International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 15 (2015) pp 35027-35038

Intelligent Classifier model employing Hybrid ELMAN Neural Network Architecture and Biogeography Based Optimization for Data Classification

N. Mohana Sundaram
Professor
Department of Computer Science and Engineering,
Karpagam Institute of Technology
Coimbatore
itismemohan@gmail.com

S. N. Sivanandam
Professor Emeritus,
Department of Computer Science and Engineering
Karpagam College of Engineering,
Coimbatore
sns12.kit@gmail.com

Abstract- Classification plays a major role in the data mining approach and it aims to build a classifier model to describe and identify the important data classes and provides a better understanding of data at large. In this paper, a novel intelligent classifier model is designed employing ELMAN neural network architecture hybridized with biogeography based optimization technique. ELMAN neural model being recurrent neural network architecture explores the classification process with its faster convergence and biogeography based optimization (BBO) is designed to tune the weights incorporated into the ELMAN neural model. These optimal weights determined employing BBO approach facilitates the ELMAN network to perform effective classification avoiding the trap from local and global optima. The proposed hybrid ELMAN – BBO classifier model is tested and validated in this paper by implementing it for various test benchmark datasets from UCI machine learning repository. Simulation results prove the efficiency of the proposed classifier model performing better than that of the earlier classifier models designed and available in the literature.

Keywords: Elman Neural networks, Classifiers, Bio geography based optimization, Hybrid Elman-BBO Classifier.

Introduction

The growing trends in information technology sector proves the advancements in handling large volume of data and the methodology adopted for collecting the data. The data belonging to large sectors like hospitals, universities, weather details over cumulative period of time, population statistics, income tax details, power plant details, stock market variations and so on are to be maintained in large storage media and to be retrieved as and when required for effective utilization. At this juncture, data classification attempts to build classifier models to describe the important data classes and provide a better understanding of the considered data at large. Several engineering and science domain employ data classification approaches for gene selection, pattern recognition and association, image processing, medical diagnosis, computer networks, bio – informatics and so on.

Several statistical and machine learning approaches were widely used in the past decades for carrying out the data classification process. Few of the statistical approaches include – naïve Bayesian classifier model, hidden Markov classifier models, nearest neighbor classification approach, logistic regression [1-4]. The earlier developed statistical models produce an inflexible data classification system and get differed from the predefined classes. Also, these statistical approaches are not scalable in nature. Further to this, is the development of machine learning approaches and case based reasoning approaches [5, 6], wherein they focus only towards the classification performance and not on the measure of in – depth understanding of the problem. Also, the major drawback with these approaches is that when large size of data are utilized, these classifiers takes a long time for searching and causes high computational cost.

With the above noted limitations of statistical and machine learning classifiers in the literature, in this research paper attempt is taken to develop an effective and efficient classifier employing the features and noted advantages of ELMAN neural network and biogeography based optimization. Various literature studies on the applicability and concept of ELMAN neural models and BBO approach is carried out and are presented here for their versatility of use in data classification problems.

Harikumar & Sukanesh [9] analyzed classification of epilepsy risk levels from EEG (Electroencephalogram) signals using Elman neural network and Multilayer Perceptron (MLP) Feed forward neural network. The result confirms that the Elman neural network superior with 97.87 % of PI and 23.31 QV compared to fuzzy classifier, MLP neural network.

A bidirectional Elman-type recurrent neural network with multiple output layers (MOLEBRNN) in order to predict the β-turns, β-turn types, and secondary structure is presented in [10]. A squashing function based construction of a fuzzy rule based classifier is developed wherein the structure for the rules evolved by GA, fuzzy membership functions are fine tuned by gradient based optimization and derivatives of membership functions are effectively computed by squashing function [11].

Tong in the year 2009 [12], carried out a work to detect the intrusion based on hybrid RBF/ Elman neural network wherein the memory of the past events are restored by Elman network and real time pattern classification is done by RBF. Elman recurrent neural networks based new generative method is developed for EEG classification and the result proved that the presented method obtain RMSE of 0.110, 93.3% of classification rate for two- task problem decision taken with 38.7 bits per minute [13].

A behavior of Biogeography-Based Optimization (BBO) over various terrain features of a satellite image and clusters of various land-cover features is presented in [14].

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 15 (2015) pp 35027-35038

Compared to conventional classifiers the presented method is noted to achieve better accuracy. Elman network based Gamelan music onset detection was developed and this was applied to Short-time Fourier Transform adapted to build spectrogram. The suggested model achieves 93 % of F- measure [15].

A Recurrent Neural Network (RNN) with feature grouping for intrusion detection is presented in [16] and the results confirm that compared to MLP and Elman-based intrusion detectors the presented approach False Alarm Rate (FAR) does not degrade and this obtains an improved detection rate and better cost per example and classification rate specifically in R2L attacks.

Biogeography based optimization (BBO) for simultaneous feature selection and feature weighting with k value selection of K-NN rule was developed [17]. The pointed out hybrid (BBO-KNN) algorithm obtain effective reduction of data dimension and better classification rate than other 6 evolutionary and 14 non- evolutionary algorithms. A feature selection based on multi-objective binary biogeography based optimization is developed and applied for 60 top gene expression data selection done by Fisher-Markov selector and SVM adapted as a classifier with LOOCV [18]. Ganeshkumar [19] carried out a work on an Enhanced Particle Swarm Optimization (EPSO) based formation of fuzzy if-then rules and membership function. The result proves that the presented approach produce better classification accuracy and convergence for all the data sets. A dynamic neural network was designed in order to recognition of breath patterns and predicts apnea [20].

A partial least squares (PLS) and cluster analysis (CA) incorporated with optimized Elman neural network algorithm is carried out, correlative and repetitive factors of the feature and samples are avoided by adapting PLS and CA respectively [21]. A Cuckoo Search Levenberg Marquardt Elman Network (CSLMEN) was designed in order to eliminate local minima problem and improve the convergence rate for data classification [22].

Wang [23] performed load forecasting analysis based on Elman neural network with load classification carried out by wavelet clustering algorithm. Sundaram [24] presented a solution to detect the Epileptic attack in Electroencephalogram (EEG) signals using Elman neural network adapting Approximate Entropy (ApEn). Liu [25] carried out a work to choose the good subset of informative gene relevant to classification using discrete biogeography based optimization. The fixed number of gene data is selected by fisher-markov selector, exploration and exploitation ability balanced by adapting discrete migration and mutation model.

Based on the above literature study on ELMAN models and BBO concept, in this research paper both these biologically originated model and evolutionary optimization technique is hybridized to perform classification maintaining the learning and generalization capability.

