Cloud-based electricity consumption analysis using neural network

Nand Kumar

Department of Computer Science,
Birla Institute of Technology and Science Pilani, Dubai Campus,
Dubai, 345055, UAE
Email: nandkumar@dubai.bits-pilani.ac.in

Vilas H. Gaidhane*

Department of Electrical and Electronics Engineering,
Birla Institute of Technology and Science Pilani, Dubai Campus,
Dubai, 345055, UAE
Email: vilasgd612@gmail.com
*Corresponding author

Ravi Kant Mittal

Department of Computer Science,
Birla Institute of Technology and Science Pilani, Pilani Campus,
Pilani, 333031, India
Email: rkm.bits@gmail.com

Abstract: In recent years, optimisation of the resource usages is necessary to analyse and understand the energy consumption pattern. In the literature, analysis has been carried out using the algorithms, which needs many assumptions, and meeting all the assumptions in practice is a very difficult task. However, there are other methods available to analyse and understand the energy consumption. In this paper, an efficient approach for energy consumption pattern analysis is proposed. It is based on the Levenberg-Marquardt algorithm-based Neural Network (LMNN) and clustering technique. The energy consumption data is collected from the educational institute building using smart system. The various experimentations are carried out on the collected real time database. The experimental results illustrate that the proposed approach is effective and computationally efficient for consumption pattern classification. The performance of the presented approach is found superior to existing clustering approaches.

Keywords: educational institute building; Levenberg-Marquardt algorithm; neural network; classification; confusion matrix; ROC curve.

Reference to this paper should be made as follows: Kumar, N., Gaidhane, V.H. and Mittal, R.K. (2020) 'Cloud-based electricity consumption analysis using neural network', *Int. J. Computer Applications in Technology*, Vol. 62, No. 1, pp.45–56.

Biographical notes: Nand Kumar is a Lecturer/Research Scholar in the Department of Computer Science at Birla Institute of Technology and Science, Pilani, Dubai Campus, DIAC, Dubai, UAE. His research interests includes cloud setup, cloud security, database efficiency, adaptive website designing. He has published research papers in international refereed journals and conferences.

Vilas H. Gaidhane received his BE degree in 2000 from the Department of Electronics, Nagpur University, Nagpur and MTech degree in 2010 from the Department of VLSI Design, UPTU Lucknow, India. He also received his PhD degree in Instrumentation and Control Engineering, University of Delhi in 2013. His research interest includes image processing, pattern recognition, computer vision, control system and microelectronics.

Ravi Kant Mittal is a Senior Professor at Birla Institute of Technology and Science, Pilani, India. He has held several administrative positions at BITS Pilani in last 40 years. He has guided five Doctoral students and numerous Masters' dissertations and projects. His research interests include software testing, robust robot design, robot path planning, MEMS, nanotechnology and e-waste. He has published over 40 papers in international and national peer reviewed journals and conferences.

1 Introduction

The economic development of any country largely depends on the efficient use of energy. In recent years, all the luxury amenities at homes and electronic devices used by humans are mostly driven by electrical energy. As the population is growing, the use of these devices has also increased and in turn the demand-consumption of energy has gone up continuously (Li et al., 2010). As per the recent literature on energy demand-consumption in buildings, approximately 90% of the total energy has been consumed by the building sectors in Gulf countries (particularly UAE) (Kumar et al., 2018a). Therefore, it is necessary to understand the utilisation of energy demand and consumption patterns.

In last decade, the various techniques have been presented to analyse and identify the energy consumption pattern within the buildings. Amin-Naseri and Sorroush (2008) proposed hybrid model for daily electrical peak load forecast using unsupervised neural network (Self-Organiaing Map) along with a feed forward neural network. Further, a logarithmic-linear model using genetic algorithm and artificial neutral network was presented for Iranian agricultural sector data of 25 years from 1981 to 2005 (Azadeh et al., 2007). Moreover, an artificial intelligence-based technique in the application of building energy systems has been presented in the past (Dounis, 2010). Further, Zhao and Magoulès (2012) reviewed the modelling and forecasting of building energy consumption related work done in recent years, which includes artificial intelligence methods like ANN and Support Vector Machines (SVM) techniques.

One of the important factors to determine building energy consumption is climatic conditions like temperature, rain, humidity, wind speed, etc. The problem with these climatic conditions is that they vary over time. Noussan and Nastasi (2018) made an effort to use the data-driven approach to analyse the building energy consumption considering the temperature factor. Further, Li et al. (2017) proposed machine learning-based stacked encoders to predict the energy consumption in a building. Moreover, Zhou et al. (2016) discussed the building energy power system, architecture and functional models of home energy management system to improve the efficiency, energy conversation, reliability and economics. Furthermore, Pan et al. (2015) presented a study on smart energy in buildings using Internet of Things (IOT) along with intelligent building applications, which can be used for comprehensive analysis of energy consumption observations and administration.

Recently, Macedo et al. (2015) discussed the use of Demand Side Management (DSM), to simulate the data collected from digital meters and created load curve patterns. Training and validation are done for ANN using these patterns to classify the new data. Further, McLoughlin et al. (2015) investigated the widely used unsupervised clustering methods to clusterise the individual household segments based on the electricity uses pattern. Furthermore, Wang et al. (2015) proposed a novel approach for energy consumption clustering by applying the time-based Markov model on reduced dimensions of the input data set. Wang et al. (2016) summarised the price and incentive-based load profiling to demand response application. It is implemented based on a state-of-the-art and comprehensive review of the data mining techniques. Moreover, Jokar et al. (2016) presented an energy theft detector based on the energy consumption patterns, which helps in identifying and predicting the normal or malicious energy consumption pattern. Chen et al. (2016) demonstrated the energy efficiency optimisation using dataflow for Deep Neural Network (DNN).

From the literature, it is concluded that the worldwide trend is the forecasting of future demand growth, enhancement of efficiency, reliability, and sustainability. The fulfilment of these trends is possible through smart grids and data analysis (Touzene et al., 2019).

After a detailed study and analysis, it is observed that the self-organising maps, Support Vector Machine (SVM), Hidden Markov Model (HMM) and ANN-based techniques (Amin-Naseri and Sorroush, 2008; Azadeh et al., 2007; Dounis, 2010; Zhao and Magoulès, 2012; Macedo et al., 2015; McLoughlin et al., 2015) are extensively used for energy consumption pattern classification. However, SOM technique results in slow training if the data set is large. Moreover, SVM-based approaches are designed originally for binary classification. As the data set increases, the training for these methods becomes more complicated and results in a false classification. Similarly, the various methods based on HMM are unable to deal with dependencies among the large data set. It is observed that in ANN-based approaches, the classification of energy consumption pattern analysis is depending on the nature of data set and the error function (cost function) used for the analysis and still remains the challenge for the researchers. Therefore, it is necessary to propose an efficient technique to acquire the data in real-time from the system as well as the cost function to obtain the accurate energy consumption pattern classification results.

The motivations of the presented work come from the fact that the ANN can perform better if its error function (cost function) is efficient. Levenberg-Marquardt (LM) approach is most suitable for large data sets and accurate classification. Therefore, in this present work, LM-based Neural Network (LMNN) method is proposed for electricity consumption pattern classification.

The rest of the paper is structured as follows: Section 2 presents a brief overview of existing and basic methods and Section 3 explain the detailed proposed system. Experimental results and discussion are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 Overview of existing approaches

In recent years, the various methods are proposed for classification of energy consumption patterns in the buildings. These existing methods are discussed in the following sub-sections.

