

$$RMSE = \sqrt{\frac{1}{m}\left(\sum_{v=1}^{m}\left(t_v - y_v\right)\right)} \tag{11}$$

$$MAPE = \frac{1}{m}\left(\sum_{v=1}^{m}\left|\frac{t_v - y_v}{t_v}\right|\right) \times 100 \tag{12}$$

$$R = \frac{\sum_{v=1}^{m}\left(t_v - \overline{t}\right)\left(y_v - \overline{y}\right)}{\sqrt{\sum_{v=1}^{m}\left(t_v - \overline{t}\right)^2 \sum_{v=1}^{m}\left(y_v - \overline{y}\right)^2}} \tag{13}$$

where $\overline{t} = \frac{1}{m}\sum_{v=1}^{m} t_v$ and $\overline{y} = \sum_{v=1}^{m} y_v$ are the average values of t_v and y_v, respectively and t_v is the actual value, y_v is the predicted value produced by the model, and m is the total number of observations.

4. EXPERIMENTAL SETUP

The Table.2 shows the initially obtained features out of the administered questionnaires and how they were abbreviated. The qualitative response from respondents was first coded using dummy variables. Afterwards, all the 350 received responses were queued with Microsoft Excel into a comma-separated values (CSV) file format. The implementation of the proposed model was carried in Python using the Anaconda framework and Scikit-learn libraries.

Table.2. Initially selected features

Features	Abbreviation
Residential location	RL
Age of family head	AFH
Employment sector of the family head	ESFH
Type of employment of family head	TEFH
Nature of Employment	NE
Apartment Type	APT
Marital status	MS
Living with family	LWF
Family size	FS
Family male-dominated or female-dominated	FMFD
The monthly salary of the family head	MSFH
Own personal vehicle	OPV
Number of vehicles Owen	NVO
Vehicle type	VT
Electricity consumption per month	ECPM

4.1 RESULTS AND DISCUSSION

This section presents the results and discussion of the proposed ANN prediction framework for forecasting electricity demand of residential users based on their household lifestyle.

4.1.1 Feature-Selection:

Feature-selection is the process of selecting a subset of features from a given set of features, by so doing selecting the feature with less noise, which interned increase the accuracy performance of the model and reduce training and prediction time. All fourteen features were input to the first phase of our proposed model discussed in section 3.3 of this paper. The feature importance ranking (Fig.6) reveals that the vehicle type (VT) (saloon or sport utility vehicle) owned by the head of the household was the social-psychological that highly determines the electrical consumption of the household. The vehicle one use is an indication of one's economic growth, and as argued in [33], economic growth raised power consumption. Followed by the monthly salary of the family head (MSFH), age of family head (AFH), marital status (MS), and nature of employment (NE).

The results affirm the study by Guo et al. [2] and Nishida et al. [9] that the age of the household is a huge determinant of household energy consumption. The results reveal a link between the VT and MSFH, which can be interpreted as the amount of monthly salary one received has a bearing on the car, he/she drives, which in a way also determine the number of electrical appliance and equipment used in the household. On the other hand, it can be deduced that the socio-economic status of the country and its people determine the amount of electrical energy consumed. The results affirm the findings presented by [20] [34].

Consequently, the results disagree with [21] report that concluded that there is no positive relationship between household electricity consumption and the household socio-economic status. The VT was the top-most relevant feature revealed from this study, implies that the state authorities can partially estimate electricity demand based on the number of the vehicle its citizens register with their vehicle and licenses authority. The Table.3 shows the summary of a traditional stepwise linear regression model that is set up to verify the output of the proposed ANN feature selection model. The result affirms VT as the most significant feature followed by MS. However, MS, which our proposed found to be the fourth most significant, was third; nonetheless, there is an agreement between the two models.