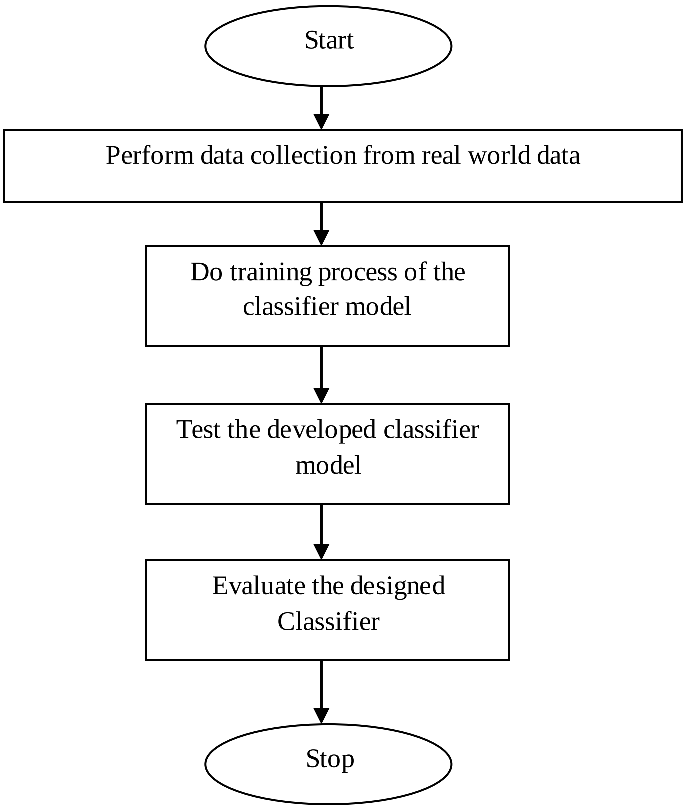

Fig.1. Data Classification system model

Data Classification System Modeling

Data classification, a supervised learning approach which takes data samples and results in a classifier model for classification of data samples in various predefined classes. The process of data classification begins with collecting real time data from the current real world problems and then the data is separated into training and testing data. Figure 1 shows the process of data classification modeling.

During the training process, percentage of training data and a priori knowledge related to the problem domain is employed to adjust parameters and train or learn the structure of the classifier. During the testing phase, the designed trained classifier is evaluated for the new testing data and results in a classification decision for each of the input pattern. The characteristics of the classifier model to be possessed include – classification accuracy, robustness and reliability, speed of the classifier, scalability and interpretability. On a classifier model satisfying the said characteristics, then the designed classifier model classifies the data samples for which the class label is not known.

Performance metrics for Data Classification Modeling

The key objective of data classification is the mechanism wherein the data are perfectly classified into its respective class category or does not get matched into the said class category. This paper computes the classification accuracy, specificity, sensitivity and area under cure of

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 15 (2015) pp 35027-35038
© Research India Publications. http://www.ripublication.com

receiver operating characteristics of the benchmark datasets from the UCI repository. These parameters determine the learning and generalization performance of the classifiers. The parameters are defined as follows:

- Sensitivity is the probability that a diagnostic test is positive, given that the specified sample belongs to the category.

$$\text{Sensitivit y} = \frac{TP}{TP + FN} \tag{1}$$

- Specificity is the probability that a diagnostic test is negative, given that the specified sample does not belongs to the category.

$$\text{Specificit y} = \frac{TN}{TN + FP} \tag{2}$$

- Classification Accuracy is defined as the probability that a diagnostic test is correctly performed.

$$\textit{Classification Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{3}$$

where,

$$\begin{array}{l} \text{True Positives (TP) - Correctly classified positive cases} \\ \text{True Negative (TN) - Correctly classified negative cases} \\ \text{False Positives (FP) - Incorrectly classified negative cases} \\ \text{False Negative (FN) - Incorrectly classified positive cases.} \end{array} \tag{4}$$

- A plot between the sensitivity and specificity with specificity on the x axis and sensitivity on the y- axis forms the receiver operating characteristics (ROC). On ROC curve rising rapidly in the upper right hand corner of the plot or if the value of area under the curve is high, then the classification test belongs to the category that performs well. When the area of curve is near to 1.0 the modeled classifier is highly reliable and when the area of the curve is 0.5, then the classifier is unreliable. The accuracy of the classifier is validated with that of the area under ROC (AUC) curve. AUC lies in the range of 0 to 1.

Thus, in this paper these above said four characteristics are employed as performance metrics to carry out the data classification using the proposed hybrid ELMAN – BBO classifier.

Proposed Hybrid ELMAN – BBO classifier model for data classification

This section proposes a hybridization of ELMAN neural network model and Biogeography based optimization approach to perform data classification for the datasets considered. ELMAN performs the classification process and the weights of ELMAN network are tuned to achieve minimum square error and faster convergence using Biogeography based optimization.

A. ELMAN Neural network model

Elman neural network [7] is a recurrent neural network model wherein the recurrent links are added into the hidden layer as feedback connection. ELMAN neural network is designed with input, hidden, recurrent link and output layer. The recurrent layer is found to copy one step delay of the hidden layer. The output of the ELMAN neural network model is taken from the hidden layer. A recurrent link layer stores the detailed information of the hidden layer and this is found to retain the memory. Hyperbolic tangential sigmoidal activation function is adopted for hidden layer and purelin activation function is adopted in output layer.

B. Proposed ELMAN Neural Network for data classification

Neural network design plays a major role in classifying and determining the accuracy of the data classification process. The data employed as inputs for the ELMAN neural network model are the samples pertaining to the respective class attributes of the dataset. The available benchmark datasets possess higher value and these values tend to suppress the influence of smaller variable during the training process. Henceforth, it is required to perform scaling operation for the datasets. The data are scaled in the range of [0, 1]. ELMAN network is trained to compute the mean square error of the problem and the iteration process is carried out till the mean square error reaches a set value. The improvement in the classification accuracy is obtained since the applicability of data scaling before training. After the scaling process is done, the parameters are assigned for the neural network design and then the learning rule selection is carried out. The trained network is tested for its performance. The classified output with the feature of the datasets and errors are computed. Figure 2 shows the block diagram of the proposed model of ELMAN network for data classification.

In case of the data classification problem considered, the inputs will be the attributes of the considered datasets. Henceforth, the number of input neurons will depend on the number of attributes of respective dataset. As well as the output neurons in this case will be the respective output class categories specified for the datasets. The output class category identification will be output from the output layer neurons. The proposed ELMAN architecture for performing data classification is as shown in Figure 3.

