# Load Forecasting using Machine Learning Methods: Review

Ebaa Jaafar Al Nainoon [1], Ibrahim Omar Habiballah [2]

Electrical Engineering Department

King Fahd University of Petroleum and Minerals

Dhahran, Saudi Arabia

*Abstract*—**Load forecasting is a vital field regarding the planning and studies of power systems. This Paper is a review for the topic of demand load prediction and forecasting using machine learning methods. it will summarize, discuss the developments, and review the recent research made so far. Three methods will be discussed in specific: neural network, recurrent networks with particle swarm optimization as well as support vector machine. Different methods of predicting load patterns and expectations are going to be presented and deliberated. Accuracy, perks, and constraints of using these methods of forecasting will also be presented. In addition, this paper will talk about using several machine learning methods side by side.**

*Keywords—Machine learning; Neural Networks; Fuzzy logic; SVM; Particle Swarm; Genetic Algorithms; Load forecasting; Power system; Planning; and Studies.*

## I. INTRODUCTION

Electricity is the spinal cord of our modern lives and commodities. It is also, extremely crucial to the national security, social and economic growth of any country. Therefore, the security and sustainability of electricity sources and the continuity of generation are vital topics, and for that, electrical load forecasting can be tremendously helpful.

Load forecasting (LF) is the expectation and prediction of the demand load that is calculated using a systematic procedure to adjust future expectation based on available parameters and information in order to determine future needs and system requirements. Precise load prediction or forecasting can serve energy and power producers to aid in ensuring the security and continuity of the power supply with no (or minimal) interruptions. As well as, in scheduling and reducing energy waste, given that electrical power is challenging to store.

Electrical load forecasting can be used in managing generation capacity, scheduling maintenance and outages, management of grid and transmission, peak reduction, and reserve management. In addition, it is important for market evaluation and assessment of the power trade capacities and capabilities of interconnections and planning reserve to surpass any deficiencies in power for any connected utilities [1], [2].

Electrical load forecasting could be grouped into 3 sections:
1. Long term (1-20 years)
2. Medium term (1 week – 1 year)
3. Short term (1 hour – 1 week)

Each of these categories can have different methods of compiling as well as different benefits and different constrains. Therefore, would require different machine learning methods in order to obtain the forecast. The different benefits for each category are organized in Table 1.

TABLE 1. TIME BASED APPLICATIONS OF LOAD FORECASTING [1]

| Time scope | Application |
|---|---|
| LTF (1-20 yrs) | Planning for the power system involves load forecasting, network planning, and generation planning. |
| MTF (1w-1y) | arranging repairs and maintenance |
| STF (1min-1w) | Analysis of unit commitment |
| | Economic load dispatch (ELD) and OPF analysis |
| | Automatic generation control |
| Milliseconds to seconds | Power system dynamic analysis |
| Nanoseconds to micro seconds | Power system transient analysis |

As summarized in Table 1, long term forecasting (LTF) is crucial to the economic planning of the upcoming future additions to the generation system as well as additional transmission planning. The medium-term forecasting (MTF) is helpful for other aspects such as, mostly, setting of tariffs, arrangement of maintenance and repairs, fuel supply scheduling and financial management. Meanwhile, the short term forecasting (STF) is utilized to supply the core data for scheduling the start-up and shutdown of the generators, preparing spinning reserves, and conducting a thorough analysis of transmission restrictions. The STF is also used to evaluate the security of systems and ELD. [1]. Forecasting is related to the time frame, the longer the period the less accurate the forecasting.

## II. METHODS OF LOAD FORECASTING

There are plenty of methods that are being currently used in order to estimate and predict the load demand. Each country, utility or bulk power producers use whatever method suits them or any method they are used to working with. Most of these methods are mostly likely outdated and have been used for a long time, in this section of the paper some classical methods of load forecasting are going to be mentioned in brief before approaching the more advanced methods using machine learning, to get a basic understanding of how these methods approach load forecasting.