2.1 Hidden Markov Model (HMM)

For statistical pattern analysis the Hidden Markov Model (HMM) has proven to be a useful tool. It has been used in a variety of applications, including biological sequence analysis, software piracy detection, speech recognition, malware detection, etc. Markov model is a statistical model that has known states and the fixed state transition probabilities. The known states are directly observable in this process. In contrast, HMM states are not directly observable. HMM (represented as λ) can be trained on a given observed sequence. It can be denoted as

$$\lambda = (A, B, \pi) \tag{1}$$

where A is a square matrix of order $N \times N$ having transition probability, B is also a matrix of order $N \times M$ having observation probability and π is a one dimensional matrix of order $1 \times N$ having the initial state distributions. However, the information about the hidden states can be obtained indirectly via the observations O (having observation sequence $O_1, O_2, ..., O_n$). First the HMM is trained and then used to score the samples, which are subsequently used as class-conditional densities in a standard Bayes classification paradigm. This is the well-known Maximum-Likelihood (ML) classification rule. The typical HMM-based classification approach adopts the ML criterion, where an unknown sequence Q is assigned to the class showing the highest likelihood, i.e.

$$Class(Q) = \arg\max_i p(Q|\lambda_i) \tag{2}$$

where λ_i is the HMM corresponding to the *i*-th class. This is known as the ML one per class approach (ML_{OPC}), where C HMMs will require C-class problems to get trained. In another arrangement, instead of one HMM for each class, train one modal for each training sequence and can be represented as

$$Class(Q) = \arg\max_k \left(\max_i p(Q|\lambda_i^k)\right) \tag{3}$$

where λ_i^k denote the HMM modal trained on sequence Q_i^k of class *k*. This approach is known as the ML one per sequence (ML_{OPS}).

The detailed HMM mathematical model has been used in many applications (Annachhatre et al., 2015; Stamp, 2017; Candanedo et al. 2017; Kumar et al., 2018b). However, HMM cannot deal with the dependencies in observations. Moreover, it has much less efficiency around 76% (Dong et al., 2010; Yang et al. 2016; Akkaya et al., 2015; Voyant et al., 2017).

2.2 Support vector machine (SVM)

Support Vector Machines (SVM) is a machine learning method or deep learning method. This method is widely used for classification or finding the separating hyperplane, which is based on statistical learning concept. This process helps in converting the original input space into a higher-dimensional feature space (Bui et al., 2016). Recently, SVM is also used as a tool for data analysis and classification (Olatomiwa et al., 2015). In this method, first it builds a decision plane to separate a set of objects with different membership class. It distinguishes between these classes membership by maximising the margin between them. The best classification on the training data set can be determined by the best generalisation abilities which can be obtained by the maximal margin hyper-planes. In the second major step, it selects the kernel function (γ) of the algorithm. The training vectors x_i and x_j are prepared for two groups *i* and *j*. Further, these training vectors are mapped to a higher dimensional space by the kernel function (φ). SVM equations approximates this function given by the equation as (Olatomiwa et al., 2015)

$$f(x) = w\varphi(x) + b \tag{4}$$

where x is the input space vector which is mapped to the high dimensional space feature $\varphi(x)$, w is a vector and b is a scalar coefficient. The Lagrange multiplier and optimal constraints can be introduced in equation 4 to obtain a generic function as

$$f(x) = \sum_{i=1}^{N} \left(\beta_i - \beta_i^*\right) K\left(x_i, x_j\right) + b \tag{5}$$

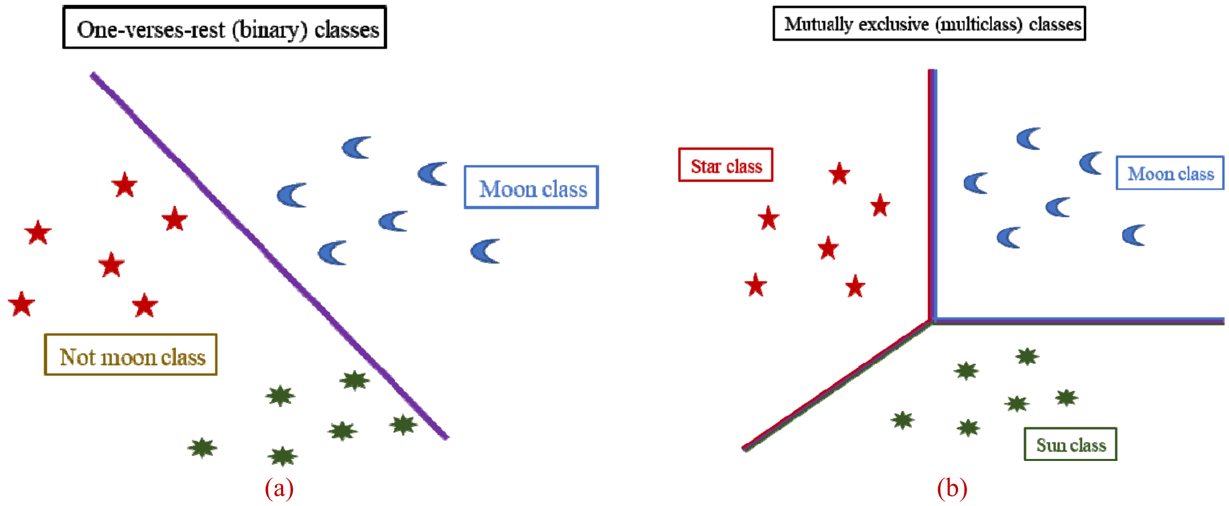
where $K(x_i, x_j)$ is called the kernel function and is the inner product of input vectors x_i and x_j. The input vectors (x_i, x_j) belong to feature spaces $\varphi(x_i)$ and $\varphi(x_j)$, respectively.

The SVM classification can be divided into binary or multiclass classification. In binary SVM classification only two classifications, either desired class or not desired class are decided. While in multiclass SVM classification, multiple classes mutually exist. The binary and multiclass SVM classification is represented as shown in Figure 1. However, the algorithmic complexity and extensive memory requirements are the problems with SVM which are computationally more expensive (Voyant et al., 2017; Basu et al., 2015; Lajnef et al., 2015).

2.3 Bayesian network

A Bayesian network is a probabilistic graphical model. This is one type of statistical model that represents a set of random variables and their conditional dependencies via a Directed Acyclic Graph (DAG) (Voyant et al., 2017).

Figure 1 The SVM classification: (a) Binary classification, (b) Multiclass classification



The Bayes theorem can be used to estimate the probability that at time the series is in the state y_k given as

$$X_t = \arg\max_k \left(P(X_t = y_k) \cdot P(\aleph | X_t = y_k)\right) \tag{6}$$

where $\aleph$ represents the initial and known measured, k the class and y_k the value of class, and P the conditional probability. The assumption of conditional independence can be used to solve this equation practically given by following equation as

$$X_t = \arg\max_k \left(P(X_t = y_k) \cdot \prod_{j=1}^{J} P\left(X_{t-j} | X_t = y_k\right) \right) \tag{7}$$

However, in such kind of predictors the conditional probability table is required to quantify the last term of the equation (Voyant et al., 2017; Paoli et al., 2010).

2.4 *Outlier detection-based classification*

In recent years, some researchers have worked on energy consumption classification models (Amin-Naseri and Sorroush, 2008; Azadeh et al., 2007; Dounis, 2010; Zhao and Magoulès, 2012). However, the desired and efficient results could not be achieved. The reason is over fitting or outliers of the input data.