From Figure 3, it can be noted that, each of the layers make independent computation on data that it receives and passes the results to the subsequent layers and finally the output is determined for the network. The input arguments for the respective datasets are transmitted through the hidden layer, where the weights are multiplied with the hyperbolic tangential sigmoidal function. The recurrent neural network learns the function based on current input along with the record of previous state output. Also, the value is transmitted through the second connection multiplied with that of the purelin activation function. During the training progress of the neural network, the earlier information gets reflected to the ELMAN neural network model. Table 1 shows the parametric notations used in this proposed ELMAN model for data classification application.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 15 (2015) pp 35027-35038

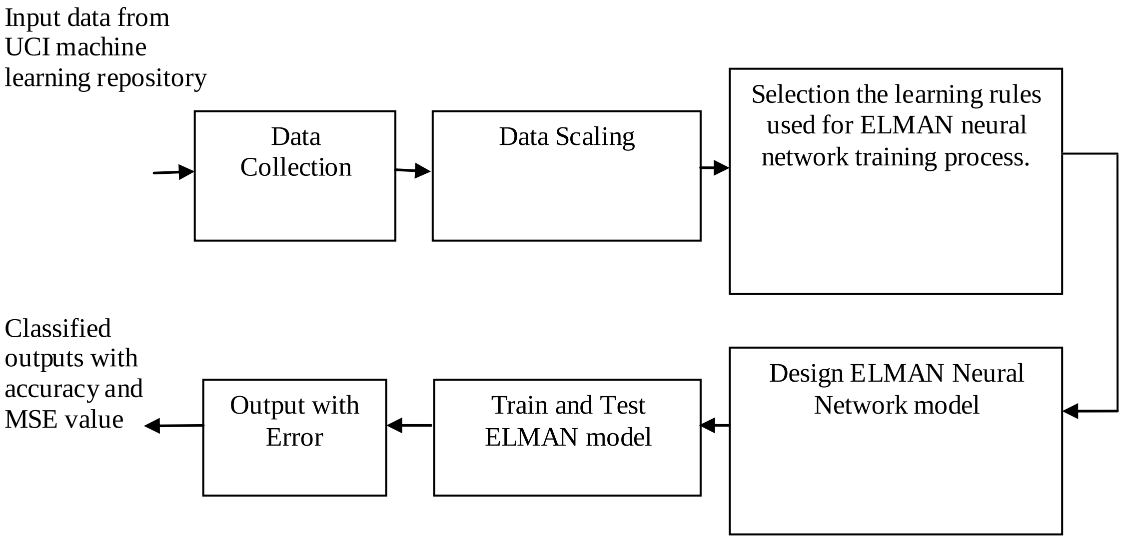

Fig.2. Proposed block diagram of ELMAN model for data classification

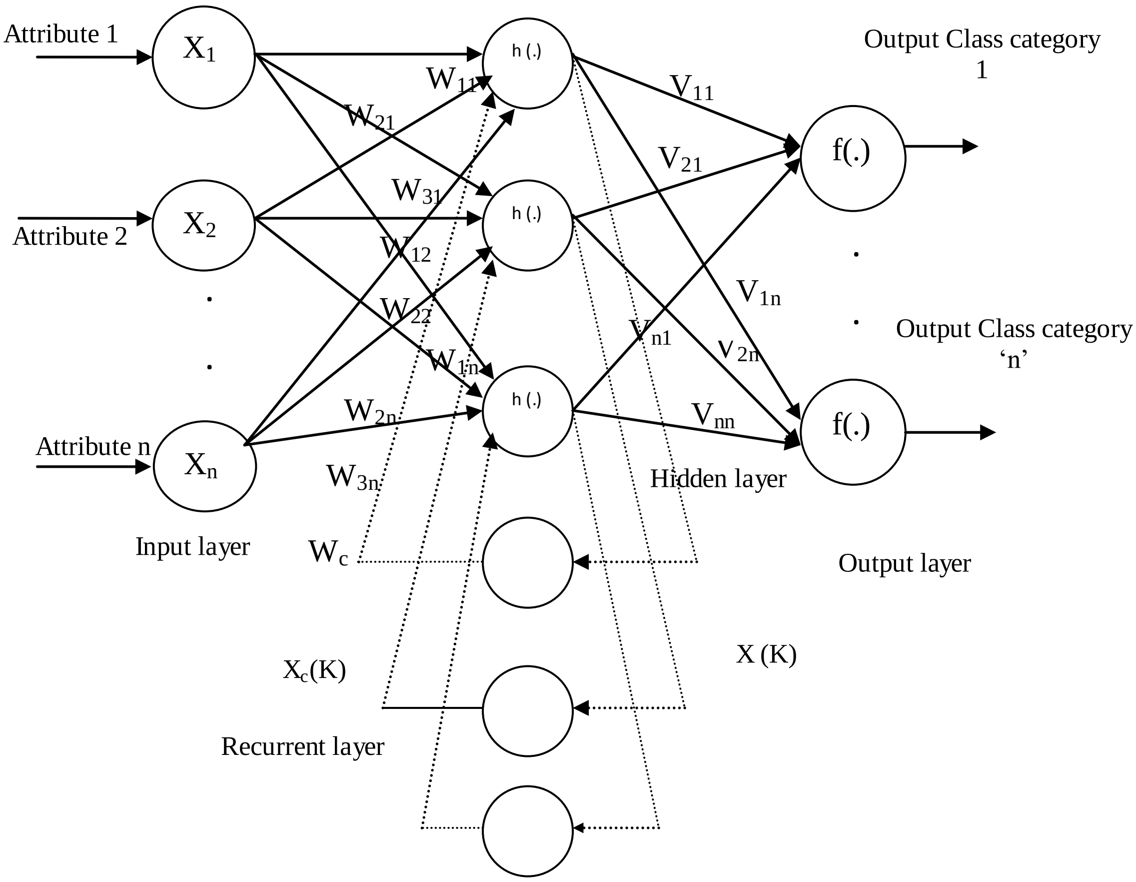

Fig.3. Proposed Elman network architecture model for data classification

C. *Proposed Methodology employing ELMAN Neural Network*

The algorithmic procedure adopted to perform data classification employing ELMAN neural network is as follows:

Step 0: Initialization

Initialize the necessary parameters like learning rate, initial weights and activation functions of the ELMAN neural model.