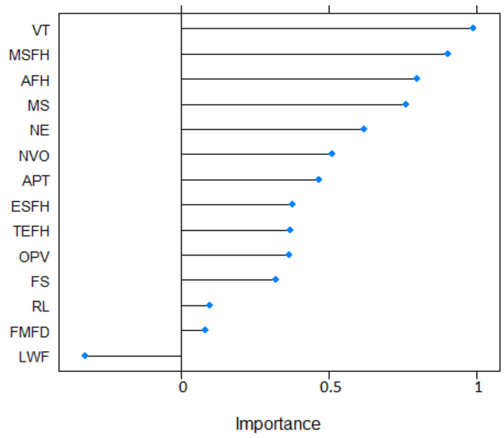

Fig.6. Feature Importance Ranking

4.1.2 *Model Training and Testing:*

Out of the feature ranking (Fig.5), five features (VT, MSFH, AFH, MS, NE, and NVO) were selected for the next stage of the proposed ANN framework. The dataset (DS) was partitioned into two, thus training (85%) and testing (15%). The training and testing of the model were performed in two-phase. *Phase 1*: all features were served ad input to BPANN_2. *Phase 2*: only the selected feature from BPANN_1 was served as input parameters into PBANN_2. In both phases, the outcome of the proposed model was benchmarked with traditional models such as decision trees (DT) and support vector regression (SVR). The results are discussed as follows:

Models performance without feature selection: The Table.4 shows the error metric measure of models without feature selection. The results reveal that the proposed forecast model (PBANN_2) without feature selection recorded a RMSE of 0.006569 and MAE 0.006132 compared with (RMSE = 0.0657 and MAE = 0.05714) for DT. The results reveal that even without a feature selection, the ANN algorithms offer a better fit in electricity load forecasting than the DT and SVR. The model's prediction compared with actual household consumption (Fig.7) affirms the error metrics in Table.4. This outcome affirms the reason why several studies in this field using AI prefer the ANN over other algorithms.

Table.3. Linear Stepwise correlation analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.549[a]	.301	.285	65.480	.301	18.975	1	44	.000
2	.627[b]	.394	.366	61.701	.092	6.554	1	43	.014

a. Predictors: (Constant), VT b. Predictors: (Constant), VT, MS c. Dependent Variable: ECPM

Table.4. Models performances without feature selection

Models	RMSE	MAE	R^2	Training Time	Testing Time
BPANN_2	0.006569	0.006132	0.02863	0.4	0.001
DT	0.0657	0.05714	0.06919	0.001	0.00001
SVR	0.08816	0.06911	-0.92494	0.002	0.00001

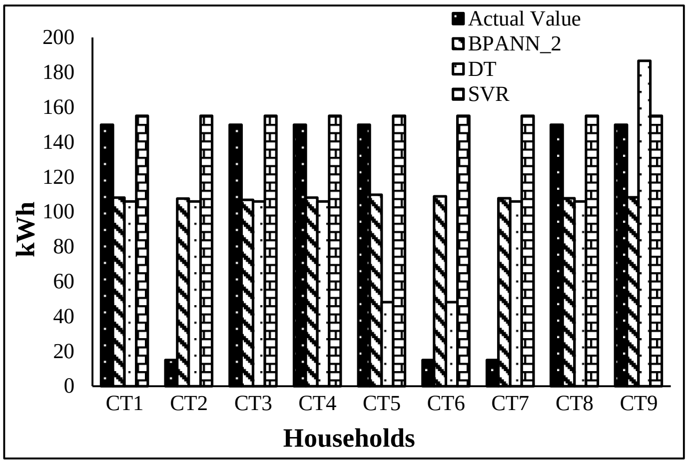

Fig.7. Models predicted output against actual values

Models performance with feature selection: The Table.5 shows the error metrics values of the models with the feature selection model (BPANN_1) been implemented. With the feature selection in place, the training time of the proposed model, reduce to 0.017 seconds as compared to 0.4 seconds without feature selection. The outcome reveals that applying feature selection can reduce the training time and testing time of a model by 95.75% with data size. The outcome affirms [31] reports that feature selection offers quicker training and testing time in addition to increasing the model's accuracy. The RMSE of 0.00725 and MAE of 0.00976 for the proposed model (BPANN_1) reveal enhancement in the accuracy of the proposed model with feature selection techniques.