The outliers are those data points which are distant from the rest of the data. These outliers will make the results unpredictable. Therefore, the outliers need to be detected and removed for efficient classification. Li et al. (2010) proposed an intelligent analysis and prediction approach for daily electricity consumption in buildings. They have carried out the classification using four steps algorithm. The features are extracted from data in the first step. In the second step the outlier is detected and removed. The formatted time-series data is used in third step. The fourth step is used to form the clusters. The energy consumption behaviour can be predicted, after the study of the clusters formed.

In feature extraction step, the mean of daily-energy consumption and peak daily consumption features are extracted. To analyse this data, the auto regression model is applied. Considering this model, prediction of the current sample $I(t)$ can be expressed as (Li et al., 2010)

$$I(\mathrm{t}) = \omega + \sum_{i=1}^{N} a_i I(t-i) + e \tag{8}$$

where $I(t)$ is the current sample's value, ω is the intercept variable, N is total number of samples, a_i is the i-th coefficient, and e is a noise parameter.

Further, the outlier detection is carried out on the daily-energy consumption by separating the inconsistent data from the rest of data set. This method is based on ordered statistics. Furthermore, Canonical Variate Analysis (CVA) tool is used, it estimates the space vectors, which maximise the difference between groups and minimised the separation within groups (Li et al., 2010). Suppose, F represents the difference between each group's mean, then it can be represented as

$$F = A_v V + r \tag{9}$$

where A_v is a singular matrix for multi-collinear data and r represents residual matrix. The directions for each group in eigenvector W are the columns of V. Partial least-squares approach can be applied to solve the equation (9). Further, the CVA analysis is applied which projects the original data into Canonical Variables (CV), also known as axes. Furthermore, a Linear Discriminate Analysis (LDA) is applied to classify the data into groups. The definition of discrimination function is given as

$$\begin{aligned} Y_i(D) = \log(\pi_i) - 1/2\left(D - \overline{D}_i\right)^T C_{W,CV}^{-1}\left(D - \overline{D}_i\right) \\ + \log\left|D_{W,CV}\right| \end{aligned} \tag{10}$$

With the maximum value of Y_i, the i-th group is chosen as a classified group (Li et al., 2010). However, it is observed that the LDA has many assumptions and restrictions. In real-time applications, it is very difficult to meet all these assumptions. Therefore, the motivation of presented work is to propose an efficient and simple approach. An artificial neural network-based approach may be an alternative solution which has least assumptions and restrictions. Therefore, a LM-based neural network has been used in the proposed approach for the classification of energy consumption patterns. However, the performance verification of classification models can be summarised in a table or matrix format commonly known as confusion matrix or error matrix. It is a two-dimensional contingency matrix, with identical sets of classes in both dimensions. It is easy to represent the classification of different classes using confusion matrix. Therefore, it is used for the representation of the different classes.

3 Proposed system

In this work, an efficient approach for energy consumption pattern classification is presented as shown in Figure 2. The presented approach consists of various blocks such as supply of the energy, current sensor, cloud storage, extraction of the data from cloud, data processing unit, neural networks classification and clustering and graphical representation of the output. Cloud storage has been used here for energy consumption data storage and application hosting. The use of cloud brings the computational flexibility and mobility to the techniques (Kumar and Mittal, 2012). This will enable the building management to understand the energy consumption patterns of the days and will help them to classify and cluster the energy consumption data in different meaningful groups. Each group can represent a unique feature of the energy consumption pattern after applying the ANN technique on the data for classification and clustering. Each block of the proposed system is explained in detail in the following subsections.

3.1 Data collection

The energy consumption data has been collected from an educational institute building located in Dubai International Academic City, Dubai and considered for the case study. The institute building has different blocks like academic block, boys' and girls' hostel blocks, library block, mechanics block, etc. For the energy consumption data collection, all these blocks have been considered. A government authority, Dubai Electricity and Water Authority (DEWA) provides the electricity in the institute building. The electricity from the main powerhouse in the institute building is distributed in different blocks as per the predefined requirement. The current sensors (best 126818/4/029, CT/CCTH/250A024) have been connected to the power consumption cables, which sense the current and send it to the smart power meter (Eniscope 8). The smart power meter is connected to internet and sends these energy consumption data to a cloud server for storage and further processing. The complete cloud-based data acquisition system is shown in Figure 3. In this study, the data is collected from the cloud storage for the duration of one year (1 Oct 2016 to 30 Sep 2017) is presented in Table 1.

Figure 2 Block diagram of the proposed system

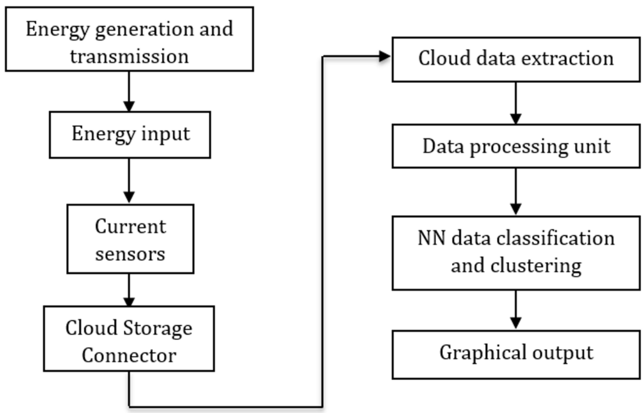
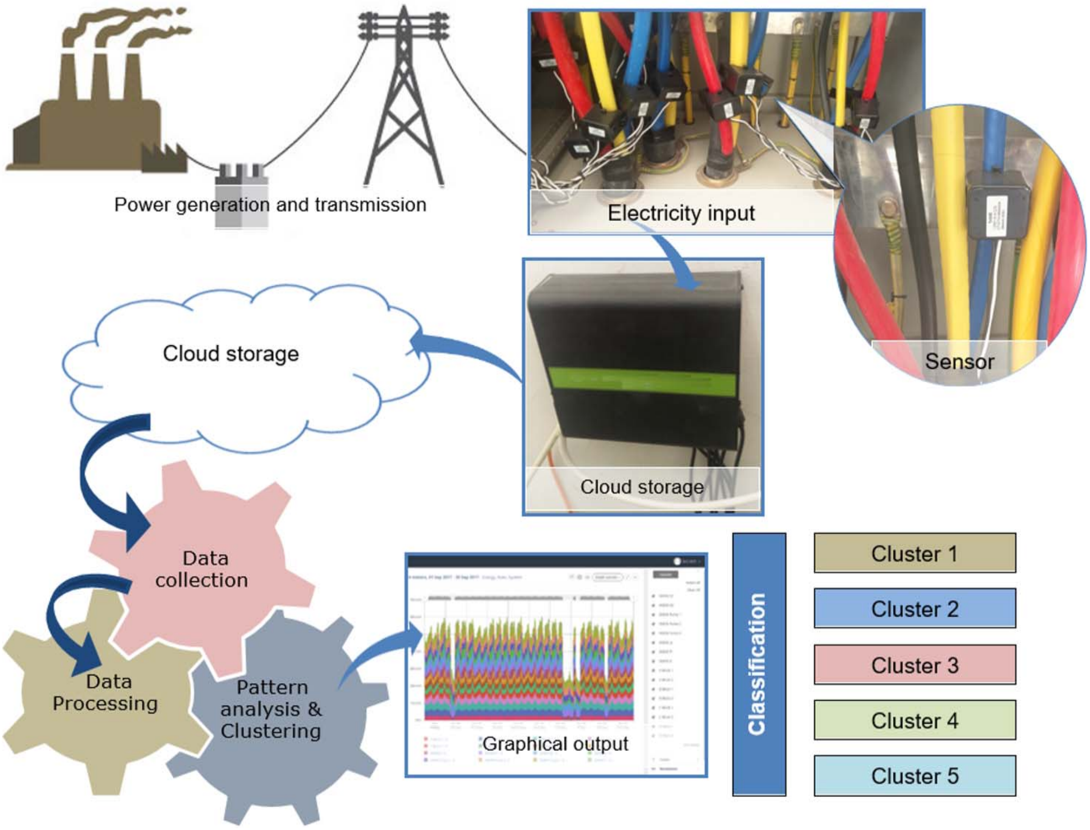


