Modeling, Simulation and Optimization of Power Plant Energy Sustainability for IoT Enabled Smart Cities Empowered With Deep Extreme Learning Machine

SAGHEER ABBAS[1], **MUHAMMAD ADNAN KHAN**[1,2], **LUIS EDUARDO FALCON-MORALES**[3], **ABDUR REHMAN**[1], **YOUSAF SAEED**[4], **MAHDI ZAREEI**[3], **(Member, IEEE), ASIM ZEB**[5], **AND EHAB MAHMOUD MOHAMED**[6,7], **(Member, IEEE)**

[1]Department of Computer Science, National College of Business Administration and Economics, Lahore 54660, Pakistan
[2]Department of Computer Science, Lahore Garrison University, Lahore 54810, Pakistan
[3]Tecnologico de Monterrey, Escuela de Ingenieria y Ciencias, Jalisco 45138, Mexico
[4]Department of Information Technology, University of Haripur, Khyber Pakhtunkhwa 22620, Pakistan
[5]Department of Information Technology, Abbotabad University of Science and Technology, Khyber Pakhtunkhwa 1168873, Pakistan
[6]Electrical Engineering Department, College of Engineering, Prince Sattam Bin Abdulaziz University, Wadi Addawasir 11991, Saudi Arabia
[7]Electrical Engineering Department, Faculty of Engineering, Aswan University, Aswan 81542, Egypt

Corresponding author: Mahdi Zareei (m.zareei@ieee.org)

ABSTRACT A smart city is a sustainable and effective metropolitan hub, that offers its residents high excellence of life through appropriate resource management. Energy management is among the most challenging problems in such metropolitan areas due to the difficulty and key role of energy systems. To optimize the benefit from the available megawatt-hours, it is important to predict the maximum electrical power output of a baseload power plant. This paper explores the method of a deep extreme learning machine to create a predictive model that can predict a combined cycle power plant's hourly full-load electrical output. An intelligent energy management solution can be achieved by properly monitoring and controlling these resources through the internet of things (IoT). The universe of artificial intelligence has produced many strides through deep learning algorithms and these methods were used for data analysis. Nonetheless, for further accuracy, deep extreme learning machine (DELM) is another candidate to be investigated for analyses of the data sequence. By using the DELM approach, a high level of reliability with a minimum error rate is achieved. The approach shows better results compared to previous investigations since previous studies could not meet the findings up to the mark and unable to predict power plant electrical energy output efficiently. During the investigation, it is shown that the proposed approach has the highest accuracy rate of 98.6% with 70% of training (33488 samples), 30% of test and validation (14352 examples). Simulation results validate the prediction effectiveness of the proposed scheme.

INDEX TERMS DELM, ANN, feedforward, power plant, prediction, smart city, IoT.

NOMENCLATURES

u	Inputs
P	Outputs
G	Points of moving average
u_x	Starting value input
N	Natural number
α	Dimension of the input matrix

v	Input layer nodes
β	Dimension of the target matrix
l	Output layer nodes
σ	I/p features
B	Target matrix
K	Input layer node, weights between I/p and hidden layer
W	Weights between hidden neurons and output layer neurons

The associate editor coordinating the review of this manuscript and approving it for publication was Shagufta Henna.

R	Hidden layer output
S	Resultant matrix
γ	Make the matrix more generalized
A_{FE}	Output of the hidden layer
hl	Hidden layer
$R_{i,j}^{hl}$	weight and bias among the output and the hidden layer
tp_j	Calculated output
S'	Transpose of the resultant matrix
$R\gamma$	Hidden layer output
A_l	Weight matrix of lth hidden layer
R	Bias of hidden layer neurons
B_l	Estimated output of the first hidden layer
R_l	Estimated output of the lth hidden layer
$A_{FE}R_E$	Net hidden layer value

I. INTRODUCTION

The smart city is a new concept, well defined and used by many researchers and institutions [1]. In a simple term, the smart city is intended to tackle or minimize problems created by rapid urbanization and population growth, such as energy consumption, waste, and mobility, through the highest efficiency and resource optimization. In the literature, one can be finding a lot of classifications of the smart city intervention areas [2]. The main disadvantage of these categories is that they classify energy mainly based on the infrastructure, ignoring other appropriate energy aspects such as a smart grid. Urban life is an important challenge already in our daily lives [3]. The Population Fund of the United Nations forecasts that by 2030 around 60% of the world's population will live in urban areas whereas 27 large cities with over 10 million people are projected to exist [4]. Thus, urgent solutions for sustainable living and urban development are being sought. The energy demands of cities are complicated and extensive [5]. As a result, modern cities should enhance current technologies and execute new alternatives in a timely and effective manner, taking advantage of the interactions between all these energy approaches [6]. The intermittent use of renewable resources, increased demand, and the need for energy-efficient systems represent important energy challenges that, as usual, are more addressed in general than separately.

Simulation models were designed to help participants understand city dynamics and mitigate the influence of energy policy substitutes [7]. However, these attempts very often tackle individually the energy regions, where the whole image is missing and therefore inadequate alternatives are produced [8]. To satisfy the growing energy requirements of current and future cities effectively, an extensive smart city model that involves all energy operations whilst maintaining the volume and complexity of the model manageable is extremely desirable [9]. Energy management performs an important part in identifying an appropriate and efficient alternative to reduce peak demand and attain energy conservation [10]. The Internet of Things provides numerous advanced and ubiquitous applications for smart cities [11]. IoT applications are increasing their energy needs, while IoT devices are growing both in terms of numbers and requirements [12]. Therefore, Smart city solutions must be capable of efficient energy use and of addressing the challenges associated with it.

On the other hand, it can help us to look at the light at the end of the tunnel by developing huge information and communications technologies, and by using the internet power [13]. Even now, without the internet, someone can hardly imagine our lives [14]. It is even difficult to think that with the increasing accessibility of linked items; the internet today is to achieve individuals to discover data and to help us in our regular life with creative facilities is not expanded to create value from such enhanced accessibility of linked items [15]. This research is about helping people to move from the Internet of People to the Internet of Things (IoT), in particular in future smart cities. According to the European Research Cluster for the Internet of Things, IoT is an auto-configuring vibrant, worldwide network infrastructure where physical and virtual items can be identified and communicated by normal and interoperable procedures. While IoT is expected to be the perfect hope of viable city lives [16]. On the other hand, in order to make IoT smarter, lots of analysis technologies are launched into IoT; in which few of the most valuable technologies are Data Mining, Artificial Intelligence, Cloud Computing and Neural Networks [17].

In this research work, managing energy is regarded as the main prototype for the implementation of smart city complicated energy systems. This research work provides a short summary of energy management and difficulties in smart cities, then specify a cohesive structure for energy sustainability for IoT based smart cities. A deep extreme learning machine (DELM) approach will be used to make smart cities energy efficient with IoT enabled sensors with improved performance. Energy-efficient systems are entering future smart cities with superior equipment, control structures, and demand response policies [18].

The Internet of Things (IoT) transmits the prospective of transform communities around all over the world into "smart cities," actually by creating a new lifestyle of urban living [19]. The major benefits include increased safety, Health, improve education and living houses environments, Energy consumption, more efficient of climate and ecosystem, green economy and improve the employment system [3]. Even though the basic idea of smart cities has existed for almost more than 10 years, but the concept has come a long way since it is introduced at the beginnings, now the posed to radically alter city life that is an emergence of different serious enablers like the internet of Things. Reference [20] stated that the vision of smart cities is a safe, secure, eco-friendly and effective metropolitan center of the future because of all the structures like power, water, and transportation.

Reference [21] stated that a smart city is an innovation-intensive and improved city that links individuals, data and

urban features through innovative technologies, to build a sustainable, safer city featuring competition and innovative trade, and improved quality of living. As citizens shape such a city through ongoing relationships, they are the main component of smart cities. That is why a smart population is acknowledged as the main driving force of smart cities; therefore, in smart cities, learning, education, and communication are significant strategies. Reference [22] argued that a smart city is an intelligent community in which different components such as individuals, climate, mobility, democracy and the economy are constructed within an intelligent structure.

Fog Computing is an alternative term to cloud computing that places certain transactions and resources at the edge of a network, rather than creating channels for cloud storage and use [23]. Fog computation can decrease the bandwidth requirements by providing no data on cloud channels and aggregating it instead at certain entry points, such as routers. This enables more strategic data compilation that may not be necessary for cloud storage. Fog computer extends the normal cloud computing to the brink and is thus also known as edge computing [24].

In this paper, a deep extreme learning machine for the prediction of power plant electrical energy output is investigated to achieve the highest accuracy. In the training and testing of estimation of electrical energy output with deep learning, a data set named combined cycle power plant with 47840 data instances are used, so that each instance includes different and diverse characteristics. Consequently, the examination and comparison with state-of-the-art techniques in the same field are made.

The remainder of this paper is organized as follows. Section 2 briefly describes the related work. Section 3 presents the method to carry out a comprehensive evaluation for the prediction of power plant electrical energy output. Section 4 discusses the simulation and results of the DELM approach. Sections 5 discuss the conclusions from the study.

II. LITERATURE REVIEW

The smart city relates to the convergence, the generation of big data and artificial intelligence systems, and the evolving IoT and smart city systems [25]. IBM utilizes a concept of cognitive computing to define technologies that can learn from various information sets, provide explanations for interaction and acquire knowledge with individuals in natural languages. Google Now is also a service to make customer suggestions and to provide the customer with the most helpful data at the correct moment. This scheme is taught by users' previous behaviors and inputs in Google accounts including Calendar, Chrome, Gmail, Search and YouTube [26]. It is nearer to the cognitive age with the use of natural language comprehension embedded with other facilities like search engines.

