Asian Journal of Information Technology 17 (3): 202-211, 2018
ISSN: 1682-3915

Enhanced Learning Approach for Diseases Diagnostic

[1]Khaled M. Fouad, [1]Tarek El Shishtawy and [2]Aymen A. Altae
[1]Department of Information Systems, Faculty of Computers and Informatics,
Benha University, Benha, Egypt
[2]Department of Information Systems, Faculty of Computers and Informatics,
Baghdad University, Baghdad, Iraq

Abstract: Medical disease diagnostic is a very important problem in the medical domain and data mining approach. Early detection of the diseases is very highly important for treating it in early stages. The challenges among machine learning methods are very important to focus on the effective tool to improve the diagnoses problem by indicating the performance of neural network classifiers. This research aims to create a new hybrid method (mRMR-FLN) by exploiting the potential performance of Fast Learning Network (FLN) classifier after integrating it with efficient feature selection algorithm, maximum Relevance Minimum Redundancy (mRMR), to achieve better diagnoses on different diseases. The components of the proposed hybrid method (mRMR-FLN) will be (mRMR) algorithm as a feature selection method and Fast Learning Network (FLN) as a neural classifier. The performance of the new model has been examined and recorded with benchmark measurement on seven evaluation measures. The proposed hybrid method (mRMR-FLN) has achieved very promising classification accuracy using 10-fold Cross-Validation (CV).

Key words: Fast Learning Network (FLN), Extreme earning Machine (ELM), (mRMR) algorithm, hybrid method, medical diagnosis, diseases

INTRODUCTION

The diseases are a particular abnormal condition that effects on patient organisms caused physical injuries that are included in public health surveillance, monitoring and reporting (Hauschild *et al*., 2017). In medical domain, data mining techniques can lead to understanding the level of associated risk (Dominic *et al*., 2015). Data mining is one of the most motivating research areas that is becoming increasingly popular in health organization (Tomar and Agarwal, 2013). The disease's diagnostic faces challenging by using different classifiers on different diseases (Hariharan and Arulanandham, 2017). Especially, for poorly understood diseases, different researchers might use significantly different diagnoses types. Without an agreed-on diagnoses, these researchers may have reported different numbers of cases and characteristics of the diseases (Linden and Yarnold, 2016). However by dependent on prehistoric data and the probability of diagnosing disease many researchers demonstrate that with a technique based on (BP) algorithm they tried to support doctors, especially for asthma disease to prevent the disease during the early itself (Rajan, 2017). Data mining provides the researchers to alter these mounds of data into useful information for medical decision making. The literature analysis for diagnosing disease reveals to how much importance of machine learning and neural network have gained in this domain (Foster *et al*., 2014). The machine learning algorithms reside in the powerfulness for obtaining and presenting the practical application value by obtaining the recognition ability and generalization, the non-linear problem will be solved by neural network methods which identifies system models. Extreme learning machine is one certain machine learning types that involve training one hidden layer neural network. Several versions of ELM have been developed with a wide interest in applying it to different areas and applications. Many data mining techniques have been employed to identify diseases based on non-invasive clinical parameter tests (Verma and Srivastava, 2016). The premature detection of disease is very necessary in order to avoid the risk of the diseases. The hybrid method of adopted a Hybrid Feature Selection (HFS) as feature selection approaches with Weighted Least Squares Twins Supports Vector Machine (WLST SVM) (Tomar and Agarwal, 2015) was proposed as a classification approach to evaluate the significance of each feature, this hybrid model recorded better results of

Corresponding Author: Khaled M. Fouad, Department of Information Systems, Faculty of Computers and Informatics, Benha University, Benha, Egypt

three diseases. The high-dimensional dataset challenge and uncertainties have been handled (Long *et al*., 2015) by using the rough sets model based attribute reduction using Chaos Firefly Algorithm (CFARS-AR) is investigated to explored on optimal reduction, the results have been compared with other machine learning methods namely Naive Bayes (NB), Support Vector Machines (SVM) and Artificial Neural Network (ANN). The proposed model was useful model as a decision support system for heart disease diagnosis. The online learning program was designed (De Lannoy *et al*., 2012) to form a personalized diagnosing model for patients by presented a parallel general regression neural network (GRNN) to diagnosing the heartbeat. Naive Bayes (NB) and Support Vector Machine (SVM) was used (Ananthapadmanabhan and Parthiban, 2014) to predict the early detection of eye disease diabetic retinopathy. The study that aimed to compare different classifiers with breast cancer disease was presented (Diz, 2014) to find the best methods in predicting this disease, among the different tests classifiers, Naive Bayes (NB) was the better to recognize masses texture and random forests was the first or second best classifier for the majority of tested groups. There were studies (Almouti and Bayat, 2016) that were related to data mining techniques for early detection of hepatitis diseases with Support Vector Machine (SVM) algorithm, according to many studies done by various scientists around the world. The hybrid method of early diagnosis of Parkinson's Disease (PD) was developed (Chen *et al*., 2016) by combining the feature selection; maximum Relevance Minimum Redundancy (mRMR) with Kernel extreme learning machine KELM. This proposed method (mRMR-KELM) achieved a very promising classification accuracy using10-fold Cross-Validation (CV).