Table 1 Sample of recorded electricity consumption data (in kW)

Time	*Days*								
	1	*2*	*3*	*4*	*5*	*6*	*7*	⋯	*365*
0:00	188.9	195.9	190.2	187.2	192.2	190.0	190.9	⋯	169.9
1:00	183.2	188.5	179.2	179.6	183.7	181.5	184.5	⋯	171.8
2:00	179.1	182.3	171.5	171.9	177.3	174.8	179.6	⋯	172.2
3:00	173.1	178.1	167.7	169.2	170.4	173.9	175.5	⋯	172.4
4:00	170.2	171.3	167.3	168.7	170.4	174.3	171.3	⋯	178.3
5:00	170.4	172.0	171.4	170.8	171.9	175.1	173.5	⋯	175.6
6:00	173.4	171.0	182.5	177.9	182.7	187.0	172.9	⋯	174.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
23:00	188.6	191.4	192.6	195.3	192.6	196.7	185.9	⋯	181.7

Figure 3 Complete cloud-based data acquisition system



The acquired data are further cleaned and pre-processed for any errors or missing data. The data in the cloud has been stored in the SQL (Structured Query Language) databases, which has been exported in the Comma-Separated Values (CSV) file format for every hour and for the duration of 12 months. This data in CSV file is used for energy consumption pattern recognition and classification. On the other hand, to process this data, another cloud server is used to host MATLAB application. This server helps in connecting the client systems to MATLAB and it provides the different options of neural network over the internet on demand. The entire data collection process is represented in Figure 3.

3.2 *Pattern classification*

In this work, the Levenberg-Marquardt (LM)-based Neural Network (NN) has been used for energy consumption pattern classification. In recent years, a SVM has also been used for the classification (Gaidhane et al., 2016; Demuth et al., 2014), However, only binary decisions can be made using SVM. The one-to-all method can be used to achieve the multiclass classification. Therefore, it is required to propose a new approach for efficient multiclass classification. In literature, the neural networks are used widely for pattern classification (Azadeh et al., 2007; Dounis, 2010; Zhao and Magoulès, 2012; Yelampalli et al., 2019). The neural network can be used in supervised as well as unsupervised manner. Under supervised machine learning, the Levenberg-Marquardt algorithm has been considered. The LM-based neural network is explained in detailed in the next section.

3.3 *Levenberg-Marquardt algorithm*

Levenberg-Marquardt can be understood as the modified Newton's method which is specially designed to minimise the non-linear function. The error function is a sum of squares error. In neural network, the mean squared error is considered as an index of performance. The Newton's method for performance optimisation is given as

$$x_{k+1} = x_k - S_k^{-1} g_k \tag{11}$$

where $S_k = \nabla^2 Y(x)$ is called the Hessian matrix and $g_k = \nabla Y(x)$ is the gradient. The square error function $Y(x)$ can be defined as (Gaidhane et al., 2014)

$$Y(x) = \sum_{i=1}^{N} e_i^{\,2}(x) = \mathrm{e}^T(x) e(x) \tag{12}$$

Then, the *j*-th element of the gradient would be

$$\left[\nabla Y(x)\right]_j = 2\sum_{i=1}^{N} e_i(x) \frac{\partial e_i(x)}{\partial x_j} \tag{13}$$

Therefore, the matrix form of the gradient can be written as

$$\nabla Y(x) = 2J^T(x) e(x) \tag{14}$$

where J is the Jacobian matrix. Suppose, x_k is the value of *k*-th iteration, then x_{k+1} for $(k+1)$-th iteration can be calculated as

$$
\begin{aligned}
x_{k+1} &= x_k - \left[2J^T(x_k)\, J(x_k) \right]^{-1} 2J^T(x_k)\, e(x_k) \\
&= x_k - \left[J^T(x_k)\, J(x_k) \right]^{-1} J^T(x_k)\, e(x_k)
\end{aligned} \tag{15}
$$

Equation (15) is termed as Gauss-Newton method. The second derivative calculation is not required in Gauss-Newton method and that is the advantage over the conventional Newton's method. Further, the Hessian matrix S_k^{-1} can be approximated as

$$S_k^{-1} = \nabla^2 Y(x_k) + \mu_k I \tag{16}$$

Equation (16) leads to the Levenberg-Marquardt algorithm as

$$
\begin{aligned}
x_{k+1} &= x_k - \left[J^T(x_k)\, J(x_k) + \mu_k I \right]^{-1} J^T(x_k)\, e(x_k) x_{k+1} - x_k \\
&= -\left[J^T(x_k)\, J(x_k) + \mu_k I \right]^{-1} J^T(x_k)\, e(x_k)
\end{aligned} \tag{17}
$$

$$\Delta x_k = -\left[J^T(x_k)\, J(x_k) + \mu_k I \right]^{-1} J^T(x_k)\, e(x_k) \tag{18}$$

The μ_k element of LM algorithm has very distinct feature, as it increases the algorithm approaches the steepest decent with small learning

$$x_{k+1} = x_k - \frac{1}{\mu_k} J^T(x_k)\, e(x_k) = x_k - \frac{1}{2\mu_k} \nabla Y(x) \tag{19}$$

where μ_k decreases and approaches to zero the algorithm becomes Gauss-Newton as termed in equation (15). In the LM algorithm μ_k generally starts with value 0.01. Another factor σ is used in this algorithm, which has values greater then equal to 1. This factor will multiply (results into slow convergence) or divide (results into faster convergence) μ_k depending on a steps value for $Y(x)$, whether it is a large or small respectively.

In the present work, multilayer LM-based neural network has been used. It prevents the back-propagation to overcome the local minima problem and makes the network learning efficient (Wang et al., 2007). The Figure 4 shows the 3-layer architecture of the LM-based neural network.

