Load Forecasting for Demand Side Management in Smart Grid using Non-Linear Machine Learning Technique

C. Bharathi[1] and D. Rekha[2*]

[1]Research Scholar, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India. bharathi.c2013@vit.ac.in, Orcid: https://orcid.org/0000-0002-0511-4221

[2*]Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India. rekha.d@vit.ac.in, Orcid: https://orcid.org/0000-0003-3227-9168

Received: December 26, 2022; Accepted: January 27, 2023; Published: March 30, 2023

Abstract

In recent times, as a result of the significant growth in population, the usage of energy consumptions is increasing rapidly. Analysis and design of the forecasting model for energy consumptions are challenging task because it is a complex and non-linear problem. Several methods have been proposed for forecasting the energy consumptions in smart grid, the non-linear relationship between the factors is not addressed. Therefore, there is a need for efficient, reliable and accurate forecasting methods to handle non-linearity for effective planning and management of energy consumption. In this paper, a novel hierarchical non-linear machine learning technique Multivariate Adaptive Regression Splines with Genetic Algorithm is proposed to manage the energy consumption demand in smart grid. Experimental results show that the proposed hierarchical approach is much more accurate for the forecasting of energy consumption in smart grid than other approaches such as Auto-Regressive Integrated Moving Average, Complex Neural Network and a Simple Regression Models. The evaluation of forecast accuracy measurement gives the least error value based on the performance metrics of the Mean Absolute Percentage Error and Root Mean Square Error are 0.0651 and 201.2381 respectively.

Keywords: Smart Grid, Load Forecasting, Non-Linearity, Multivariate Adaptive Regression Splines, Genetic Algorithm.

1 Introduction

Recently, one of the most substantial applications of Internet of Things (IoT) is contemplated as Smart Grid (SG). SG consists of computer, automation, control and advanced digital system which ensures a duplex communication (two-way communication) between the providers and consumers. Demand Side Management (DSM) is a technique which uses the utility of the users and electricity cost to develop the power efficiency and stability by scheduling flexible loads based on the demand. Reliability of the SG will be enhanced by managing and forecasting the power demand to attain load balancing and it is an important phase in the smart grid context along with IoT. For an effective DSM, it is proposed that envisioned smart grid will use dynamic pricing-based approach and Load Forecasting (LF). LF has a significant role in power system by scheduling, control and operation. In terms of time, LF can be

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 1 (March), pp. 200-214. DOI: *10.58346/JOWUA.2023.I1.016*

[*]Corresponding author: Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

classified as follows: long-term (more than a year), medium-term (a few weeks, perhaps a year), or short-term (several hours). Here we focus on medium- and long-term forecasting.

Linear and non-linear techniques play a pivotal role in the process of synchronization and reliability in SG. A non-linear system is the system in that the output is inversely proportional to the input. Consequently, the precision will also be higher in the non-linear method (Senthil kumar et al., 2018). The following models are used for forecasting load which is based on the linear method: Exponential smoothing, Auto-Regressive Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), Multiple Linear Regression (MLR) and Semi-Parametric Model (SPM).

The following technique are used for the non-linear systems: Artificial Intelligence (AI) methods, exact feedback linearization, Non-Linear Programming (NLP) and Linear Programming (LP). AI methods include Genetic Algorithm (GA), Artificial Neural Network (ANN), Ant Colony (AC), Fuzzy Logic (FL), Differential Evolution (DE) and Particle Swarm Optimization (PSO).

ARIMA is one of the models for predicting or forecasting the data based on the time series and it contains more error. In the sequel, SG will be immensely complicated and non-linear power network system. It is identified that algorithm which consider the non-linearity with interaction between the independent and dependent variables are not applied to forecast the energy consumptions. Multivariate Adaptive Regression Splines (MARS) is a non-linear regression medialization system that supports to identify the interaction between the variables (Senthilkumar and Paulraj, 2013). MARS produces accurate prediction though in complex conditions where the association between the features and the class variables are non-monotonous and hard to imprecise with the models of parametric (Kumar et al., 2016). GA, a heuristic method for search an error component achieve the same function by reducing the number of best fits. To improve the efficiency of forecasting model in the SG environment a hierarchical of MARS and GA in SG is proposed.

Hierarchical algorithm are used for STLF for smart cities (Elattar et al., 2020). Here, the proposed algorithms Locally Weighted Support Vector Regression optimized by the modified Grasshopper Optimization Algorithm (LWSVR-MGOA) method, these prediction results are correlated with the Local Support Sector Regression (Local SVR) (Elattar et al., 2009), Locally Weighted SVR (LWSVR) (Elattar et al., 2010) and Local SVR optimized by MGOA (Local SVR+MGOA) and Generalized Locally Weighted Group Method of Data Handling (GLWGMDH) (Elattar et al., 2012). Those algorithms are also compared with the existing methods (Elattar et al., 2020) such as Genetic Algorithm (GA), Flower Pollination Algorithm (FPA), Particle Swarm Optimization (PSO), Salp Swarm Algorithm (SSA), Firefly Algorithm (FA), Modified Firefly Algorithm (MFA), States Of Matter Search (SMS), Bat Algorithm (BA), AFSA Aided by Ocean Current Power (AFSAOCP), Artificial Fish Swarm Algorithm (AFSA), Improved AFSA (IAFSA) and the conventional Grasshopper Optimization Algorithm (GOA). Here MARS algorithm is not applied because it is mainly for non – linearity problems. Hence it is identified that the proposed approach is the first and foremost approach, deals with MARS algorithm for forecasting load in smart grid environment.

The remaining of this article proceeds as follows: Related works are discussed in Section 2. The elaboration of the proposed methodology is in Section 3. Experimental setup and evaluation criteria are described in Section 4. Experimental results and discussions can be found in Section 5. A brief conclusion is delivered in Section 6.

2 Related Works

The foremost objectives of SG are given as follows: efficiently utilizing the power system, two-way information flow between consumer and utility, reduce the greenhouse gas emission, minimal the production cost, energy production optimization, integration of renewable energy resources (RES), Battery Electric Vehicle (BEV) and Plug-in Hierarchical Electric Vehicle (PHEV) utilization then implementation of real time billing and pricing. Smart cities deal with multiple hurdles to achieve the load demand without increasing the production volume (Ahmad et al., 2018). The power consumption of the users can be modeled to generate a preferred configuration of DSM. The safety operation of the power grid and efficiency improvement is based on the lower power fluctuation. Using DSM, load balancing can be attained by flattening the power consumption (Bharathi et al., 2019). With the population growth and a global expansion of energy framework, the demand for electricity increases rapidly. To effectively handle this growing demand, smart grids are being utilized. The fundamental component of DSM in SG is forecasting the load, as it permits smart grid operators to produce on effective and efficient decisions (Fallah et al., 2018).