The goal of this research is presenting a new method based on fast learning neural machine as a tool for diseases diagnostic. The proposed method is a hybrid method (mRMR-FLN) which is performed by exploiting FLN classifier and maximum relevance minimum redundancy (mRMR) as a feature selection method, for diseases diagnosis.

MATERIALS AND METHODS

FLN basic Fast Learning Network (Basic FLN): The fast learning network FLN is based on a single layer feedforward neural network and a three-layer feedforward neural network; input layer, hidden layer and output layer with a parallel connection. Given N arbitrary distinct samples (X_i, Y_i) where Xi = $[X_{i1}\ X_{i2}, ..., X_{in}]^T$ is the training sample of a-dimensional vector quantity, $Y_i = [Y_{i1}, Y_{i2}, \ldots, Y_{in}]^T$ is the target vector. The fast learning network has hidden nodes, W^{in} is the m×n input weight matrix, $b = [b_1, b_2, b_3, ..., b_m]^T$ (Fig. 1) is the bias matrix of hidden layer, W^{oh} is the 1*M matrix which connects hidden layer with output layer. W^{oi} is the I*N connective weight matrix between the output layer and the input layer. C = [C1, C2, …, $C_l]^T$ is the bias matrix of output layer. The j_{th} output target vector of the FLN could be written as shown in Eq. 1:

$$y_j = f\left(w^{ai}x_j + c + \sum_{k=1}^{m} W_k^{oh} g\left(W_k^{in} x_j + bk \right) \right), \quad j = 1, 2, ..., N \tag{1}$$

where, f (.) and g (.) are the active functions of output layer and hidden layer, separately. Equation 1 could be rewritten as shown in Eq. 2. Where:

$$Y = f\left(W^{oi}X + W^{oh}G + c\right) = f\left(\begin{bmatrix} W^{oi} & W^{oh} & c \end{bmatrix}\begin{bmatrix} X \\ G \\ I \end{bmatrix}\right) = f\left(W\begin{bmatrix} X \\ G \\ I \end{bmatrix}\right) \tag{2}$$

$$G = \begin{bmatrix} g\left(W_1^{in}X_1 + b_1\right) & \cdots & g\left(W_1^{in}X_N + b_1\right) \\ \vdots & \ddots & \vdots \\ g\left(W_m^{in}X_1 + b_m\right) & \cdots & g\left(W_m^{in}X_N + b_m\right) \end{bmatrix} m \times N \tag{3}$$

The matrix W = [W^{oi} W^{ho} C] could be called as output weights. G is called the hidden layer output matrix of FLN, the ith row of G is the ith hidden neuron's output vector with respect to inputs X_1, X_2, ..., X_n. FLN could randomly generate the input weights and biases b = [b_1, b_2, b_3, ..., b_m] of the hidden layer. After that, the FLN could be thought of a linear system and the output weights W could be analytically determined through the following form:

$$\left\| f\left(W\begin{bmatrix} X \\ G \\ I \end{bmatrix}\right) - Y \right\| = \min_w f\left(W\begin{bmatrix} X \\ G \\ I \end{bmatrix}\right) - Y \tag{4}$$

For an invertible activation function (.) the output weights are also analytically determined by Eq. 5:

$$\left\| W\begin{bmatrix} X \\ G \\ I \end{bmatrix} - f^{-1}(Y) \right\| = \min_w \left\| W\begin{bmatrix} X \\ G \\ I \end{bmatrix} - f^{-1}(Y) \right\| \tag{5}$$

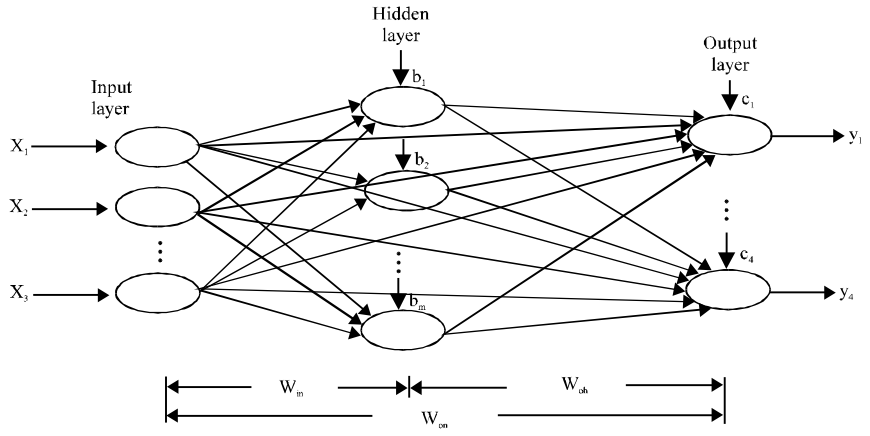

Fig. 1: Procedures structures of the Fast Learning Network (FLN)