IoT is the latest communication paradigm that will emerge in the future when all items used in daily life are integrated with electronic transmission and interaction microcontrollers, detectors and transceivers and protocol servers. This type of

protocol can be used to prepare them for a great opportunity to interact with one another and with customers and to become an essential part of the Internet [27]. The smart city project is most popular now a day. The latest estimates and research forecast of the World Health Organization that the population of the world will live in municipal regions in the future. Reference [28] supposes that, by 2050, 70% of the population will live in cities. The smartness project is used for creating luxury in life. The smart city is a project which uses the latest technology to enhance the quality of urban life. Smart cities also improve the performance of the environment and provide people with superior facilities [29]. Transmission and Information Technology is important to support and change urban cities to smart cities [30].

Reference [31] decided to explore new ways of tackling the problems of energy management. Among the approaches discussed he chose to focus in particular on methods for 'machine learning' to manage repeated in-patient real-time scenarios through the use of historical data. It was at a period when the revival of artificial neural networks was pursued with great enthusiasm.

In order to address nonlinear relations in the ability of artificial neural networks, environmental conditions are considered as inputs of the model and energy generated as outputs of the model. We can forecast the plant's output power based on environmental conditions using this model [32]. A GA method for the production of MLP for CCPP power output estimates has been proposed in this study [33]. MLP model selection based on the GA Multiple- layers perceptron is performed using an individual mutation, three different crossover processes, and two distinct fitness functions for several hidden layers.

The Internet of Things (IoT) connects multiple physical gadgets, cars, mobile and dissimilar stuff that can interact and exchange data with hardware, software, sensors, actuators and network [34]. Since it was founded, IoT has played a crucial part from a variety of traditional equipment to prevalent household objects and in latest years has attracted the attention of educational, academic, technological and industrial scientists as well as the public. The main aim is to control, manage, and monitor everything readily at a key stage, and to automatically identify other stuff which is interconnected and can even make choices on their own.

In [35], For different local ambient conditions, the ANN model is employed to predict the working parameters and efficiency of a gas turbine. Intelligent systems are also used to model a fixed gas turbine. In [36], the identification methods for the ANNs have been established and the findings have shown that the identification system for the ANN is perfectly suited for estimating gas turbine conduct from full speed to full load conditions in many different operating points.

In addition, the research on electricity consumption with machine learning instruments has also been carried out several times [37], [38]. Some research, such as [39], has been found to be a similar study on the total electric power produced by a cogeneration power station with three gas

turbines, one steam turbine, and a district heating system in this article.

Similar to waste storage procedures and norms, procedures and mechanisms for the recycling of information are needed in metropolitan towns with the generation of hundreds or thousands of gigabytes of information per second. Methods of data analytics and algorithms of machine learning should be prepared to obtain information and knowledge to decrease digital waste [40]. Although computing and storage techniques have recently advanced, most data-analytical approaches take advantage of sampling techniques that are time-efficient but overlook a big number of data that can contain significant information models not presented in the samples. Data sets with lots of constraints can be taken to obtain perceptive analysis by means of the use of deep neural networks (DNNs) [41]. The most important thing, is the IoT platform actually designing and development, requires a perfect solution that is called middleware-level solution to enable the seamless interoperability between Machine-to-Machine based application and existing internet-based service [42].

There are several works in the research community that suggest cognitive solutions, that are appropriate to the requirements of IoT devices. Reference [43] suggested a cognitive management structure that allows smart objects to communicate and thus create end-users more conscious. The focus in this work was on the reuse of available object functionality and services across three levels, comprising virtual objects (VOs), virtual composites (CVOs) and level of service. The type of service originates from the features that a stakeholder or a specific implementation requires of the desired service. The CVOs will be responsible for these functionalities. The writers demonstrated that in their suggested structure the service delivery period is reduced, leading to reduced operating costs.

An additional research project carried out by Wu et al. to develop a cognitive structure for IoT applications is Cognitive IoT (CIoT) [44]. The structure provides relationships of five behavioral functions: a perception-action cycle, mass analysis of information, semantic derivation and discovery of information, smart decision making and supply of on-demand services. They acknowledged two areas for the understanding and learning of objects in a cognitive setting. They derived the semantics from the analyzed data and found useful outlines and directions as knowledge.

Reference [45] outlined a cognitive framework for smart homes on the basis of cognitive dynamic, IoT systems and utilized the Bayesian, Bayesian filter and RL models in the core of their cognitive memory. The Bayesian model ranks above the environmental control unit. The Bayesian filter measures the status of the scheme and RL offers the method for determining the finest feasible activities based on the complete awards.

Reference [46] have opted for integrating artificial intelligence in fog computing to promote intelligent large data exploration. In comparison to centralized cloud intelligence and assessment. They implemented a hierarchical fog computing model for big data analysis for intelligent city applications. This model enhances general efficiency by decreasing the communication bandwidth because raw information is not transmitted to the cloud, and because the fog is close to the information source in real-time analysis. In their model, they used a hidden Markov (HMM) model to assist big data analysis in an intelligent pipeline surveillance scheme.

III. EXPERIMENTAL

A. SYSTEM MODEL

The smart city manages and controls resources through smart data systems. To ensure optimum supplies and effective use of urban assets, the development of IoT-based technologies to tackle these issues generated by those technologies is essential. Furthermore, intelligent alternatives are the primary premises of a smart city for transport, healthcare, comfort, farming, and public. In this research work, the energy efficiency model has been introduced for smart cities, which show how deep extreme learning machine techniques are used to provide the inhabitants of the cities with superior facilities. The general intelligence context in smart cities is presented in this paper. This structure provides four intelligence concentrations: smart city and IoT facilities, deep extreme learning machine, fog computation, and cloud computation.

FIGURE 1 shows the general situation of the deep extreme learning machine approach within the smart city structure hierarchy where the intelligent software manager that is implemented in fog or in the cloud depends on the features of the analytics needed. The raw information can, therefore, be transmitted to the cloud or the fog. The operational analytical agent based on the approach of a deep extreme learning machine then gives a proper response based on predictions for infrastructure devices (e.g. adjust the consumption of energy on the basis of the information). The reason behind this architecture is to deepen the abstraction of information and understanding while transiting the information through the smart city infrastructure. A city-wide abstraction is required at the highest levels for the long-term management of city resources and services. On the other side, sensor or intelligent object-generated information is used at the smallest stage for short-term management of assets and facilities. In addition, fog-driven analytics promote local activities in predefined situations, while cloud-driven analysis can cover broader geographical areas with diverse situations. Fog Computing further divided into two sub-phase training and validation phase. In the training phase, backpropagation is used to train data and then trained model export to the cloud. Respectively in validation phase trained model import from the cloud and predict electricity production of a power plant from real data.

IoT infrastructure levels are used to detect surroundings by the sensors and resource-contracted equipment. The limitation of the resources of those devices impedes the use of complex and extensive learning models. However, in order to take analytics and intelligence nearer to the information source (for example, end-users, IoT-resource resilient equipment), contemporary and sophisticated designs

of learning like deep extreme learning machines are required.

A new path of research is to overcome this resource limit, so that deeper neural model networks can be used. During the past few years, various methods have also been suggested for compressing or planting deep neural networks in order to load them into IoT systems, wearable electronics, and smartphones that are restricted by resources [47]. The raw information is collected and transferred to the Cloud computing level on the fog computing stage. The DELM technique can be used at this stage because the resources at this stage are lower than the IoT assets. Lightweight intelligence must also be provided at this stage to the IoT gates and representatives in order to efficiently integrate facilities in the assistance of smart city applications [48]. At the cloud computing stage, mechanisms and methodologies can be incorporated with semantical teaching and ontologies to obtain high-level ideas and trends from the gathered information. Deep learning models are extremely appropriate at this stage since they can provide a more profound overview of data.

The purpose is to make the infrastructure more intelligent so that resources can efficiently utilize it. They want modern technology, including water wastes, electricity consumption, transport traffic congestions and so on, to address many pressing problems. Building intelligent cities help them address all of these issues, resulting in favorable economic results. This will make the living atmosphere more effective and sustainable.

The assumption of smart cities appears interesting, but how viable is it. This must obtain information in each industry in order to be effective. It is hard to implement alternatives at this level because every city is distinctive and therefore each moment a fresh set of issues arises. If someone wants to construct a scalable model, this needs to discover models that can constantly be used by many distinct cities to implement the intelligent city model. The focus is on intelligent data analysis and building solid information collection capabilities, communication protocols, inter-operability between computers, data storage systems, intelligence levels, etc.

In this research work, the deep extreme learning machine (DELM) technique is being unified to make smart cities energy efficient with IoT enabled sensors. FIGURE 3 demonstrates that in deep extreme learning machine (DELM) diverse amounts of hidden layers, different hidden neurons, and numerous kinds of activation functions have been used to attain the finest structure of DELM for energy efficiency. The proposed technique comprises of three diverse layers, namely data acquisition, pre-processing, and application layer. In the application layer, there are two sub-layers namely the prediction layer and performance evaluation. Real data from sensors are collected and actuators for experimental analysis. The data collected are provided for the acquisition layer as an input. Various data cleaning and preparation systems were implemented to extract anomalies from the information in the pre-processing layer. In the application layer, Deep extreme learning machine (DELM) have

been used for energy efficiency. The DELM takes the benefits of both extreme learning and deep learning techniques [49].

The complete system procedure is shown in FIGURE 2. In which layer of data acquisition contains the parameters of input, they will go to the neural system, where an algorithm has been trained to predict power plant electrical energy output. Nowadays, artificial neural networks can be used in all sectors. The artificial neural network comprises of a set of neurons which are characterized by special arrangement.

The main parts of an artificial neural network are neurons and connections between them. A neuron is the fundamental unit of processing information that forms the foundation for the performance of ANN. Neurons are conjunct processing elements that work together to solve a problem.