Step 1: Data collection

In this paper, the datasets considered for implementing the algorithm is the benchmark datasets from the UCI machine learning repository. Around 10 datasets are considered in this research paper, with the number of instance and number of samples as available in the repository. The data sample are trained with different percentages of training and testing data, and for 70% of training data and 30% of testing data, the classification accuracy and mean square error (MSE) is noted to be better than the other variations. As a result, this percentage is maintained for performing training and testing performance of the considered datasets.

TABLE.1. Simulation Parameters

Parametric Entities	Descriptions
X_1, X_2,X_3,…,X_n	Input Vector with the specified number of arguments with respect to the dataset.
Y_1, Y_2, ….Y_n	Output class category
$W_1 = [W_{11,} W_{12}, ...W_{1n}, W_{21}, W_{22}, ...W_{2n}, W_{31}, W_{32}, W_{3n}]$	Weight vector between the input and the hidden layer.
V=[V_{11},V_{12},…,V_{1n},V_{21},V_{22},V_{23},…,V_{2n},…V_{n1},V_{n2},…V_{nn}]	Weight vector of hidden to output link vector
$W_c = [W_{c11}, W_{c12}, ...W_{c1n}, W_{c21}, W_{c22}, ...W_{c2n}, W_{c31}, W_{c32}, ...W_{c3n}]$	Weight vector of recurrent link layer to input vector
$X_c(K) = X(K-1)$	Input of recurrent link layer
The neurons are connected one to one between hidden and recurrent link layer $f(.)$ represents the purelin activation function and $h(.)$ represents the hyperbolic sigmoid activation function	

Step 2: Scaling of Data
For data scaling, Min-Max technique is used which scales within the range [0 1]. The Min-Max technique is used for scaling of input data. The scaling of data is computed by the following eqn. (5),

$$\phi_i' = \left(\frac{\phi_i - \phi_{\min}}{\phi_{\max} - \phi_{\min}}\right)(\phi'_{\max} - \phi'_{\min}) + \phi'_{\min} \quad (5)$$

where ϕ_i, $\phi_{\min}$, $\phi_{\max}$ be the actual input data, minimum and maximum input data and , be the minimum and maximum desired target value.

Step 3: Designing the Elman network for data classification
The parameters like number of input, hidden and output neuron are to be designed. The numbers of hidden neurons is taken to be half that of the input neuron. The input arguments get transmitted through the hidden layer that gets multiplied with the weights by hyperbolic sigmoid function. Also, the output gets transmitted through the second connection and gets multiplied with weights by purelin activation function. Henceforth during the training of the network, past information gets reflected into the Elman neural network. The terminating condition for the neural network includes reaching the minimum error point or a specified number of iterations. Figure 4 shows the flowchart depicting the proposed methodology of data classification employing ELMAN neural network model.

Step 4: Training and evaluate performance of ELMAN network
The benchmark data is divided into training and testing of network. The training data is used to develop classifier models and testing data is used to validate performance of models from the training data. 70% data is used for training and 30% data is used for testing for all the benchmark datasets. The testing data is used to evaluate the performance of ELMAN network. MSE is used as the criteria for classification performance accuracy.

Thus this ELMAN network designs a classifier model to implement on the considered datasets. Further, in the ELMAN model Biogeography Based optimization technique is hybridized to tune the weights of the output layer and hidden layer and as well, the weights of hidden layer and input layer. The forthcoming section presents an overview of BBO approach employed to tune the weights of ELMAN neural model.

D. Proposed Methodology employing ELMAN Neural Network

BBO approach is based on how the species migrate from one island to another, how new species are generated, and how species become extinct. Basically, a habitat is called as any Island (area) that is geographically isolated from other Islands in nature [8]. In BBO approach, the key parameter of importance is Habitat Suitability Index (HSI). Habitats which possess high HIS have a large number of species; on the other hand those with a low HSI possess small number of species. Low species immigration rate is noted for habitats with a high HSI since they are saturated with that of the species. In pursuance, high HSI habitats have a high emigration rate. Habitats with a low HSI tend to have high species immigration rate due to their sparse populations. The emigration rate works in a similar manner. Emigration in BBO does not specify the emigrating island will lose a feature. The worst solution is assumed to have the worst features; hence it has a very low emigration rate and a low chance of sharing its features. The solution that has the best features possesses the highest probability of sharing them. Mathematically the concepts of emigration and immigration rate are represented by a probabilistic model. Consider the probability that the habitat contains exactly species at . changes from time to time as follows:

i) there were S species at time t, and no immigration or emigration occurred between t and $(t+\Delta t)$;

ii) there were $S-1$ species at time t, and one species immigrated;

iii) there were $S+1$ species at time t, and one species emigrated.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 15 (2015) pp 35027-35038