where, $f(.)^{-1}$ is the invertible function of f (.). According to the moore-penrose generalized inverse, the minimum norm least-squares solution of the linear system could be written as:

$$\hat{W} = f^{1}(Y)\begin{bmatrix} X \\ G \\ I \end{bmatrix}^{+} = f^{-1}(Y)H^{+} \tag{6}$$

If rank(H) = N, then Eq. 6 could be rewritten as:

$$\hat{W} = f^{1}(Y)\left(\left[X^{T}G^{T}I^{T}\right]\begin{bmatrix} X \\ G \\ I \end{bmatrix}\right)^{-1}\begin{bmatrix} X \\ G \\ I \end{bmatrix}^{7} = f^{1}(Y)\left(X^{T}X+G^{T}G+I^{T}1\right)^{-1}H^{7} \tag{7}$$

As shown in above learning process, the output layer nodes in FLN classifier (Li *et al*., 2014) not only get the recodification of the external information on the hidden layer nodes but also gets the external information itself directly through the input layer nodes. In addition, many researches of literature (Li *et al*., 2017) have shown that a Multilayer Feedforward Neural Networks (MFNNs) can be solving the linear problem with higher efficiency than single layer feedforward neural network and used it in simulated approximation problem with the approximation of nonlinear mappings from the inputs to the outposts. With this advantage, the FLN with the same or a smaller number of hidden units can achieve much better generalization performance and stability than ELM and KELM.

Feature selection algorithm: This study gives a description of mRMR, mRMR is a filter type feature selection method that looks up to choose features which they are relevant to the target class (maximum relevance) and holding up with the feature subset containing as non-redundant features as possible (minimum redundancy) (Peng *et al*., 2005) (Table 1).

Table 1: The feature subset obtained by mRMR filter

Size	Feature subset
1	F1
2	F1 F2
3	F1 F2 F3
4	F1 F2 F3 F4
5	F1 F2 F3 F4 F5
n	F1 F2 F3 F4 F5, ..., Fn

Proposed hybrid method for disease diagnostic: The main aim of the proposed hybrid method is to provide an efficient and accurate diagnostic tool. In this proposed hybrid methods, feature selection is primarily applied to identify the integrated features in each disease (dataset), after then several feature subsets with higher features ranking are fed to the ELM, KELM and FLN models for evaluating the performance (Fig. 2). We can see that two main measures of the ELM, FLN based method are a selection of hidden neurons and activation functions while the main measure of KELM based method is the choice of the parameter pair. The three hybrid methods are comprehensively evaluated on five diseases and they will be estimated by using k-fold CV.

Experiments design

Diseases descriptions: All the following five datasets are taken from UCI machine learning repository (Lichman, 2013), the UCI machine learning repository is a large set of databases, domain theories and data generators which are utilized by the machine learning community to the experimental analysis of machine learning algorithms.

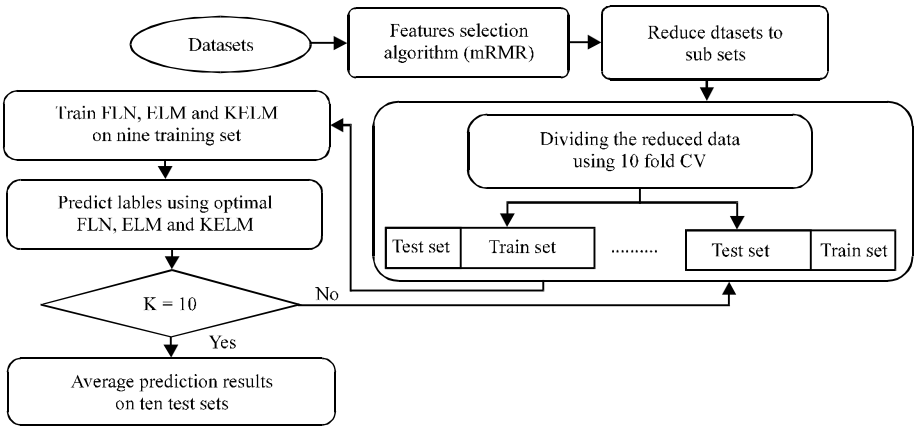

Fig. 2: Overall procedure of the proposed ELM, KELM and FLN based diagnosis method