B. DEEP EXTREME LEARNING MACHINE

The deep extreme learning machine (DELM) is a well-known method used in various areas for predicting health problems, energy consumption predictions, transportation and traffic management, etc. [49]. The traditional ANN algorithms require more samples and slow learning times and can overfit the learning model [50]. The idea of ELM was first specified by [51]. The DELM is used widely in various areas for classification and regression purposes because DELM learns fast and it is efficient in the cost of computational complexity. Extreme learning machine is feedforward neural network which means data only goes one way through the series of layers but we have used backpropagation method in this proposed model during training phase where information flows back through the network and in backpropagation method network adjust the weights to achieve high accuracy with minimum error rate. During validation phase weights of the network are constant in which we import the trained model and predict the real data. There are three layers included in the DELM model the input layer, multiple hidden layers, and an output layer. The structural model of a DELM is shown in FIGURE 3, where np represents input layer nodes, ℓ represents hidden layer nodes, and O_p indicates output layer nodes.

In order to make energy sustainable cities, the prediction of electricity production in a power plant is a major real-life problem. For the effectiveness and cost-effectiveness of a power plant, the complete energy production of a baseload power plant should be properly predicted. It is helpful to increase the revenue from the accessible megawatt-hours (MWh). Turbine reliability and durability rely heavily on the forecast of its energy generation, especially where elevated profitability and contractual liabilities are restricted. This research examines the deep extreme learning machine (DELM) method for developing a predictive model to estimate the complete energy production of a power plant hourly.

C. SERVICE-ORIENTED ARCHITECTURE

The main advantage of smart grids lies in the capacity of incorporation of energy sources into the network and control of energy consumption and generation, as is shown in III-D.

Energy generation is the first step of the smart grid value chain, it involves power sources and relieves broad-based technology monitoring and control to communicate with the next step known as power distribution. This is focused on a proximity network connecting consumers to the grid and transmitting data via advanced metering facilities. Power consumption is the final step in the smart network value chain, involving both residential and industrial electricity users. Therefore, to maximize the operation, it is very important to supervise their consumption and production.

D. DATASET

A power plant's baseload function is affected by four primary parameters that are used in the dataset as input factors, such as temperature (T), ambient pressure (AP), relative humidity (RH), and exhaust vacuum (V). These parameters influence the production of electrical power, which is termed the target variable such as electrical power output (EP). The temperature (T) input variable is measured in the range between 1.81 °C and 37.11 °C. Ambient pressure (AP) is the input variable and is measured in the range between 992.89 and 1033.30 millibar. The variable relative humidity (RH) is evaluated as a percentage in the range from 25.56% to 100.16%. The variable exhaust vacuum (V) is evaluated with the spectrum between 25.36 to 81.56 cm Hg. The data set uses electrical energy output (EP) as a target variable. It is calculated in megawatt with the range of 420.26–495.76 MW. In the data acquisition layer, inputs will be taken from the collection of data such as temperature(T), ambient Pressure, (AP), relative humidity (RH) and exhaust vacuum(V). These variables affect the output of electrical energy, which is termed as the target variable (EP). In the pre-processing layer, cleaning abnormalities in data and data reduction will be used for quality data in machine learning. In the application layer, a deep extreme learning machine method will be considered for energy management. In the evaluation layer, three parameters mean absolute error (MAE), root means square error (RMSE), means absolute percentage errors (MAPE) and mean square error (MSE) will be observed for calculating full load electrical energy output to improve the efficiency of energy management for smart cities. In this article, DELM was used to train and fit 47840 sets of data. This data arbitrarily divides into 70% of training (33488 samples), 30% of data is used for validation and testing (14352 samples). In Table 1 a pseudocode of proposed deep extreme learning machine of a powerplant electrical energy output is described.

In Eq. (1), a mathematical representation of the moving average filter is considered. Where u ( ) denotes the inputs, P [ ] denotes the output and G represents the point of moving average.

$$P[x] = \frac{1}{G}\sum_{T=0}^{G-1} u(x+T) \quad (1)$$

In the modeling of machine learning algorithms to increase predictability and to improve the training process, complete

TABLE 1. Pseudocode of proposed DELM based training model of powerplant electrical energy output.

Sr No.	Steps
1	Start
2	Initialization σ , B , K, W, Error (*F*)= 0 and the number of epochs £ = 0
3	For each training pattern p
	a) do the feedforward phase to i) Calculate AF_{hl} for each hidden layer, where, hl=1,2…,6. ii) calculate $R_{i,j}^{hl}$ for each hidden layer, where, hl =1,2…,6.
	b) Compute output error signals and hidden layer error signals.
	c) do the backpropagation phase to i) Calculate the $\Delta R_{i,j}^{hl}$ ii) Then adjust/update the weights $R_{i,j}^{hl}$ of all hidden layers, where, hl =1,2…,6.
	£ = £ + 1
5.	Test stopping criteria: if no stopping criterion is satisfied, go to step 3.
6.	Stop

sample data have been standardized to fit in the interval [0, 1] by using Eq. (2) given below:

$$C = \frac{u_x - u_{min}}{u_{max} - u_{min}}, \quad \text{where } x = 1, 2, 3 \dots N \quad (2)$$

Take a training sample at first $[\sigma,B] = [\sigma_v, \beta_v]$ (x = 1, 2, ..., Z), and input sample $\sigma = [\alpha_{v1}\alpha_{v2}\alpha_{v3}... \alpha_{vz}]$ and a targeted matrix $B = [\beta_{11}\beta_{12}\beta_{13}... \beta_{1z}]$ was composed of samples from training, the matrices σ and B can then be represented, respectively, as in Eqs. (3) and (4), where the dimensions α and β are the input matrix and output matrix. The ELM will then adjust weights arbitrarily among input and the hidden layer. Where k_{D1} represents the weights between the kth input node and l^{th} hidden layer node as shown in Eq. (5).

$$\sigma = \begin{bmatrix} \alpha 11 & \alpha 12 \cdots & \alpha 1n \\ \alpha 21 & \alpha 22 \dots & \alpha 2n \\ \alpha 31 & \alpha 32 \dots & \alpha 3n \\ \vdots & \vdots & \vdots \\ \alpha_{D1} & \alpha_{D2} & \alpha_{Dn} \end{bmatrix} \quad (3)$$

$$B = \begin{bmatrix} \beta 11 & \beta 12 \cdots & \beta 1n \\ \beta 21 & \beta 22 \dots & \beta 2n \\ \beta 31 & \beta 32 \dots & \beta 3n \\ \vdots & \vdots & \vdots \\ \beta_{D1} & \beta_{D2} & \beta_{Dn} \end{bmatrix} \quad (4)$$

$$K = \begin{bmatrix} \kappa 11 & \kappa 12 \cdots & \kappa 1n \\ \kappa 21 & \kappa 22 \ldots & \kappa 2n \\ \kappa 31 & \kappa 32 \ldots & \kappa 3n \\ \vdots & \vdots & \vdots \\ \kappa_{D1} & \kappa_{D2} & \kappa_{Dn} \end{bmatrix} \quad (5)$$

$$W = \begin{bmatrix} \varpi 11 & \varpi 12 \cdots & \varpi 1n \\ \varpi 21 & \varpi 22 \ldots & \varpi 2n \\ \varpi 31 & \varpi 32 \ldots & \varpi 3n \\ \vdots & \vdots & \vdots \\ \varpi_{D1} & \varpi_{D2} & \varpi_{Dn} \end{bmatrix} \quad (6)$$

Next, the ELM has randomly selected the biases of the hidden layer nodes, as in Eq. (7). Extreme Learning Machine also preferred an f(u) function that was the network activation function. Data acquisition layer in Fig. 1, the resulting matrix is shown in Eq. (8). Respectively the column vector of the resulting matrix S shown in Eq. (9) [49].

$$B = [\beta_1, \quad \beta_2, \quad \beta_3 - - - - - - - \beta_D] T \quad (7)$$

$$S = [s_1, \quad s_2, \quad s_3 - - - - - - s_z] \, m * n \quad (8)$$

$$s_1 = \begin{bmatrix} s_{1j} \\ s_{2j} \\ s_{3j} \\ \vdots \\ s_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{e=1}^{i} M_{c1} f \left(\forall_c i_e + \beta_c \right) \\ \sum_{e=1}^{i} M_{c2} f \left(\forall_c i_e + \beta_c \right) \\ \sum_{e=1}^{i} M_{c3} f \left(\forall_c i_e + \beta_c \right) \\ \vdots \\ \sum_{e=1}^{i} M_{cn} f \left(\forall_c i_e + \beta_c \right) \end{bmatrix}$$

$$(e = 1, 2, 3, \ldots., n) \quad (9)$$

Then we can obtain Eq. (10) with regard to Eqs. (8) and (9). The outcome of the hidden layer is R and transposition of S as S' and values of weight matrix γ are calculated in Eq. (11) with the least square method.

$$R\gamma = S' \quad (10)$$

$$\gamma = R^+ S' \quad (11)$$

The γ regularization term was used to increase the network's overall stability [52]. Deep learning emerges and now a day the very famous subject for researchers. A system with a minimum of four layers with inputs/outputs fulfills the needs of a deep learning system. The neurons of every layer are trained in a deep neural network on a diverse set of parameters with the result of the previous layer. This allows extensive datasets to be processed by the deep learning networks. Deep learning captured numerous researchers' attention because it is very effective in solving real-world issues. DELM is used in our proposed work to capture both ELM advantages and deep learning. DELM's model comprised of a single input layer with four neurons, six hidden layers, with Ten neurons each hidden layer, and FIGURE 3 shows one layer of output with one neuron. The test and error method for selecting the number of nodes from hidden layers was used because of the lack of any special mechanism for specifying hidden layer neurons. The second hidden layer output is achievable as;

$$R_l = S\gamma^+{}_l \quad \textbf{Where } l = 1, 2, 3, \ldots 6 \quad (12)$$

where γ^+ is the general inverse of a matrix γ. Thus, second hidden layer values can be easily attained through Eq. (11).

