




Article

Towards Short Term Electricity Load Forecasting Using Improved Support Vector Machine and Extreme Learning Machine

Waqas Ahmad ¹, Nasir Ayub ¹, Tariq Ali ², Muhammad Irfan ² , Muhammad Awais ³ , Muhammad Shiraz ¹ and Adam Glowacz ^{4,*} 

¹ Department of Computer Science, Federal Urdu University of Arts, Science and Technology, Islamabad 44000, Pakistan; saqaw@ymail.com (W.A.); nasir.ayubse@gmail.com (N.A.); drmuhammadshiraz@fuuastisb.edu.pk (M.S.)

² College of Engineering, Electrical Engineering Department, Najran University, Najran 61441, Saudi Arabia; tariqhsp@gmail.com (T.A.); irfan16.uetian@gmail.com (M.I.)

³ School of Computing and Communications, Lancaster University, Bailrigg, Lancaster LA1 4YW, UK; m.awais11@lancaster.ac.uk

⁴ Department of Automatic Control and Robotics, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Kraków, Poland

* Correspondence: adglow@agh.edu.pl

Received: 30 April 2020; Accepted: 2 June 2020; Published: 5 June 2020



Abstract: Forecasting the electricity load provides its future trends, consumption patterns and its usage. There is no proper strategy to monitor the energy consumption and generation; and high variation among them. Many strategies are used to overcome this problem. The correct selection of parameter values of a classifier is still an issue. Therefore, an optimization algorithm is applied with deep learning and machine learning techniques to select the optimized values for the classifier's hyperparameters. In this paper, a novel deep learning-based method is implemented for electricity load forecasting. A three-step model is also implemented, including feature selection using a hybrid feature selector (XGboost and decision tree), redundancy removal using feature extraction technique (Recursive Feature Elimination) and classification/forecasting using improved Support Vector Machine (SVM) and Extreme Learning Machine (ELM). The hyperparameters of ELM are tuned with a meta-heuristic algorithm, i.e., Genetic Algorithm (GA) and hyperparameters of SVM are tuned with the Grid Search Algorithm. The simulation results are shown in graphs and the values are shown in tabular form and they clearly show that our improved methods outperform State Of The Art (SOTA) methods in terms of accuracy and performance. The forecasting accuracy of Extreme Learning Machine based Genetic Algo (ELM-GA) and Support Vector Machine based Grid Search (SVM-GS) is 96.3% and 93.25%, respectively. The accuracy of our improved techniques, i.e., ELM-GA and SVM-GS is 10% and 7%, respectively, higher than the SOTA techniques.

Keywords: electricity load forecasting; smart grid; feature selection; Extreme Learning Machine; Genetic Algorithm; Support Vector Machine; Grid Search

1. Introduction

Load forecasting has a huge impact on routine electric functions including fuel resource management and for accurate decision making to stabilize the demand and supply of electricity. After revolutionary overhauling of the electricity market internationally, the importance of load forecasting has increased multifold and also encompassed other areas of significance, e.g., financial

planning and energy trading, etc. Precision in load forecasting lays the foundation of a system to establish a spot pricing, which in turn benefits the system in acquiring minimum purchasing cost of electricity in the local and regional markets. Load forecasting has a great significance in demand-side management planning, storage planning to include its maintenance and schedule, integrating renewable energy sources and multiple other utilities involving smart grid implementations. Load forecasting also enables the electric consumers in creating a link between price and demand for electricity and get benefits by adjusting the usage pattern in line with cost advantages. Electric grids are defined as a network of electric power generators in which users are interconnected by transmission lines and well-synchronized through control centres. Moreover, the power grid usually refers to the transmission system of electric power. Similarly, Traditional Grids (TGs) are defined as the grids which connect power providing system to distribution network further. Whereas, a TG is operated on a centrally controlled system as shown in Figure 1.

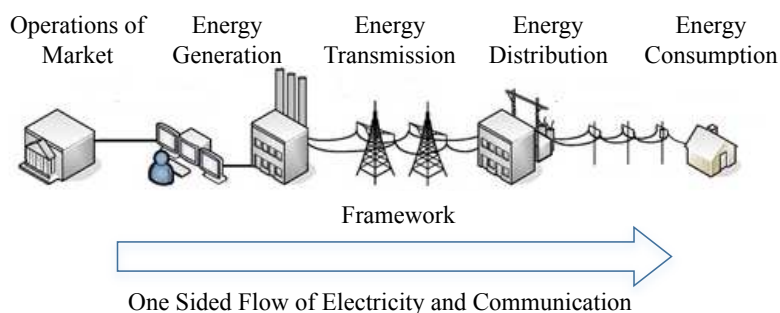


Figure 1. Traditional Grid.

Besides, Smart Grid (SG) is defined as an enhanced electric grid that operates in two-ways between a provider and consumer. It incorporates artificial intelligence, intercommunication between and smartly operating distributor network among them. SG brings a revolution by efficiently managing the power generation. Moreover, it distributes the electricity as per the requirement and consumption of the end-users. The latest technologies and types of equipment are also incorporated to establish SGs. Furthermore, it is a requirement in today’s energy-starved world due to an acute shortage of energy in summers and other conditions like bad weather, etc. SG incorporates the latest technologies and techniques from power generation to power consumption as shown in Figure 2. In SG, the work mainly focuses on infrastructure, management system and protective structure [1].

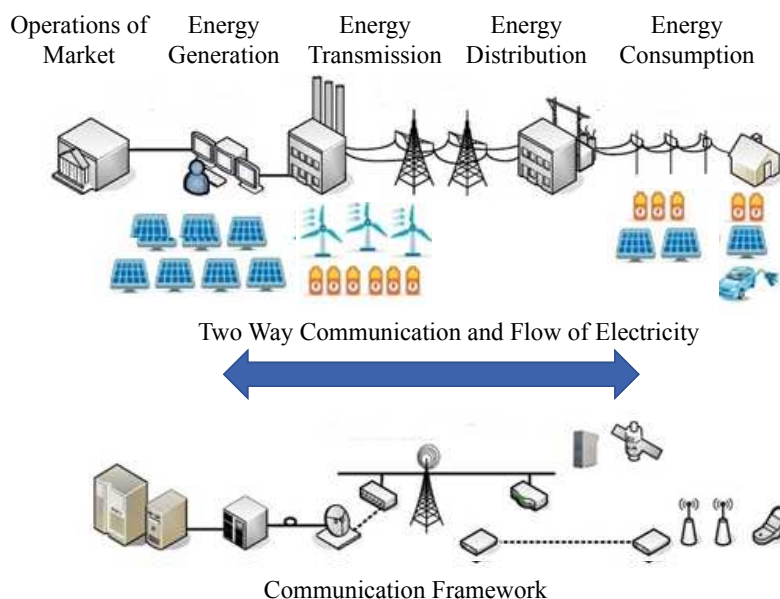


Figure 2. Smart Grid.

As per the aforementioned literature, the power grid is one of the most complex structures consisting of a myriad of transformers, substations, generator sets and long-distance power lines. The main purpose of the power grid is to supply power to end-users. It extends from a small local design to meet the daily needs of hundreds or millions of consumers through ultra-long high-voltage and low-voltage lines and makes an excellent interconnect structure. The basic components of this structure are shown in Figure 3, i.e., the power plant, transmission line and the distribution system.

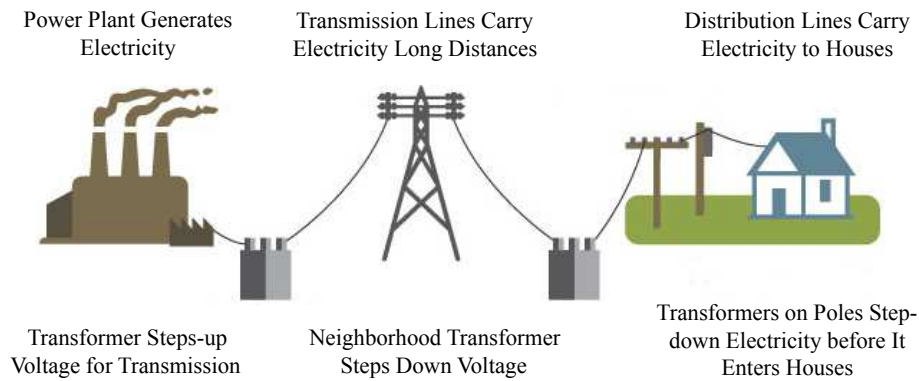


Figure 3. Electricity generation, transmission and distribution.