Power cannot be stored in enormous amount. Based on the demand only it will get generated. Hence LF plays a vital role for the utilities which the system load should be prophesied beforehand. For several years in industrial applications, linear system has been functioning. Now for the nonlinear featured system, linear scheme is not effective. The majority of the real-world problems are non-linear. Focusing on non-linear problem in SG is actually less. Few differences between two different systems are error will be high in linear system and error will be very low in use non-linear system (Elattar et al., 2020). Short term load forecasting is a highly complex model as the accuracy of the model is influenced by various parameters like season, economic, time, etc., those relationships are typically nonlinear (Elattar et al., 2010). Typically, linear modelling techniques like and Partial Least Squares (PLS), Multiple Linear Regression (MLR) and Principal Component Regression (PCR) are used (Deconinck et al., 2007). Such information can be quite complex and the technique of linear modelling frequently model only a little bit of the information as the input data. No algorithm exists for load forecasting based on the non-linear system.

Based on the literature, a large amount of traditional and AI algorithms are applied to forecast the energy consumption. The energy forecast model is highly complex and also with the characteristics of nonlinearity. ARIMA is mainly applied for load forecasting but it will not focus non-linearity explicitly. Here, nonlinearity, which cannot be accurately signified by the traditional algorithms (Papalekopulos and Hesterberg, 1990; Taylor et al., 2006; Hippert et al., 2001). To overcome this problem, few AI techniques were used. Artificial Neural Network (ANN) is among the one (Hippert et al., 2001). ANN model will focus non-linear model but explicit rules will switch. Therefore, in this paper, an algorithm hierarchical MARS-GA which supports explicit rule is proposed.

MARS is a challenger to neural networks and it has none of the boundaries of neural networks like nonlinearities, missing data, and interactions (Francis, 2003). MARS models are more flexible, easy to comprehend and infer compared with the neural network (Senthilkumar and Paulraj, 2013). MARS models perform predictions very fast compared with SVM. MARS identifies important predictor variables and does not require more time for modelling even for the high dimensional (Lu et al., 2011). MARS produces better results in various studies compared with artificial neural network (Adamowski et al., 2012). Hierarchical of MARS with other AI techniques will be used for the betterment of the result. Here GA is used as a hierarchical algorithm with MARS. GA works as follows (Bharathi et al., 2019): Step 1: i) It begins with the initial population creation. ii) Evaluate the fitness of each individual.

Step 2: Check the convergence criterion, if it is ok then the algorithm is converged or else follow the next steps. Step 3: i) Select NP (from the preceding population). ii) Random pool is created. iii) Execute crossover and mutation. iv) Replace the last generated solutions with the new solutions. v) Repeat step 2. To advance the precision of the LF model, hierarchical technique is proposed with MARS and GA.

The existing methods (Jyothsna et al., 2021), used the various models Seasonal Autoregressive Integrated Moving Average (SARIMA), Autoregressive Moving Average (ARMA), Simple LSTM and Complex LSTM in ERCOT dataset for linear technique. The models of ARMA are an amalgamation of Auto Regressive (AR) and Moving Average (MA) models, where the actual value of the temporal series is given linearly related with the prior values and also related with the actual and prior residual series. The models of SARIMA models amalgamate recurrent differencing by the model of an ARIMA. Those models are used to model time series data with periodical features.

3 Proposed Method

Multivariate Adaptive Regression Splines (MARS)

In this study, a novel hierarchical MARS with GA for forecasting the energy consumption in smart grid has been developed, as shown in figure 1. The MARS algorithm is a procedure of versatile regression that selects the basic functions to approximate the reply through the selection of backward/forward steps. It deals with multidimensional data, evaluating each factor and possible interaction among them. It eliminates a certain number of predictors if they do not contribute to increasing the performance of the final model. MARS model (Senthil et al., 2008) of the form is given in the equation (1)

$$\boldsymbol{\hat{y}} = \boldsymbol{\hat{f}(x)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \boldsymbol{c_i B_i(x)} \tag{1}$$

The model is a weighted sum of basis functions $\boldsymbol{B_i(x)}$. Each $\boldsymbol{c_i}$ is a constant coefficient. Each basis function $\boldsymbol{B_i(x)}$ takes one of the following three forms: i) a constant 1. There is just one such term, the intercept. ii) a hinge function has the form max(0, $\boldsymbol{x - c}$) or max(0, $\boldsymbol{c - x}$), where c is a constant called the knot. iii) a product of two or more hinge functions. MARS uses the basis function to establish the relationship between the independent and dependent variables. Generalized Cross Validation (GCV) (Craven and Wahba, 1979) is used to identify the most significant predictor. It is used to identify the most significant predictor, rank the predictor and eliminate insignificant predictor of the model (Salford System, 2013). If a variable receives a score 100, it is the most significant predictor and the variable receives a score 0 which is not used in the MARS model (Steinberg et al., 1999). GCV is defined in the equations (2) & (3)

$$\boldsymbol{GCV(M)} = \frac{\frac{1}{N}\sum_{i=1}^{N}[y_i - \hat{f}_M(x)]^2}{[1 - \frac{C(M)^*}{N}]^2} \tag{2}$$

$$\boldsymbol{C(M)}^* = \boldsymbol{C(M) + \delta M} \tag{3}$$

N is the no. of observations. *C(M)** is a complexity cost function of the model generating *f*, the default is to set equal to a function of the effective number of parameters. *M is* the number of non-constant basis functions in the MARS model and $\boldsymbol{\delta}$ is a cost for each basis-function optimization.