The models predicted values against the actual household energy consumption (Fig.8) show that the proposed ANN model (BPANN_2) once again outperformed traditional DT and SVR. However, it is noticed from the results (Fig.8) that both the DT and BPANN_2 prediction accuracy have increased after feature selection. The results reveal that the proposed forecast model can effectively forecast feature residential electricity demand based on household lifestyle at an accuracy of 94.3%. That is, the error rate for the proposed system is 5.7%. The results imply that residential users of electricity can effectively project their electricity demand in feature based on the lifestyle they wish to live in the future. Thus, it would offer them the opportunity to plan their budgets to prevent any unexpected shortage of electricity due to their purchasing power.

Table.5. Models error metric measure with feature selection techniques

Models	RMSE	MAE	R^2	Training Time	Testing Time
BPANN_1 + BPANN_2	0.000726	0.000976	0.000252	0.017	0.0001
BPANN_1 + DT	0.00168	0.0036	0.013504	0.002	0.00001
BPANN_1 + SVR	0.008365	0.00555	-0.78633	0.001	0.00001

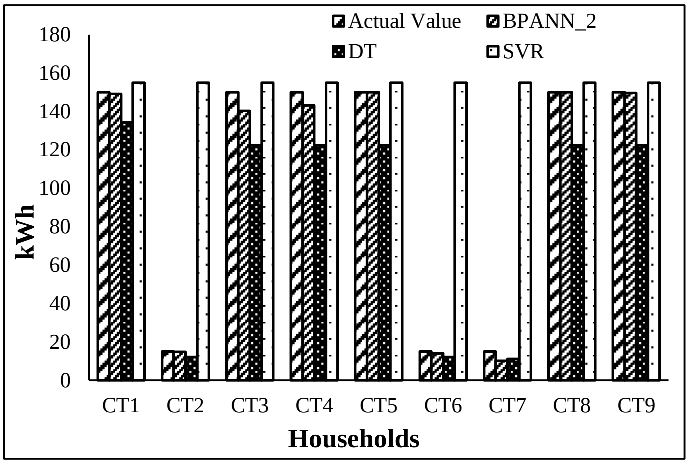

Fig.8. Models performance after feature selection

5. CONCLUSION AND DIRECTION FOR FUTURE RESEARCH

Electricity is an essential commodity in everyone's life and a conduit for economic development in every country. The growth rate of electricity demand in residential facilities keeps increasing over the years. However, several studies based on electricity forecasting overlooks this area, and the few studies that were focused in this area based their prediction on weather parameters as independent variables to their forecast models. On the other hand, residential electricity consumption is mainly based on household lifestyle. Therefore, the current study sought to forecast residential electrical energy consumption based on household lifestyle using an artificial neural network.

The experimental setup with 350 household lifestyle data randomly selected towns in the Sunyani Municipality in the Bono region of Ghana revealed that the type of vehicle used by heads of households was the most significant lifestyle feature in residential electricity load forecasting. Furthermore, the proposed model's accuracy (94.3%) as compared to the traditional machine learning such as SVR and DT shows that the proposed model can effectively forecast residential electricity consumption at an error rate of 5.7%.

Also, the use of actual electricity consumption (kWh) as the target variable compared to the amount (Ghana cedis) spent on electricity monthly by a household, makes the outcome of this study independent of the country's socio-economic factors such as gross domestic product, exchange rate, and inflation. Since the actual units (kWh) consumed by a household dependent on the household energy management lifestyle and not the country's socio-economic factors. In summary, the current study contributions are as follows: (i) An ANN-based feature-selection technique for optimal selection of the most significant independent variables. (ii) An empirical comparison of the proposed ANN model compared with other state-of-the-art forecast models. (iii) We add to the scarcity in the literature on residential load forecasting based on household lifestyle. (iv) To the best of our knowledge, this study is the first load forecasting of residential electricity consumption based on household lifestyle in Ghana.

However, the accuracy of the forecast model is one of the evaluation metrics for a good forecast model. Therefore, despite the achieved accuracy measured in this study, we believe that there is more room for improvement. Hence, in future, we look forwarded to add diverse independent parameters such as weather data and more household lifestyle data (such as the number of electrical appliances and the average time of use) using hybridization of several state-of-the-art machine learning algorithms to optimize feature selection of parameters for better accuracy measure.