### A. Classical Methods

Classical methods approach load forecasting in several ways and as mentioned in [3] – [6] could be divided in 2 categories which have been listed in Figure 1.
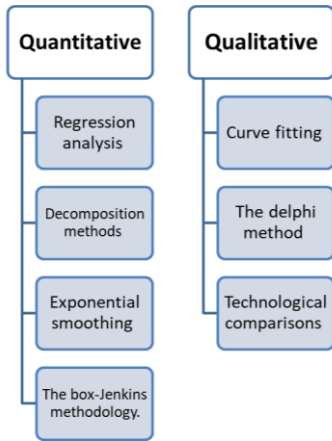
Figure 1. The base categories of classical forecasting mthods



Figure 2. simple representation of neural networks [8].

Exemplified in figure 1 above are the two basic groups for classical methods, choosing between these categories mainly depends on the available information and data. In qualitative methods, no historical data are used to make a prediction, instead the opinion of experts in a structurally developed approach is used to forecast loads, some of which are represented in figure 1.

The quantitative methods, to the contrary, are based on objective forecasting methods using mathematical and statistical approaches on the available historical data while assuming the continuity of some numerical past features (such as temperature). Quantitative methods are plentiful, and each has a variety of costs, errors and features, and choosing one over the other depends on the availability of data and desired use and output among other considerations [3].

Recently, a lot of new methods and procedures have been investigated and presented to tackle the shortcomings of some of these prior techniques of LF as many of the parameters these approaches require are hard to acquire.

### B. Machine Learning Methods

In modern times, classical approaches have become outdated due to their limitations and machine learning has emerged as a field and has evolved in application variety and become quite known in the power analysis and forecasting field for its many merits. Machine Learning is a subsidiary branch of the expanding field of artificial intelligence, and it centers around employing data and algorithmic principles to improve the accuracy of machines and models progressively through analysis and recognition of patterns, almost replicating the human learning process.

Machine learning methods are numerous, each of them serving a certain purpose or better fitting for certain circumstances, input features, and type of data. theses machine learning approaches include the following:

### 1) Artificial Neural Networks (ANN)

It is a model architecture that tries to imitate the information processing way of a human brain. This aims towards understanding, identifying, and categorizing data intelligently. Data is injected through neural network layers that can assign weights after processing data which is then sent to the proceeding layer of nodes in the network [7].
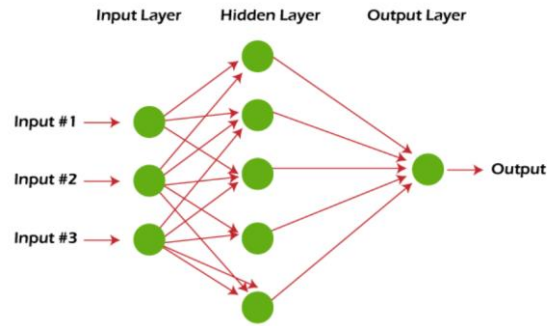
Neural networks can adapt to the changes in input and still produce the best possible output without needing to redesign the model or output characteristics which makes it a decently accurate model of machine learning [9].

This approach is extensively used in predicting load patterns as it is an extremely flexible one, and allows for accurate forecasting even with minimal parameters. Standard artificial neural networks (see Figure 2) include an input layer, a hidden layer, and an output layer at its simplest for. ANN are used for load forecasting in several research including [10] – [13] some of which date back to the initial emerging of machine learning as a field in the 1990s.

### 2) Support Vector Machine (SVM)

It is an ML approach that employs statistical learning theory (SLT), it is regularly used for categorization applications and regression assessment due to SLTs capabilities in recognizing patterns and analyzing data [14]. This technique aims at locating a discriminant function based on the independent training data set, that can suitably forecast labels for newly obtained occurrences. It requires fewer training data and less computational resources than generative approaches. Since SVM follows an analytical approach to optimization problems, it returns the same optimal parameters of solution at all times, contrary to other ML methods like genetic algorithms. SVM have a good generalization ability and is robust, which makes it one of the most broadly known supervised ML approaches yet one of the simplest. Moreover, SVMs possess a fast convergence speed as well as a solid nonlinear processing capability. However, they perform poorly for huge sums of data [15]. Presented in Figure 3 is a general structure for an SVM approach.
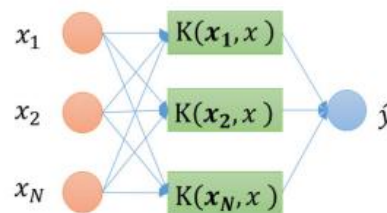


Figure 3. General structure for SVM model [15].