$$f (A_l R + B_l) = R_l \quad (13)$$

In the Eq. (13), the parameters A_l, R, B_l, and R_l represent the first two hidden layers' weight matrix, preference of the first neurons in the hidden level, the first hidden layer assessed output, and the second hidden layer estimated output respectively.

$$A_{F_E} = f^{-1} (l) F_E^+ \quad (14)$$

F_E^+ is the inverse of F_E and in order to compute Eq. (5), the active function $f(\mathrm{x})$ is being used. Hence the required outcome of the second hidden layer is revised as follows by indicating the correct $f(\mathrm{x})$ activation function:

$$R_{l+1} = f (A_{FE} R_E)$$
$$\textit{where } A_{FE} R_E = Nh_{l+!}$$
$$R_{l+1} = f (Nh_{l+1}) \quad (15)$$

Updating the weighted matrix γ among the second layer and third layer as per Eq. (16). R_{l+1}^+ is the inverse of R_{l+1}. The estimated layer's results are thus shown as Eq. (17).

$$\gamma_{l+1} = R_{l+1}^+ S \quad (16)$$

$S\gamma^+ new$ is the inverse of the weight matrix μ_{l+1}. The DELM then sets the matrix $F_{ME} l = [B_{l+1}, A_{l+1}]$. Eqs. (10) and (11) allow the output of the further layer to be achieved.

$$f (x) = \frac{1}{1 + e^{-x}} \quad (17)$$

The back-propagation algorithm incorporates, weight initialization, feedforward, back error propagation, and weight and bias update is subject to distinctive developments. An activation function like $g (x) = sigmoid$ exists on each neuron in the hidden layer. The sigmoid input function and the hidden layer of the DELM can be composed in this way;

$$F = \frac{1}{2} \sum_{j} (rj - tp_j) 2 \quad (18)$$

r_j = Desired output
tp_j = calculated output
Eq. (18) shows the backpropagation of error, which can be calculated by the sum of the square of the desired output from the calculated output divided by 2.

To reduce the overall error the adjustment of weight is required. Eq. (19) shows the rate of change in weight for the output layer.

$$\Delta R_{i,j}^{hl=6} \propto - \frac{\partial H}{\partial R^{hl=6}}$$
$$\textit{where } i = 1, 2, 3 \ldots \ldots \ldots 10 \, (Neurons)$$
$$\text{and j} = \text{output Layer} \quad (19)$$

$$\Delta R_{i,j}^{hl=6} = -\text{const} \frac{\partial H}{\partial R^{hl=6}} \quad (20)$$
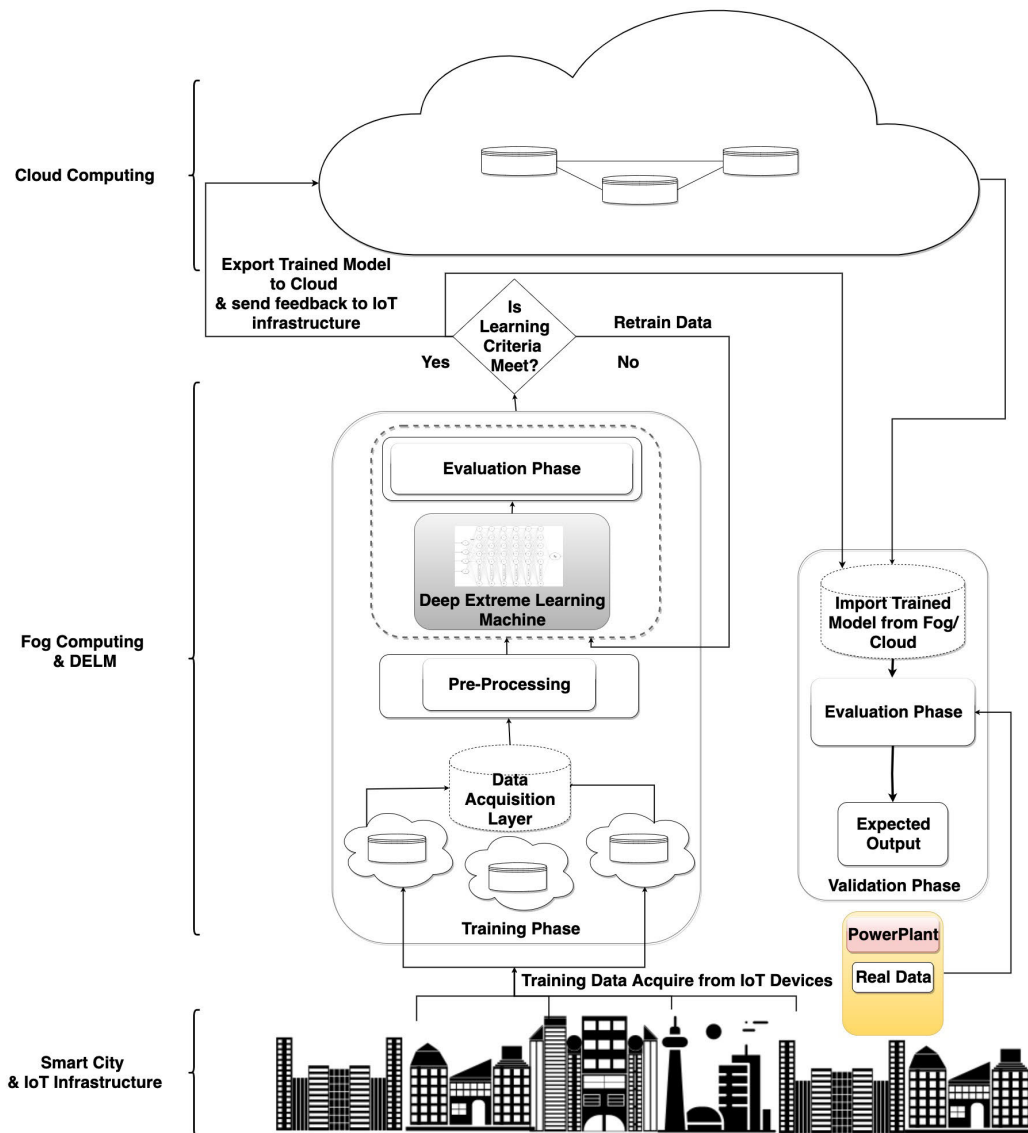
FIGURE 1. Proposed model for energy sustainability in the smart city.

writing Eq. (20) by using the chain rule method

$$\Delta \mathrm{R}_{i,j}^{hl=6} = -\text{const} \frac{\partial H}{\partial tp_j^{hl}} \times \frac{\partial tp_j^{hl}}{\partial NhR_j^{hl}} \times \frac{\partial NhR_j^{hl}}{\partial R_{i,j}^{hl}} \quad (21)$$

The value of change weight can be achieved after substituting the values in Eq. (21) as shown in Eq. (22).

$$\Delta \mathrm{R}_{i,j}^{hl=6} = \text{const}\left(r\mathrm{j} - tp_j\right) \times \left(tp_j^{hl}\left(1 - tp_j^{hl}\right) \times tp_j^{hl}\right)$$

From tp to R_6

$$\Delta \mathrm{R}_{i,j}^{hl=6} = \text{const} \ni_j tp_j^{hl} \quad (22)$$

The calculation to determine appropriate weight change to the hidden weight is shown in the following procedure. It is more complex because the weighted connection can lead to errors at all nodes.

From R_6 to R_1 or R_n
Where $n = 5, 4, 3, 2, 1$

$$\Delta \mathrm{R}_{i,n}^{hl} \propto -\left[\sum_j \frac{\partial H}{\partial tp_j^{hl}} \times \frac{\partial tp_j^{hl}}{\partial NhR_j^{hl}} \times \frac{\partial R_j^{hl}}{\partial tp_n^{hl}}\right] \times \frac{\partial tp_n^{hl}}{\partial NhR_n^{hl}}$$
$$\times \frac{\partial NhR_n^{hl}}{\partial R_{i,n}^{hl}}$$