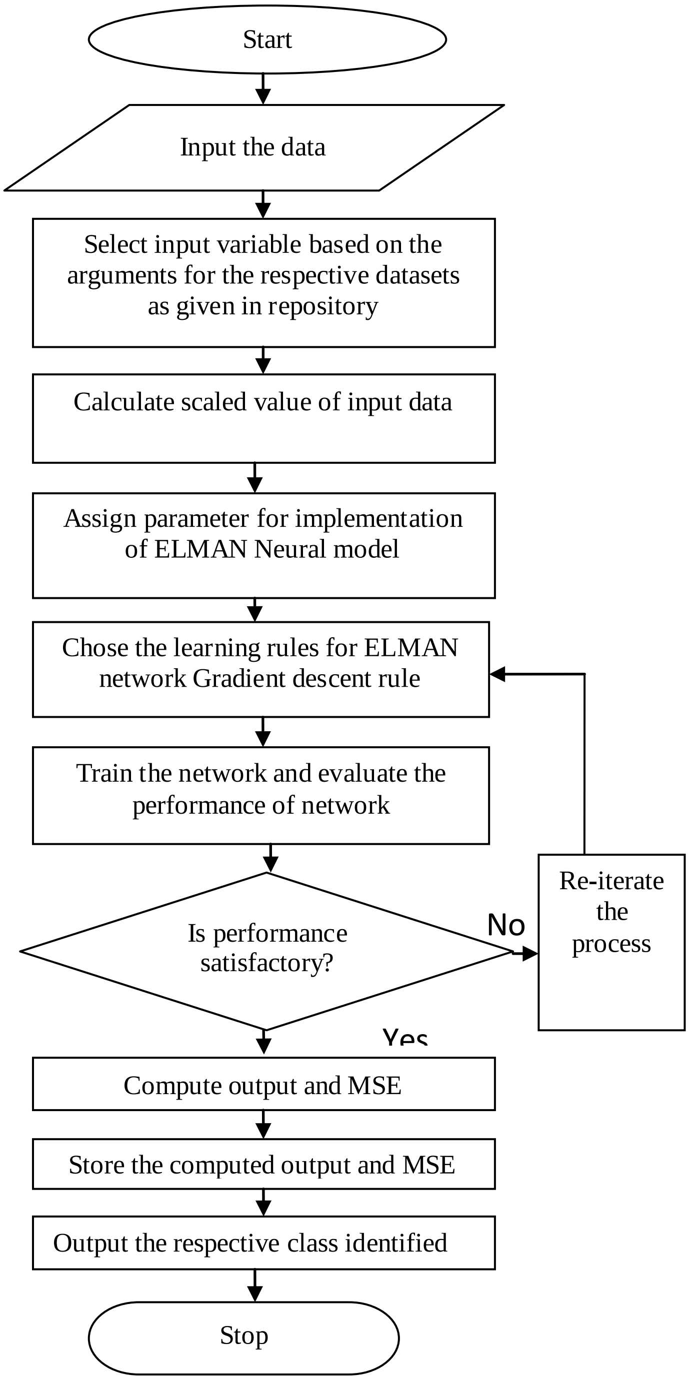

Fig.4. Flowchart for the proposed flow for data classification employing ELMAN neural network model

If time Δt is small enough so that the probability of more than one immigration or emigration can be ignored then taking the limit of (6) as $\Delta t \to 0$ gives the following equation:

$$\dot{P}s = \begin{cases} -(\lambda_s + \mu_s)Ps + \mu_{s+1}Ps+1 & S=0 \\ -(\lambda_s + \mu_s)Ps + \lambda s-1Ps-1 & 1 \le S \le S\max -1 \\ +\mu_{s+1}Ps+1 & \\ -(\lambda_s + \mu_s)Ps + \lambda s-1Ps-1 & S = S\max . \end{cases} \tag{7}$$

Considering the straight-line graph of migration, the equation for emigration rate μ_k and immigration rate λ_k for k number of species is given by:

$$\mu_k = \frac{E_k}{n} \tag{8}$$

$$\lambda_k = I(1 - \frac{k}{n}) \tag{9}$$

On the value of $E = I$ and combining (8) and (9), the equation is given by,

$$\lambda k + \mu k = E \tag{10}$$

The two main operators in BBO are the migration and mutation operators. In this case, H_i is the i^{th} SIV set of habitat H. With the migration operator, BBO can share the information among solutions. Basically, poor solutions tend to accept more useful information from good solutions. Henceforth this operation makes BBO good at exploiting the information of the current population.

Thus, it can be noted that the process of migration and mutation in BBO results in an improved exploration and exploitation; as a result the process of flow of BBO is hybridized into the proposed ELMAN model for data classification to tune their optimal weights.

E. Proposed hybrid ELMAN – BBO classifier for data classification

In ELMAN neural network model, the initial weights between all the layers including the recurrent layer are randomly initialized. This random initialization and random updation of weights results in getting trapped with local and global minima of the network. As well, due to random process, this leads to premature convergence of the network model. Henceforth, this paper focuses on hybridizing ELMAN neural model with that of the biogeography based optimization approach to tune for the optimal weights resulting in faster convergence and avoiding trap of local and global minima. The algorithm developed for the proposed hybrid ELMAN – BBO classifier is as given below:

TABLE.2. Description of Datasets for Proposed hybrid ELMAN – BBO Algorithm

Name of the Dataset	Number of samples	Number of attributes	Number of output classes	Type of each of attribute	Number of samples per class
Iris	150	4	3	All attributes are Real	50, 50 ,50
Ecoli	336	8	8	All attributes are Real	143,77,52,35,20,5,2,2
Yeast	1484	8	10	All attributes are Real	463, 429, 244, 163, 51, 44, 37, 30, 20, 5
Glass Identification	214	9	6	All attributes are Real	70,17,76, 13,9,29
Wine	178	13	3	Integer - 1,6,13 Real – 2,3,4,5,7,8,9,10,11,12	59, 71, 48
Credit Approval	690	14	2	Integer - 3,11,14 Real - 2,8 Categorical - 1,4,5,6,7,9,10,12,13	307,383
Page Blocks Classification	5473	10	5	Integer - 1,2,3,8,9,10 Real - 4,5,6,7	4913, 329, 28, 88, 115
Wisconsin Breast Cancer	699	9	2	All attributes are integers	458, 241
Pima Indians Diabetes	768	8	2	Integer - 1,2,3,4,5,8 Real - 6,7	500, 268
Hepatitis	155	19	2	Categorical 1, 3 – 14, Integer - 2, 16, 17, 19 Real - 15, 18	123, 32

Step 1: Initialize the population P randomly.

Step 2: Evaluate the fitness of the problem considered and sort the population from best to worst.

Step 3: Initialize species count probability for each of the Habitat

Step 4: While the stopping condition criteria is not met do

Step 5: Store the best habitats in a temporary array.

Step 6: For each habitat, map the HSI to number of species S, λ and μ

Step 7: Probabilistically choose the immigration island based on μ

Step 8: Perform migration process. Randomly selected SIV (Suitability Index Variable) based on the selected island in Step7.

Step 9: Mutate the population probabilistically.

Step 10: Evaluate the fitness and sort the population from best to worst. Reject the unfit population.

Step 11: Input the values related to best fitness to tune the weights.

Step 12: With the tuned weights, perform learning rule process of ELMAN network

Step 13: Compute the error of the network.

Step 14: Perform weight updation and compute the training performance.

Step 15: Sort the population.

Step 16: Check for feasibility for similar habitats

Step 17: Stop the algorithmic process.

The above presented hybrid ELMAN Neural Network – BBO algorithm is used to perform data classification for the considered benchmark datasets from UCI machine learning repository with modified weighting factors. This ensures better solution region and will not be over run and will also result faster convergence towards classification of the problem domain.

Datasets for implementation of proposed hybrid classifier

The proposed hybrid ELMAN – BBO technique is validated for its classifier performance on ten benchmark dataset classification problems from UCI Machine Learning Repository [26]. The complete details of these ten real benchmark datasets belonging to various categories of life sciences, finance, information technology and physical science employed in simulation of this data classification work is presented in Table 2. For each of the considered dataset from Iris to Hepatitis, the number of samples, the number of attributes, the number of output classes, type of each of the attributes and the number of samples per class are given as in Table 2. These considered datasets are found to possess multivariate characteristics.

Each of these datasets possesses their distinct attributes as specified in column 3 of Table 2. The input to the proposed hybrid ELMAN – BBO model will be the number of these arguments pertaining to the respective data.