### 3) Fuzzy Logic

Fuzzy logic is a way of describing or viewing information and data that is not based on the usual classical logic which is typically binary (either 0 or 1). Fuzzy logic acknowledges the

grey area between the black and white. As in, objects can have "degrees" or "grades" of truths to them rather than the "yes" to "no" range of computer reasoning. This concept was introduced initially in the 1960s by Zadeh in his paper "Fuzzy sets," [16] for the purpose of better understanding and analyzing data that doesn't fall in a classical logic frame. Generally, most objects in the real world don't fall into a "yes" or "no" frame. Uncertainties and partial degrees of truths play a big rule in human reasoning which is what artificial intelligence and consequentially machine learning is aiming to replicate. This is where fuzzy sets have emerged as a concept which Zadeh suggested in [16] to deal with class criteria that aren't Sharpley defined.

These grades of relations of objects or data in a fuzzy set have a range that extends from 0 to 1 (in between values). This logic is defined by a membership function that designates the "grades" to each fuzzy set of objects. This means that any "object" or "data point" in a set has a value relating it to the membership function of the said "set" or "class", the value of this membership function at the calculated point denotes exactly the strength of this point at this set (ea. the closer the number is to 1, the stronger the grade of membership of this particular point is to the set).

In the field of electrical load forecasting, this can be very useful in representing the vagueness in some of the parameters contributing to the load pattern where the effect of each parameter can be defined as its own membership function, such application can be observed in [17], [18] where a combination of fuzzy sets logic and convolution neural networks is used to develop a forecasting model.

### 4) Genetic Algorithms (GA)

Genetic algorithm is a machine learning concept that focuses on mimicking the same principle as biological genetic evolution and natural selection. It is very effective for search optimization that includes huge and unorganized data. It is generally used to solve constrained and unconstrained complex issues. Genetic algorithms are a subset of the field of evolutionary computation, and they have proven to be very adaptable [19].

In GA, a variety of solutions are provided to a problem. These solutions would encounter processes of "mutation" and recombination to produce new offspring and the previous process is then repeated for several "generations" of solutions. This concept is similar to genetic reproduction in humans, and likewise, the new solutions would "mate" and "combine" based on the fittest match to produce the optimum "offspring". This procedure aims to guarantee the creation of fitter or better solutions in subsequent generations and it continues until it reaches the desired outcome criterion [19]. Some of these pre-mentioned operations, which get performed on the solutions to create new generations, are:

(1) selection

(2) Crossover

(3) Mutation

Therefore, it can be considered as the analog version of chromosomic biology. Moreover, due to the flexibility of this method it has been used also to generate load forecasts, in [20] a function for the long-term load forecasting problem utilizing genetic algorithm has been designed and optimized.

Plentiful other methods are available in literature, however, these are the most popular. Some methods merge classical and modern methods to come up with hybrid approaches to load forecasting. In addition, a combination of different machine learning methods could be used together in one forecasting model like [21] that suggests an ANN model associated with genetic algorithm for optimization of weights and biases. All these different methods have their own perks and flaws and most of them can be used to predict loads at varying efficiencies and computation time. Some of these applications will be discussed in details in the following sections of this paper. Such as a deep neural networks approach, recurrent neural network with Elman's and particle swarm optimization along with an SVM application employing a kernel function and fuzzy logic.

## III. LOAD FORECASTING WITH DEEP NEURAL NETWORKS

Some research applications have used deep neural networks in forecasting applications. Similarly, [2] proposes a deep neural network composition to provide hourly day ahead predictions for electric loads. It uses different types of NN components, models and layers to model different factors that may affect load consumption and behavior.

### A. Factors affecting Load forecasting.

Each factor influences the load forecasting in different proportions and each model may include and consider different factors, some of which are mentioned below.