Table 2: The confusion matrix

----------Predicted class----------			Description
Actual class	Yes	No	(True positive) indicates to correct prediction of having the disease
Yes	TP	FN	(True negative) indicates to correct prediction of non-having the disease
No	FP	TN	(False negative) indicates to false prediction of non-having the disease
			(False positive) indicates to false prediction of having the disease

Experimental setup: The entire experiment is implemented in the MATLAB platform which runs on Windows 7 ultimate operating system with intel® on 64, Core i5 processor 3340 MB (2.70 Ghz 2.70 GHz) and 4 GB of RAM. mRMR program can be obtained from.

Performance measures: The classification accuracy, AUC, sensitivity, specificity, Precision, F-measure, G-mean. Are generally used as performance metrics for evaluation the performance of the binary classification task, especially for the disease diagnostic task. The confusion matrix must be introduced to define these measures as shown in Table 2.

The AUC measure is the area under the Receiver Operating Characteristic (ROC) curve represented by (AUC) measure with true positives versus false positive rates in the area of the plots. A classifier that achieves a higher AUC has the advantage among other classifiers that achieves small AUC, AUC is the best method for comparing classifiers (Fawcett, 2006; Fawcettm, 2003).

RESULTS AND DISCUSSION

In this research the results firstly showing the best-indicated parameters for each classifier in all 5 diseases as shown in Table 3, secondly this research investigating with detail for the performance of (mRMR-FLN) with five experiments, each experiment will be specific for one disease. This disease (dataset) will be reduced with (mRMR) algorithm by ranking the dataset attributes to feature subsets, these feature subsets will be processed with neural classifiers (FLN, ELM and KELM). Then the seven evaluation measures (AUC, accuracy, sensitivity, specificity, precision, F-measure, G-main) for each classifier will be obtained. Our focus will be on the classification performance for the highest average subset to each classifier. Then, the performance comparison for these three proposed neural classifiers (FLN, ELM and KELM) will clearly indicated to adopt the better classifier performance among them and clearly can see the efficiency for the new hybrid method (mRMR-FLN) that which recorded better performance in all feature subsets, even in elapsed time performance on each disease (FLN) needs less execution time to accomplished its diagnosing.

Experiment 1; The performance classification on heart disease: The best performance of FLN is obtained on the feature subset size 13 with average AUC 0.9597, accuracy 0.9007, sensitivity 0.8599, specificity 0.9310, precision 0.9074, F-measure 0.8819 and G-mean 0.8801 over 10 Runs of 10-fold CV. The Elapsed time for this model (mRMR-FLN) in this disease is 5.9868 sec.

The best performance of ELM is obtained on the feature subset size 13 with average AUC 0.9244, accuracy 0.8462, sensitivity 0.8140, specificity 0.8704, precision 0.8276, F-measure 0.8191 and G-mean 0.8175 over 10 Runs of 10-fold CV. The Elapsed time for this model (mRMR-ELM) in this disease is 8.0602 sec.

Table 3: Best Parameters for FLN, ELM and KELM on five diseases

	FLN		ELM		KELM	
Datasets	Activation function	No. of hidden neurons	Activation function	No. of hidden neurons	Gamma γ	C
Heart disease	Sin	4	Hardlim	152	1	1
Breast cancer	Tribas	1	Sig	44	1	1
Hepatitis	Tribas	11	Hardlim	25	1	1
Diabetes	Sig	169	Sig	189	1	1
Parkinson	Hardlim	13	Sig	104	1	1

Table 4: Classification performances of FLN, ELM and KELM on heart disease dataset

Classifier	Feature subset	AUC	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
ELM	1	0.7595	0.6546	0.8166	0.8444	0.8452	0.8246	0.2488
	5	0.7968	0.7385	0.6906	0.7735	0.6981	0.6912	0.6881
	10	0.9023	0.8325	0.8125	0.8475	0.8029	0.8058	0.8038
	13	**0.9244**	**0.8462**	**0.8140**	**0.8704**	**0.8276**	**0.8191**	**0.8175**
KELM	1	0.6062	0.8504	0.5139	0.8963	0.5595	0.8224	0.2300
	5	0.8630	0.7653	0.6548	0.8469	0.7672	0.7047	0.6986
	10	0.9197	0.8347	0.8047	0.8571	0.8131	0.8070	0.8051
	13	**0.9476**	**0.8626**	**0.8231**	**0.8920**	**0.8554**	**0.8376**	**0.8359**
FLN	1	0.8990	0.8356	0.8197	0.8477	0.8054	0.8107	0.8088
	5	0.9331	0.8630	0.8412	0.8793	0.8427	0.8405	0.8390
	10	0.9522	0.8755	0.8605	0.8867	0.8545	0.8559	0.8543
	13	**0.9597**	**0.9007**	**0.8599**	**0.9310**	**0.9074**	**0.8819**	**0.8801**