In a nutshell, the TG system is not well-equipped to fulfil the requirements of the end-users. Whereas, in SG, these grids are utilized in a manner to fulfil the need of the consumer. At this end, Smart Meters (SMs) are used to handle the appliances. Information is transferred between consumer and utility using these SMs, which benefit both the provider and the consumer [2]. The main techniques employed in SG are energy scheduling and management. Algorithms and subsystems exchange records and their statistics are used to further improve the energy generation and consumption in a closed-loop [3].

SM is an advanced energy meter that is used to get the consumption pattern information from the user as shown in Figure 4. So, there are two different types of SMs depending upon different communication technologies, i.e., power line carrier and radiofrequency [3–5].

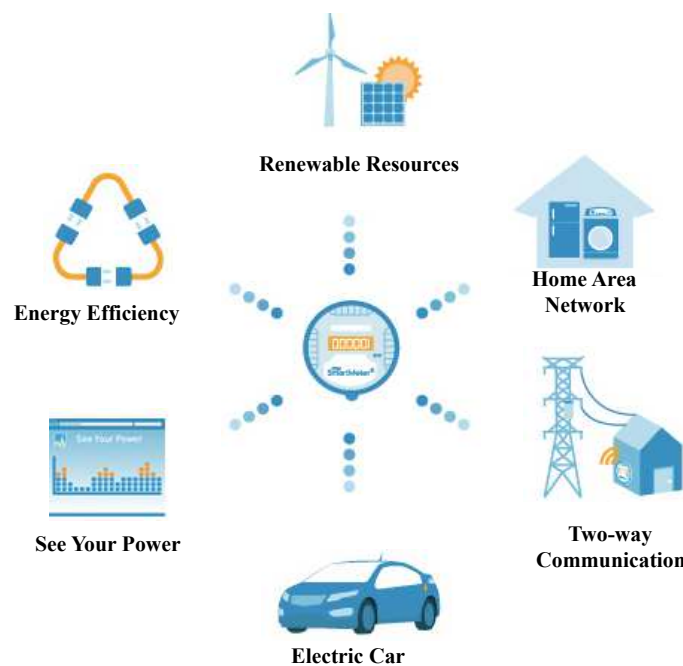


Figure 4. Smart Meter.

With the help of electricity load forecasting, the utility company can plan and have a better understanding of future consumption and load demand. The minimum risk for the utility company and understanding the pattern of the electricity help to fulfil the load demand. It helps the utility in understanding consumer demand and to define a proper period for the maintenance of electricity in residential areas [5–9]. Moreover, it helps the utility to avoid the generation of energy, in operation and supply chain management. The utility aims to seek for economical generation and maintains reliability in terms of an increase in load demand.

Transmission planning is also very important for the utility. It identifies the areas that need to be expanded in terms of the load to maintain reliability and growth rate. It also ensures that power can be delivered from generators [9–11]. Similarly, distribution planning is also a very important process that helps to meet the requirements, if required, i.e., determine location, size and time of installation of distribution equipment.

Nowadays, energy forecasting is being implemented through various models in the last few decades. However, increased industrialization and urbanization have led to an increased energy demand with already depleting energy resources. Hence, there is a strong need to significantly alter the existing models or to establish the latest forecasting models to incorporate the rising customer awareness and concerns and to make them an active participant in an SG environment [12].

There are few challenges which may be encountered during the development of the newer models, i.e., availability of greater volumes of electric power consumption data, hybrid grouping making profiling, correct identification of the leading factors contributing to electricity consumption, identifying major factors contributing to consumption, particularly in off-peak hours and quantifying and incorporating outside environmental parameters and incorporating in a climate system or other resources of energy, and scenario-based feedback collection, i.e., non-availability of demand and corresponding cost history data.

Therefore, many techniques are proposed to address these aforementioned issues; however, there still exist some challenges, i.e., fluctuation in energy generation and consumption to control the fluctuating behaviour between the energy consumption pattern and generation pattern, technique accuracy and tuning the hyperparameters for the prediction of electricity load data. To address these issues, a machine learning and deep learning-based model is proposed. The main contributions of this work are:

- a machine learning and deep learning-based model is proposed, i.e., Extreme Learning Machine based Genetic Algorithm (ELM-GA) and Support Vector Machine based Grid Search (SVM-GS),
- the hyperparameter values are tuned using an optimization algorithm to obtain maximum accuracy,
- DT, XGboost and RFE are used in the feature engineering process for removing the redundancy and cleaning the data,
- the GA and GS optimization algorithms are applied to the ELM and SVM to calculate the optimum hyperparameter values.

2. Related Work

Electricity load forecasting is important in terms of production and transmission of energy. The basic work of any system which deals with power or energy is to keep checking the system load at all possible times. This tracking is for different periods. Many factors can cause variations in the load. So, a proper watch must be done on power generation as these sources are not much flexible [12]. The fuel must be available in large quantities for power generation. The following are the three types of load forecasting mainly. The duration of one to years are considered in long term forecasting. Medium-term load forecasting which covers months to weeks, whereas short term forecasting is done on the scheduling of days or usually an hour of a day [13–16]. A similar day lookup approach method is also used for electricity load forecasting and is based upon historical data or previous data for any period, which has the same characteristics, for example, someday of a

week, different dates, weather, humidity, etc. This previous data can be used to make trend analysis using regression [17]. The regression-based approach in this technique, linear regression is used to examine dependent variation to specify independent variation. First, consider the independent variable, because most changes occur in the independent variable. In load forecasting, the dependent variable is usually electricity demand and depends on electricity production. Independent variables are usually related to weather, such as humidity, temperature, and wind. The future value of the dependent variable can be estimated by this method [18–21]. The artificial neural network performs nonlinear modeling and adaptation. There is no need to presume a functional relationship between weather and load variables as they are not required in advance. The ANN can adapt the new data when it is exposed. ANN is currently being used in power system problems such as alarm processing, topological observability, fault diagnosis and security assessment, etc. [22]. Time series analysis is done on properly sequenced data which is at uniform intervals for the desired results. TSA is used to know the pattern of the data to predict or know future results based on the previous events that occurred. This method is usually applied to the short time span for the prediction of future events [23].

Expert systems are very intelligent. It can increase their knowledge and work when new data is given to the system. In expert systems, knowledge engineers are added to acquire knowledge and used to create new predictive models for load prediction [24]. Support Vector Machines (SVMs) are a powerful way to provide solutions to regression and classification problems. Linear functions are used in SVM to create linear decision boundaries. In the neural network, there is a problem in choosing the correct architecture, whereas, in SVM, difficulty occurs in selecting suitable kernel [25]. Fuzzy logic is a method that uses the same logic as Boolean logic. Boolean accepts input truth values as 0 and 1. In fuzzy logic, inputs are based on comparisons. In this technique, mathematical models are not used for mapping the value of the input and the result. Fuzzy logic is not affected by noise. When processing is done by this technique “defuzzification” is used to obtain an accurate output or result [26]. The time factor method is completed with short-term load forecasting since it is used every hour for a short time [27]. The weather factor technique uses the load pattern of different weather conditions as the load varies season to season. These loads can be of different appliances heavily used in a different season. Temperature is an important factor in terms of the effect created in this method [28]. The random disturbance is the method in which load demand depends on the sudden disturbance like the suddenness of Television when Minister of any country goes public or come in any program. During this time, the power system contains electrical machines that are being turned on or turned off, which has an impact on sudden changes in the loads. This change is called a random disturbance [29]. Cultural events also cause random disturbance on load. Economic factors have a big influence on LTLF and less effect on STLF. The economic factors have a clear effect on energy consumption. Different factors like industries’ heavy appliances generally economic trends have an impact on load growth [30].