### 1) Weather

Weather is one of the biggest factors affecting load variations during the day or the year. As significantly low temperatures can cause an increase in the usage of heaters while huge rises can cause a growth of air-conditioning usage. Furthermore, extreme weather conditions increase time spent indoors. This explains how weather plays an important role in shaping electricity consumption patterns. Alternative weather parameters like wind speed and humidity also might prove beneficial for load predictions, however we lack the necessary information about such factors within the available data set, and their effect is less prominent than that of temperature, except in the case of extreme conditions.
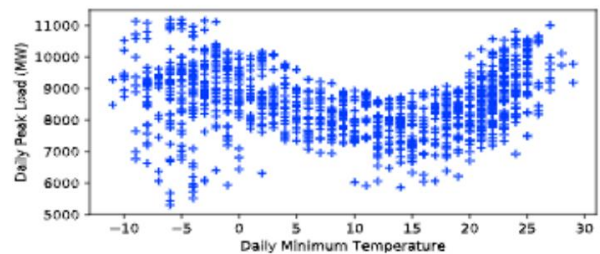


Figure 4. The effect of temperature in daily load patterns

Additionally, the time itself and the week day of the target time are clearly useful indicators for predicting hourly load. The variation that time of day and temperature can cause are displayed in Figure 4.

*2) Holidays*

On long holidays and public/national vacations, people's consumption patterns vary as it is increased due to home gatherings or decreased due to travel. It is variable depending on the location as well.

*3) Electricity price and energy policies*

Some countries have energy regulations that can cause the load to follow a certain pattern. For example, in some countries the price of electricity can be higher during specific times of day to encourage people to conserve energy at usually high/peak load hours.

The above-mentioned patterns are some of the features that were utilized as inputs to the NN specifically as inputs to the deep feed forward component of the model. This model in [2] aims to predict hourly loads for a specific day and hour based on weather, time of day and holiday for this specific location and time using the historical information going back 24 hours.
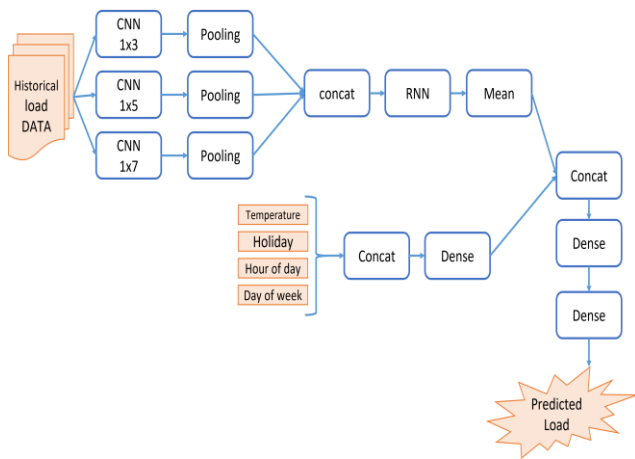


Figure 5. The layers and model for the proposed CNN method [2]

Reference [2] proposes the use of multiple Convulsion Neural Networks (CNN) layers to adapt and interpret historical data for load pattern. According to [22] CNNs have proven their efficiency in performing tasks such as extracting and learning features from input data. Therefore, in figure 5 one can see that 3 parallel CNN components with different filter sizes have been employed to transform the historical load pattern into a set of variable features. These features and input components that are time-series in nature are then processed through recurrent neural networks (RNNs) due to their efficiency in modeling sequential dynamic data. RNNs are paired with a long-short term memory unit (LSTM) to avoid gradient issues that occur with simple RNNs, the neural component of LTSM is incorporated to model variability of the historical load data. In addition, other factors such as temperature and national vacation days are also included using feed-forward components to make a vector representation. All these features are then concatenated and feed as inputs to the DNN portion of this model and the model will compute and learn from these raw data by itself.

The data set is divided into 3 sets: training, validation and testing sets. The evaluation is conducted on 3 years' worth of hourly load data and the method has been proven effective from experimental results.

The proposed method is compared with support vector regression (SVM), linear regression and a 3rd method (similar to the proposed but with differences in layers). The proposed method is seen to outperforms the other methods in load forecasting.