REFERENCES

[1] I.K. Nti, S. Asafo Adjei and M. Agyemang, "Predicting Monthly Electricity Demand using Soft-Computing Technique", *International Research Journal of Engineering and Technology*, Vol. 6, No. 6, pp. 1967-1973, 2019.

[2] Z. Guo, K. Zhou, C. Zhang, X. Lu, W. Chen and S. Yang, "Residential Electricity Consumption Behavior: Influencing Factors, Related Theories and Intervention Strategies", *Renewable and Sustainable Energy Reviews*, Vol. 81, No. 1, pp. 399-412, 2018.

[3] N.I. Nwulu and O.P. Agboola, "Modelling and Predicting Electricity Consumption using Artificial Neural Networks", *Proceedings of 11th International Conference on Environment and Electrical Engineering*, pp. 1059-1063, 2012.

[4] A. Bogomolov, B. Lepri, R. Larcher, F. Antonelli, F. Pianesi and A. Pentland, "Energy Consumption Prediction using People Dynamics Derived from Cellular Network Data", *EPJ Data Science*, Vol. 5, No. 3, pp. 1-15, 2016.

[5] L.G. Swan and V.I. Ugursal, "Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques", *Renewable and Sustainable Energy Reviews*, Vol. 13, No. 8, pp. 1819-1835, 2009.

[6] Indexmundi, "Electricity Consumption Per Capita-World", Available at: https://www.indexmundi.com/map/?v=81000, Accessed on 2019.

[7] IEA, "IEA's Electricity Information", Available at: https://www.iea.org/statistics/electricity/#targetText=In 2017%2C world total final,increase of 4.6%25 from 2016, Accessed on 2019.

[8] Ghana Energy Commission, "National Energy Statistics Report 2017", Available at: http://energycom.gov.gh/files/ENERGY_STATISTICS_2017.pdf, Accessed on 2017.

[9] K. Nishida, A. Takeda, S. Iwata, M. Kiho, I. Nakayama, "Household Energy Consumption Prediction by Feature Selection of Lifestyle Data", *Proceedings of International Conference on Smart Grid Communications*, pp. 235-240, 2017.

[10] A. Veit, C. Goebel, R. Tidke, C. Doblander and H. Jacobsen, "Household Electricity Demand Forecasting-Benchmarking State-of-the-Art Methods", *Proceedings of International Conference on Future Energy Systems*, pp. 1-10, 2014.

[11] P. Huber, M. Gerber, A. Rumsch and A. Paice, "Prediction of Domestic Appliances usage based on Electrical Consumption", *Energy Informatics*, Vol. 1, pp. 265-428, 2018.

[12] W. Kong, Z.Y. Dong, Y. Jia, D.J. Hill, Y. Xu and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network", *IEEE Transactions on Smart Grid*, Vol. 10, No. 1, pp. 841-851, 2019.

[13] A. Soares, A. Gomes and C.H. Antunes, "Categorization of Residential Electricity Consumption as a Basis for the Assessment of the Impacts of Demand Response Actions", *Renewable and Sustainable Energy Reviews*, Vol. 30, pp. 490-503, 2014.

[14] L. Wilson, "Average Household Electricity Use around the World", Available at: http://shrinkthatfootprint.com/average-household-electricity-consumption, Accessed on 2019.

[15] H. Ritchie, "Number of People in the World without Electricity Falls below One Billion", Available at: https://ourworldindata.org/number-of-people-in-the-world-without-electricity-access-falls-below-one-billion#targetText=Number of people in the world without electricity falls below one billion&targetText=Global access to electricity has, risen to over 87 p, Accessed on 2019.

[16] The World Bank, "Access to Electricity (% of population)", Available at: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS, Accessed on 2019.

[17] Electricchoice.com, "Residential Electricity for Houses, Apartments, Condos, Trailers, etc.,", Available at: https://www.electricchoice.com/residential-electricity.