An intelligent modular approach is a technique, which is used on STLF. This method has a great impact on the day ahead load forecasting models such as fuel purchases and planning for energy. It aims to increase forecast accuracy without spending more execution time by using a hybrid artificial neural network [31]. To improve the accuracy of electricity load forecasting, Restricted Boltzmann Machine (RBM) and Rectified Linear Unit (ReLU) are implemented. For short term electricity load forecasting, a neural network with the ReLU activation function is implemented. RBM process and train the data, while ReLU performs electricity load forecasting [32]. Deep Long Short-Term Memory (DLSTM) with DNN is applied for the forecasting of electricity load and price. The combination of DLSTM with DNN improves the processing, which can easily be carried out on big data [33]. Deep Auto Encoders (DAEs) provide excellent results in obtaining or achieving accuracy and understanding data. DAE has improved performance in obtaining the accuracy of the results [34]. Gated Recurrent Units (GRU) method is applied to forecast the electricity price [35]. Enhanced SVM and enhanced CNN are implemented using the feature engineering model for classification [36]. Neural network techniques, i.e., the layers of Long Short-Term Memory (LSTM) is added in the layers of Convolutional

Neural Network (CNN) for forecasting the load [37–39]. CNN and LSTM perform better when they are combined rather than performing separately with several other models. The hybrid of CNN and LSTM gives better results [40]. Stack berg is a theoretical game used for the minimization of operational cost and scheduling appliances [41,42]. Feature engineering is used for selection and extraction [43]. There are many techniques used for feature engineering in electric loads [44–47]. Residential load management uses meta-heuristic and machine learning techniques to check energy costs and achieve maximum benefits or comfort during peak hours [48]. In [49], author tuned the hyperparameters of SVM and achieve better accuracy. Wind speed and power prediction is performed using machine learning algorithms in [50].

3. Problem Statement and Motivation

In machine learning, every technique has pros and cons. However, better performance and accuracy are the key concerns in forecasting the electricity load. At this end, a huge amount of data produces hindrance to achieve accuracy during forecasting. As a result, many techniques are designed and modified to address these issues within the limited period; however, there still exist some challenges, i.e., fluctuation in energy generation and consumption to control the fluctuating behavior between the energy consumption pattern and generation pattern [32], technique accuracy and tuning the hyperparameters for the prediction of electricity load data [26,49] and computational complexity during the unclear information, e.g., irrelevant and redundant features in the data, which increase the computation time of the training process and decrease the accuracy of electricity load forecasting. To address these issues, a machine learning and deep learning-based model is proposed. Moreover, the hyperparameter values are tuned using an optimization algorithm to obtain maximum accuracy. Furthermore, DT, XGboost and RFE are used in the feature engineering process for removing the redundancy and cleaning the data. In the end, the GA and GS optimization algorithms are applied to the ELM and SVM to calculate the optimum hyperparameter values.

4. Proposed Model

To address the mentioned problems, we have proposed a model in which firstly, all the features/data from the dataset are imported. Secondly, a hybrid feature selector; XGboost and DT are applied to select useful features. Thirdly, the more relevant features are extracted using RFE method. After working and preprocessing of data, ELM classifier with GA and SVM classifier with GS is applied for classification and forecasting load data as shown in Figure 5. Figure 6 shows the flow chart of the proposed model.

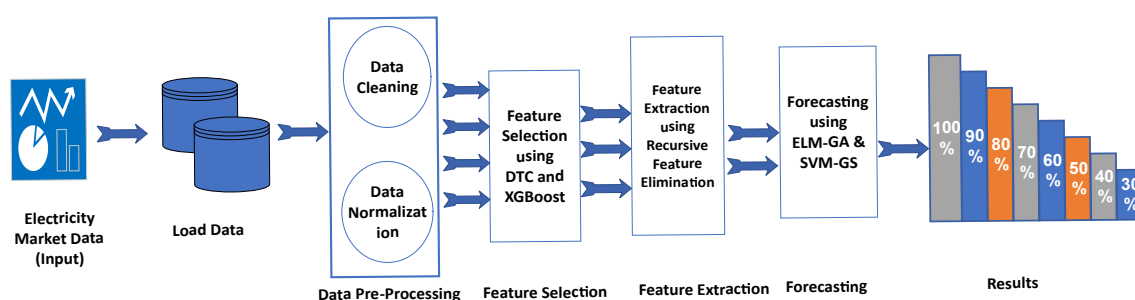


Figure 5. Proposed model.

4.1. Dataset

The daily electricity load data of three years, i.e., January 2017 to December 2019 are used in this paper, which is taken from Independent System Operator New England (ISO NE) (<https://www.iso-ne.com>). It supplies electricity to different cities in England. The dataset contains dependent and independent data, i.e., weather, temperature, humidity, etc. A column named “electricity load” is our target data. All the features other than the target features have an impact on the target data.

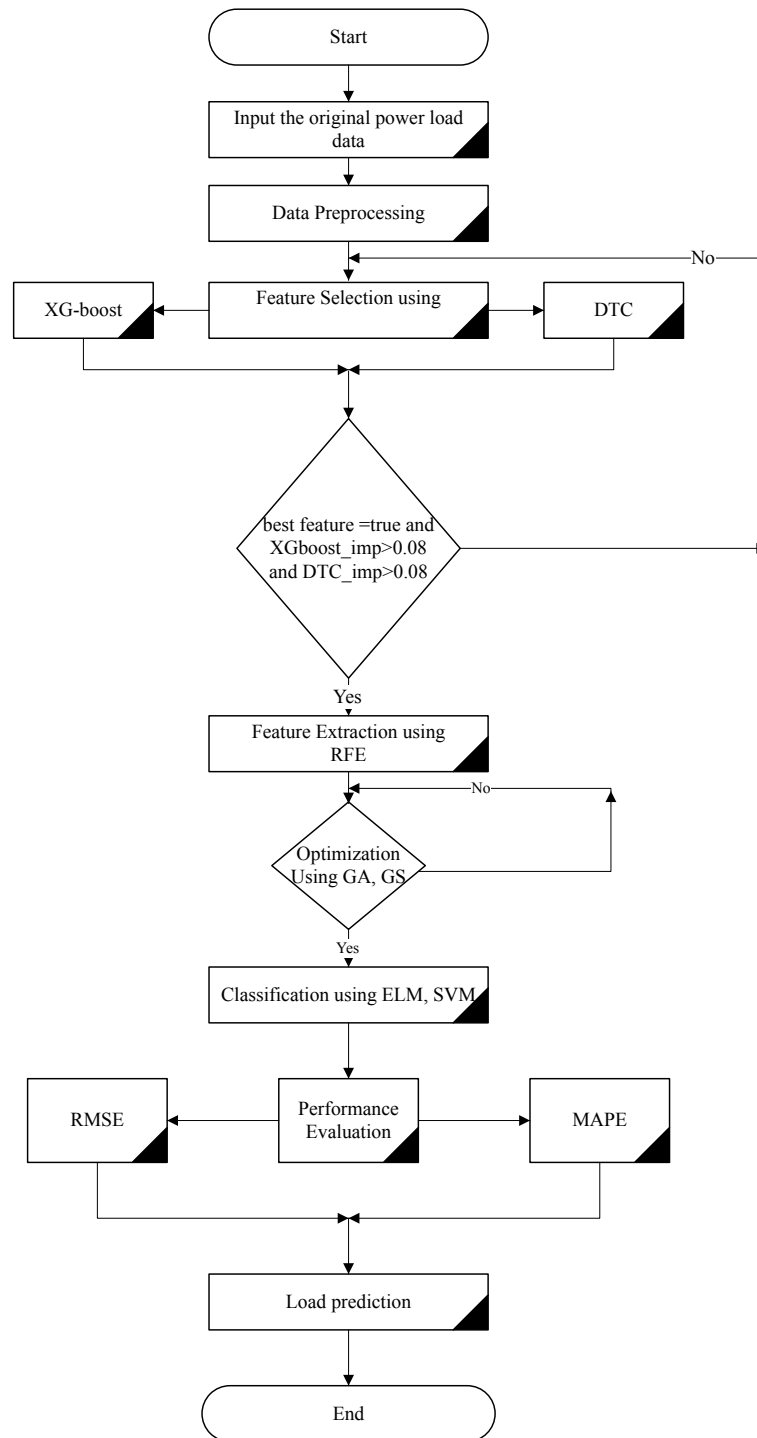


Figure 6. Flow Chart.