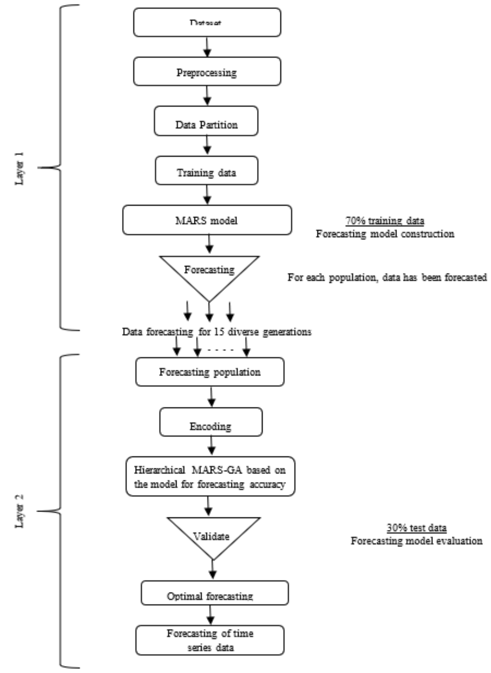

Figure 1: Process of Hierarchical MARS with GA

GA to Measure Forecast Accuracy

AIC (Banks and Joyner, 2017) is a sophisticated approach which is based on the fit of in-sample to evaluate the probability of a technique to forecast the values of future. The one with minimal AIC value than others will considered to be a good model. In the practice of standard machine learning technique (time series or lesser data), if one cannot simply evaluate the performance of the model using test dataset then AIC is often used. The working procedure of AIC is by estimating the model's fit on the training dataset and including a consequence term to the model's complexity. AIC can be used when the same data are using between models, when the same outcome variable are measured between the models and when an infinite size of sample. The objective function is used by the GA to give a best solution and is given in the following equation of (4)

$$\text{Minimize} \quad P = \exp\left(\left(AIC_{min} - AIC_i\right)/2\right) \tag{4}$$

where exp = "e" to the power of parenthesis and AIC_{min} = lowermost score of AIC in the scores series. From the formula, for any given AIC_i, the probability can be calculated in which the model of "i" diminishes the loss of data.

Stage 1: Initial Population

Randomly selects the initial population. This parameter plays major role in increasing the GA performance. The population size refers to the amount of chromosomes per generations. There is no norm to define the size.

Stage 2: Generation and Selection

In evolutionary algorithms to attain the best chromosomes, selection process is vitally important. As stated in the working principle of GA, "the fittest individuals have a greater chance of survival than weaker ones" chromosomes are chosen by the selection operator from the mating pool. The probabilistic selection is done with the help of Roulette wheel, primarily for evaluating the pedicted expected data using the fitness function.

Stage 3: Fitness Function

Intention of this work is to do a forecast based on the MARS model as given in the equation (1) with least error. As a matter of fact, we stab to diminish the discrepancies from actual data. The fitness function relies upon the MARS model for the prediction of individual time series load data objects, as shown in the following equation of Akaike information criterion (AIC). In general, the best stability of the model fit is designated by the preferred result to discover the lowermost probable AIC. The equation (Banks and Joyner, 2017) of fitness function is given as follows in equation (5)

$$AIC = -2\ln(L)+2k \tag{5}$$

where, L = Likelihood value, k = No. of estimated parameters. For a model's measure of fit, maximum log – likelihood is used. The data fits the best by a method with maximum likelihood.

Stage 4: Crossover and Mutation

Here, this study uses the probability of fixed crossover and one-point crossover. The preconception of outcomes are reduced by this probability across various groups due to the enormous values of data. Object values are interconnected and need to be swapped to generate two new descendants. Two points are chosen to build more value for the best fit. Arbitrarily, a modest amount of chromosomes are mutated when new chromosomes are arrived. Two opposite data values are swapped for the load object values. The motivation behind this low percentage of mutation is keeping the forecast changes stable across various generations.

Stage 5: Next Generation

Repeat steps 2 – 4 constantly for 5 generations to attain the next generations. Since the fitness functions repeat themselves after 5 generations, those 5 generations are sufficient for this data. In second level, for each object, the validation of the prediction accuracy is done by using the selected generations. This phase provides completely correlated data for covering the forecasts for some months.

4 Experimental Setup and Evaluation Criteria

This section illustrates the experimental setup, dataset description and evaluation criteria of proposed algorithm of hierarchical MARS-GA.

Experimental Setup

This research mainly focuses on the design of decision support system for predicting the energy consumption to manage the load efficiently in the SG environment. In this experiment, to evaluate the effectiveness of the proposed hierarchical MARS-GA, ERCOT dataset is used. Furthermore, the proposed hierarchical MARS-GA is compared with various regression models like MLR, PCR, PLS, Elastic Net and MARS.

In machine learning, dataset is divided into two parts such as training dataset and test dataset for prediction model. Model will be constructed using the given training dataset and it will be evaluated using the test dataset. 70% of data will be used as a training dataset. Using the training dataset prediction model is constructed and the constructed model is evaluated using 30 % of the test dataset and also by using various performance metrics. If the accuracy of the prediction model is acceptable, use the prediction model to forecast the real time data otherwise fine tune the model parameter to improve the accuracy of the prediction model.

Dataset Description

In this study, the energy consumption ERCOT power grid market data prepared by the Central Operations Coordinator for Texas is selected to evaluate the proposed method hierarchical MARS-GA. ERCOT be operations with prior statistics data with the footprint between 2006 and 2010. ERCOT offers a unique framework for the impact of price assessment of incorporate recurrent renewable energy into the major electricity grid for a variety of reasons. Predominant intention is because amongst all the organized wholesale power market and it then became a nodal market in the United States (U.S.) in 2010. It controls the power flow of 24 million Texas users' covers almost 90% of the state's power load (MA et al., 2018).

The features enlisted in this ERCOT dataset includes time stamp such as Time in minutes, Day, Month and Year and four different parameters that influence the energy consumption namely temperature of the dry bulb, temperature of the wet bulb, air saturation temperature, and relative humidity, electricity price. The data was collected for every half an hour of a day (Date). Dry bulb (DryBulb) is the temperature of the dry bulb which is obtained from the air but not subjected to the humidity or the radiation of the solar. Wet bulb (WetBulb) is the temperature of the wet bulb which is obtained using a thermometer where the measurement device of the bulb is moistened by a damp rag. Dew point (Dewpnt) corresponds to the air saturation temperature. When the temperature falls with a stable quantity of water, the relative humidity (Humidity) rises. The amount of the electricity (ElecPrice) and the power load consumption (SYSLoad) (Johannesen et al., 2019).