TABLE.3. Parameters for the proposed ELMAN – BBO Classifier

Parameters	ELMAN Neural Model	Parameters	BBO Algorithm
Output Neuron	Based on number of classes of the datasets	Habitat Size	50
No. of hidden layer	1	Habitat Modification probability	1
Input Neurons	Based on number of attributes of the respective datasets	Immigration Probability bounds per gene	[0,1]
No. of Epochs	500	Step size for numerical integration	1
Threshold	1	Maximum Immigration	1
		Migration rate for each island	1
		Mutation probability	0.005
		Maximum Iteration	500

Simulation Results of proposed ELMAN – BBO classifier model

The developed novel hybrid ELMAN – BBO model is implemented for data classification process considering the ten datasets as presented in Table 2 [26]. The performance metrics for performing the classification are those presented in section 3 of this paper. The proposed hybrid ELMAN – BBO model is employed in this research paper to analyze the learning performance and generalization ability. The parametric details considered for implementation of the proposed ELMAN – BBO is as given Table 3. The proposed hybrid ELMAN – BBO classifier algorithm is simulated in MATLAB R2013b environment for the ten datasets and executed in a PC with Intel core i5 processor with 3.5 GHz speed and 10GB RAM with 64 bit operating system.

A. Learning and Generalization ability of proposed model

The performance of the proposed ELMAN – BBO approach is noted for the learning performance metrics – classification accuracy, sensitivity and specificity and for the generalization metrics – area under curve. Table 4 shows the computed performance metrics to validate the learning and generalization performance of the proposed classifier. This table also presents the metrics evaluated for the already existing classifier models available in the literature. From Table 4, it can be noted that the proposed ELMAN – BBO approach implemented for performing data classification for the considered ten datasets performs better classification rate than that of the earlier available classifier models. The approach of hybridizing ELMAN neural model and BBO algorithm has intended to avoid premature convergence of the system as well overcoming the local and global minima problem.

With respect to each of the datasets, it can be observed that the classification accuracy is increased in large and as well the performance for correctly classified data and incorrectly classified data. The Wisconsin breast cancer dataset is found to achieve higher classification accuracy of about 93.8 rather than that of the other considered datasets. On simulation process, it is noted that the time elapsed for the system to converge with that of ELMAN model alone is higher than that of the hybridized model. There exists a noticeable improvement on implementing the hybrid model with respect to the learning and generalization metrics rather than individual ELMAN model considered separately. Figure 5 show the variation of the performance metrics obtained using simulation for Ecoli dataset, which proves that the proposed ELMAN – BBO approach is better in comparison with that of the other methods. The plot can be computed for the other datasets as well in a similar manner.

The generalization ability of the proposed ELMAN – BBO model is achieved by evaluating the area under curve of the receiver operating characteristics (AUC). This generalization analysis is a statistical analysis. When the value of the area under curve is noted to be closer to 1, then the designed ELMAN – BBO is highly reliable in nature and when it is low it proves the classifier model loses its classification nature. Table 4 presents the AUC values calculated for each of the considered benchmark datasets. From this AUC values, it is noted that the proposed hybrid ELMAN – BBO approach possess the ability to provide a richer measure of AUC with respect to the variation in the optimal weights computed. In Table 4 compared to all the other available classifier models in the literature the proposed ELMAN – BBO approach results in AUC values for the ten datasets tending near to 1. This proves the reliability nature of the proposed approach.

Performance analysis and Discussion

A hybrid ELMAN neural network model with biogeography based optimization approach s proposed in this research paper to perform effective and efficient data classification of benchmark datasets from UCI machine learning repository. Simulation is carried out to execute data classification process and the simulated learning and generalization ability of the classifier for each of the ten datasets is tabulated in Table 4. From Table 4, it is inferred that the proposed ELMAN – BBO approach provides better classification accuracy than the other approaches from the literature [11, 27 – 28]. Figure 5 shows the variation of learning performance metrics of the proposed classifier with high performance than that of the other classifiers proposed.

TABLE.4. Performance metrics of proposed ELMAN – BBO model

Approach	Performance Metrics	Iris	Ecoli	Yeast	Glass	Wine	Credit	Page	Breast	PIMA	Hepatitis
Hybrid Genetic Based Machine Learning (HGBML) [27]	Classification Accuracy (%)	83.4	82.4	81.2	84.3	87.2	86.2	84.2	85.3	84.2	81.6
	Sensitivity (%)	87.5	82.3	86.5	83.4	85.9	87.2	89.6	82.5	88.7	80.8
	Specificity (%)	80.3	84.6	82.8	85.1	86.0	86.5	88.4	81.9	84.1	81.8
	AUC	0.379	0.518	0.889	0.847	0.705	0.458	0.308	0.652	0.362	0.457
Genetic Algorithm-Gradient Approach (GAGA) [11]	Classification Accuracy (%)	87.2	84.1	83.4	86.5	89.4	87.4	85.3	88.2	89.4	85.8
	Sensitivity (%)	88.1	83.9	87.4	84.0	87.5	88.9	90.1	85.6	90.3	83.7
	Specificity (%)	81.9	85.2	84.7	87.1	86.8	87.4	90.6	82.3	86.9	83.8
	AUC	0.379	0.543	0.634	0.887	0.715	0.418	0.268	0.612	0.374	0.659
Mixed Genetic Algorithm (MGA) [28]	Classification Accuracy (%)	90.1	85.7	84.2	88.2	91.3	91.2	87.4	92.4	90.2	88.9
	Sensitivity (%)	89.6	85.1	88.6	85.2	88.1	88.9	87.3	89.6	90.5	87.1
	Specificity (%)	82.6	87.9	86.3	88.9	87.7	89.3	91.3	84.5	87.3	84.2
	AUC	0.385	0.964	0.584	0.699	0.731	0.319	0.275	0.588	0.414	0.710
Proposed Hybrid ELMAN – BBO approach	Classification Accuracy (%)	91.9	88.9	87.4	90.5	92.6	93.4	89.6	93.8	91.6	90.9
	Sensitivity (%)	90.1	87.8	89.3	87.4	89.3	89.7	89.4	90.7	91.7	89.0
	Specificity (%)	84.3	89.7	88.8	89.7	89.0	90.8	91.9	85.7	88.9	86.4
	AUC	0.657	0.970	0.715	0.809	0.856	0.624	0.567	0.789	0.804	0.821

For the proposed methodology, other performance metrics observed during the simulation process of training and testing employing the proposed ELMAN – BBO approach and with ELMAN model with random weight initialization is presented in Table 5. In Table 5, one can note that the developed ELMAN – BBO approach possess least MSE value, higher accuracy, training and testing efficiency and minimal iterations than that with ELMAN model alone. Also, the developed ELMAN – BBO approach yields minimal MSE and takes less computational training time to implement classification process for the ten benchmark datasets.