$$\Delta \mathrm{R}_{i,n}^{hl} = -H\left[\sum_j \frac{\partial H}{\partial tp_j^{hl}} \times \frac{\partial tp_j^{hl}}{\partial NhR_j^{hl}} \times \frac{\partial R_n^{hl}}{\partial tp_n^{hl}}\right] \times \frac{\partial tp_n^{hl}}{\partial R_n^{hl}}$$
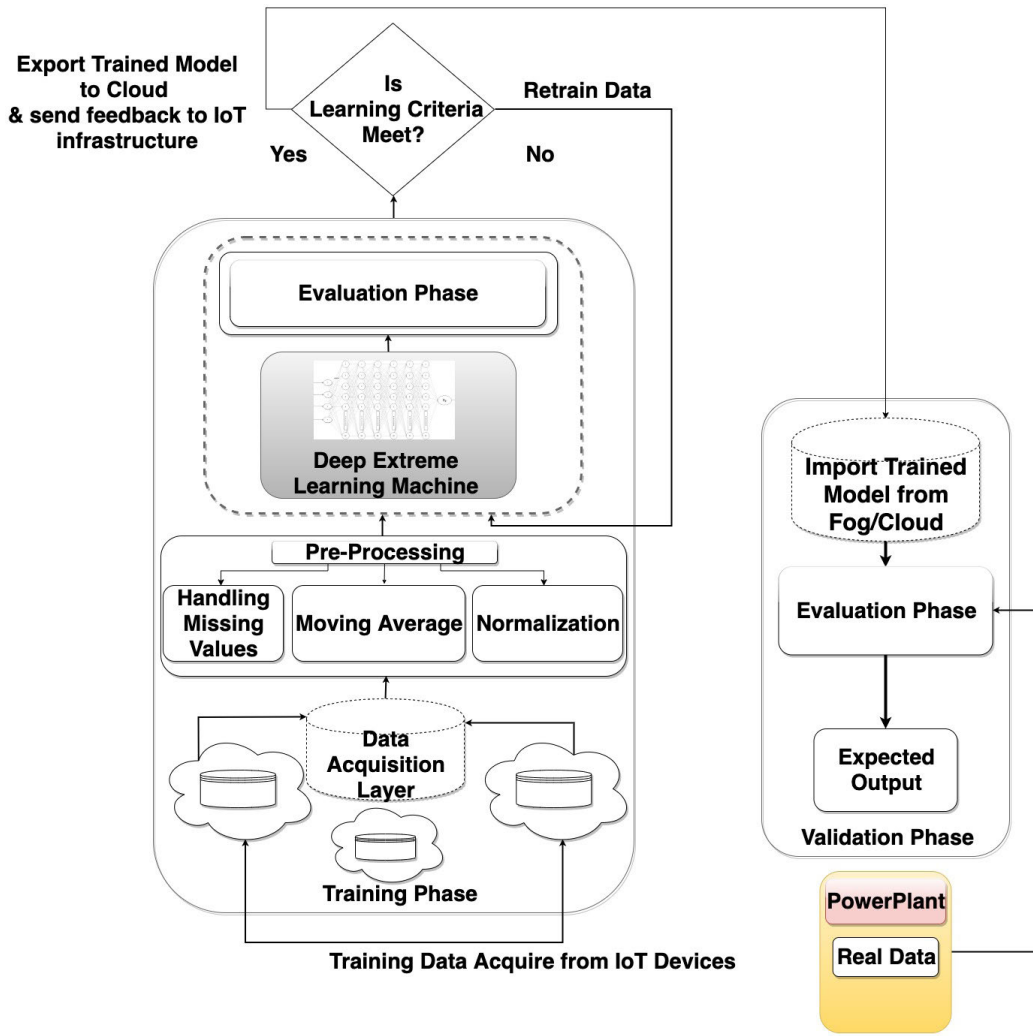$$\times \frac{\partial NhR_n^{hl}}{\partial R_{i,n}^{hl}}$$

FIGURE 2. Proposed diagram for the prediction of power plant electrical energy output using the DELM approach.

$$\Delta \mathrm{R}_{i,n}^{hl} = H \left[\sum_j \left(rj - tp_j^{hl} \right) \times tp_n^{hl} \left(1 - tp_j^{hl} \right) \times \left(R_{n,j} \right) \right]$$
$$\times tp_j^{hl} \left(1 - tp_j^{hl} \right) \times X_{i,n}$$

$$\Delta \mathrm{R}_{i,n}^{hl} = H \left[\sum_j \left(rj - tp_j^{hl} \right) \times tp_n^{hl} \left(1 - tp_j^{hl} \right) \times \left(R_{n,j} \right) \right]$$
$$\times tp_j^{hl} \left(1 - tp_j^{hl} \right) \times X_{i,n}$$

$$\Delta \mathrm{R}_{i,n}^{hl} = H \left[\sum_j \left(rj - tp_j^{hl} \right) \times tp_n^{hl} \left(1 - tp_j^{hl} \right) \times \left(R_{n,j} \right) \right]$$
$$\times tp_n^{hl} \left(1 - tp_n^{hl} \right) \times X_{i,n}$$

$$\Delta \mathrm{R}_{i,n}^{hl} = H \left[\sum_j \ni_j (R_{n,j}^{hl}) \right]$$
$$\times tp_n^{hl} \left(1 - tp_j^{hl} \right) \times X_{i,n}$$

$$\Delta \mathrm{R}_{i,n}^{hl} = H \ni_n X_{i,n}$$

where

$$\ni_n = \left[\sum_j \ni_j (R_{n,j}^{hl}) \right] \times tp_n^{hl} \left(1 - tp_n^{hl} \right)$$

The process of upgrading the weight and bias among the output and the hidden layer is shown in Eq. (23).

$$R_{i,j}^{hl=6} (t) = R_{i,j}^{hl=6} (t) + \lambda \Delta R_{i,j}^{hl=6} \qquad (23)$$

Eq. (24) shows how updating the weight and bias among the input and the hidden layer.

$$R_{i,n}^{hl} (t) = R_{i,n}^{hl} (t+1) + \lambda \Delta R_{i,j}^{hl} \qquad (24)$$

IV. RESULTS AND DISCUSSION

In the proposed article, the deep extreme learning machine algorithm has been applied to data [53] and in this regard,
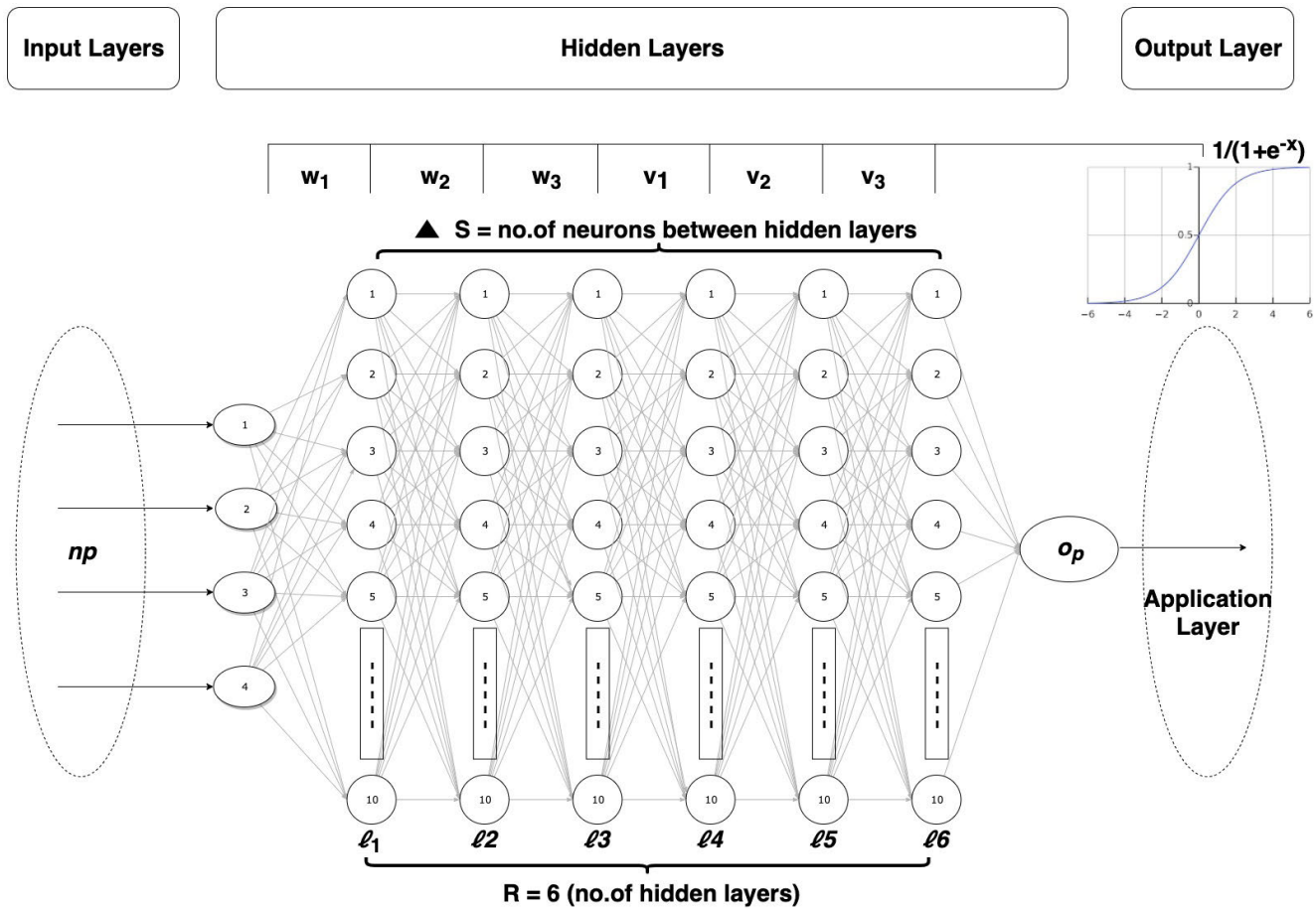
FIGURE 3. Structural diagram of a deep extreme learning machine.

the Matlab tool has been performed for simulation. In Matlab, a python script was implemented to train data [53]. In this article, DELM was used to train and fit 47840 sets of data. This data arbitrarily divides into 70% of training (33488 samples), 30% of data is used for validation and testing (14352 samples). Data were previously processed to remove data abnormalities and free the data from error. DELM has attempted to discover the finest configuration model for power plant electrical energy output prediction in different hidden layers, hidden neurons and combinations of Activation Functions. Therefore, we have tried the same number of neurons, different types of active functions in hidden layers. In this work, we used the proposed DELM for prediction to properly test the effectiveness of this algorithm. In order to measure the performance of this DELM algorithm together with the counterpart algorithms, we used different statistical measures written in Eqs. (25,26).