## IV.  RECURRENT NEURAL NETWORKS FOR LOAD FORECASTING

Reference [23] suggests a recurrent neural network model that can be utilized for load forecasting. Recurrent neural network (RNN) is a prospect of neural networks where components are connected in a direct cycle. RNNs can sort out arbitrary patterns of inputs by utilizing their internal memory. Simple RNNs use the concept of back propagation (BP) for model training. It is a type of neural networks model that falls under supervised learning as the network continually adjusts weights and biases until it matches the historical data which is then assigned set numerical values that matches the input features with the output predictions using training data.

RNN is made up of a minimum of a single feedback circle. It could comprise of a single level of neurons where the output of every neuron back-feeds the input of all the other neurons. The outcome of the neurons in output layer and hidden layers are calculated through taking the weighted total amount for inputs through the tangent sigmoid activation function and pure linear transfer functions respectively. Arithmetically, the weights are revised using a learning rate and the partial derivation of the variable, indicating the direction of change with respect to a local or global minimum, with respect to the output error which are updated every loop to adjust the best fitted solution of weights to the current problem.

Elman's recurrent neural network was chosen in [23] to be modeled to process the load forecasting data as it has proven to be an efficient architecture of RNNs. As shown in Figure 6, ERNN consists of recurrent links in the hidden layer of neurons that connects to a delay unit which stores the output of these neurons for a single time period before it re-feeds them to the input of the hidden layer.
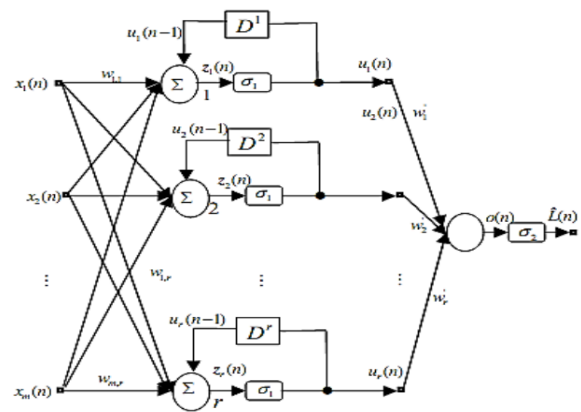


Figure 6. The topology of an Elman's RNN [15]

**IJERTV11IS120094**
      **www.ijert.org**
      **151**

## A. Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is a means which offers a population-based search process. In a PSO scheme particles (input data) travel in a search space that has several dimensions. During the movement of these particles, each one of them undergoes their own tunings or modifications (position modifications) based on their own experience in the multidimensional space or the experience of a nearby particle. This way, the particles adjust themselves by using their own optimum position and using the help of a neighbor for better fitted results. Therefore, it can be said that PSO takes advantage and merges local as well as global search approaches, to create a balance of exploration and exploitation.
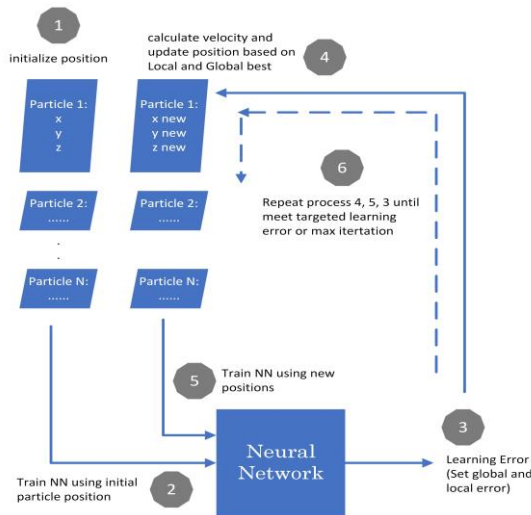


Figure 7. The training process for this PSO-ERNN network [15]

In the application discussed in [23], PSO was used as a tool to adjust the best fitted weights and positions in the Elman's RNN and then the optimum solution to this learning process is used to predict and forecast loads. Figure 7 summarizes the learning process for this network.

It was deduced in [23] that particle swarm optimized Elman recurrent neural network (PSO-ERNN) is an optimum prototype for 168 hours prior load prediction. In terms of forecasting future loads, this kind of network can be quite efficient. Though, the simulations looked to be extremely prosperous, data from other sources still need to be utilized to test the developed models in order to verify the consistency of these models for other data samples. This model can be improved by considering other parameters that affect the daily load forecast.