BBO technique is a diversified search technique which explores and exploits in an effective manner with the specified constraints and tunes the weight values for the ELMAN network. ELMAN network acting a recurrent network with the tuned weights yields faster convergence with minimal MSE and less iterative process. Each of the datasets for the considered training and testing samples with the number of attributes acting as number of input neurons performs its classification training process without getting trapped in local minima problem. From Table 5 and Figures 6 to Figure 8, it is inferred that the proposed ELMAN – BBO network results in better classification of the data in comparison with that of the other approaches available in literature.

TABLE.5. Other performance metrics of proposed hybrid ELMAN – BBO approach

Approach	Performance Metrics	Iris	Ecoli	Yeast	Glass	Wine	Credit	Page	Breast	PIMA	Hepatitis
ELMAN Model	MSE Error	0.901	0.675	1.098	0.995	1.237	1.034	1.573	0.871	0.885	0.708
	Training Efficiency % Mean	80.92	82.76	81.72	83.90	82.95	83.77	82.29	86.75	84.71	85.11
	Testing Efficiency % Mean	86.26	85.70	90.65	85.62	86.99	82.99	81.33	80.96	82.53	85.15
	Accuracy%	80.98	76.95	80.29	89.67	84.34	87.31	86.24	90.27	85.99	82.87
	Time in seconds	159.5	176.9	165.8	174.9	169.7	164.5	156.4	167.8	170.9	198.2
	No. of Iterations	500	500	432	490	500	494	500	476	493	417
Proposed Hybrid ELMAN – BBO approach	MSE Error	0.625	0.418	0.596	0.714	0.739	0.957	0.909	0.436	0.729	0.631
	Training Efficiency % Mean	89.36	90.90	90.57	88.45	89.09	91.56	88.15	89.17	92.67	87.37
	Testing Efficiency % Mean	89.51	88.70	93.76	88.95	89.04	87.42	87.50	83.20	89.76	86.59
	Accuracy%	91.9	88.9	87.4	90.5	92.6	93.4	89.6	93.8	91.6	90.9
	Time in seconds	146.4	157.9	152.8	138.1	150.8	158.9	149.1	153.9	167.9	173.4
	No. of Iterations	403	467	394	398	500	478	500	324	456	398

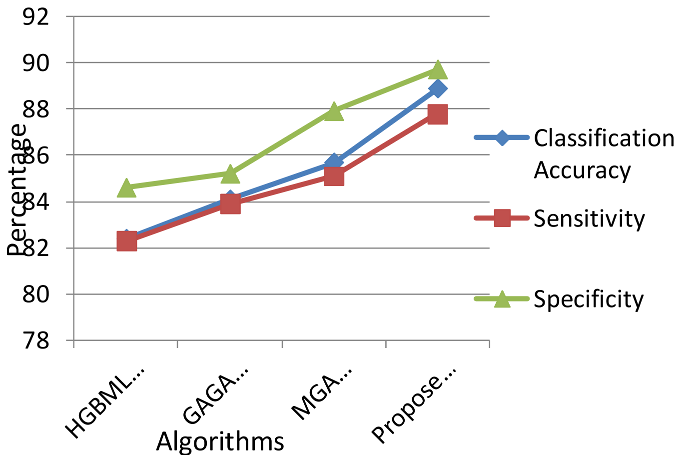

Fig.5. Plot of performance metrics for Ecoli dataset

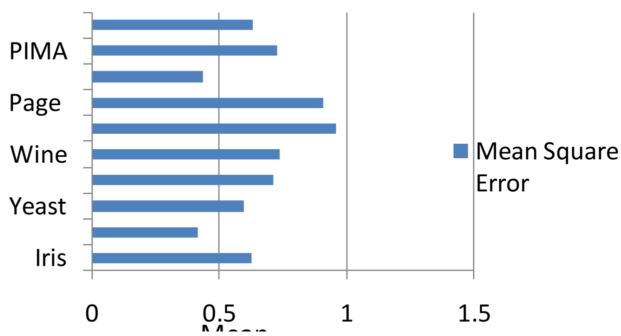

Fig.6. MSE value computed for the ten datasets with ELMAN – BBO model

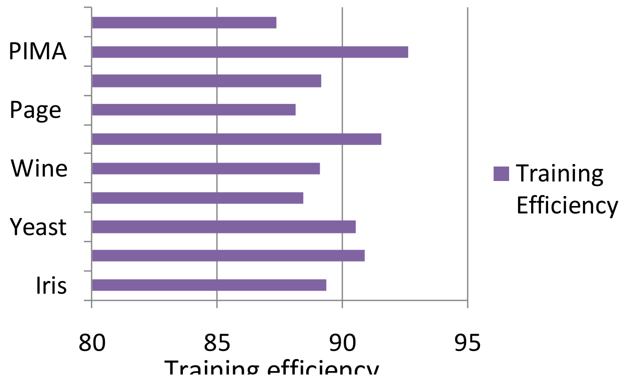

Fig.7. Percentage of training efficiency for the ten datasets with ELMAN – BBO model

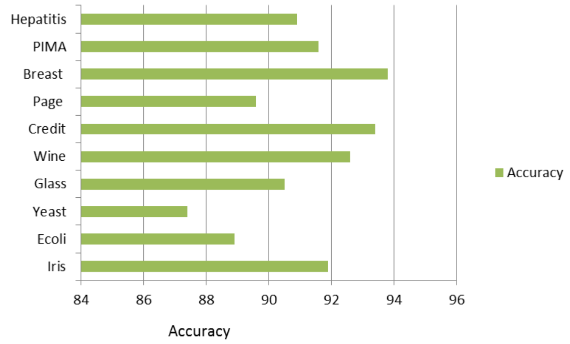

Fig.8. Accuracy Percentage for the ten datasets with ELMAN – BBO model

Conclusion

In this research paper, a novel hybrid ELMAN – BBO based algorithm is proposed that hybridizes the ELMAN recurrent neural network model with Biogeography based Optimization to perform data classification process for the benchmark datasets. BBO technique is employed in the hybridization process to optimize the input weights and hidden weights and biases and as well to determine the output weights. The ELMAN neural model is improved for better performance employing BBO technique and is enhanced to achieve better classification accuracy with minimal MSE, minimal iterations and less computational time. The performance of the proposed ELMAN model employing BBO was better than the performance of the other available methods reported in the literature for ten benchmark datasets from the UCI Machine Learning Repository. The results prove that the proposed model achieves better tuned weights reducing the mean square error, increasing the classification accuracy and as well improving the generalization capability.