As the consumption pattern of the electricity load of the similar month in every year is approximately the same. Therefore, we have taken three years of data, i.e., 36 months. At this end, we have divided the dataset into sets, i.e., training and testing. Therefore, 90% of data is used as training while 10% of the data is used for testing because more the data is provided for training, more will be the learning rate of the model. Furthermore, the data of similar months of every year like January 2017, January 2018, and January 2019 are combined to perform a short-term load forecast for December 2019.

The month-wise arrangement of data in the dataset helps us in better training of our model to find out the load pattern of months. For weekly forecasting, all the data except the first week of December 2019, i.e., 1 December 2019 to 7 December 2019 are taken as training. The training model is tested in the first week of December. Moreover, for the first five months of the year 2019 are also considered for training and testing. Likewise, all data except January 2019 are also taken for training and testing purposes. Moreover, the same scenario is followed for February 2019, March 2019, April 2019 and May 2019. Simulation and results show the efficacy of the proposed models. Figure 7 shows the data overview and feature names.

Date	DA_Demand	RT_Demand	DA_LMP	DA_EC	DA_CC	DA_MLC	RT_LMP	RT_EC	RT_CC	RT_MLC	Dry_Bulb	Dew_Point	System_Load	Reg_Capacity_Price	Reg_Service_Price
1/1/2017	16679	16391	63	62	0	1	45	45	0	0	30	9	16688	6	0
1/2/2017	16924	16649	57	56	0	1	22	22	0	0	31	15	16934	7	0
1/3/2017	15999	16624	53	52	0	0	58	58	0	0	27	25	16958	12	1
1/4/2017	14404	15765	41	40	0	0	77	76	0	1	41	40	16058	21	1
1/5/2017	18036	18557	83	82	0	1	63	62	0	1	19	-2	18867	8	3
1/6/2017	18504	19259	82	82	0	1	130	130	0	1	15	10	19554	28	0
1/7/2017	19409	19883	115	114	0	1	127	126	0	1	10	-11	20159	24	0
1/8/2017	19192	19813	107	106	0	1	90	89	0	1	19	0	20143	11	0
1/9/2017	17996	18155	85	85	0	1	60	60	0	1	23	8	18454	5	0
1/10/2017	17734	17620	102	101	0	1	81	80	0	1	17	-3	17899	8	0
1/11/2017	17029	17601	113	113	0	1	87	86	0	1	27	12	17874	11	2
1/12/2017	17489	17694	97	96	0	1	85	84	0	1	34	32	17993	17	0
1/13/2017	19463	19034	116	115	0	1	104	101	3	1	15	-2	19307	11	1
1/14/2017	18440	18901	97	96	0	1	93	93	0	1	21	15	19213	9	1
1/15/2017	17852	18212	107	106	0	1	99	98	0	1	26	17	18513	19	1
1/16/2017	18037	17913	114	112	0	1	88	87	0	1	20	-1	18219	23	1
1/17/2017	17640	17321	104	103	0	1	60	60	0	0	20	0	17574	9	0
1/18/2017	15825	15683	53	53	0	0	7	7	0	0	44	41	15929	13	0
1/19/2017	17233	17185	73	72	0	0	49	49	0	0	35	23	17443	17	0
1/20/2017	17716	17683	72	71	0	1	56	56	0	0	30	9	17917	17	0

Figure 7. Dataset Overview.

4.2. Feature Engineering

The statistical mechanics of the feature selection process of the proposed model is applied to the dataset and the most relevant features are selected from the dataset by calculating feature importance. We achieved accurate results by combining DT and XG boost as shown in Figure 8. By setting the feature selection threshold, select features whose importance is equal to or higher than the threshold while deleting other features. The features are selected according to the following formula:

$$f(s) = \begin{cases} \text{reserve } i, & IXG[f] + IDT[f] \geq t, \\ \text{drop } i, & IXG[f] + IDT[f] < t. \end{cases} \quad (1)$$

where, variable IXG [i] indicates the calculated feature importance by XGB, and IDT [i] by the DT method. The symbol t represents the threshold for selection of features, and the f symbol represents the feature.

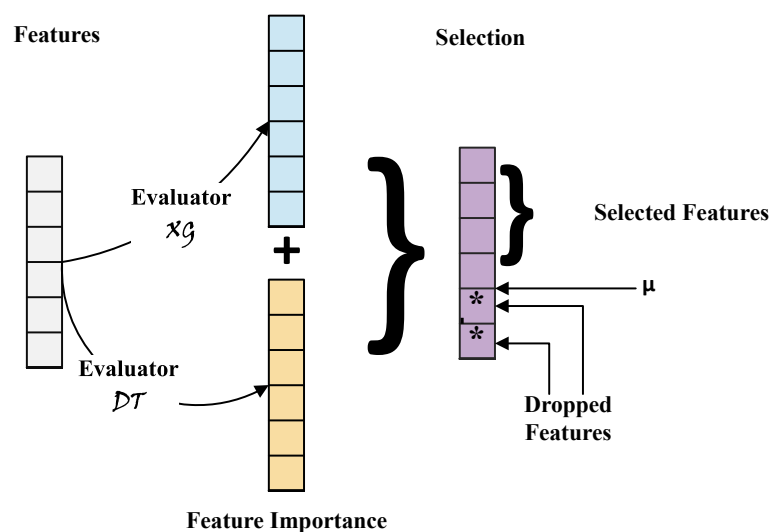


Figure 8. Hybrid feature selector.

Where evaluator XG and DT show the feature' importance calculated by XGboost and Decision Tree Technique. μ represents the threshold set for the selection of features.

After the feature selection step, the feature extraction is carried out by the RFE method. The feature extraction technique extracts non-redundant features and also those features that have a high impact on the target feature in the dataset. RFE, as its name suggests that it removes features recursively and builds a model with the help of extracted attributes and calculates or gives the accuracy of the model. RFE can work on the combination of different attributes that can help in predicting the target variable.

4.3. Classification and Forecasting

ELM with GA algorithm and SVM with the GS algorithm is used as a classifier in the proposed model. The ELM technique is a feed-forward NN used for classification, clustering, regression and compression. ELM can produce better generalization performance when tuned with GA. These models can learn faster than networks based on backpropagation. ELM-GA models can give better performance than SVM-GS. ELM is used for classification and regression problems. ELM is used to learn about a parameter of hidden nodes, input weights that are randomly used and are not tuned. ELM provides extremely fast learning with greater speed, better performance with less involvement of humans [48]. There are different variants of ELM, i.e., voting-based ELM, incremental ELM, error-minimized ELM, pruning ELM, and evolutionary ELM, etc. The main applications of ELM are pattern recognition, classification, image processing, regression, and forecasting, etc.