The exhaustive search is applied in this study to the original dataset consisting of four parameters as input variables and a target parameter as a response. The goal is to choose a minimal model with the best subset that predicts the response correctly [54]. To this end, we applied an exhaustive search to the original dataset after collecting preliminary statistical data to find the best subset by comparing all the competing

candidate subsets ($2^4 - 1 = 15$) in the experiments. In addition, we divided the experiments into four groups, applying the sub-sets with one, two, three, and four parameters to the regression methods. We calculated the best subset of each test in these experiments by analyzing and comparing the results of all regression methods for the candidate subsets, which are shown in TABLE 2. In Eq. (25), O represent the predictive output of megawatt-hour and T represent the actual output. O_0 and T_0 represent that there is no change in predictive and powerplant electricity production output respectively from the previous cycle. O_1 and T_1 represent the change in electricity production from the previous cycle of the predictive and actual demand of a powerplant respectively. O_0/T_0 and O_1/T_1 represents predictive and actual output are same. Similarly, O_1/T_0 and O_1/T_0 represents error, which predictive and actual demand of electicity are varied.

$$\text{Miss rate} = O_1/T_0 + O_0/T_1 \qquad (25)$$

$$\text{Accuracy} = \frac{\left(O_0/T_0 + O_1/T_1\right)}{T_0 + T_1} \qquad (26)$$

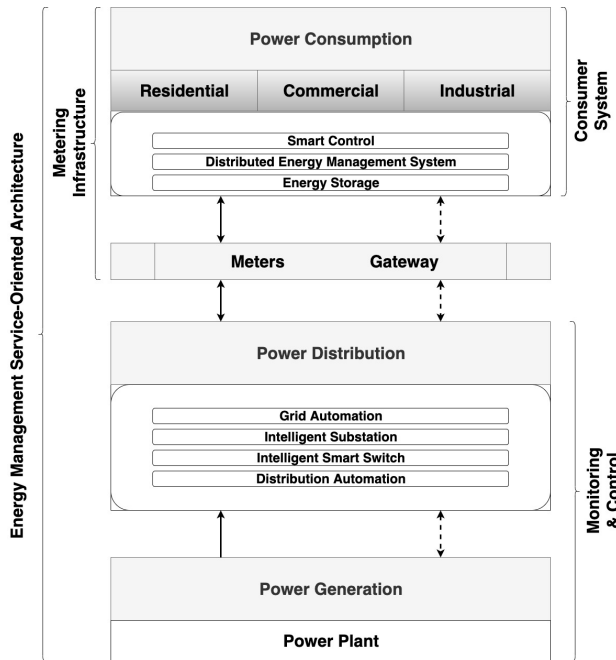$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}\lceil y_j - y\rceil \qquad (27)$$

FIGURE 4. Power plant service-oriented architecture.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_j - y)^2}{n}} \qquad (28)$$

$$MAPE = \frac{100\%}{n} \sum \frac{|y - y|}{y} \qquad (29)$$

$$MSE = \sum_{i=1}^{n} (y_j - y)^2 \qquad (30)$$

For the occupancy data set [53], the DELM approach was used and the results obtained can be seen in TABLE 3. After comparing with expected output and result that got after applying the proposed approach, Table 3 shown that the result of our proposed approach gives 98.6% accuracy and 1.6% miss rate during training.

We take 30% of data (14352 samples) for testing and validation from the dataset [53]. After comparing with the expected output and result that got after applying the proposed approach it can be shown in Table 3. Table 3 also shows that the proposed approach accuracy during testing and validation is 93.9% and the miss rate is 6.1%.

It is observed that the overall performance of the proposed method (DELM) during training was 98.6% accurate as shown in

TABLE 4, while the miss rate of training is 1.6%. But in testing and validation, the overall performance of the proposed method (DELM) was 93.9% accurate, while the miss rate of training is 6.1% as shown in Table 3. It also observed from TABLE 3 in the training phase results accuracy increases with the minimum miss rate as compared to the testing and validation phase.

TABLE 2. All possible subsets of the dataset.

Subsets	AT	V	AP	RH
With 1 parameter				
AT	1	0	0	0
V	0	1	0	0
AP	0	0	1	0
RH	0	0	0	1
With 2 parameters				
AT–V	1	1	0	0
AT–AP	1	0	1	0
AT–RH	1	0	0	1
V–AP	0	1	1	0
V–RH	0	1	0	1
AP–RH	0	0	1	1
With 3 and 4 parameters				
AT–V–AP	1	1	1	0
AT–V–RH	1	1	0	1
AT–AP–RH	1	0	1	1
V–AP–RH	0	1	1	1
AT–V–AP–RH	1	1	1	1

According to Table 4 [55] utilized combined cycle power plant data set [53] consisting of 47840 data samples, the mean square error in each round is increasing. In the [55] approach, there is only one hidden layer and an increasing number of neurons.

TABLE 5 shows that during training the mean square error of [55] methodology is decreasing respectively but It is observed that MSE of proposed DELM approach with the same number of neurons is less than [55]. Moreover, with the increase in the number of hidden layers, the error is further mitigated. However, the number of neurons is taken 10 in all the cases, whereas in [55], the number of neurons is significantly high with a single layer. It can lead to the conclusion that having more layers with a smaller number of neurons is a better idea compared to having more neurons in a single layer.

For subsequent tests, in order to compare the performance of different learning methods, 5 × 2 cross-validation [56] was applied. In this scheme, the dataset is randomly shuffled 5 times and each of them used in the 2-fold CV. The resulting validation set performances of size 10 are used for statistical significance tests.

Further, it is observed that in [55], with a single hidden layer when the neuron count increases the performance of the system also increases as shown in

TABLE 3. Mean square error (MSE) and accuracy comparison of all subsets during training and testing.

Subset	Training (Proposed)		Testing (Proposed)	
	Mean Square Error (MSE)	Accuracy	Mean Square Error (MSE)	Accuracy
With 1 parameter				
AT	7.54775	0.986508	7.45470	0.986750
V	7.38847	0.984047	7.40622	0.983997
AP	7.31204	0.971651	7.50462	0.973106
RH	7.57501	0.947354	7.93714	0.946686
With 2 parameters				
AT–V	6.70690	0.986732	6.50564	0.988277
AT–AP	7.07385	0.986050	7.05461	0.988037
AT–RH	6.97621	0.986620	6.57197	0.988354
V–AP	7.48162	0.980422	7.41285	0.984317
V–RH	7.49845	0.982414	7.44146	0.985534
AP–RH	7.54875	0.975087	7.84028	0.979201
With 3 and 4 parameters				
AT–V–AP	6.42667	0.985389	6.50157	0.989091
AT–V–RH	6.30749	0.985395	6.07163	0.989327
AT–AP–RH	6.74485	0.986460	6.87200	0.988155
V–AP–RH	7.73745	0.978328	7.50817	0.984677
AT–V–AP–RH	5.86300	0.986384	6.46653	0.939105

TABLE 5. While, in the proposed DELM solution as the number of hidden layers increases with the same (10) number of neurons, the performance increases significantly. This means that the performance of the system is increased by the increased number of neurons but not much as in the proposed DELM system.

TABLE 5 enlists MSE performances of local models (LM). Here, LM1 is k-NN + ANN model with k=100, LM2 is K-Means + ANN with K=20, LM3 is a K-Means + ANN ensemble of population 3 with K=20, LM4 is a K-Means + ANN ensemble of population 5 with K=20 and LM5 is a K-Means + ANN ensemble of population 3 with K=10. As can be seen local models with clustering yield better results with an ensemble.

TABLE 5, the prediction of power plant electrical energy output is carried out by all conventional findings with data from the combined cycle power plant data set. The proposed approach with DELM outperforms in terms of accuracy by other prototypes, such as backpropagation [55]. The overall results for [55] were 89.56%, but the proposed DELM system performance is 98.6% and was higher than the previously proposed methods in terms of accuracy rate. The values of the statistical measures suggest that DELM performance is much

TABLE 4. Training and testing accuracy of the proposed DELM system with varying hidden layers during the prediction of power plant electrical energy output.

	Training	Testing
Mean Square Error (MSE)	5.86300	6.46653
Accuracy (%)	98.6	93.9
Miss Rate (%)	1.4	6.1

higher than the other approaches. So, the proposed DELM is a considerable choice for the power plant electrical energy output prediction.

TABLE 5. Mean square error (MSE) comparison of the proposed DELM system with [55].

|	LM1 [55]	LM2 [55]	LM3 [55]	LM5 [55]	LM6 [55]	DELM (Proposed)
F11	20.95	20.08	17.86	15.33	15.04	6.50564
F21	19.77	20.51	15.71	14.98	14.98	7.05461
F31	20.03	18.43	15.94	15.35	14.94	6.57197
F41	20.08	19.09	16.59	14.97	14.46	7.41285
F51	19.36	19.04	16.10	15.38	15.11	7.44146
F12	19.46	27.50	15.65	15.23	15.21	6.50157
F22	19.72	18.09	16.27	15.59	16.04	6.07163
F32	19.77	19.34	16.04	15.50	15.27	6.87200
F42	20.57	18.68	16.88	16.09	18.32	7.50817
F52	20.21	20.94	16.11	15.90	15.50	6.46653
Mean	19.99	20.17	16.32	15.43	15.49	6.840643
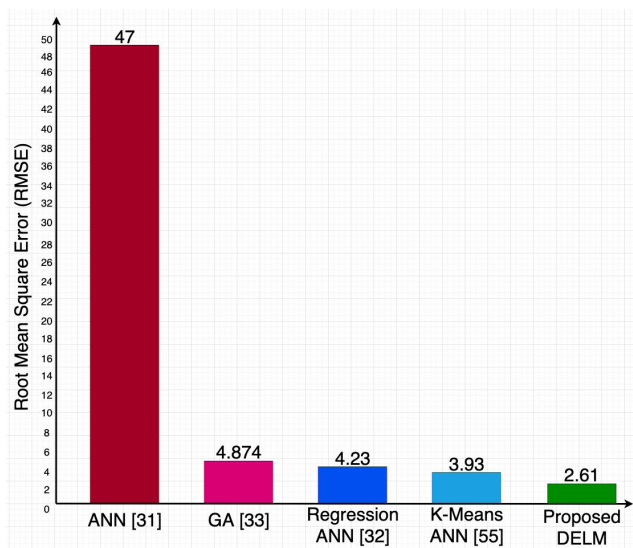
FIGURE 5. Root mean square error (RMSE) comparison of the proposed DELM system with state-of-the-art methods.