The main purpose of LM-based neural network is to understand the relations between a set of input vectors d_i and target output vectors t_j pairs $\left[(d_1,t_1),(d_1,t_2),\ldots,(d_1,t_j)\right]$, $\left[(d_2,t_1),(d_2,t_2),\ldots,(d_2,t_j)\right],\ldots,\left[(d_i,t_1),(d_i,t_2),\ldots,(d_i,t_j)\right]$. The neural network performance measure is the mean square error (MSE). If each target output t occurs with equal probability, the mean squared error is proportional to the sum of squares error over the q targets in the training set. The square error function $Y(x)$ can be defined as

$$
\begin{aligned}
Y(x) &= \sum_{i=1}^{q} (t_i - d_j)^T (t_i - d_j) \\
&= \sum_{i=1}^{q} l_q^T l_q = \sum_{i=1}^{m} \sum_{j=1}^{q} (l_{j,i})^2 = \sum_{i=1}^{N} e_i^2(x)
\end{aligned} \tag{20}
$$

where $(l_{j,i})$ is the *j*-th element of the error for the *i*-th input/target (d_i,t_j) pair, $e^2(x)$ is the square error between the input vector d_j and target output t_i of the neural network, and $N = 1,2,3,\ldots,i \times j$. Here, x is the vector of all the weights and threshold. The square error function $Y(x)$ is to be minimised with respect to the parameter vector x. The error vector is the combination of error elements and can be written as

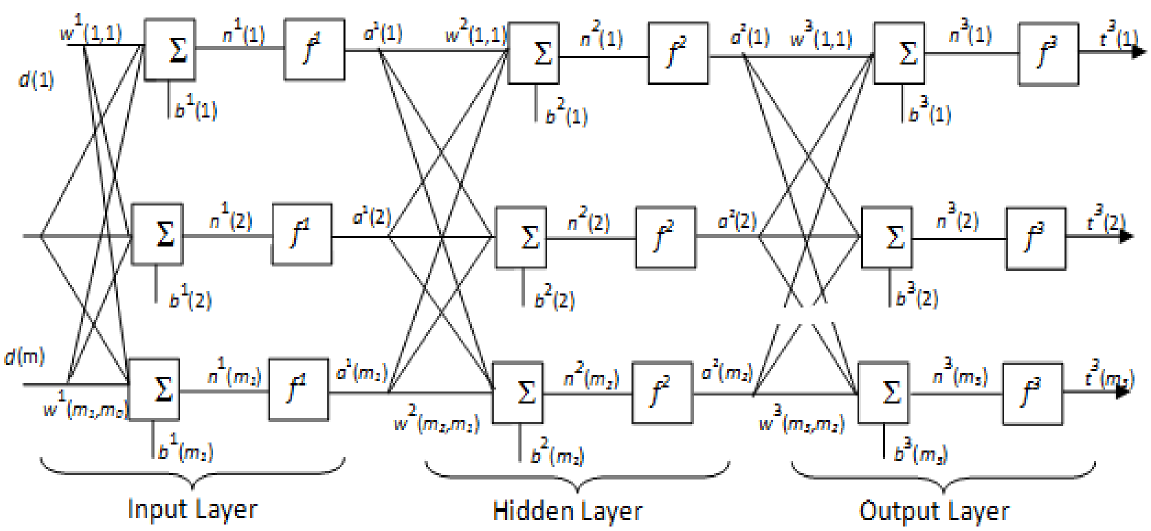
$$E^T = \left[l_{1,1}, l_{1,2}, \ldots, l_{1,j}, l_{2,1}, \ldots, l_{i,j} \right] = \left[e_1, e_2, \ldots, e_N \right] \tag{21}$$

Similarly, the parameter vector can be written as

$$P^T = \left[w_{1,1}^1, w_{1,2}^1, \ldots, w_{i,j}^1, b_1^1, \ldots, b_j^1, w_{1,1}^2, \ldots, b_i^j \right] = \left[x_1, x_2, \ldots, x_n \right] \tag{22}$$

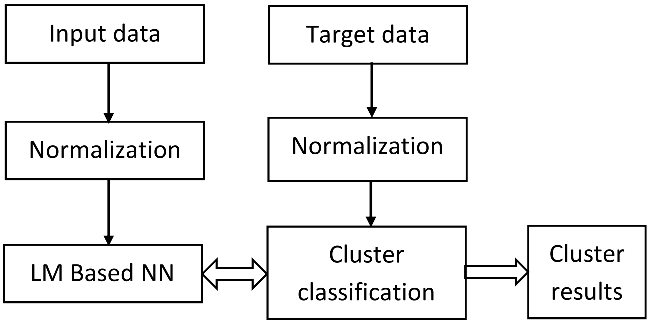
The backpropagation algorithm variant helps in computation of Jacobian matrix, which is the key step in LM algorithm (Singh et al., 2007).

Figure 4 3-layer LM-based neural network's architecture



The LM algorithm gets training from the input data to produce the output. The training set is populated from the input data. From the training set the ψ target tuples are extracted. The sum of squares error over the target tuples in the training set is calculated commonly called as MSE. The highly non-linear nature of the error function $Y(x)$ is best suited for the optimisation of the error. It can help the LM algorithm to adapt straightforward approach for the training. Thus, this algorithm can be a correct solution for the optimisation of the non-linear function. Therefore, it is a natural choice to accept LM algorithm for training the system. The neural network which uses LM-based algorithms for classification, can be used to evaluate the performance of such networks (Gaidhane et al., 2016). The whole process of energy consumption classification and clustering algorithm is shown in Figure 5.

Figure 5 Process of energy consumption classification and clustering algorithm



4 Experimental results and discussion

The proposed system is tested using the series of experiments on the energy consumption data. The experimental results are described and analysed in the following subsections. The LMNN-based proposed method is implemented using MATLAB R2017a on a cloud server. This cloud server has been set up with the following details to host the MATLAB software and web application. The server hardware is a Dell EMC machine with dual processors having 3.6 Ghz frequencies and 16 cores each, 256 GB RAM with 2 v100 Graphical Processing Units (GPUs). The server platform is a windows server family Operating Server (OS) (windows server 2012 R2), it is used as host OS. The type-1 hypervisor is used on top of the host OS to virtualise the server for multiple virtual server entities where software or web applications are hosted. The Internet Information Services (IIS) are enabled for this server and connected with a 10 GBPS speed of Ethernet. This server setup is used to host MATLAB R2017a software and web application for energy consumption data collection.

4.1 Energy consumption data collection and system setup

To determine the accuracy, performance and efficiency of the proposed system, the energy consumption data set of 12 months duration has been prepared. This data has been collected for an educational institute building, which is in Dubai International Academic City (DIAC). The data collected is on a per hour basis and considered for different patterns and classifications like working days, non-working days, holidays, summer and winter break, week days, weeks in months, per hour, per month, etc. as shown in Figure 2. The energy consumption varies from as low as 24 kW to as high as 270 kW in 12 months.

4.2 Data classification using LM-based neural network classifier (LMNN)

The LM-based neural network classifier (LMNN) is applied for classification of various energy consumption clusters. Each hour data collection for 24 hours in a day has suggested a 24-input neuron layer neural architecture. Therefore, various experimentations have been carried out using 24-input neuron at the input layer, 10-neuron hidden layer and 5-output neuron at the output layer in neural network architecture (24–10–5 NN architecture). This data of one year has been classified into five clusters and therefore, the output neuron grid is linked to these 24-input neurons, where a particular weight is assigned to each one of them to connect to other neurons.

In the recent work (Kumar et al., 2018a), SOM clustering method has been discussed and applied on the energy consumption data collected from an academic institute building. One of the resultant clusters of SOM is presented in Figure 6 for working days. A hexagon in a figure represents the classification of the neurons, which falls under same input pattern. The bigger hexagon represent the larger the patterns assigned to a neuron. For simplicity and understanding the energy consumption patterns, analysed further i.e., for workability, month wise, and weekday wise, etc. To verify the results of SOM, k-means clustering algorithm is employed and checked the authenticity of these results. Based on the k-means clustering, an algorithm has been employed to find out the optimal number of clusters for the energy consumption (Kumar et al., 2018a). The stepwise clustering patterns decision algorithm is summarised in Algorithm 1. The algorithm works on the workability of the input day of any month.

Algorithm 1 Clustering patterns decision algorithm

Step 1: Consider the input (day of any month).