Evaluation Criteria

The effectiveness of the proposed hierarchical MARS-GA is evaluated using the performance metrics Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). MAPE is one of the best frequently used Key Performance Indicators for measuring the accuracy of the forecast. MAPE is the sum of errors divided by the demand (each period individually). MAPE equation is represented in (6)

$$MAPE = \frac{1}{n}\sum\frac{|e_t|}{d_t} \tag{6}$$

where *n* is the no. of periods where we have both a forecast and a demand, e_t is the error value and d_t is the demand value. RMSE is an eccentric Key Performance Indicators but really useful one. RMSE is stated as the square root of the average squared error. RMSE equation is represented in the equation (7)

$$RMSE = \frac{\sqrt{\frac{1}{n}\sum e_t^2}}{\frac{\sum d}{n}} \tag{7}$$

where n is the no. of periods where we have both a forecast and a demand and e_t is the error value and d is the demand. MAPE is easily interpreted. RMSE deals with the values of large error. It is quite easy to distinguish. Errors are squared means that significantly more weight is attributed to the more significant errors. Therefore, an error of 10, is hundred times worse than an error of 1. These are the main reasons for using these metrics MAPE and RMSE in this non-linear regression model.

5 Experimental Results and Discussions

This section illustrates the results and discussions of the proposed hierarchical MARS-GA and also comparison between existing and proposed methods. In this study, experiment were conducted to forecast the energy consumption in two different dimension: First model focus on the time series data and the second forecasting model is constructed using the parameters that influence the energy consumption such as temperature of the dry bulb, temperature of the wet bulb, air saturation temperature, and relative humidity, electricity price. MARS algorithm sample result for the time series forecasting model and second forecasting model with various parameters are presented in table 1 and 2 respectively.

Table 1: First Model Sample Results

Basis functions	Coefficients
(Intercept)	1652.59543
h(Month-4)	654.92280
h(7-Month)	235.02216
h(Month-7)	-1200.25453
h(Month-9)	605.29027
h(Minutes-270)	20.45121
h(Minutes-510)	-15.88679
h(750-Minutes)	8.75819
h(Minutes-750)	-2.39818
h(Minutes-1170)	-6.85962
h(Month-4) * h(Minutes-510)	0.41102
h(Month-4) * h(510-Minutes)	-0.38584
h(7-Month) * h(Minutes-540)	1.25042
h(7-Month) * h(540-Minutes)	-0.61265
h(7-Month) * h(Minutes-750)	-1.88383
h(Month-7) * h(Minutes-750)	-1.24680

From table 1 it is identified that the MARS algorithm considers only the Minutes and Month as significant predictor for forecasting the energy consumption and it eliminates the other two-time series factors day year. It indicates that energy consumption trend is based on month only. Also, the results indicate that there is a relationship between the usage of energy in minutes and month. In particular, MARS algorithm creates various basis function for month and minutes based on the trend. From the table 2, the experimental result indicates that Humidity, Electricity Price, temperature of Wet Bulb and Dry Bulb plays a significant role in the trend of energy consumption. Also, it is identified that there is a non-linear relationship between DryBulb and ElecPrice; WetBulb and Humidity; and WetBulb with ElecPrice.

Table 2: Second Model Sample Results

Basis functions	Coefficients
(Intercept)	19025.2345
h(16.35-WetBulb)	121.8087
h(WetBulb-16.35)	250.4510
h(ElecPrice-18.11)	133.3147
h(ElecPrice-26.45)	-181.2719
h(219.16-ElecPrice)	-58.1308
h(ElecPrice-219.16)	47.9546
h(20.2-DryBulb) * h(219.16-ElecPrice)	0.4305
h(DryBulb-20.2) * h(219.16-ElecPrice)	0.2087
h(WetBulb-16.35) * h(Humidity-67.5)	-5.1416
h(WetBulb-16.35) * h(67.5-Humidity)	-1.6538
h(10.25-WetBulb) * h(219.16-ElecPrice)	-1.0744
h(16.35-WetBulb) * h(ElecPrice-42.56)	0.0189
h(16.35-WetBulb) * h(42.56-ElecPrice)	-5.7296
h(64.5-Humidity) * h(219.16-ElecPrice)	0.0277
h(Humidity-64.5) * h(219.16-ElecPrice)	-0.1453

In general, the experimental results of MARS algorithm explain that the rules generated by MARS algorithm are explicit and easily understandable by the decision maker compared with the other algorithms available in the literature. Also, Month, Humidity, Electricity Price, temperature of Wet Bulb and Dry Bulb are the most significant factor to design the decision support system to forecast the energy consumption in the smart grid environment. Furthermore, the advantage of MARS algorithm is: it creates various basis function for every parameter based on the trend and also it identifies non-linear relationship between the features (input parameters) and the target variables.

ERCOT dataset is divided into two parts such as training dataset and testing dataset. The algorithms MLR, PCR, PLS, Elastic Net and MARS are applied on training. The error rates of various regression models such as MLR, PCR, PLS, Elastic Net and MARS based on the performance metric MAPE are 0.1006, 0.1013, 0.1006, 0.1006 and 0.5765 respectively. Likewise, the error rate based on the performance metric RMSE are given as follows: MLR: 1221.7645, PCR: 1230.5759, PLS: 1221.7644, Elastic Net: 1221.7535, MARS: 761.2377.

To validate the trained model the same algorithms are applied on test data set. Furthermore, GA is incorporated into the MARS algorithm to improve the performance. The actual and forecasted load which is forecasted on March10, 2009 for each and every hour on a particular day by comparing regression models such as MLR, PCR, PLS, Elastic Net, MARS with Hierarchical MARS – GA are presented in the figures 2a, 2b, 2c, 2d and 2e respectively. They are compared with the actual load which is represented in black color. By comparing all the predicted loads with the actual load, evidently discerned that Hierarchical MARS – GA, which is represented in black color, is highly correlated with the actual load.

To verify the effectiveness of the proposed hierarchical MARS-GA is evaluated with three different performance metrics MAPE and RMSE. The comparison of hierarchical MARS-GA with the regression models MLR, PCR, PLS, Elastic Net and MARS are depicted in Table 3. From Table 3, the hierarchical MARS-GA achieved lowest error rate compared with the other five models. It is clearly depicted in the given figures 3a and 3b. The value of hierarchical MARS-GA is lesser than the other five models. For the performance metric of MAPE the error rate of MLR is 0.1165, PCR is 0.1173, PLS is 0.1162, Elastic Net is 0.1169, MARS is 0.1193 and hierarchical MA RS–GA is 0.0651 which is the lowest error rate among all. Likewise, RMSE performance metric MLR error rate value is 1221.7744, PCR is 1228.7745, PLS is 1221.7723, Elastic Net is 1221.8811, MARS is 851.8348 and hierarchical MARS – GA is 201.2381 which consists of lower error rate than the other models. When compared with the training dataset, testing dataset are more likely to be lowest error rate.