In Figure 5, the prediction of powerplant electrical energy output is carried out by all conventional findings with data from combined cycle powerplant [53] other than the Artificial Neural Network [31]. The proposed approach with DELM outperforms in terms of accuracy by other prototypes, Artificial Neural Network [31], GA base Multilayer Perceptron [33], Regression ANN Model [32] and K-Means + ANN [55]. Among these approaches, the worst approach is an Artificial Neural Network [31] with an RMSE of 47. Moreover, K-Means + ANN [55] is the best approach during the training phase compared to Regression ANN Model [32]. In addition, the accuracy of the GA base Multilayer Perceptron [33] and Regression ANN Model [32] is quite close. The proposed DELM system RMSE is 2.61 and was lower than the previously proposed methods in terms of accuracy. The values of the statistical measures suggest that DELM performance is much higher than the other approaches. So, the proposed DELM is a considerable choice for powerplant electrical energy output prediction.

V. CONCLUSION

Modeling, analysis, and prediction of power plant electrical energy output is a challenging task. In this research, a model for power plant electrical energy output prediction has been proposed to improve the prediction accuracy. The proposed model is an expert system based on an artificial neural

system (ANN) with a deep extreme learning machine (DELM) possessing a high level of potential to predict power plant electrical energy output. Various numbers of the hidden layer neurons were defined, and diverse activation functions and features were used for the ideal arrangement of different DELM parameters to obtain an optimized structure.

For measuring the performance of the proposed approach, various statistical measures have been used. These measuring figures show that proposed DELM in contrast to other algorithms is way better in terms of accuracy. Compared to past approaches, the proposed DELM technique produces attractive results. The proposed technique exhibits 98.6% accuracy which much better than the existing techniques. Moreover, it is observed that the proposed approach exhibits an affordable computational complexity. DELM has been used in the proposed work to encapsulate the benefits of ELM as well as deep learning. We are confident in initial results and intended to expand this work in the future by investigating different datasets, learning machines, structures, and algorithms.

FUTURE WORK

In this segment, we explored how the technique of a deep extreme learning machine used to predict electric power plant energy output. Compare this issue to the systematic experiments of simulated datasets, in particular, because the neural network is only supplied with partial information: the neural network is blind to certain topological changes, for example, changes irrespective of the presence or absence of power lines. The deep extreme learning machine performs better in these more complicated cases than the neural network baseline. This allows approximation of flows without complete information about the topology of the power plant compared to most models in use today. This model performs fairly well on data similar to those on which it has been trained. This is an important result: deep extreme learning machine can be used for grid operations. Still, the flows on completely unseen topologies have been struggling. This more refers to an evaluation metric issue. Transmission system operators do not really want to estimate flows; they want to know whether or not the power grid will be unsafe.

Further research could be done on this dimension. The aim of future studies will be to identify and quantify this metric more accurately. The neural network would be more often retrained to improve the performance of the new topologies. An algorithm that can be learned within a few hours and that has the property of being "fine-tuned" has been tested over a span of more than a month. Another way of improving the results would be by learning not only from snapshots but also from the power plant that operators are observing.

REFERENCES

[1] A. Atta, S. Abbas, M. A. Khan, G. Ahmed, and U. Farooq, "An adaptive approach: Smart traffic congestion control system," *J. King Saud Univ.-Comput. Inf. Sci.*, pp. 1–8, Oct. 2018.

[2] H. Chourabi, T. Nam, S. Walker, J. R. Gil-Garcia, S. Mellouli, K. Nahon, T. A. Pardo, and H. J. Scholl, "Understanding smart cities: An integrative framework," in *Proc. 45th Hawaii Int. Conf. Syst. Sci.*, Jan. 2012, pp. 2289–2297.

[3] A. Fatima, M. A. Khan, M. Abbas, S. Waqas, M. Anum, L, and M. Asif, "Evaluation of planet factors of smart city through multi-layer fuzzy logic (MFL)," *ISC Int. J. Inf. Secur.*, vol. 11, no. 3, pp. 51–58, 2019.

[4] M. Naphade, G. Banavar, C. Harrison, J. Paraszczak, and R. Morris, "Smarter cities and their innovation challenges," *Computer*, vol. 44, no. 6, pp. 32–39, Jun. 2011.

[5] K. Iqbal, M. Adnan, S. Abbas, Z. Hasan, and A. Fatima, "Intelligent transportation system (ITS) for smart-cities using mamdani fuzzy inference system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 2, pp. 94–105, 2018.

[6] M. AsadUllah, M. A. Khan, S. Abbas, A. Athar, S. S. Raza, and G. Ahmad, "Blind channel and data estimation using fuzzy logic-empowered opposite learning-based mutant particle swarm optimization," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–12, Dec. 2018.

[7] S. Karnouskos and T. N. D. Holanda, "Simulation of a smart grid city with software agents," in *Proc. 3rd UKSim Eur. Symp. Comput. Modeling Simulation*, 2009, pp. 424–429.

[8] A. Schaffers, H. Komninos, N. Pallot, M. Trousse, B. Nilsson, and M. Oliveira, "Smart cities and the future Internet: Towards cooperation frameworks for open innovation," in *The Future Internet Assembly*. Berlin, Germany: Springer, 2011, pp. 431–446.

[9] A. Q. Huang, M. L. Crow, G. T. Heydt, J. P. Zheng, and S. J. Dale, "The future renewable electric energy delivery and management (FREEDM) system: The energy Internet," *Proc. IEEE*, vol. 99, no. 1, pp. 133–148, Jan. 2011.

[10] H. Kanchev, D. Lu, F. Colas, V. Lazarov, and B. Francois, "Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4583–4592, Oct. 2011.

[11] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

[12] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.

[13] M. A. Khan, M. Umair, and M. A. S. Choudhry, "GA based adaptive receiver for MC-CDMA system," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 23, pp. 2267–2277, 2015.

[14] S. Abbas, M. S. Khan, G. Ahmed, M. S. Abdullah, M. S., and U. Farooq, "Bio-inspired neuro-fuzzy based dynamic route selection to avoid traffic congestion," *Int. J. Sci. Eng. Res.*, vol. 2, no. 6, pp. 284–289, 2011.

[15] N. S. Naz, S. Abbas, M. Adnan, B. Abid, N. Tariq, and M. Farrukh, "Efficient load balancing in cloud computing using multi-layered mamdani fuzzy inference expert system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 3, 2019.

[16] M. Dohler, I. Vilajosana, X. Vilajosana, and J. LLosa, "Smart cities: An action plan," in *Proc. Barcelona Smart Cities Congr.*, Barcelona, Spain, 2011, pp. 1–6.

[17] A. R. Khare and P. Shrivasta, "Data mining for the Internet of Things," in *Proc. IGI Global Exploring Converg. Big Data Internet Things*, 2018, pp. 181–191.

[18] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, "Smart cities of the future," *Eur. Phys. J. Special Topics*, vol. 214, no. 1, pp. 481–518, 2012.

[19] A. Zanella, N. Bui, A. P. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[20] R. E. Hall, B. Bowerman, J. Braverman, J. Taylor, H. Todosow, and U. V. Wimmersperg, "The vision of a smart city," in *Proc. 2nd Int. Life Extension Technol. Workshop*. Paris, France: Brookhaven National Lab., Sep. 2000, pp. 1–7.

[21] T. Bakici, E. Almirall, and J. Wareham, "A smart city initiative: The case of barcelona," *J. Knowl. Economy*, vol. 4, no. 2, pp. 135–148, Jan. 2012.

[22] R. Giffinger, H. Gudrun, and G. Haindlmaier, "Smart cities ranking: An effective instrument for the positioning of the cities," *Archit. City Environ.*, vol. 4, no. 12, pp. 7–26, Feb. 2010.

[23] S. Yi, C. Y. Li, and Q. Li, "A survey of fog computing: Concepts, applications and issues," in *Proc. Workshop Mobile Big Data*, 2015, pp. 37–42.

[24] K. Habak, C. Shi, E. W. Zegura, K. A. Harras, and M. H. Ammar, "Elastic mobile device clouds: Leveraging mobile devices to provide cloud computing services at the edge," *Fog 5G IoT*, pp. 159–163, Mar. 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119187202.ch7

[25] V. Albino, U. Berardi, and R. M. Dangelico, "Smart cities: Definitions, dimensions, performance, and initiatives," *J. Urban Technol.*, vol. 22, no. 1, pp. 3–21, Feb. 2015.

[26] P. Kaltenrieder, E. Portmann, and T. Myrach, "Fuzzy knowledge representation in cognitive cities," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Aug. 2015, pp. 1–8.

[27] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.

[28] M. Moore, P. Gould, and B. S. Keary, "Global urbanization and impact on health," *Int. J. Hygiene Environ. Health*, vol. 206, nos. 4–5, pp. 269–278, Jan. 2003.

[29] P. Hall, "Creative cities and economic development," *Urban Stud.*, vol. 37, no. 4, pp. 639–649, 2000.

[30] R. P. Dameri, "Defining an evaluation framework for digital cities implementation," in *Proc. IEEE Int. Conf. Inf. Soc.*, Jun. 2012, pp. 466–470.

[31] B. Donnot, "Deep learning methods for predicting flows in power grids: Novel architectures and algorithms," Univ. Paris-Saclay, Saint-Aubin, France, 2019.

[32] E. A. Elfaki and A. H. Ahmed, "Prediction of electrical output power of combined cycle power plant using regression ANN model," *J. Power Energy Eng.*, vol. 06, no. 12, pp. 17–38, 2018.

[33] I. Lorencin, N. Andelic, V. Mrzljak, and Z. Car, "Genetic algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation," *Energies*, vol. 12, no. 22, p. 4352, Nov. 2019.