ELM is very effective as a new training algorithm for SLFN. It is effective and very efficient. Therefore, for a single hidden layer of ELM, we suppose that the i -th node is

$$h_i(\mathbf{x}) = G(\mathbf{a}_i, b_i, \mathbf{x}). \quad (2)$$

where, a_i and b_i represents the parameters of i -th node. The output function of ELM of L hidden layer is

$$f_L(\mathbf{x}) = \sum_{i=1}^L \beta_i h_i(\mathbf{x}). \quad (3)$$

where B_i is the output of i -th hidden node

$$\mathbf{h}(\mathbf{x}) = [G(h_1(\mathbf{x}) \dots, h_L(\mathbf{x}))], \quad (4)$$

is the hidden layer of output mapping of ELM. Where N training sample and output matrix H is

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} G(\mathbf{a}_1, b_1, \mathbf{x}_1) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ G(\mathbf{a}_1, b_1, \mathbf{x}_N) & \cdots & G(\mathbf{a}_L, b_L, \mathbf{x}_N) \end{bmatrix}, \quad (5)$$

and T is training data of the target matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_N \end{bmatrix}. \quad (6)$$

ELM is a type of NN in which the parameter, i.e., mapping of hidden layers is tuned GA (an optimization algorithm). Furthermore, its main object is:

$$\text{Minimize} : \|\beta\|_p^{\sigma_1} + C \|\mathbf{H}\beta - \mathbf{T}\|_q^{\sigma_2}, \quad (7)$$

where $\sigma_1 > 0, \sigma_2 > 0, p, q = 0, \frac{1}{2}, 1, 2, \dots, +\infty$.

Different combinations of the above parameters are used and their results are used in different algorithms for classification and regression.

The SVM technique draws a hyperplane between the data. Accurate classification of data can be made by the optimal values of the hyperparameters of SVM. The optimized value of the main three parameters of SVM, i.e., intensive loss function (ϵ), cost penalty (c), and kernel parameter (k) is still a serious issue. In our proposed model, the hyperparameters of SVM are tuned with the GS algorithm, which increases the performance of SVM.

5. Simulation Setup

To simulate our proposed model, we have used anaconda spyder (a desktop application) based on python libraries, i.e., Keras v2.3.0, NumPy v1.18.4, TensorFlow v2.0, and Scikit-learn v0.23. The system specifications are Intel core i3, 512 GB of storage and 8 GB RAM. Whereas, the electricity load dataset used in our proposed model is taken from <https://www.iso-ne.com>. Simulation results show the efficacy of the proposed models, respectively.

(1) Feature Selection based on Hybrid Feature Selector

The hybrid feature selector (XG-Boost and DT) is applied to the dataset, which calculates the importance of features in numeric format. The high feature importance value is calculated via feature selection technique, which shows the dependency of the input features (high or low) on target value, i.e., in the current scenario this is named “system load” in the respective dataset. Features with high importance are elected as the best feature and features with low importance value are considered as unimportant features. Hereafter, the unimportant features are removed from the dataset. The removal of unimportant features is according to the threshold, i.e., 0.08 in the given scenario. Figure 9 shows the features’ importance calculated by feature selection techniques.

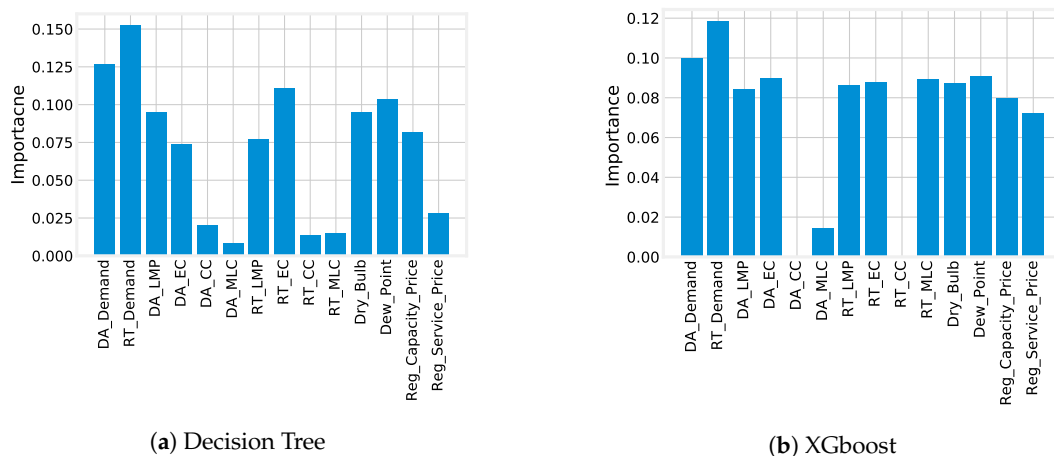


Figure 9. Feature importance.

(2) Feature Extraction using RFE

Table 1 shows the target features and other influential features of the dataset. After feature selection, RFE is used as a feature extractor, which results in dimension reduction. The features with the “TRUE” dimensions are not redundant and have a high impact on the target feature, which are selected as an input to the classifier. While the features with the “False” dimensions are redundant and rejected.

Table 1. Features Overview and Dimensions calculated by Redundancy removal using Feature Extraction (RFE).

Target Feature	Features	Short Name	Dimension
System Load	Day-Ahead Cleared Demand	DA_Demand	TRUE
	Regulation Market Service clearing price	Reg_Capacity_Price	TRUE
	Real-Time Demand	RT_Demand	TRUE
	The dewpoint temperature	Dew_Point	FALSE
	Day-Ahead Locational Marginal Price	DA_LMP	FALSE
	The dry-bulb temperature	Dry_Bulb	FALSE
	Energy Component of Day-Ahead	DA_EC	FALSE
	Marginal Loss Component of Real-Time	RT_MLC	FALSE
	Congestion Component of Day-Ahead	DA_CC	FALSE
	Congestion Component of Real-Time	RT_CC	FALSE
	Marginal Loss Component of Day-Ahead	DA_MLC	FALSE
	Energy Component of Real-Time	RT_EC	TRUE
	Real-Time Locational Marginal Price	RT_LMP	TRUE
	Regulation Market Capacity clearing	Reg_Service_Price	FALSE

(3) ELM-GA and SVM-GS performance and Comparison with SOTA Algorithms

Figure 10 shows the normal load of duration 1 January 2019 to 31 December 2019. The consumption pattern of electricity load is shown in Figure 10, which is according to the dataset provided by ISO-NE. Whereas, the variation between the days of the month is shown in Figure 10. It is due to the arrangement of data in months.

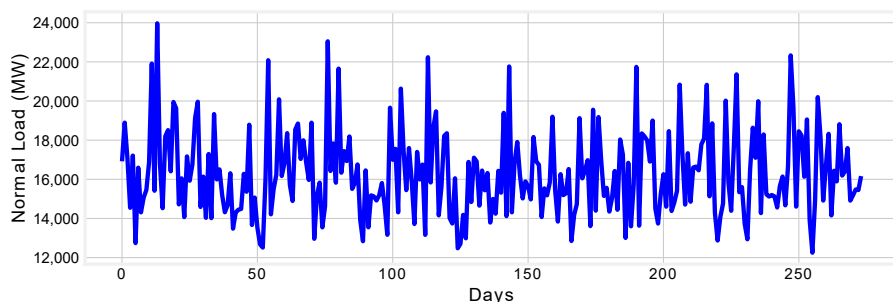


Figure 10. Normal electricity load according to the dataset.

In Figure 11, the results of the forecasted load of the first week of December, i.e., 1 December to 7 December 2019 is shown. Results validate that one of the proposed techniques, i.e., ELM-GA performs better than SVM-GS in terms of load forecasting accuracy.

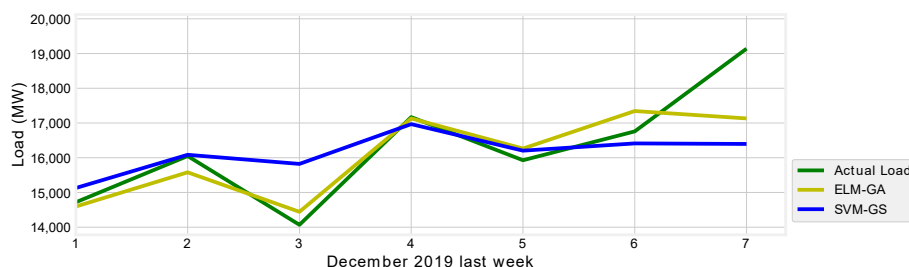


Figure 11. One week prediction (1 December 2019 to 7 December 2019).