Figure 4 illustrates the comparison of hierarchical MARS-GA with the existing methods [19] such as ARMA, SARIMA, Simple LSTM and Complex LSTM for the whole month of February, 2009 and they are represented in the figures 4a, 4b, 4c and 4d, in which it is visibly noted that hierarchical MARS-GA forecasted load (black color) is extremely correlated with the actual load (red color). Furthermore, the hierarchical MARS-GA method is compared with the results of ARMA, SARIMA, Simple LSTM and Complex LSTM and it is presented in Table 4.

Table 3: Comparison of Proposed Hierarchical MARS-GA with Various Regression Models

MODEL	**MLR**	**PCR**	**PLS**	**Elastic Net**	**MARS**	**Hierarchical MARS-GA**
MAPE	0.1165	0.1173	0.1162	0.1169	0.1133	0.0651
RMSE	1221.7744	1228.7745	1221.7723	1221.8811	851.8348	201.2381

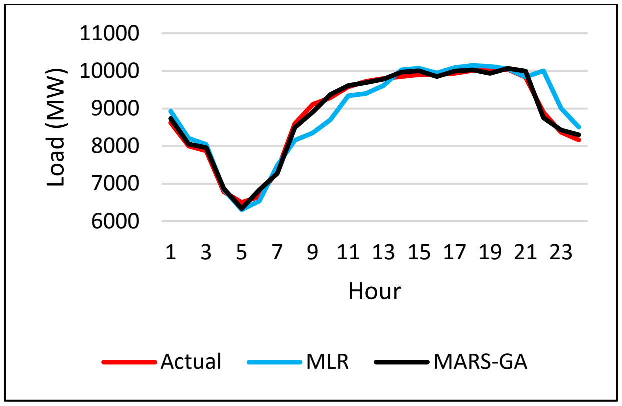

Figure 2a: Actual and Forecasted Load using MLR and MARS-GA (Hourly) on March 10, 2009

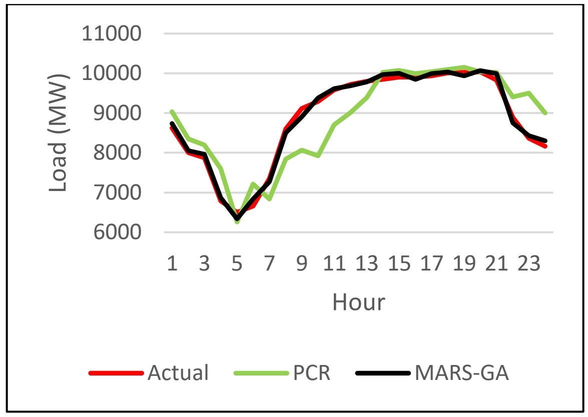

Figure 2b: Actual and Forecasted Load using PCR and MARS-GA (Hourly) on March 10, 2009

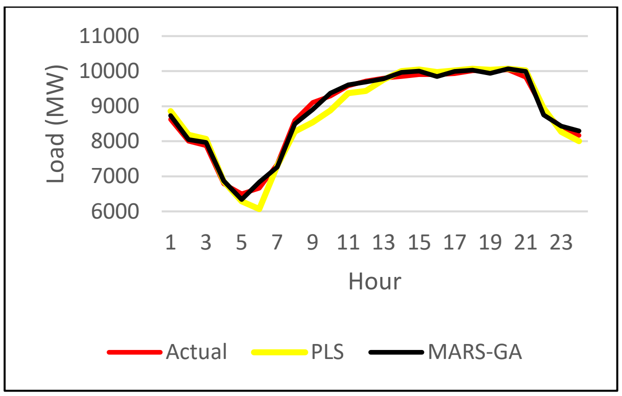

Figure 2c: Actual and Forecasted Load using PLS and MARS-GA (Hourly) on March 10, 2009

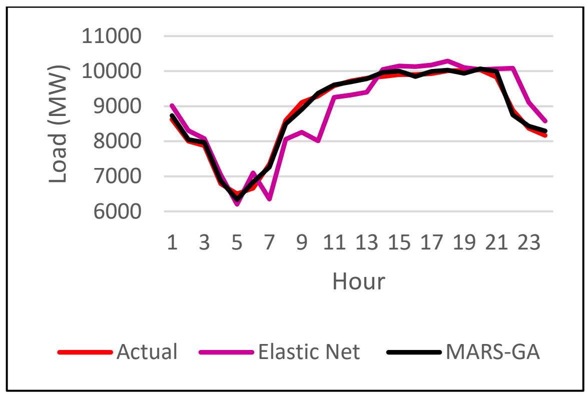

Figure 2d: Actual and Forecasted Load using Elastic Net and MARS-GA (Hourly) on March 10, 2009

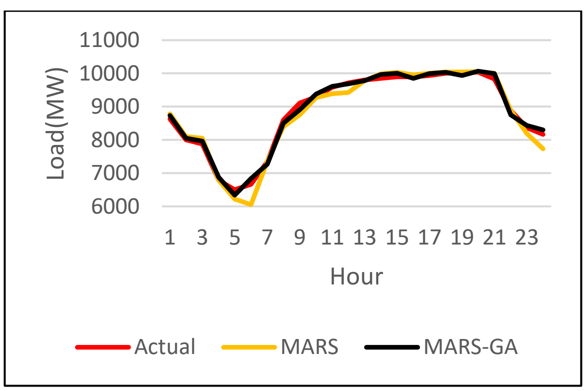

Figure 2e: Actual and Forecasted Load using MARS and MARS-GA (Hourly) on March 10, 2009