Table 5: Classification performances of FLN, ELM and KELM on breast cancer dataset

Classifiers	Feature subset	AUC	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
ELM	1	0.7747	0.7149	0.7459	0.7916	0.6278	0.7029	0.6991
	5	0.8014	0.8857	0.7991	0.8193	0.8138	0.8449	0.8434
	9	0.8533	0.8019	0.8278	0.8744	0.7724	0.7970	0.8191
	10	**0.8551**	**0.8019**	**0.8131**	**0.8796**	**0.7773**	**0.8152**	**0.8068**
KELM	1	0.7560	0.7130	0.7573	0.8056	0.6352	0.7927	0.8592
	5	0.8770	0.8569	0.8111	0.8841	0.8638	0.8079	0.8084
	9	0.8720	0.9172	0.8621	0.9499	0.9131	0.8172	0.8052
	10	**0.8780**	**0.9189**	**0.8678**	**0.9493**	**0.9125**	**0.8471**	**0.8383**
FLN	1	0.7995	0.7665	0.8166	0.8444	0.6852	0.8248	0.8406
	5	0.9193	0.8558	0.8993	0.8914	0.8895	0.7963	0.8933
	9	0.9175	0.9293	0.8713	0.9564	0.8073	0.8859	0.8151
	10	**0.9428**	**0.9301**	**0.8803**	**0.8788**	**0.9431**	**0.8842**	**0.8584**

Bold values are significant values

The best performance of KELM is obtained on the feature subset size13 with average AUC 0.9476, accuracy 0.8626, sensitivity 0.8231, specificity 0.8920, precision 0.8554, F-measure 0.8376 and G-mean 0.8359 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-KELM) in this disease is 9.298585 sec. The classification performance of FLN, ELM and KELM for seven evaluation measures on Heart disease will be shown in Table 4.

Experiment 2; The performance classification on breast cancer: The best performance of FLN is obtained on the feature subset size 10 with average AUC 0.9428, accuracy 0.9301, sensitivity 0.8803, specificity 0.8788, precision 0.9431, F-measure 0.8842 and G-mean 0.8584 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-FLN) in this disease is 7.5276 sec.

The best performance of ELM is obtained on the feature subset size 10 with average AUC 0.8551, accuracy 0.8019, sensitivity 0.8131, apecificity 0.8796, precision 0.7773, F-measure 0.8152 and G-mean 0.8068 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-ELM) in this disease is 10.255544 sec.

The best performance of KELM is obtained on the feature subset size10 with average AUC 0.8780, accuracy 0.9189, sensitivity 0.8678, specificity 0.9493, precision 0.9125, F-measure 0.8471 and G-mean 0.8383 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-KELM) in this disease is 16.0438 sec. The classification performance of FLN, ELM and KELM for seven evaluation measures on breast cancer will be shown in Table 5.

Experiment 3; The performance classification on diabetes: The best performance of FLN is obtained on the feature subset size 8 with average AUC 0.9657, accuracy 0. 9016, sensitivity 0.9775, specificity 0.9299, precision 0. 9052, F-measure 0.8828 and G-mean 0.8813 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-FLN) in this disease is 7.5276 sec.

Table 6: Classification performances of FLN, ELM and KELM on diabetes dataset

Classifiers	Feature subset	AUC	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
ELM	1	0.8198	0.8259	0.8168	0.8092	0.8057	0.8734	0.8014
	5	0.8454	0.8245	0.8788	0.8194	0.8063	0.8776	0.8743
	7	0.8086	0.8205	0.8913	0.7805	0.8075	0.8815	0.8780
	8	**0.8316**	**0.8322**	**0.9035**	**0.7908**	**0.8063**	**0.8877**	**0.8805**
KELM	1	0.8736	0.8564	0.8221	0.7913	0.8124	0.8510	0.8498
	5	0.8963	0.8444	0.8986	0.8371	0.8062	0.8461	0.8448
	7	0.8194	0.8261	0.9322	0.8230	0.8366	0.8388	0.8381
	8	**0.8762**	**0.8636**	**0.9530**	**0.8204**	**0.8129**	**0.8893**	**0.8816**
FLN	1	0.9149	0.8434	0.8472	0.8633	0.8214	0.8171	0.8151
	5	0.9128	0.8454	0.9554	0.8725	0.8287	0.8170	0.8152
	7	0.9175	0.8398	0.9626	0.8464	0.8073	0.8172	0.8151
	8	**0.9657**	**0.9016**	**0.9775**	**0.9299**	**0.9052**	**0.8828**	**0.8813**