The comparison of forecasted results of the proposed methods, i.e., ELM-GA and SVM-GS with an actual load of respective months are shown in Figure 12. Simulation results show that the performance of the proposed technique ELM-GA is nearly the same as the actual data, while the graph of proposed method SVM-GS shows a little deviation as compared to the actual values. However, SVM-GS has

better performance than SOTA approaches. In Figure 12, all data except January 2019 is taken as training data. After training the model, it is tested on January 2019. The same cases are revised for February 2019, March 2019, April 2019, May 2019 and June 2019, respectively.

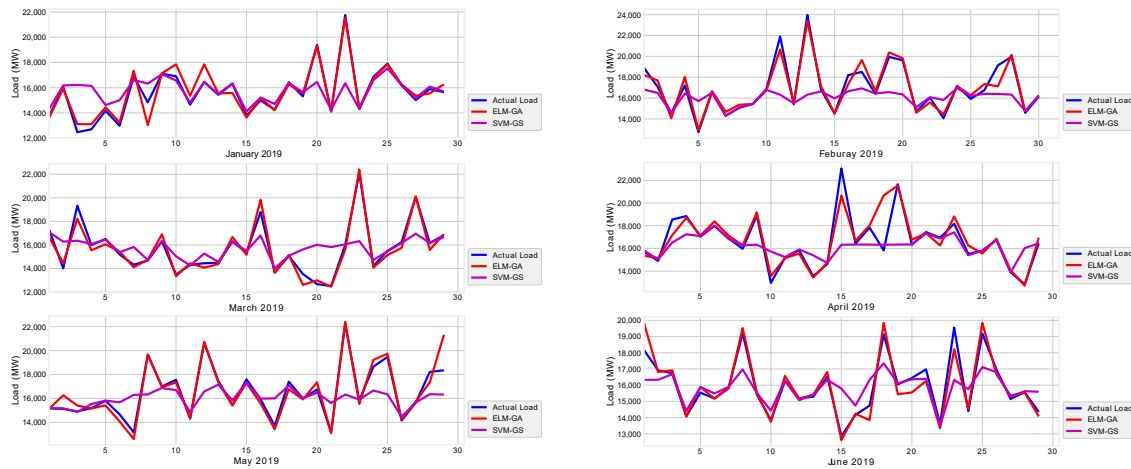


Figure 12. Comparison of forecasting results of different months with actual month load data.

Figure 13 shows the comparative analysis of the proposed models against SOTA techniques. We have used the optimization techniques to tune the parameters of the classifier and to get the prediction accuracy of the classifier that is close to the actual values given in the dataset.

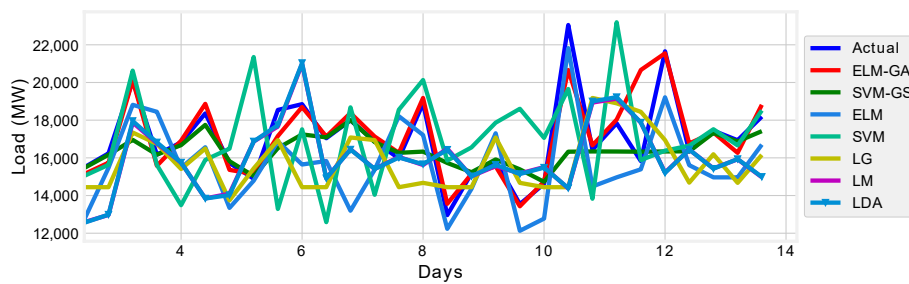


Figure 13. Forecasting of proposed technique with respect to the State of the Art (SOTA).

Figure 14 shows load prediction using optimized ELM based GA; whereas, GA is used for parameter tuning. Also, the proposed technique is compared with SVM in which GS is used for parameter tuning. Simulation results show that the proposed model is more accurate than others, as shown in Table 2.

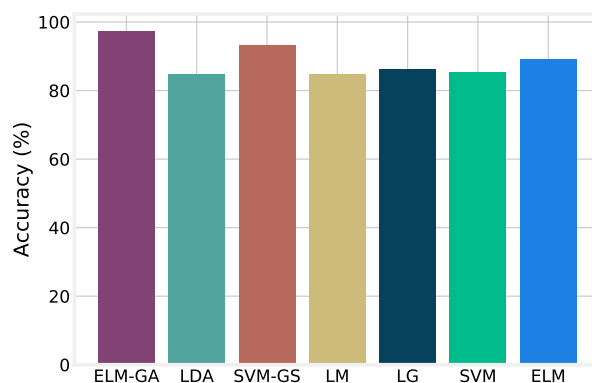


Figure 14. Accuracy of proposed technique with respect to SOTA.

Table 2. Proposed and SOTA techniques results.

Techniques	Accuracy	Performance Metrics			
		MAPE	RMSE	MSE	MAE
ELM-GA	96.42%	2.58	737.35	5.44	4.39
SVM-GS	93.25%	6.75	1811.95	32.83	10.95
LG	86.14%	13.86	2918.49	85.18	22.1
LM	84.88%	15.12	3030.06	91.81	24.46
LDA	84.9%	15.1	3028.09	91.69	24.44
ELM	89.23%	10.77	2014.66	40.59	16.12
SVM	85.31%	14.69	3034.03	92.05	22.75

In addition, performance evaluation metrics are used to demonstrate the productivity of the proposed techniques. At this end, results validate the accuracy of ELM-GA and SVM-GS as 93.25% and 96.42%, respectively. The ELM-GA and SVM-GS have 7% and 8% better results than SOTA techniques.

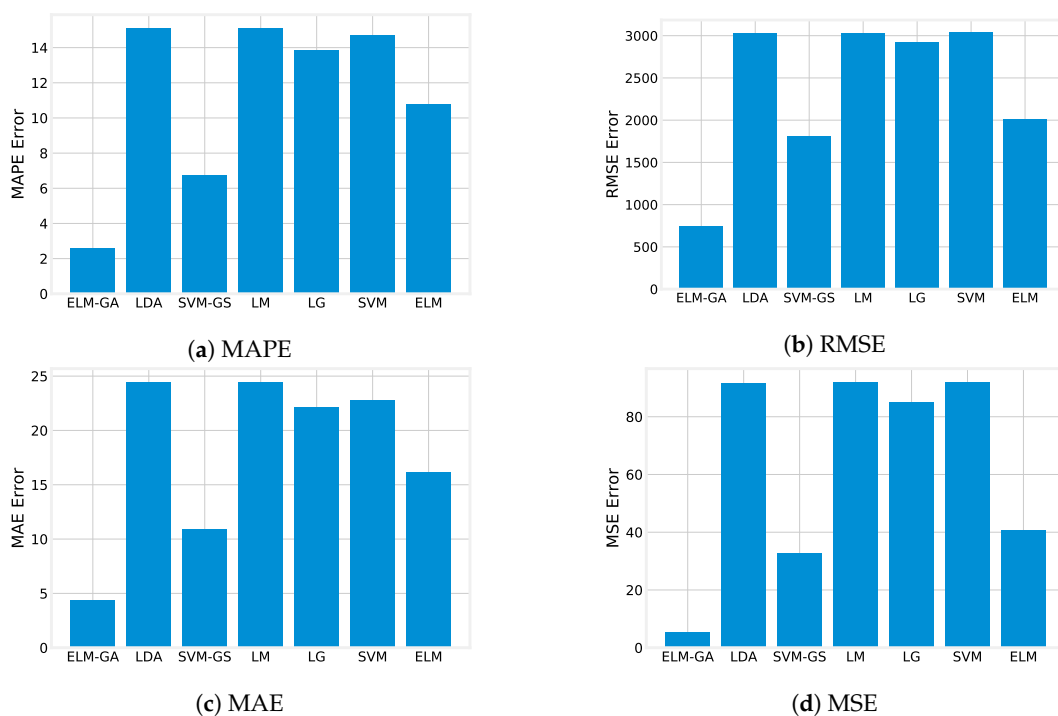
Performance Evaluation

The performance error of the proposed and SOTA methods are evaluated using performance metrics, i.e., Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The Error rate is calculated by RMSE and MAPE using Equations (8) and (9), respectively [51].