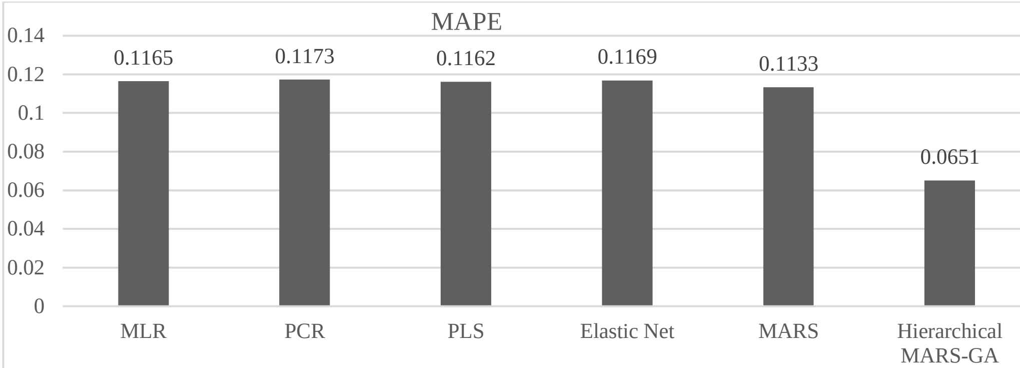

Figure 3a: Error Rate Using MAPE

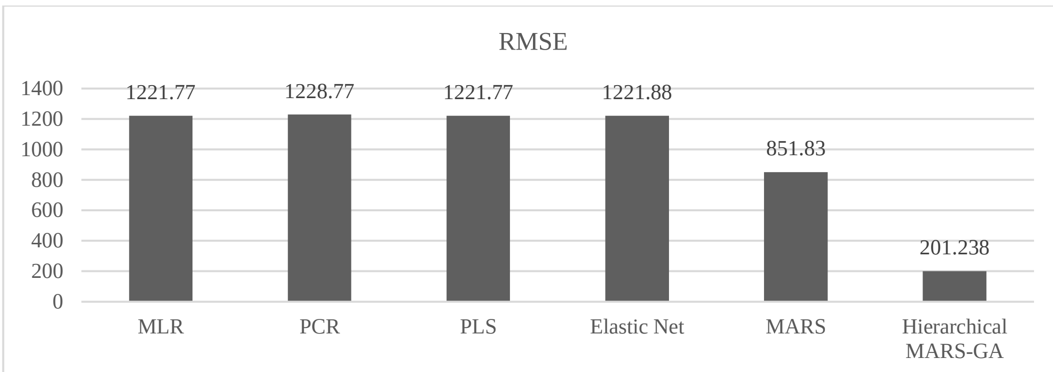

Figure 3b: Error Rate Using RMSE

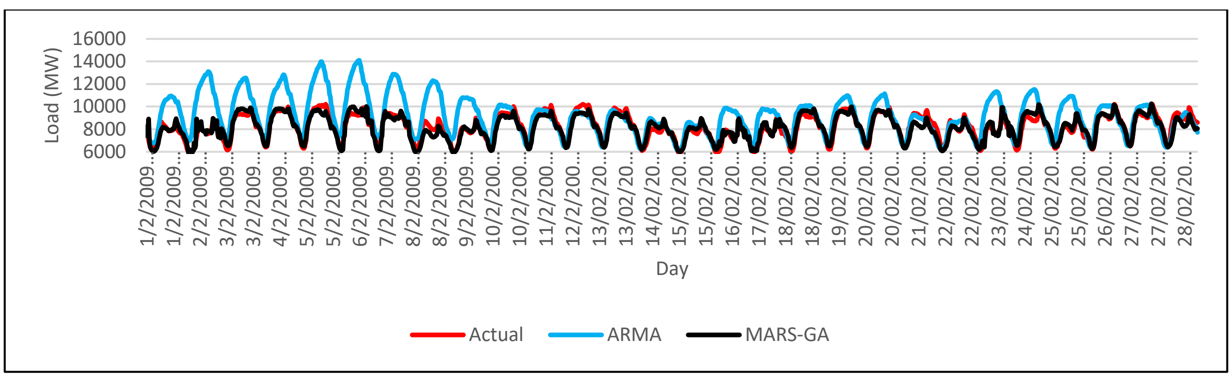

Figure 4a: Comparison of Hierarchical MARS-GA with ARMA

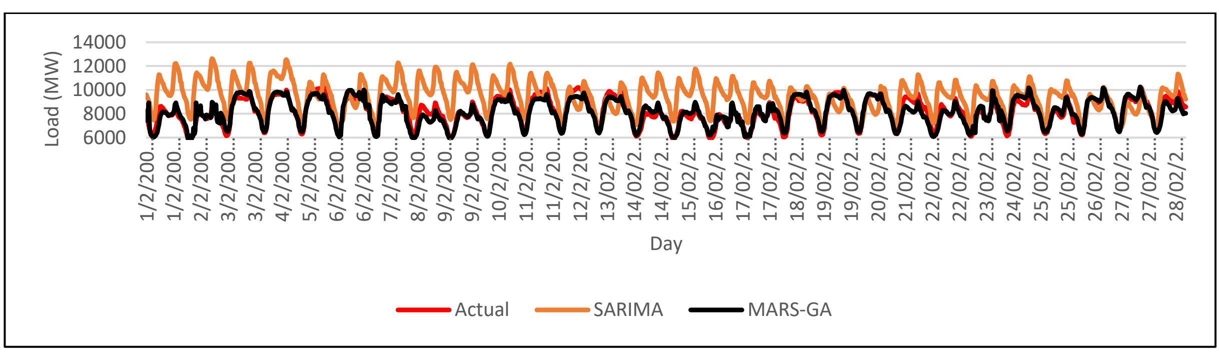

Figure 4b: Comparison of Hierarchical MARS-GA with SARIMA

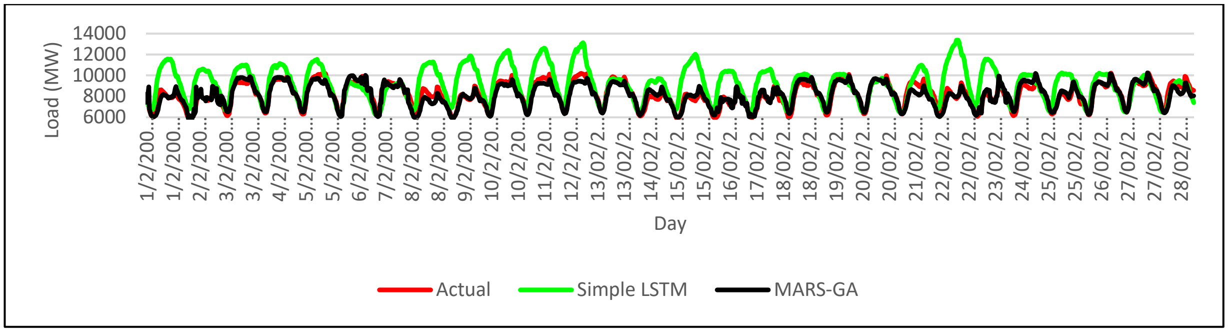

Figure 4c: Comparison of Hierarchical MARS-GA with Simple LSTM

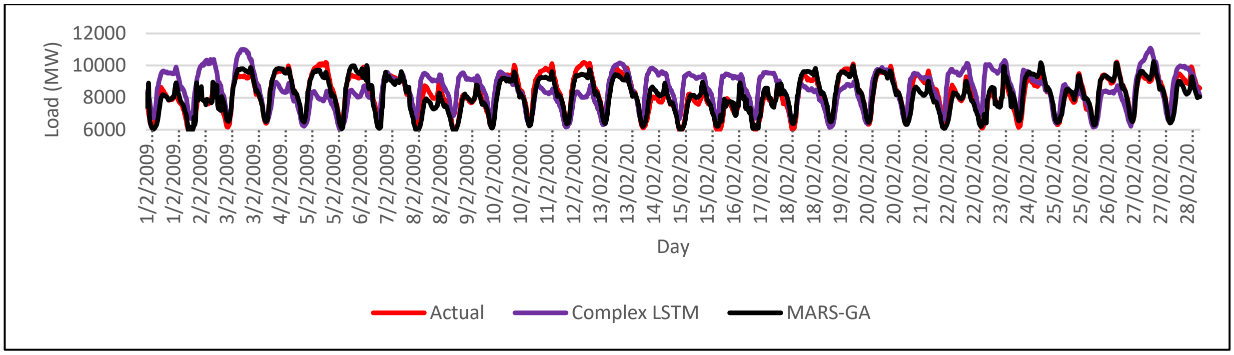

Figure 4d: Comparison of Hierarchical MARS-GA with Complex LSTM

Table 4: Comparison of Hierarchical MARS-GA with (Jyothsna et al., 2021)

	MODEL	MAPE	RMSE
EXISTING METHOD	ARMA	9.13	3451
	SARIMA	4.36	1638
	Simple LSTM	2.638	716.534
	Complex LSTM	1.664	229.63
PROPOSED METHOD	Hierarchical MARS-GA	0.065	201.238

Here, based on the MAPE performance metric, the error rate of the existing models ARMA, SARIMA, Simple LSTM and Complex LSTM are 9.13, 4.36, 2.368, 1.664 respectively and the proposed model hierarchical MARS – GA is 0.065 which is comparatively lower than the existing method and showed in figure 5a. Based on the RMSE performance metric, the error rate of the existing models ARMA, SARIMA, Simple LSTM and Complex LSTM are 3451, 1638, 716.534, 229.63 respetively and the proposed model hierarchical MARS – GA is 201.2 respectively and the proposed model hierarchical MARS – GA is 201.24 which is moderatelyzlower than the existing models which is clearly represented in figure 5b.

According to the empirical results, it is concluded that the proposed hierarchical MARS-GA outperforms existing techniques ARMA, SARIMA, Simple LSTM and Complex LSTM. The proposed method performs better than the existing algorithm due to the following advantages. It extracts a relevant subset of features from a large number of features to reduce the computational complexity. It is an efficient algorithm to deal with complicated and large data. It supports