Table 7: Classification performances of FLN, ELM and KELM on hepatitis dataset

Classifiers	Feature subset	AUC	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
ELM	1	0.7816	0.8179	0.8168	0.7750	0.8721	0.8516	0.8489
	5	0.7767	0.8773	0.8112	0.7333	0.7919	0.8792	0.8740
	10	0.7311	0.8704	0.8313	0.7167	0.8335	0.8494	0.8470
	15	0.8452	0.8050	0.8403	0.7583	0.8747	0.8529	0.8504
	19	**0.8609**	**0.8649**	**0.8548**	**0.8333**	**0.8538**	**0.8778**	**0.8756**
KELM	1	0.8142	0.9050	0.8367	0.8025	0.8388	0.8810	0.8788
	5	0.8113	0.8971	0.9771	0.8308	0.8111	0.8838	0.8799
	10	0.8125	0.8975	0.8700	0.8125	0.8098	0.8891	0.8848
	15	0.8994	0.8483	0.8362	0.8767	0.8039	0.8871	0.8824
	19	**0.9263**	**0.9054**	**0.9062**	**0.8975**	**0.8387**	**0.8858**	**0.8836**
FLN	1	0.9149	0.9389	0.9276	0.8633	0.8214	0.8171	0.8151
	5	0.8943	0.9325	0.9642	0.8527	0.8065	0.8072	0.8056
	10	0.9175	0.9293	0.9772	0.8464	0.8073	0.8172	0.8151
	15	0.9428	0.9300	0.9798	0.8788	0.8431	0.8400	0.8383
	19	**0.9640**	**0.9666**	**0.9375**	**0.9330**	**0.9100**	**0.8803**	**0.8782**

Bold values are significant values

The best performance of ELM is obtained on the feature subset size 8 with average AUC 0.8316, accuracy 0.8322, sensitivity 0.9035, specificity 0.7903, precision 0.8063, F-measure 0.8877 and G-mean 0.8805 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-ELM) in this disease is 9.5518 sec.

The best performance of KELM is obtained on the feature subset size 8 with average AUC 0.8762, accuracy 0.8636, sensitivity 0.9530, specificity 0.8204, precision 0.8129, F-measure 0.8893 and G-mean 0.8816 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-KELM) in this disease is 15.406359 sec. The classification performance of FLN, ELM and KELM for seven evaluation measures on Heart disease will be shown in Table 6.

Experiment 4; The performance classification on hepatitis disease: The best performance of FLN is obtained on the feature subset size 19 with average AUC 0.9640, accuracy 0.9666, sensitivity 0.9375, specificity 0.9330, precision 0.9100, F-measure 0.8803 and G-mean 0.8782 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-FLN) in this disease is 5.7527 sec.

The best performance of ELM is obtained on the feature subset size 19 with average AUC 0.8609, accuracy 0.8649, sensitivity 0.8548, specificity 0.8333, precision 0.8538, F-measure 0.8778 and G-mean 0.8756 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-ELM) in this disease is 8.5109 sec.

The best performance of KELM is obtained on the feature subset size 19 with average AUC 0.9263, accuracy 0.9054, sensitivity 0.9062, specificity 0.8975, precision 0.8387, F-measure 0.8858 and G-mean 0.8836 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-KELM) in this disease is 10.117522 sec. The classification performance of FLN, ELM and KELM for seven evaluation measures on Heart disease will be shown in Table 7.

Experiment 5; The performance classification on Parkinson (PD): The best performance of FLN is obtained on the feature subset size 22 with average AUC 0.8757, accuracy 0.9684, sensitivity 0.9667, specificity 0.6395, precision 0.8953, F-measure 0.9295 and G-mean 0.9281 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-FLN) in this disease is 5.620333 sec.

The best performance of ELM is obtained on the feature subset size 22 with average AUC 0.7236, accuracy 0.7807, sensitivity 0.9095, specificity 0.6371, precision

0.8227, F-measure 0.8639 and G-mean 0.8617 over 10 runs of 10-fold CV. The Elapsed time for this model (mRMR-ELM) in this disease is 10.5789 sec. The best performance of KELM is obtained on the feature subset size 22 with average AUC 0.7787, accuracy 0.8603, sensitivity 0.9219, specificity 0.6595, precision 0.9004, F-measure 0.9099 and G-mean 0.9087 over 10 runs of 10-fold CV. The elapsed time for this model (mRMR-KELM) in this disease is 12.435771 sec. The classification performance of FLN, ELM and KELM for seven evaluation measures on heart disease will be shown in Table 8.