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (F_i - A_i)^2}{n}} \quad (8)$$

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (9)$$

where A , A_t represents the actual value and F , F_t shows forecast value. The values of MAPE and RMSE of proposed techniques are much lower than SOTA as shown in Figure 15.

**Figure 15.** Performance error.

The detailed results with their performance error are shown in Table 2. It is clearly shown in the table that the proposed techniques outperform SOTA techniques. Here, the performance of ELM-GA is higher than SVM-GS in terms of execution time and forecasting accuracy. ELM-GA and SVM-GS have 7% and 8% better results than benchmark techniques.

Figure 16 describes the asymmetric loss function and accuracy of the model, which shows that the performance of the model is increasing continuously. Moreover, the loss is decreasing by training the model with respect to Epochs.

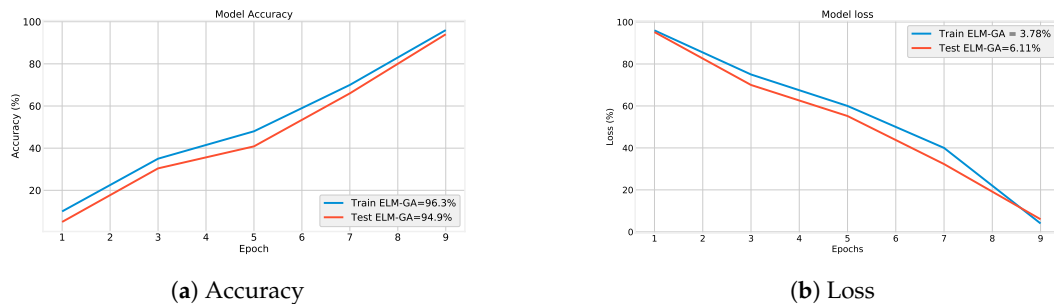


Figure 16. Performance accuracy and loss function.

6. Conclusions

In this paper, a machine learning and meta-heuristic algorithm based model is proposed for short term electrical load forecasting. The main advantage associated with this model is to achieve better performance and maximum accuracy. Therefore, a real-time electricity load dataset is taken into account, ranging from January 2017 to December 2019 from ISO-NE. Afterwards, data is converted into supervised learning sets, i.e., features (training x and test label x) and labels (training y and test label y). Furthermore, the proposed model is divided into three fundamental steps including feature selection using a hybrid feature selector (XGboost and DT), Redundancy removal using Feature Extraction (RFE) technique and classification/forecasting using ELM-GA and SVM-GS. At this end, parameters of ELM are tuned with the GA and hyperparameters of SVM are tuned with GS. In the end, a comparative analysis is performed with baseline SOTA techniques. It is evident from the results that proposed techniques outperform SOTA techniques in counterparts. In the future, the different low-performance classifiers can further be enhanced with different meta-heuristic techniques to improve forecasting accuracy.

Author Contributions: The research conceptualization and methodology were done by W.A., N.A. and M.A. The technical and theoretical framework was prepared by T.A. and M.I. The technical review and improvement were performed by M.A. and N.A. The overall technical support, guidance, and project administration was done by M.S. and A.G. Finally, responses to the queries of the reviewers were done by N.A. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TGs	Traditional Grids
SG	Smart Grid
SMs	Smart Meters
SVM	Support Vector Machine
RBM	Restricted Boltzmann Machine
ReLU	Rectified Linear Unit
DLSTM	Deep Long Short-Term Memory
DAEs	Deep Auto Encoders
GRU	Gated Recurrent Units
CNN	Convolutional Neural Network

ISO NE	Independent System Operator New England
SOTA	State Of The Art
ELM	Extreme Learning Machine
GS	Grid Search
NN	Neural Network
MAPE	Mean Average Percentage Error
RFE	Redundancy removal using Feature Extraction
RMSE	Root Mean Square Error

References

1. Fang, X.; Misra, S.; Xue, G.; Yang, D. Smart grid—The new and improved power grid: A survey. *IEEE Commun. Surv. Tutorials* **2011**, *14*, 944–980. [[CrossRef](#)]
2. Samadi, P.; Wong, V.W.; Schober, R. Load scheduling and power trading in systems with high penetration of renewable energy resources. *IEEE Trans. Smart Grid* **2015**, *7*, 1802–1812. [[CrossRef](#)]
3. Zhao, Z.; Lee, W.C.; Shin, Y.; Song, K.B. An optimal power scheduling method for demand response in home energy management system. *IEEE Trans. Smart Grid* **2013**, *4*, 1391–1400. [[CrossRef](#)]
4. Davito, B.; Tai, H.; Uhlaner, R. The smart grid and the promise of demand-side management. *McKinsey Smart Grid* **2010**, *3*, 8–44.
5. Liu, Y. Wireless sensor network applications in smart grid: Recent trends and challenges. *Int. J. Distrib. Sens. Networks* **2012**, *8*, 492819. [[CrossRef](#)]
6. Siano, P.; Sarno, D. Assessing the benefits of residential demand response in a real time distribution energy market. *Applied Energy* **2016**, *161*, 533–551. [[CrossRef](#)]
7. Aghaei, J.; Alizadeh, M.I. Demand response in smart electricity grids equipped with renewable energy sources: A review. *Renew. Sustain. Energy Rev.* **2013**, *18*, 64–72. [[CrossRef](#)]
8. Paterakis, N.G.; Erdinc, O.; Catalao, J. P. An overview of Demand Response: Key-elements and international experience. *Renewable and Sustainable Energy Reviews*. *Renew. Sustain. Energy Rev.* **2017**, *69*, 871–891. [[CrossRef](#)]
9. Pinson, P.; Madsen, H. Benefits and challenges of electrical demand response: A critical review. *Renew. Sustain. Energy Rev.* **2014**, *39*, 686–699.
10. Tabar, V.S.; Jirdehi, M.A.; Hemmati, R. Energy management in microgrid based on the multi objective stochastic programming incorporating portable renewable energy resource as demand response option. *Energy* **2017**, *118*, 827–839. [[CrossRef](#)]
11. Zheng, J.; Gao, D.W.; Lin, L. Smart meters in smart grid: An overview. In Proceedings of the 2013 IEEE Green Technologies Conference (GreenTech), Denver, CO, USA, 4–5 April 2013; pp. 57–64.
12. Bessa, R.J. Solar power forecasting for smart grids considering ICT constraints. In Proceedings of the 4th Solar Integration Workshop, Berlin, Germany, 10–11 November 2014.
13. Huang, S.J.; Shih, K.R. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. *IEEE Trans. Power Syst.* **2003**, *18*, 673–679. [[CrossRef](#)]
14. Kandil, N.; Wamkeue, R.; Saad, M.; Georges, S. An efficient approach for shorter term load forecasting using artificial neural networks. In Proceedings of the 2006 IEEE International Symposium on Industrial Electronics, Montreal, QC, Canada, 9–13 July 2006; Volume 3, pp. 1928–1932.
15. Mandal, P.; Senjyu, T.; Urasaki, N.; Funabashi, T. A neural network based several-hour-ahead electric load forecasting using similar days approach. *Int. J. Electr. Power Energy Syst.* **2006**, *28*, 367–373. [[CrossRef](#)]
16. Topalli, A.K.; Erkmen, I.; Topalli, I. Intelligent short-term load forecasting in Turkey. *Int. J. Electr. Power Energy Syst.* **2006**, *28*, 437–447. [[CrossRef](#)]
17. Mu, Q.; Wu, Y.; Pan, X.; Huang, L.; Li, X. Short-term load forecasting using improved similar days method. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 28–31 March 2010; pp. 1–4.
18. Wang, J.M.; Wang, L.P. A new method for short-term electricity load forecasting. *Trans. Inst. Meas. Control* **2008**, *30*, 331–344. [[CrossRef](#)]
19. Ruzic, S.; Vuckovic, A.; Nikolic, N. Weather sensitive method for short term load forecasting in electric power utility of Serbia. *IEEE Trans. Power Syst.* **2003**, *18*, 1581–1586. [[CrossRef](#)]