to identify the complex nonlinear relationship between the independent variable and dependent variable. It rules are easy to understand and faster compared with the NN. Its accuracy is higher when compared with the existing methods. Also, the developed GA quickly converges to the optimum solution easily due to their exploitation and exploration features.

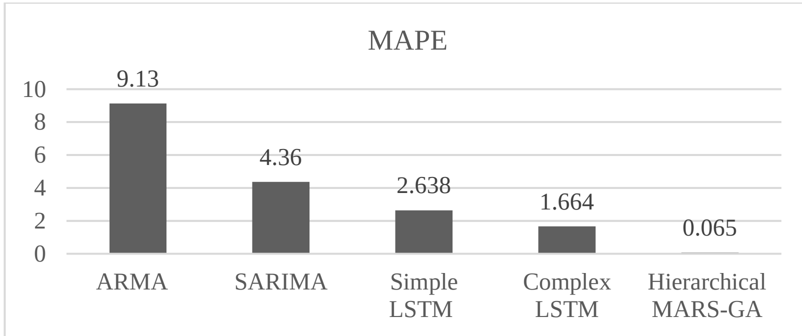

Figure 5a: Error Rate Using MAPE

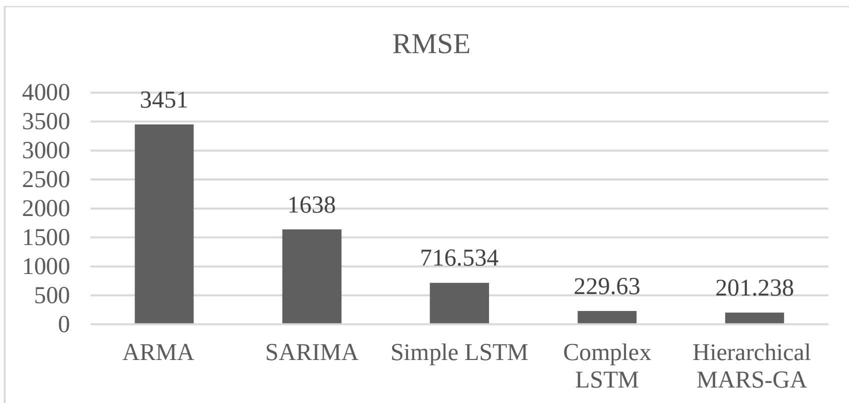

Figure 5b: Error Rate Using RMSE

6 Conclusion

In this article, a novel hierarchical non – linear machine learning technique is proposed, in order to forecast the energy consumptions, in which the model relies on MARS and GA. Considering the availability of all the energy resources, underestimate or overestimate of energy consumptions become vital. The decisions of investment on energy and the activities of planning the energy resource are utterly relied on the forecast. In this condition, conventional models alone can be inadequate for the accuracy of the forecast. So, we trained our proposed model with the various regression methods such as MLR, PCR, PLS, Elastic Net, MARS and hierarchical MARS – GA. From the result analysis, the evaluation of forecast accuracy measurement gives the least error value based on the performance metrics of the MAPE and RMSE are 0.0651 and 201.2381 respectively. From this we can conclude that the hierarchical MARS – GA has the potential to enhance the forecasting of data with a higher accuracy value when compared with the other conventional models and it performs well.

References

[1] Adamowski, J., Chan, H.F., Prasher, S.O., & Sharda, V.N. (2012). Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of hydroinformatics*, *14*(3), 731-744.

[2] Ahmad, A., Javaid, N., Mateen, A., Awais, M., & Khan, Z. A. (2019). Short-term load forecasting in smart grids: An intelligent modular approach. *Energies*, *12*(1), 1-21.