As shown in Table 9, the elapsed time for each classifier is recorded to show the duration for each classifier in diseases diagnostic. Through, the previous Table 9 and 10 for each disease, the outperforms clearly shown for the new hybrid model (mRMR-FLN) among (mRMR-ELM) and (mRMR-KELM), this challenge is

Table 8: Classification performances of FLN, ELM and KELM on Parkinson (PD) Dataset

Classifiers	Feature subset	AUC	Accuracy	Sensitivity	Specificity	Precision	F-measure	G-mean
ELM	1	0.8386	0.8564	0.9412	0.6112	0.6249	0.6489	0.6349
	5	0.8309	0.8861	0.9421	0.6721	0.6858	0.7098	0.6958
	10	0.8537	0.8975	0.9425	0.7925	0.8062	0.8302	0.8162
	15	0.7323	0.7483	0.7673	0.6993	0.7130	0.7370	0.7230
	20	0.8038	0.8506	0.8736	0.8085	0.8222	0.8462	0.8322
	22	**0.8840**	**0.9018**	**0.9248**	**0.8438**	**0.8575**	**0.8815**	**0.8675**
KELM	1	0.8453	0.8631	0.9479	0.6779	0.6916	0.7156	0.7016
	5	0.8121	0.8673	0.9233	0.5933	0.6070	0.6310	0.6170
	10	0.8605	0.9043	0.9493	0.7993	0.8130	0.8370	0.8230
	15	0.8466	0.8626	0.8816	0.8136	0.8273	0.8513	0.8373
	20	0.8692	0.9155	0.9385	0.8775	0.8912	0.9152	0.9012
	22	**0.8453**	**0.8631**	**0.9479**	**0.6779**	**0.6916**	**0.7156**	**0.7016**
FLN	1	0.8386	0.8749	0.9597	0.6897	0.7034	0.7274	0.7134
	5	0.8205	0.8757	0.9267	0.5967	0.6104	0.6344	0.6204
	10	0.8901	0.9339	0.9789	0.8289	0.8426	0.8666	0.8526
	15	0.9374	0.9534	0.9724	0.9044	0.9181	0.9421	0.9281
	20	0.9077	0.9545	0.9775	0.9165	0.9302	0.9542	0.9402
	22	**0.9278**	**0.9684**	**0.9914**	**0.9104**	**0.9241**	**0.9481**	**0.9341**

Bold values are significant values

Table 9: Execution time for each classifier on each disease

	Classifiers (sec)		
Datasets	FLN	ELM	KELM
Heart disease	5.986800	08.060200	09.298585
Breast cancer	7.527600	10.255544	16.043800
Diabetes	7.727600	09.551800	15.406359
Hepatitis	5.752700	08.510900	10.117522
Parkinson (PD)	5.620333	10.578900	12.435771

Table 10: Classification accuracies obtained our hybrid method with other methods for all 5 diseases

Dataset	Studies	Methods	Accuracy (%)
Heart disease	Chavanpatil and Sonawane (2017)	ANN	85.30
	Thomas and Princy (2016)	KNN	80.60
	Long *et al.* (2015)	CFARS-AR	88.30
	Long *et al.* (2015)	BPSORS-AR	87.00
	De Lannoy (2012)	Weighted Conditional Random Fields (WCRF)	85.00
	Foster *et al.*, 2014)	SVM	81.10
	This study	mRMR-FLN	90.07
Breast cancer	Hariraj *et al.* (2017)	KNN	92.10
	Kharya and Soni (2016)	DT	85.71
	Kharya and Soni (2016)	Weighted Associative Classifier (WAC)	90.41
	Tomar and Agarwal (2015)	LSTSVM+under sampling	92.83
	Senturk and Kara (2014)	Rapid Miner (RM)	87.11
	Lin *et al.* (2014)	SVM	87.69
	This study	mRMR-FLN	93.02
Diabetes	Muthukumar and Krishna (2016)	fuzzy soft sets (FSSs)	86.67
	Padmanaban and Parthiban (2016)	NB	86.00
	Choubey and Sanchita (2016)	GA-MLPNN	84.20
	Kang *et al.* (2015)	SVM	80.00
	Ananthapadmanabhan and Parthiban	NB	83.00
	Temurtas *et al.* (2009)	MLNN with LM	82.37
	This study	mRMR-FLN	90.16
Hepatitis	Husain *et al.* (2017)	LSSVM	93.70