## V. SUPPORT VECTOR MACHINE BASED LOAD FORECASTING

Support vector machine is a ML learning load forecasting approach that employs statistical learning theory (SLT) [14]. It works in tandem with the kernel function to create a linear map from the input data space to the high-dimensional feature region. With this, problems like nonlinearity, small sample sizes, large dimensions, and local optimization can be resolved as SVM provides a more global solution. However, a downside of traditional SVM algorithms is that they are unable to filter useless, redundant data and cannot simplify the dimensionality. In [24], an SVM based load prediction model is proposed to forecast loads employing data mining technology to overcome the shortcomings of classical SVM models as it conquers the issues in processing large quantities of data and eliminates the information redundancy. This proposal takes into account a variety of factors and parameters such as average load rang, weather circumstances, and specific date characteristics. This sequence of new data will act as the training portion of the SVM model. This simplifies the dimensions of the intake and the computations, as well as makes the historical data more regular, thus more applicably convenient.

The data mining algorithm that was integrated with this model selects similar day samples accounting for all external and internal factors (previously mentioned, which classical SVM applications fail to consider). This algorithm merges the use of similarity mining, fuzzy grouping and grey relational theory. The proposed method in succeeds in simplifying the input data and calculations. Additionally, strengthens the guidelines for the training samples. A good prediction results from this proposed model, by selecting similar-days from the initial data, samples that significantly deviate from the load traits of the forecasted day in the original set of data are removed. As a result, the model's training process is obviously faster, and the predictions produced are of higher accuracy.
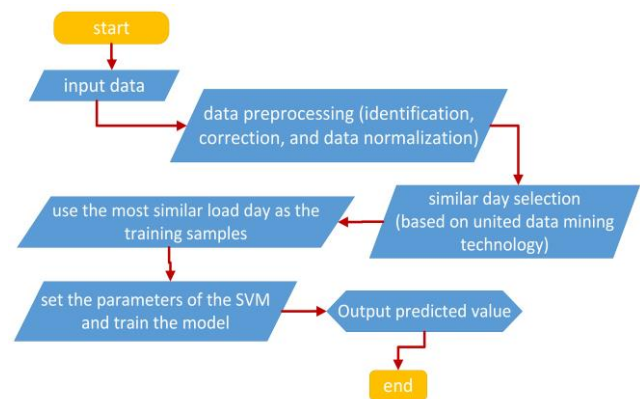


Figure 8. Flow chart of the load forecasting SVM similar days based application [24].

The kernel function of the SVM is taken to be the gaussian radial basis function. The predicted results are compared with the basic SVM method's results, and it gives an average error of 2.06% which is less than the average error for the classical SVM which turned out to be around 2.54%. This is due to the proposed method's ability to eliminate efficiently. The redundancy of data and makes calculations simpler. It's also efficient, fast, and can meet requirements.

## VI. CONCLUSION

Machine learning is a rapidly evolving field, that has left an impact in many other areas. Load forecasting is no exception, as several approaches pertaining to machine learning have been developed, investigated and tested in order to tackle the deficiencies of some of the prior techniques of load forecasting. The review of cutting-edge machine learning techniques utilized in various research for load forecasting was

the main subject of this paper. Generally, training wise, more data would yield more accurate predictions and more relevant input features to consider while predicting and forecasting loads using these methods. Machine learning methods need more studying and further research to be able to incorporate it towards creating more efficient, reliable systems and an important milestone in smart grids. It significantly outperformed several applications and offered various methodological advantages.

## ACKNOWLEDGMENT

## REFERENCES

[1] H. Malik, N. Fatema, and A. Iqbal, "Intelligent Data Analytics for Time-Series Load Forecasting Using Fuzzy Reinforcement Learning (FRL)," in Intelligent data-analytics for condition monitoring: Smart Grid Applications, London, United Kingdom: Academic Press, 2021, pp. 193–213.