[3] Al-Douri, Y.K., Al-Chalabi, H., & Lundberg, J. (2018). Time Series forecasting using genetic algorithm. In *the Twelfth International Conference on Advanced Engineering Computing and Applications in Sciences*.

[4] Banks, H.T., & Joyner, M.L. (2017). AIC under the framework of least squares estimation. *Applied Mathematics Letters*, *74*, 33-45.

[5] Bharathi, C., Rekha, D., & Vijayakumar, V. (2017). Genetic algorithm-based demand side management for smart grid. *Wireless personal communications*, *93*, 481-502.

[6] Peter, C., & Grace, W. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik*, *31*(4), 377-403.

[7] Deconinck, E., Coomans, D., & Vander Heyden, Y. (2007). Exploration of linear modelling techniques and their combination with multivariate adaptive regression splines to predict gastro-intestinal absorption of drugs. *Journal of pharmaceutical and biomedical analysis*, *43*(1), 119-130.

[8] Elattar, E.E., Goulermas, J., & Wu, Q.H. (2010). Electric load forecasting based on locally weighted support vector regression. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(4), 438-447.

[9] Elattar, E.E., Goulermas, J.Y., & Wu, Q.H. (2011). Generalized locally weighted GMDH for short term load forecasting. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(3), 345-356.

[10] El-Attar, E.E., Goulermas, J.Y., & Wu, Q.H. (2009). Forecasting electric daily peak load based on local prediction. *In IEEE Power & Energy Society General Meeting*, 1-6.

[11] Elattar, E.E., Goulermas, J.Y., & Wu, Q.H. (2010). Integrating KPCA and locally weighted support vector regression for short-term load forecasting. *In Melecon 15th IEEE Mediterranean Electrotechnical Conference*, 1528-1533.

[12] Elattar, E.E., Sabiha, N.A., Alsharef, M., Metwaly, M.K., Abd-Elhady, A.M., & Taha, I.B. (2020). Short term electric load forecasting using hybrid algorithm for smart cities. *Applied Intelligence*, *50*, 3379-3399.

[13] Fallah, S.N., Deo, R.C., Shojafar, M., Conti, M., & Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions. *Energies*, *11*(3), 1-31.

[14] Francis, L. (2003). Martian chronicles: is MARS better than neural networks? In *Casualty Actuarial Society Forum*, 75-102.

[15] Hippert, H.S., Pedreira, C.E., & Souza, R.C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, *16*(1), 44-55.

[16] Johannesen, N.J., Kolhe, M., & Goodwin, M. (2019). Relative evaluation of regression tools for urban area electrical energy demand forecasting. *Journal of cleaner production*, *218*, 555-564.

[17] Jyothsna, D., Kumar, T.S., & Prasad, T.S. (2021). Analysis and Prediction of Electricity Load Forecasting. *Annals of the Romanian Society for Cell Biology*, *25*(6), 2134-2143.

[18] Kumar, D.S., Sukanya, S., & BIT-Campus, T. (2016). Feature selection using multivariate adaptive regression splines. *International Journal of Research and Reviews in Applied Sciences and Engineering (IJRRASE)*, *8*(1), 17-24.

[19] Lu, C.J., Wu, J.Y., Lee, T.S., & Lian, C.M. (2011). Incorporating feature selection method into neural network techniques in sales forecasting of computer products. *In Advances in Neural Networks–ISNN 2011: 8th International Symposium on Neural Networks, ISNN 2011, Guilin, China, Proceedings, Part III 8*, 246-255. Springer Berlin Heidelberg.

[20] Ma, Z., Zhong, H., Xie, L., Xia, Q., & Kang, C. (2018). Month ahead average daily electricity price profile forecasting based on a hybrid nonlinear regression and SVM model: an ERCOT case study. *Journal of Modern Power Systems and Clean Energy*, *6*(2), 281-291.

[21] Papalexopoulos, A.D., & Hesterberg, T.C. (1990). A regression-based approach to short-term system load forecasting. *IEEE Transactions on power systems*, *5*(4), 1535-1547.

[22] Leybourne, S.J., Mills, T.C., & Newbold, P. (1998). Spurious rejections by Dickey–Fuller tests in the presence of a break under the null. *Journal of Econometrics*, *87*(1), 191-203.
[23] Salford System, (2013). Nonlinear Regression: Modern Approaches and Applications, 1-4.
[24] Senthil Kumar, D., & Rhymend Uthariaraj, V. (2008). A Decision Support System for Predicting Academic Performance of Candidates in Engineering Admissions using MARS. *International Journal of Learning*, *15*(3).
[25] Senthilkumar, D., & Paulraj, S. (2013). Diabetes disease diagnosis using multivariate adaptive regression splines. *AGE*, *768*, *5*(5), 3922-3929.
[26] Senthilkumar, D., K Reshmy, A., & G Kavitha, M. (2018). Non-Linear Machine Learning Techniques for Multi-Label Image Data Classification. *Applied Mathematics & Information Sciences*, *12*(6), 1139-1145.
[27] Steinberg, D., Colla, P.L., & Martin, K. (1999). MARS user guide. *San Diego, CA: Salford Systems*.
[28] Taylor, J.W., De Menezes, L.M., & McSharry, P.E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International journal of forecasting*, *22*(1), 1-16.
[29] Talegaon, S., & Krishnan, R. (2020). Administrative models for role-based access control in android. *Journal of internet services and information security*, *10*(3), 31-46.

Authors Biography

C. Bharathi is currently pursuing her Ph.D. degree from Vellore Institute of Technology, Chennai. She received her B.E. degree in Computer Science and Engineering from Anna University, Trichy, in 2011 and her MTech degree specializing in Multimedia Technologies from Anna University, Coimbatore, in 2013. From 2009 to 2013, she worked on various projects in cloud computing and network security. She joined as a full-time research scholar at the Vellore Institute of Technology, Chennai, and worked as a research associate from 2013–2017. She worked as an assistant professor at Saveetha University for a year from 2017 to 2018. She has also published papers in notable journals during this time.

Dr. D. Rekha is currently working as an Associate Professor in Vellore Institute of Technology, Chennai. She has more than 17 years of experience which includes 13 years in teaching and 4 years in Research. Her area of research includes Wireless Sensor Networks, Multihop Networks, Evolutionary Algorithms, Robotics, Artificial Intelligence, Mobile Cloud Computing and Internet of things. She published many papers in reputed journals. She served as General and Technical Program Chair of numerous conferences.