Step 2: Check the workability of the input day

Step 3: If the workability is a working day then the input day will be one of the working days set (Sunday, Monday, Tuesday, Wednesday, Thursday).

Step 4: Else-if the workability is non-working day then the input day will be either weekly holiday (Friday, Saturday) or other holidays (Ramadan, winter break, summer break).

Step 5: Check the aggregate consumption and Mean Square Error against the cluster ranges.

Step 6: The matching cluster with aggregate consumption is confirmed.

To validate the above proposed clusters by *k*-means algorithm, various experimentations are carried out using LMNN employed with five-input neuron architecture. These five input neurons are linked to the 10 hidden neuron layers. These hidden layers are in turn are connected by particular weight and five-output neuron as shown in Figure 7.

4.3 Result verification

Generally, the performance and verification of experimental results is represented in the form of confusion matrix (Gaidhane et al., 2016, 2018). The confusion matrix is prepared to determine the performance of classification and clustering. Confusion matrix is a specific square table, which allows algorithm's performance visualisation. The input energy consumption data is for 365 days, which has been supplied to the LMNN algorithm. The algorithm has divided this data for training, validation and testing the system. 70% data (255 sample records) has been considered for training, 15% (55 sample records) each has been considered for validation and testing, respectively. Total 365 (255 + 55 + 55 = 365) sample records are considered for the experimentations. Table 2 shows the results for five-class of data clustering in a confusion matrix. From Table 2, it is observed that the proposed approach performs better as the accuracy is close to 98.85%, which is higher than the existing technique (Hernández et al., 2012).

Figure 6 Electricity consumption patterns for workability – working days (Kumar et al., 2018a)

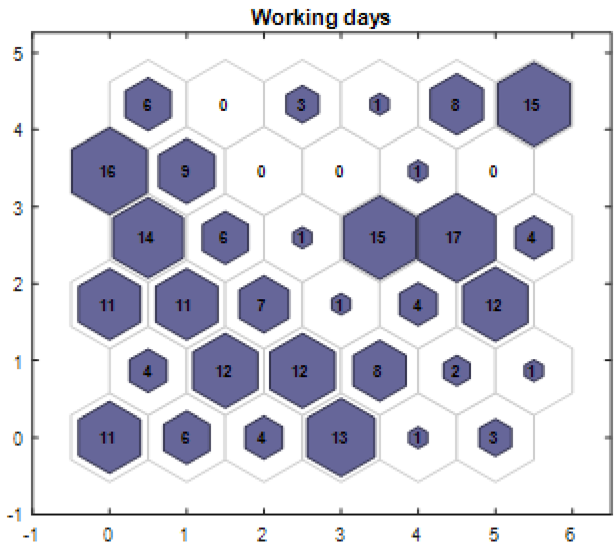


Figure 7 Basic architecture of back propagation supervised learning neural network (shown in MATLAB)

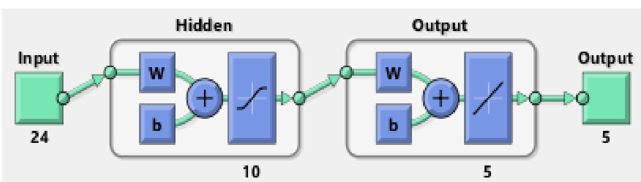


Table 2 Five-class confusion matrix for energy consumption classification using LMNN

Clusters	*C1 (%)*	*C2 (%)*	*C3 (%)*	*C4 (%)*	*C5 (%)*
Cluster 1	**96.83**	3.13	0	0	0
Cluster 2	0	**100**	0	0	0
Cluster 3	0	3.57	**96.43**	0	0
Cluster 4	0	0	4.55	**95.45**	0
Cluster 5	0	0.4	0	0.88	**99.12**

4.4 False positive and true positive test results

It is important for the classification method to correctly identify the clusters. The objective is to minimise the false positive classification and maximise the true positive classification. Therefore, it is necessary to validate the system against such inputs for faithful results (Gaidhane et al., 2018). The validation and evaluation of the proposed technique is done using false positive and true positive test. While a day's data is identified correctly for the cluster is true positive, whereas, if it is identified incorrectly for the cluster is false positive (Carter et al., 2016).

In this experiment, the energy consumption data for 365 days has been considered. This data is classified into five different clusters using Self-Organising Map (SOM). After clustering, the data has been given as input to the LMNN algorithm. The results of training (70%), validation (15%) and testing (15%) from LMNN has shown in Tables 3, 4 and 5, respectively. The overall results are shown in Table 6. It is observed from the results that the total 32 days data are classified into cluster 1 and only one-day data is falsely classified into cluster 2. Thus, this one-day data is considered as a false positive in cluster 2 (3.13% for cluster 2). Similarly, the 50 days data has been classified into cluster 2 perfectly. Moreover, the false positive and true positive test results are summarised in Table 7.

Table 3 Five-class training confusion matrix (number of records)

Clusters	*C1*	*C2*	*C3*	*C4*	*C5*
Cluster 1	**22**	0	0	0	0
Cluster 2	0	**33**	0	0	0
Cluster 3	0	0	**21**	0	0
Cluster 4	0	0	0	**21**	0
Cluster 5	0	0	0	0	**158**

Table 4 Five-class validation confusion matrix (number of records)

Clusters	*C1*	*C2*	*C3*	*C4*	*C5*
Cluster 1	**5**	0	0	0	0
Cluster 2	0	**11**	0	0	0
Cluster 3	0	0	**4**	0	0
Cluster 4	0	0	0	**1**	0
Cluster 5	0	0	0	0	**34**

Table 5 Five-class testing confusion matrix (number of records)

Clusters	*C1*	*C2*	*C3*	*C4*	*C5*
Cluster 1	**5**	1	0	0	0
Cluster 2	0	**6**	0	0	0
Cluster 3	0	0	**3**	1	0
Cluster 4	0	0	1	**0**	0
Cluster 5	0	0	0	2	**36**

Table 6 Five-class all confusion matrix (number of records)

Clusters	*C1*	*C2*	*C3*	*C4*	*C5*
Cluster 1	**32**	1	0	0	0
Cluster 2	0	**50**	0	0	0
Cluster 3	0	0	**28**	1	0
Cluster 4	0	0	1	**22**	0
Cluster 5	0	0	0	2	**228**

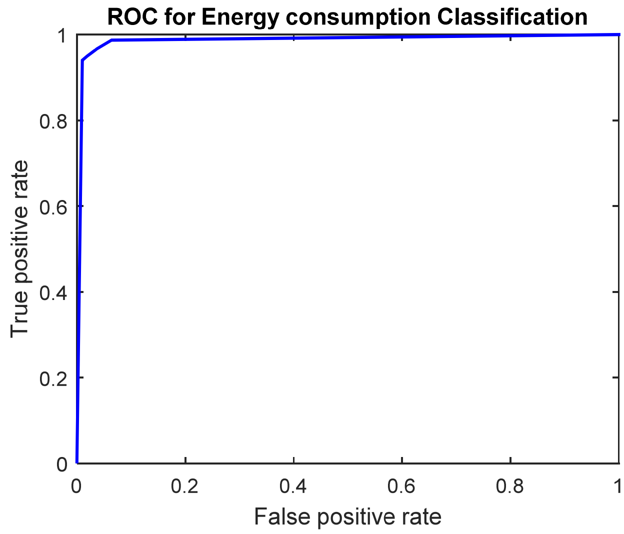
Table 7 Five-class cluster's true positive and false positive rate

Clusters	*True positive probability*	*False positive probability*
Clusters 1	0.9687	0.0313
Clusters 2	1	0
Clusters 3	0.9643	0.0357
Clusters 4	0.9545	0.0455
Clusters 5	0.9912	0.0880

In past years, Hernández et al. (2012) proposed a data processing system based on Self-Organising Map (SOM) and *k*-means clustering algorithm to analyse the electricity consumption patterns in the industrial parks. The data set used in this study has been collected offline from various resources and the larger environment like nations has been considered to carry out the analysis. Therefore, the real time clustering and analysis on such data set may not be applicable. Also, in recent work, Kumar et al. (2018a) presented a hybrid approach, which is based on SOM and *k*-means clustering algorithm. The real-time data from Dubai-based academic institute has been used to test and validate the system. It is observed that the highest accuracy for clusters finding efficiently of this system is 98% approximately. However, if the power consumptions in one cluster are very close to other clusters then the accuracy may come down. From Table 6, it is observed that false positive rate is very less and thus the proposed system performs better than the existing approaches (Kumar et al., 2018a, 2018b; Hernández et al., 2012; Tarkhaneh et al., 2018).