[18] K. Olaniyan, B.C. Mclellan, "Estimating Residential Electricity Consumption in Nigeria to Support Energy Transitions", *Sustainability*, Vol. 10, No. 5, pp. 1-22, 2018.

[19] J.C. Cramer, N. Miller, C. Paul, M.B. Hackett, M.T. Dietz, L.E. Vine, D.D. Levine and J.D. Kowalczyk, "Social and Engineering Determinants and Their Equity Implications in Residential Use", *Energy*, Vol. 10, No. 2, pp. 1283-1291, 1985.

[20] F. McLoughlin, A. Duffy and M. Conlon, "A Clustering Approach to Domestic Electricity Load Profile Characterisation using Smart Metering Data", *Applied Energy*, Vol. 141, pp. 190-199, 2015.

[21] E. Leahy and S. Lyons, "Energy Use and Appliance Ownership in Ireland", *Energy Policy*, Vol. 38, No. 8, pp. 4265-4279, 2010.

[22] D. Faggella, "What is Machine Learning?", Available at: https://www.techemergence.com/what-is-machine-learning, Accessed on 2018.

[23] I.K. Nti and J.A. Quarcoo, "Self-motivation and Academic Performance in Computer Programming Language Using a Hybridised Machine Learning Technique", *International Journal of Artificial Intelligence and Expert Systems*, Vol. 8, No. 2, pp. 12-30, 2019.

[24] A.D. Adebayo, A.F. Adekoya and T.M. Rahman, "Predicting Stock Trends Using Tsk-fuzzy Rule Based System", *Journal of Energy and Natural Resource Management*, Vol. 4, No. 7, pp. 48-55, 2017.

[25] K.R. Akshatha and K.S. Shreedhara, "Implementation of Machine Learning Algorithms for Crop Recommendation Using Precision Agriculture", *International Journal of Research in Engineering, Science and Management*, Vol. 1, No. 6, pp. 58-60, 2018.

[26] A. Crane-Droesch, "Machine Learning Methods for Crop Yield Prediction and Climate Change Impact Assessment in Agriculture", *Environmental Research Letters*, Vol. 13, No. 11, pp. 1-23, 2018.

[27] A. Choudhury, "Classification of Functioning, Disability, and Health: ICF-CY Self Care (SCADI Dataset) Using Predictive Analytics", *Proceedings of International Conference on Institute of Industrial and Systems Engineers*, pp. 1-8, 2019.

[28] A. Choudhury and B. Eksioglu, "Using Predictive Analytics for Cancer Identification", *Proceedings of International Conference on Institute of Industrial and Systems Engineers*, pp. 23-29, 2019.

[29] M. Zahid, F. Ahmed, N. Javaid, R. Abbasi, K.H. Zainab, A. Javaid, M. Bilal, M. Akbar and M. Ilahi, "Electricity Price and Load Forecasting using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids", *Electronics*, Vol. 8, No. 2, pp. 1-18, 2019.

[30] IBM, "IBM Knowledge Center", Available at: https://www.ibm.com/support/knowledgecenter/en/SS3RA7_15.0.0/com.ibm.spss.crispdm.help/crisp_overview.htm, Accessed on 2018.

[31] S. Goswami, A.K. Das, A. Chakrabarti and B. Chakraborty, "A Feature Cluster Taxonomy based Feature Selection Technique", *Expert Systems with Applications*, Vol. 79, pp. 76-89, 2017.

[32] N. Challita, M. Khalil and P. Beauseroy, "New Feature Selection Method based on Neural Network and Machine

Learning", *Proceedings of International Conference on Engineering Technology*, pp. 81-85, 2016.

[33] Enerdata, "Electricity Domestic Consumption", Available at: https://yearbook.enerdata.net/electricity/electricity-domestic-consumption-data.html, Accessed on 2019.

[34] F. Mcloughlin, A. Duffy and M. Conlon, "Characterising Domestic Electricity Consumption Patterns by Dwelling and Occupant Socio-Economic Variables: An Irish Case Study", *Energy and Buildings*, Vol. 48, pp. 240-248, 2012.