[34] E. Brown. (2016). *Who Needs the Internet of Things?* [Online]. Available: https://www.linux.com/news/who-needs-internet-things

[35] M. Fast, M. Assadi, and S. De, "Development and multi-utility of an ANN model for an industrial gas turbine," *Appl. Energy*, vol. 86, no. 1, pp. 9–17, Jan. 2009.

[36] M. Rahnama, H. Ghorbani, and A. Montazeri, "Nonlinear identification of a gas turbine system in transient operation mode using neural network," in *Proc. 4th Conf. Thermal Power Plants (CTPP)*, 2012, pp. 1–6.

[37] G. K. F. Tso and K. K. W. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761–1768, Sep. 2007.

[38] A. Azadeh, M. Saberi, and O. Seraj, "An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: A case study of Iran," *Energy*, vol. 35, no. 6, pp. 2351–2366, 2010.

[39] U. Kesgin and H. Heperkan, "Simulation of thermodynamic systems using soft computing techniques," *Int. J. Energy Res.*, vol. 29, no. 7, pp. 581–611, 2005.

[40] R. Kitchin, "The real-time city? Big data and smart urbanism," *GeoJournal*, vol. 79, no. 1, pp. 1–14, Nov. 2013.

[41] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: Approaches and challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.

[42] L. Foschini, T. Taleb, A. Corradi, and D. Bottazzi, "M2M-based metropolitan platform for IMS-enabled road traffic management in IoT," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 50–57, Nov. 2011.

[43] P. Vlacheas, R. Giaffreda, V. Stavroulaki, D. Kelaidonis, V. Foteinos, G. Poulios, P. Demestichas, A. Somov, A. Biswas, and K. Moessner, "Enabling smart cities through a cognitive management framework for the Internet of Things," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 102–111, Jun. 2013.

[44] Q. Wu, G. Ding, Y. Xu, S. Feng, Z. Du, J. Wang, and K. Long, "Cognitive Internet of Things: A new paradigm beyond connection," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 129–143, Apr. 2014.

[45] S. Feng, P. Setoodeh, and S. Haykin, "Smart home: Cognitive interactive people-centric Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 34–39, Feb. 2017.

[46] B. Tang, Z. Chen, G. Hefferman, S. Pei, T. Wei, H. He, and Q. Yang, "Incorporating intelligence in fog computing for big data analysis in smart cities," *IEEE Trans Ind. Informat.*, vol. 13, no. 5, pp. 2140–2150, Oct. 2017.

[47] S. Han, J. Pool, J. Tran, and W. J. Dally, "Learning both weights and connections for efficient neural network," in *Proc. Neural Inf. Process. Syst.*, 2015, pp. 1135–1143.

[48] A. Al-Fuqaha, A. Khreishah, M. Guizani, A. Rayes, and M. Mohammadi, "Toward better horizontal integration among IoT services," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 72–79, Sep. 2015.

[49] M. Fayaz and D. Kim, "A prediction methodology of energy consumption based on deep extreme learning machine and comparative analysis in residential buildings," *Electronics*, vol. 7, no. 10, p. 222, Sep. 2018.

[50] J. Cheng, Z. Duan, and Y. Xiong, "QAPSO-BP algorithm and its application in vibration fault diagnosis for a hydroelectric generating unit," *J. Vib. Shock*, vol. 34, no. 23, pp. 177–181, 2015.

[51] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: A survey," *Int. J. Mach. Learn. Cybern.*, vol. 2, no. 2, pp. 107–122, May 2011.

[52] H. Wei, G. Liu, F. Yan, and J. Sun, "Multi-modal deep extreme learning machine for robotic grasping recognition," in *Proc. Adaptation, Learn. Optim. ELM*, vol. 2, 2016, pp. 223–233.

[53] (Mar. 2014). *UCI Machine Learning Repository*. [Online]. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/combined+cycle+power+plant

[54] B. C. Wallet, D. J. Marchette, J. L. Solka, and E. J. Wegman, "A genetic algorithm for best subset selection in linear regression," presented at the 28th Symp. Interface Sydney, Sydney, NSW, Australia, Jul. 12, 1996.

[55] H. Kaya, P. Tüfekci, and F. S. Gürgen, "Local and global learning methods for predicting power of a combined gas & steam turbine," in *Proc. Int. Conf. Emerg. Trends Comput. Electron. Eng.*, 2012, pp. 13–18.

[56] T. G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms," *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998.

SAGHEER ABBAS received the M.Phil. and Ph.D. degrees in computer science from the School of Computer Science, NCBA&E, Lahore, Pakistan. He is currently working as an Assistant Professor with the School of Computer Science, NCBA&E. He has been teaching graduate and undergraduate students in computer science and Engineering for the past eight years. He has published about 60 research articles in international journals as well as reputed international conferences. His research interests primarily include cloud computing, the IoT, intelligent agents, image processing, and cognitive machines with various publications in international journals and conference.

MUHAMMAD ADNAN KHAN received the Ph.D. degree from ISRA University, Pakistan. Prior to joining the NCBA&E, he has worked in various academic and industrial roles in Pakistan. He is currently working as an Assistant Professor with the Department of Computer Science, Lahore Garrison University, Lahore, Pakistan. He has been teaching graduate and undergraduate students in computer science and Engineering for the past 10 years. He is currently guiding four Ph.D. scholars and eight M.Phil. Scholars. He has published about 120 research articles in international journals as well as well in reputed international conferences. His research interests primarily include MUD, channel estimation in multicarrier communication systems, as well as image processing and medical diagnosis using soft computing with various publications in journals and conferences of international repute.

LUIS EDUARDO FALCON-MORALES received the M.Sc. degree in mathematics from UNAM, in 1994, and the Ph.D. degree in computer vision from CINVESTAV, in 2007. During his Ph.D., he has spent several months with GRASP Laboratory, University of Pennsylvania, and also with the Computer Vision Laboratory of Berkeley, University of California. His current researches focus on computer vision, image processing, as well as machine and deep learning. He has been a full-time Professor with the Department of Computer Science School, Tecnologico de Monterrey, since 1998. He has worked and published papers in several topics, such as medical image processing and machine learning for detection of diabetes based on retina images; digital analysis of blood patterns for the detection of several diseases; leader of a project of computer vision for the assistance of visually impaired people; as well as sentiment analysis of a social networking platform.

ABDUR REHMAN received the B.S. degree in computer sciences from PAKAIMS, Lahore, Pakistan. He is currently pursuing the M.Phil. degree with the School of Computer Science, NCBA&E, Lahore. His research interests primarily include cloud computing, the IoT, intelligent agents, and smart city with various publications in international journals and conferences.

YOUSAF SAEED received the Ph.D. degree in cognitive VANETs from NCBA&E, Lahore, Pakistan, and the M.S. degree in broadband and high-speed communication networks from the University of Westminster, London, U.K., where he achieved distinction in Research Thesis on IPv6. He is currently working as an Assistant Professor with the Department of Information Technology, University of Haripur, Pakistan. He received the International Students Award at the College of North West London, U.K. He acquired seven Research Projects from the National Grassroots ICT Research Initiative (NGIRI) Program from the Ministry of Information Technology and Telecom, Government of Pakistan. His achievements include eighteen publications, including three monographs, ten journal articles, and five conference papers. His patent is under review regarding emergency vehicles-based traffic lights control system.

MAHDI ZAREEI (Member, IEEE) received the M.Sc. degree in computer networks from the University of Science, Malaysia, in 2011, and the Ph.D. degree from the Communication Systems and Networks Research Group, Malaysia–Japan International Institute of Technology, University of Technology, Malaysia, in 2016. In 2017, he joined the School of Engineering and Sciences, Tecnologico de Monterrey, as a Postdoctoral Fellow. He is currently a Research Professor with the Tecnologico de Monterrey. His research interests include wireless sensor and ad hoc networks, energy harvesting, cognitive radio networks, and performance optimization. He is a member of the Mexican National Researchers System (level I). He is also serving as an Associate Editor for IEEE Access and the *Ad Hoc & Sensor Wireless Networks journals*.

ASIM ZEB received the B.Sc. and M.Sc. degrees in computer science from the University of Peshawar, Pakistan, in 2002 and 2005, respectively, and the Ph.D. degree in computer science from the University of Technology, Malaysia, in 2016. He has served as a Research Fellow at the Nagoya Institute of Technology, Japan, from 2014 to 2015. He has been serving as an Assistant Professor of computer science with the Department of Physical and Numerical Science, Qurtuba University of Science and Information Technology, Peshawar, Pakistan, since 2016. He is supervising students in the field of self-organized networks, networks architectures and protocols, wireless BANs, information centric networks, text mining, and cognitive radio networks. He has received the MJIIT-Malaysia Scholarship (2013–2014) and the JASSO-Japan Scholarship (2014–2015).

EHAB MAHMOUD MOHAMED (Member, IEEE) received the B.E. degree in electrical engineering from South Valley University, Egypt, in 2001, the M.E. degree in electrical engineering from South Valley University, Egypt, in 2006, and the Ph.D. degree in information science and electrical engineering from Kyushu University, Japan, in 2012. From 2013 to 2016, he joined Osaka University, Japan, as a Specially Appointed Researcher. Since 2017, he has been an Associate Professor with Aswan University, Egypt. He has also been an Associate Professor with Prince Sattam bin Abdulaziz University, Saudi Arabia, since 2019. His current research interests are 5G, B5G, and 6G networks, cognitive radio networks, millimeter wave transmissions, Li-Fi technology, MIMO systems, and underwater communication. He is the General Chair of the IEEE ITEMS'16 and the IEEE ISWC' 18. He is a Technical Committee Member in many international conferences and a reviewer in many international conferences, journals, and transactions.

• • •