[2] W. He, "Load forecasting via Deep Neural Networks," Procedia Computer Science, vol. 122, pp. 308–314, 2017.

[3] M. A. Hammad, B. Jereb, B. Rosi, and D. Dragan, "Methods and models for Electric Load Forecasting: A comprehensive review," Logistics & Sustainable Transport, vol. 11, no. 1, pp. 51–76, Feb. 2020.

[4] S. A.-h. Soliman and A. M. Al-Kandari, Electrical Load Forecasting: Modeling and Model Construction, 1st ed. Butterworth–Heineman, 2010.

[5] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 2nd ed. OTexts: Melbourne, Australia, 2018.

[6] N. Abu-Shikhah and F. Elkarmi, "Medium-Term Electric Load Forecasting Using Singular Value Decomposition," Energy Conversion and Management, vol. 36, no. 7, pp. 4259-4271, 2011.

[7] "What Is Machine Learning (ML)?", UCB-UMT, 2020. [Online]. Available: https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/. [Accessed: 15- Sep- 2022].

[8] "Machine Learning Techniques", www.javatpoint.com, 2022. [Online]. Available: https://www.javatpoint.com/machine-learning-techniques. [Accessed: 14- Sep- 2022].

[9] B. Mahesh, "Machine Learning Algorithms -A Review", International Journal of Science and Research (IJSR), vol. 9, no. 1, pp. 381-386, 2020. Available: DOI: 10.21275/ART20203995 [Accessed 15 September 2022].

[10] M. Hayati and Y. Shirvany, "Artificial neural network ap-proach for short term load forecasting for illam region," World Academy of Science, Engineering and Technology, vol. 28, pp.280–284, 2007.

[11] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," International journal of forecasting, vol. 14, no. 1, pp. 35–62, 1998.

[12] N. Kandil, R. Wamkeue, M. Saad, and S. Georges, "An efficient approach for short term load forecasting using artificial neural networks," International Journal of Electrical Power& Energy Systems, vol. 28, no. 8, pp. 525–530, 2006.

[13] D. C. Park, M. El-Sharkawi, R. Marks, L. Atlas, and M. Damborg, "Electric load forecasting using an artificial neural network," IEEE transactions on Power Systems, vol. 6,no. 2, pp. 442–449, 1991.

[14] kumar singh, Ibraheem, S. Khatoon, M. Muazzam, and D. Chaturvedi, "Load forecasting techniques and methodologies: A review," in 2012 2nd International Conference on Power, Control and Embedded Systems, 2013.

[15] Z. Wang, J. Li, S. Zhu, J. Zhao, S. Deng, S. Zhong, H. Yin, H. Li, Y. Qi, and Z. Gan, "A review of load forecasting of the Distributed Energy System," IOP Conference Series: Earth and Environmental Science, vol. 237, p. 042019, 2019.

[16] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338–353, 1965.

[17] P. Mukhopadhyay, G. Mitra, S. Banerjee, and G. Mukherjee, "Electricity load forecasting using fuzzy logic: Short term load forecasting factoring weather parameter," 2017 7th International Conference on Power Systems (ICPS), Jun. 2018.

[18] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, and M. H. Lee, "Short-term load forecasting by using a combined method of convolutional neural networks and Fuzzy Time Series," Energy, vol. 175, pp. 365–377, 2019.

[19] A. Lambora, K. Gupta, and K. Chopra, "Genetic algorithm- A literature review," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Oct. 2019.

[20] "Applications of genetic algorithms to load forecasting problem," Studies in Computational Intelligence, pp. 383–402, 2008.

[21] P. Ray, S. K. Panda, and D. P. Mishra, "Short-term load forecasting using genetic algorithm," Advances in Intelligent Systems and Computing, pp. 863–872, 2018.

[22] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[23] S. Quaiyum, Y. I. Khan, S. Rahman, and P. Barman, "Artificial Neural Network based Short Term Load Forecasting of Power System," International Journal of Computer Applications, vol. 30, no. 4, pp. 1–7, Sep. 2011.

[24] X. Liu, W. Zhang, S. Ye, and Y. Han, "Short-term load forecasting approach with SVM and similar days based on United Data Mining Technology," 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), Oct. 2019.