References

[1] Domingos Pedro, M. and Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss", Machine Learning, Vol.29, pp. 103–¬137, 1997.

[2] Li, J., Najmi, A. and Gray, R.M. "Image classification by a two dimensional hidden Markov model", Journal of IEEE, Transactions on Signal Processing, Vol.48, No.2, pp.517-33, 2000.

[3] Domeniconi, C.; Peng, J. and Gunopulos, D. "Locally Adaptive Metric Nearest-Neighbor Classification", Journal of IEEE, Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.9, pp.1281-1285, 2002.

[4] Fort, G. and Lacroix, S.L. "Classification using partial least squares with penalized logistic regression", Bioinformatics, Vol.21, No.7, pp. 1104-1111, 2005.

[5] Hsu, C.W., Chang, C.C. and Lin, C.J. "A Practical Guide to Support Vector Classification", Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, 2003.

[6] Arshadi, N. and Jurisica, I. "Data mining for case-based reasoning in high-dimensional biological domains", Journal of IEEE, Transactions on Knowledge and Data Engineering,Vol.17, No.8, pp. 1127-1137, 2005.

[7] Lin, WM & Hong, CM 2011, 'A new Elman neural network-based control algorithm for adjustable-pitch variable-speed wind-energy conversion systems', IEEE Transactions on power electronics, vol. 26, no. 2, pp. 473-481.

[8] D. Simon, "Biogeography-based optimization," IEEE Trans. Evol. Comput., vol. 12, no. 6, pp. 702– 713, Dec. 2008.

[9] Sukanesh, R., & Harikumar, R. (2007). A Comparison of Genetic Algorithm & Neural Network (MLP) In Patient Specific Classification of Epilepsy Risk Levels from EEG Signals. Engineering Letters, 14(1), 43 - 60.

[10] Kirschner, A., & Frishman, D. (2008). Prediction of β-turns and β-turn types by a novel bidirectional Elman-type recurrent neural network with multiple output layers (MOLEBRNN). Gene, 422(1), 22-29.

[11] Dombi, J., & Gera, Z. (2008). Rule based fuzzy classification using squashing functions. Journal of Intelligent and Fuzzy Systems, 19(1), 3-8.

[12] Tong, X., Wang, Z., & Yu, H. (2009). A research using hybrid RBF/Elman neural networks for intrusion detection system secure model. Computer physics communications, 180(10), 1795-1801.

[13] Forney, E. M., & Anderson, C. W. (2011, July). Classification of EEG during imagined mental tasks by forecasting with Elman recurrent neural networks. InNeural Networks (IJCNN), The 2011 International Joint Conference on (pp. 2749-2755). IEEE.

[14] Panchal, V. K., Bhugra, D., Goel, S., & Singhania, V. (2011, April). Study on the behaviour of BBO over natural terrain features. In Electronics Computer Technology (ICECT), 2011 3rd International Conference on (Vol. 4, pp. 28-32). IEEE.

[15] Wulandari, D. P., Suprapto, Y. K., & Purnomo, M. H. (2012, July). Gamelan music onset detection using Elman network. In Computational Intelligence for Measurement Systems and Applications (CIMSA), 2012 IEEE International Conference on (pp. 91-96). IEEE.

[16] Sheikhan, M., Jadidi, Z., & Farrokhi, A. (2012). Intrusion detection using reduced-size RNN based on

feature grouping. Neural Computing and Applications, 21(6), 1185-1190.

[17] Kardan, A., Kavian, A., & Esmaeili, A. (2013, May). Simultaneous feature selection and feature weighting with K selection for KNN classification using BBO algorithm. In Information and Knowledge Technology (IKT), 2013 5th Conference on (pp. 349-354). IEEE.

[18] Li, X., & Yin, M. (2013). Multiobjective binary biogeography based optimization for feature selection using gene expression data. NanoBioscience, IEEE Transactions on, 12(4), 343-353.

[19] Ganeshkumar, P., Rani, C., & Deepa, S. N. (2013). Formation of Fuzzy If-Then rules and Membership function using enhanced particle swarm optimization.International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,21(01), 103-126.

[20] Dmitry, D., Olga, G., & Berestneva, O. G. (2014). Sleep Apnea Detection Based on Dynamic Neural Networks. In Knowledge-Based Software Engineering(pp. 556-567). Springer International Publishing.

[21] Jia, W., Zhao, D., Shen, T., Tang, Y., & Zhao, Y. (2014). Study on optimized Elman neural network classification algorithm based on PLS and CA.Computational intelligence and neuroscience, 2014, 12.

[22] Nawi, N. M., Khan, A., Rehman, M. Z., Herawan, T., & Deris, M. M. (2014). CSLMEN: A New Cuckoo Search Levenberg Marquardt Elman Network for Data Classification. In Recent Advances on Soft Computing and Data Mining (pp. 173-182). Springer International Publishing.

[23] Wang, S. F., & Yang, H. C. (2014, July). Application of Combined Intelligent Algorithm in Load Forecasting. In Advanced Materials Research (Vol. 945, pp. 2895-2899).

[24] Sundaram, S., Arivazhagan, D., Ganeshkumar, K. A recurrent Elman neural network - based approach to detect the presence of epileptic attack in Electroencephalogram (EEG) signals. International Journal of Engineering and Technology, Vol. 6, No. 5, pp. 2388-2391, 2014.

[25] Liu, B., Tian, M., Zhang, C., & Li, X. (2015). Discrete Biogeography Based Optimization for Feature Selection in Molecular Signatures. Molecular Informatics, 34(4), 197-215.

[26] Frank Andrew and Arthur Asuncion. "UCI Machine Learning Repository [http://archive. ics. uci. edu/ml]. Irvine, CA: University of California." School of Information and Computer Science 213 (2010).

[27] Ishibuchi, H., Yamamoto, T. and Nakashima, T. “Hybridization of fuzzy GBML approaches for pattern classification problems”, IEEE Transaction on System Man and Cybernetics Part B, Vol.35, No.2, pp.359-365, 2005.

[28] Devaraj, D. and Ganeshkumar, P. “Mixed Genetic Algorithm approach for Fuzzy Classifier Design”, International Journal of Computational Intelligence and Applications, Vol.9, No.1, pp.49-67, 2010.