Furthermore, to evaluate the proposed classification algorithm, Receiver Operating Characteristic (ROC) curve is used. A binary classification rule is the base principle ROC follows and results into four possible outputs i.e. true positive or negative and false positive or negative (Carter et al., 2016). The false positive and true positive rates from Table 7 are used to plot the ROC curve as shown in Figure 8. It shows the ROC curve for the energy consumption classification pattern. It is observed from Figure 8 that the resultant curve is closer to the upper left corner, which indicates the higher overall accuracy of the test (Gaidhane et al., 2016).

Figure 8 ROC for energy consumption classification



For fair comparison and validation, more experimentations were carried out in the same manner on the next year (September 2017 to August 2018) data. The results are shown in Table 8.

Table 8 Five-class confusion matrix (number of records) for Sep. 2017 to Aug. 2018

Clusters	*C1*	*C2*	*C3*	*C4*	*C5*
Cluster 1	**28**	1	0	0	0
Cluster 2	0	**49**	0	0	0
Cluster 3	0	0	**31**	1	0
Cluster 4	0	0	1	**27**	0
Cluster 5	0	0	0	2	**225**

Further, different neural network architectures along with their accuracy are listed in the Table 9. It is observed from the table that the 24-10-5 neural network architecture results in best accuracy.

4.5 *Comparison of proposed approach and existing methods*

In this experiment, the comparison between *k*-means, SOM, *k*-means + SOM, BP-ANN and proposed approach has been carried out in the form of MSE and accuracy is summarised in Table 10. It is observed that the MSE using the proposed approach is much less and it is approximately 1.2076, which is lower than all other existing methods. Moreover, the accuracy is better (98.7924) than the existing methods. Further, the computational time for the *k*-means, SOM, *k*-means + SOM, BP-ANN and proposed approach has been computed and summarised in Table 10. It is observed from

Table 10 that the proposed approach takes less computational time as compared to the other existing methods.

Table 9 Neural network architectures' accuracy

S. no.	*Neural network architecture*	*Epochs*	*Mean square error (MSE)*	*Accuracy (%)*
1	24-2-5	7	1.6481	98.3519
2	24-3-5	6	1.5978	98.4022
3	24-4-5	8	1.9648	98.0352
4	24-5-5	7	1.8790	98.1210
5	24-6-5	4	1.4033	98.5967
6	24-7-5	12	1.7176	98.2824
7	24-8-5	6	1.4574	98.5426
8	24-9-5	3	1.3868	98.6132
9	**24-10-5**	**5**	**1.2076**	**98.7924**
10	24-11-5	7	1.7807	98.2193
11	24-12-5	7	1.7964	98.2036
12	24-13-5	7	1.8441	98.1559
13	24-14-5	5	1.8358	98.1642
14	24-16-5	5	2.3988	97.6012
15	24-18-5	8	1.8902	98.1098
16	24-20-5	5	1.8028	98.1972

Table 10 Comparison of proposed approach and existing methods (*k*-means, SOM, *k*-means + SOM and BP-ANN)

Method	*Mean square error (MSE)*	*Accuracy (%)*	*Computational time (Seconds)*
k-means	2.9562	97.0438	2.0419
SOM	2.7725	97.2275	2.0175
k-means+SOM	1.9284	98.0716	2.3618
BP-ANN	4.7204	95.2796	2.5609
Proposed approach	1.2076	98.7924	1.9123

5 Conclusions

In this paper, an efficient approach is proposed to implement a cloud-based application for electricity consumption analysis. A cloud-based web application has been used to collect the real-time electricity consumption data of an academic institute. The acquired data is stored on a cloud server. Further, the electricity consumption data has been tested and analysed using the LM-based artificial neural network. The various experimentations are carried out and the different clusters are formed based on possibilities for all the days of a year such as days with minimum or maximum electricity consumption in a week or month etc.

The results of the proposed approach are compared with the existing methods such as *k*-means, SOM and ANN. It is observed that the proposed approach performs better in the form of electricity consumption pattern classification as compared to the existing methods. Thus, the proposed approach can provide the meaningful information about the electricity consumption in the institute building. This will help the building management to plan for the electricity demand, supply, budgets, etc.

An alternative application of the proposed approach is in the field of improving the feedback provided to consumers about the consumption of the electricity. The feedback can be based on the comparison of the real-time electricity consumption with the desired consumption pattern in the past. This type of feedback will make well-suited suggestions that provide the awareness among the consumers about the electricity consumption and its optimisation. The presented approach can be applied on microgrid environments, small industrial parks, residential and commercial buildings, etc.

References

Akkaya, K., Guvenc, I., Aygun, R., Pala, N. and Kadri, A. (2015) 'IoT-based occupancy monitoring techniques for energy-efficient smart buildings', *Proceedings of the Wireless Communications and Networking Conference Workshops (WCNCW)*, *IEEE*, pp.58–63.

Amin-Naseri, M.R. and Sorroush, A.R. (2008) 'Combined use of unsupervised and supervised learning for daily peak load forecasting', *Energy Conversion and Management*, Vol. 49, No. 6, pp.1302–1308.

Annachhatre, C., Austin, T.H. and Stamp, M. (2015) 'Hidden Markov models for malware classification', *Journal of Computer Virology and Hacking Techniques*, Vol. 11, No. 2, pp.59–73.

Azadeh, A., Ghaderi, S.F., Tarverdian, S. and Saberi, M. (2007) 'Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption', *Applied Mathematics and Computation*, Vol. 186, No. 6, pp.1731–1741.

Basu, K., Debusschere, V., Bacha, S., Maulik, U. and Bondyopadhyay, S. (2015) 'Nonintrusive load monitoring: A temporal multilabel classification approach', *IEEE Transactions on Industrial Informatics*, Vol. 11, No. 1, pp.262–270.

Bui, D.T., Tuan, T.A., Klempe, H., Pradhan, B. and Revhaug, I. (2016) 'Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree', *Landslides*, Vol. 13, No. 2, pp.361–378.

Candanedo, L.M., Feldheim, V. and Deramaix, D. (2017) 'A methodology based on Hidden Markov Models for occupancy detection and a case study in a low energy residential building', *Energy and Buildings*, Vol. 148, pp.327–341.

Carter, J.V., Pan, J., Rai, S.N. and Galandiuk, S. (2016) 'ROC-ing along: evaluation and interpretation of receiver operating characteristic curves', *Surgery*, Vol. 159, No. 6, pp.1638–1645.