20. Haida, T.; Muto, S. Regression based peak load forecasting using a transformation technique. *IEEE Trans. Power Syst.* **1994**, *9*, 1788–1794. [[CrossRef](#)]
21. Charytoniuk, W.; Chen, M.S.; Van Olinda, P. Nonparametric regression based short-term load forecasting. *IEEE Trans. Power Syst.* **1998**, *13*, 725–730. [[CrossRef](#)]
22. Amjady, N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Trans. Power Syst.* **2001**, *16*, 498–505. [[CrossRef](#)]
23. Park, D.C.; El-Sharkawi, M.A.; Marks, R.J.; Atlas, L.E.; Damborg, M.J. Electric load forecasting using an artificial neural network. *IEEE Trans. Power Syst.* **1991**, *6*, 442–449. [[CrossRef](#)]
24. Kandil, M.S.; El-Debeiky, S.M.; Hasanien, N.E. Long-term load forecasting for fast developing utility using a knowledge-based expert system. *IEEE Trans. Power Syst.* **2002**, *17*, 491–496. [[CrossRef](#)]
25. Mohandes, M. Support vector machines for short-term electrical load forecasting. *Int. J. Energy Res.* **2002**, *26*, 335–345. [[CrossRef](#)]
26. Ayub, N.; Javaid, N.; Mujeeb, S.; Zahid, M.; Khan, W.Z.; Khattak, M.U. Electricity Load Forecasting in Smart Grids Using Support Vector Machine. In Proceedings of the 33rd International Conference on Advanced Information Networking and Applications, Matsue, Japan, 27–29 March 2019; Volume 926, pp. 1–13.
27. Chu, W.; Keerthi, S.S.; Ong, C.J. A general formulation for support vector machines. In Proceedings of the 9th International Conference on Neural Information Processing, Singapore, 18–22 November 2002; Volume 5, pp. 2522–2526.
28. Kumar, V.; Pal, S. A Literature Survey of Load Forecasting Methods and Impact of Different Factors on Load Forecasting. *Int. J. Res. Appl. Sci. Eng. Technol.* **2017**, *5*, 469–472. [[CrossRef](#)]
29. Salkuti, S.R. Short-term electrical load forecasting using radial basis function neural networks considering weather factors. *Electr. Eng.* **2018**, *100*, 1985–1995. [[CrossRef](#)]
30. Aggarwal, S.K.; Saini, L.M.; Kumar, A. Electricity price forecasting in deregulated markets: A review and evaluation. *Int. J. Electr. Power Energy Syst.* **2009**, *31*, 13–22. [[CrossRef](#)]
31. Ahmad, A.; Javaid, N.; Mateen, A.; Awais, M.; Khan, Z.A. Short-term load forecasting in smart grids: An intelligent modular approach. *Energies* **2019**, *12*, 164. [[CrossRef](#)]
32. Wang, K.; Xu, C.; Zhang, Y.; Guo, S.; Zomaya, A.Y. Robust big data analytics for electricity price forecasting in the smart grid. *IEEE Trans. Big Data* **2017**, *5*, 34–45. [[CrossRef](#)]
33. Mujeeb, S.; Javaid, N.; Ilahi, M.; Wadud, Z.; Ishmanov, F.; Afzal, M.K. Deep long short-term memory: A new price and load forecasting scheme for big data in smart cities. *Sustainability* **2019**, *11*, 987. [[CrossRef](#)]
34. Zahid, M.; Ahmed, F.; Javaid, N.; Abbasi, R.A.; Kazmi, Z.; Syeda, H.; Javaid, A.; Bilal, M.; Akbar, M.; Ilahi, M. Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids. *Electronics* **2019**, *8*, 122. [[CrossRef](#)]
35. Fan, C.; Xiao, F.; Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. *Appl. Energy* **2017**, *195*, 222–233. [[CrossRef](#)]
36. Khan, Z.A.; Zafar, A.; Javaid, S.; Aslam, S.; Rahim, M.H.; Javaid, N. Hybrid meta-heuristic optimization based home energy management system in smart grid. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 4837–4853. [[CrossRef](#)]
37. Moghaddass, R.; Wang, J. A hierarchical framework for smart grid anomaly detection using large-scale smart meter data. *IEEE Trans. Smart Grid* **2017**, *9*, 5820–5830. [[CrossRef](#)]
38. Samuel, O.; Javaid, S.; Javaid, N.; Ahmed, S.H.; Afzal, M.K.; Ishmanov, F. An efficient power scheduling in smart homes using Jaya based optimization with time-of-use and critical peak pricing schemes. *Energies* **2018**, *11*, 3155. [[CrossRef](#)]
39. Ryu, S.; Noh, J.; Kim, H. Deep neural network based demand side short term load forecasting. *Energies* **2017**, *10*, 3. [[CrossRef](#)]
40. Zhao, J.; Dong, Z.; Li, X. Electricity price forecasting with effective feature preprocessing. In Proceedings of the 2006 IEEE Power Engineering Society General Meeting, Montreal, QC, Canada, 18–22 June 2006.
41. Javaid, N.; Ahmed, A.; Iqbal, S.; Ashraf, M. Day ahead real time pricing and critical peak pricing based power scheduling for smart homes with different duty cycles. *Energies* **2018**, *11*, 1464. [[CrossRef](#)]
42. Luo, F.; Ranzi, G.; Wan, C.; Xu, Z.; Dong, Z.Y. A multistage home energy management system with residential photovoltaic penetration. *IEEE Trans. Ind. Inform.* **2018**, *15*, 116–126. [[CrossRef](#)]
43. Khalid, R.; Javaid, N.; Rahim, M.H.; Aslam, S.; Sher, A. Fuzzy energy management controller and scheduler for smart homes. *Sustain. Comput. Inform. Syst.* **2019**, *21*, 103–118. [[CrossRef](#)]

44. Ertugrul, Ö.F. Forecasting electricity load by a novel recurrent extreme learning machines approach. *Int. J. Electr. Power Energy Syst.* **2016**, *78*, 429–435. [[CrossRef](#)]
45. Khan, M.A.; Javaid, N.; Mahmood, A.; Khan, Z.A.; Alrajeh, N. A generic demand-side management model for smart grid. *Int. J. Energy Res.* **2015**, *39*, 954–964. [[CrossRef](#)]
46. Bilalli, B.; Abelló, A.; Aluja-Banet, T.; Wrembel, R. Intelligent assistance for data pre-processing. *Comput. Stand. Interfaces* **2018**, *57*, 101–109. [[CrossRef](#)]
47. Fallah, S.N.; Deo, R.C.; Shojafar, M.; Conti, M.; Shamshirb, S. Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. *Energies* **2018**, *11*, 596. [[CrossRef](#)]
48. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
49. Rojas-Domínguez, A.; Padierna, L.C.; Valadez, J.M.C.; Puga-Soberanes, H.J.; Fraire, H.J. Optimal hyper-parameter tuning of SVM classifiers with application to medical diagnosis. *IEEE Access* **2017**, *6*, 7164–7176. [[CrossRef](#)]
50. Li, Z.L.; Xia, J.; Liu, A.; Li, P. States prediction for solar power and wind speed using BBA-SVM. *IET Renew. Power Gener.* **2019**, *13*, 1115–1122. [[CrossRef](#)]
51. Morley, S.K.; Brito, T.V.; Welling, D.T. Measures of model performance based on the log accuracy ratio. *Space Weather* **2018**, *16*, 69–88. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).