Chen, Y.H., Emer, J. and Sze, V. (2016) 'Eyeriss: a spatial architecture for energy-efficient dataflow for convolutional neural networks', *Proceedings of the ACM SIGARCH Computer Architecture News*, IEEE Press, Vol. 44, No. 3, pp.367–379.

Demuth, H.B., Beale, M.H., De Jess, O. and Hagan, M.T. (2014) *Neural Network Design*, Martin Hagan.

Dong, B. andrews, B., Lam, K.P., Höynck, M., Zhang, R., Chiou, Y.S. and Benitez, D. (2010) 'An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network', *Energy and Buildings*, Vol. 42, No. 7, pp.1038–1046.

Dounis, L. (2010) 'Artificial intelligence for energy conservation in buildings', *Advances in Building Energy Research*, Vol. 4, pp.267–299.

Gaidhane, V.H., Hote, Y.V. and Singh, V. (2014) 'A computationally efficient approach for template matching-based image registration, *Sādhanā*, Vol. 39, No. 2, pp.317–331.

Gaidhane, V.H., Hote, Y.V. and Singh, V. (2016) 'Emotion recognition using eigenvalues and Levenberg–Marquardt algorithm-based classifier', *Sādhanā*, Vol. 41, No. 4, pp.415–423.

Gaidhane, V.H., Hote, Y.V. and Singh, V. (2018) 'An efficient similarity measure approach for PCB surface defect detection', *Pattern Analysis and Applications*, Vol. 21, pp.277–289.

Hernández, L., Baladrón, C., Aguiar, J., Carro, B. and Sánchez-Esguevillas, A. (2012) 'Classification and clustering of electricity demand patterns in industrial parks', *Energies*, Vol. 5, pp.5215–5228.

Jokar, P., Arianpoo, N. and Leung, VC. (2016) 'Electricity theft detection in AMI using customers' consumption patterns', *IEEE Transactions on Smart Grid*, Vol. 7, No. 1, pp.216–226.

Kumar, G., Kaushik, M. and Purohit, R. (2018) 'Reliability analysis of software with three types of errors and imperfect debugging using Markov model', *International Journal of Computer Applications in Technology*, Vol. 58, No. 3, pp.241–249.

Kumar, N. and Mittal, R.K. (2012) 'Cloud computing setup for a campus environment', *Proceedings of the International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM'12)*, pp.88–91.

Kumar, N., Gaidhane V.H. and Mittal R.K. (2018) 'Clustering and analysis of electricity consumption patterns in educational building', *Journal of Engineering Technology*, Vol. 6, No. 2, pp.537–547.

Lajnef, T., Chaibi, S., Ruby, P., Aguera, P.E., Eichenlaub, J.B., Samet, M., Kachouri, A. and Jerbi, K. (2015) 'Learning machines and sleeping brains: automatic sleep stage classification using decision-tree multi-class support vector machines', *Journal of Neuroscience Methods*, Vol. 250, pp.94–105.

Li, C., Ding, Z., Zhao, D., Yi, J. and Zhang, G. (2017) 'Building energy consumption prediction: an extreme deep learning approach', *Energies*, Vol. 10, No. 10, pp.1525–1544.

Li, X., Bowers, C. P. and Schnier, T. (2010) 'Classification of energy consumption in buildings with outlier detection', *IEEE Transactions on Industrial Electronics*, Vol. 57, No. 11, pp.3639–3644.

Macedo, M.N.Q., Galo, J.J.M., De Almeida, L.A.L. and Lima, A.D.C. (2015) 'Demand side management using artificial neural networks in a smart grid environment', *Renewable and Sustainable Energy Reviews*, Vol. 41, pp.128–133.

McLoughlin, F., Duffy, A. and Conlon, M. (2015) 'A clustering approach to domestic electricity load profile characterisation using smart metering data', *Applied Energy*, Vol. 141, pp.190–199.

Noussan, M. and Nastasi, B. (2018) 'Data analysis of heating systems for buildings: a tool for energy planning, policies and systems simulation', *Energies*, Vol. 11, No. 1, pp.233–247.

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petković, D. and Sudheer, C. (2015) 'A support vector machine–firefly algorithm-based model for global solar radiation prediction', *Solar Energy*, Vol. 115, pp.632–644.

Pan, J., Jain, R., Paul, S., Vu, T., Saifullah, A. and Sha, M. (2015) 'An internet of things framework for smart energy in buildings: designs, prototype, and experiments', *IEEE Internet of Things Journal*, Vol. 2, No. 6, pp.527–537.

Paoli, C., Voyant, C., Muselli, M. and Nivet, M.L. (2010) 'Forecasting of preprocessed daily solar radiation time series using neural networks', *Solar Energy*, Vol. 84, No. 12, pp.2146–2160.

Singh, V., Gupta, I. and Gupta, H.O. (2007) 'ANN-based estimator for distillation using Levenberg-Marquardt approach', *Engineering Applications of Artificial Intelligence*, Vol. 20, No. 2, pp.249–259.

Stamp, M. (2017) *Introduction to Machine Learning with Applications in Information Security*, Chapman and Hall/CRC.

Tarkhaneh, O., Isazadeh, A. and Khamnei, H.J. (2018) 'A new hybrid strategy for data clustering using cuckoo search based on Mantegna levy distribution, PSO and *k*-means', *International Journal of Computer Applications in Technology*, Vol. 58, No. 2, pp.137–149.

Touzene, A., Al Yahyai, S. and Oukil, A. (2019) 'Smart grid resources optimisation using service oriented middleware', *International Journal of Computer Applications in Technology*, Vol. 59, No. 1, pp.53–63.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F. and Fouilloy, A. (2017) 'Machine learning methods for solar radiation forecasting: a review', *Renewable Energy*, Vol. 105, pp.569–582.

Wang, R.L., Zhang, C. and Okazaki, K. (2007) 'A new method of learning for multi-layer neural network', *International Journal of Computer Science and Network Security (IJCSNS)*, Vol. 7, No. 5, pp.86–89.

Wang, Y., Chen, Q., Kang, C. and Xia, Q. (2016) 'Clustering of electricity consumption behavior dynamics toward big data applications', *IEEE Transactions on Smart Grid*, Vol. 7, No. 5, pp.2437–2447.

Wang, Y., Chen, Q., Kang, C., Zhang, M., Wang, K. and Zhao, Y. (2015) 'Load profiling and its application to demand response: a review', *Tsinghua Science and Technology*, Vol. 20, No. 2, pp.117–129.

Yang, J., Santamouris, M. and Lee, S.E. (2016) 'Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings', *Energy and Buildings*, Vol. 121, pp.344–349.

Yelampalli, P.K.R., Nayak, J. and Gaidhane, V.H. (2019) 'A novel binary feature descriptor to discriminate normal and abnormal chest CT images using dissimilarity measures', *Pattern Analysis and Appl*ication, pp.1–10, https://doi.org/10.1007/s10044-018-00771-2 (In press)

Zhao, H.X. and Magoulès, F. (2012) 'A review on the prediction of building energy consumption', *Renewable and Sustainable Energy Reviews*, Vol. 16, No. 6, pp.3586–3592.

Zhou, B., Li, W., Chan, K.W., Cao, Y., Kuang, Y., Liu, X. and Wang, X. (2016) 'Smart home energy management systems: Concept, configurations, and scheduling strategies', *Renewable and Sustainable Energy Reviews*, Vol. 61, pp.30–40.