Table 10: Continue

Dataset	Studies	Methods	Accuracy (%)
	Almouti and Bayat (2016)	NB	87.0000
	Muthukumar and Krishna (2016)	Fuzzy Soft Sets (FSSs)	80.0000
	Almouti and Bayat (2016)	Rough set	89.0000
	Vijayarani1 (2015)	SVM	80.0000
	Tomar and Agarwal (2015)	HFS+WLSTSVM	87.5000
	This study	mRMR-FLN	96.6600
Parkinson (PD)	Rao and Kaladhar (2016)	NB	83.0000
	Suganya and Sumathi (2015)	PSO	87.0213
	Luukka (2011)	Fuzzy entropy measures+similarity	85.0300
	Shahbaba and Neal (2009)	Dirichlet process mixtures	87.7000
	Ozcift and Gulten (2011)	CFS-RF	87.1000
	Chen *et al.* (2016)	(mRMR-KELM)	95.9700
	This study	mRMR-FLN	96.8400

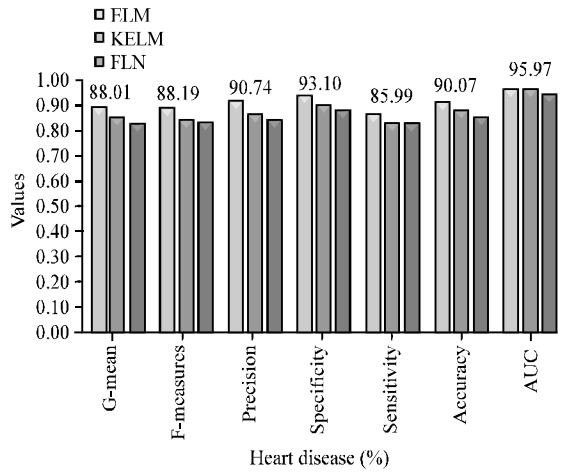

Fig. 3: Results for last feature subset on heart

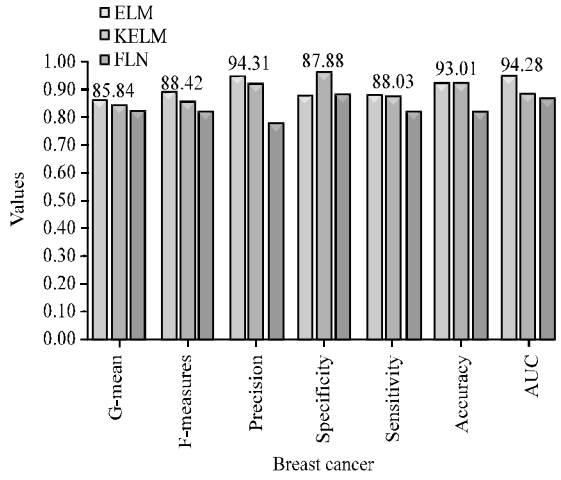

Fig. 4: Results for last feature subset on B.C

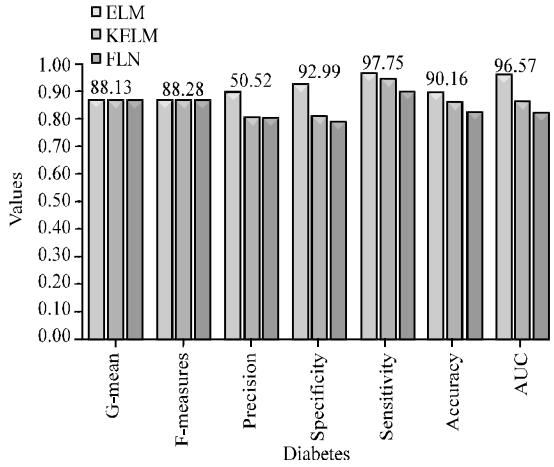

Fig. 5: Results for last feature subset on diabetes

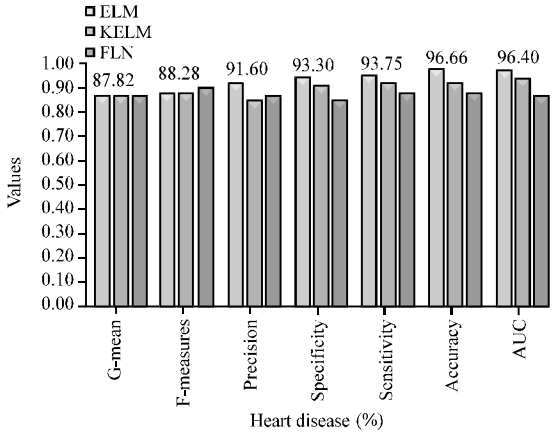

Fig. 6: Results for last feature subset on hepatitis

evident in the last feature subset to each classifier for each disease, it recorded highest average in the seven evaluation measures. Therefore, this study focused on the last feature subset for each disease when the evaluation measures to each classifier for each disease which will be presented as shown in Fig. 3-6. Finally, from the above, these Fig. 7 shows that FLN outperforms KELM and ELM in most of cases. Also, FLN achieves the best results in the last feature subset for these five diseases.

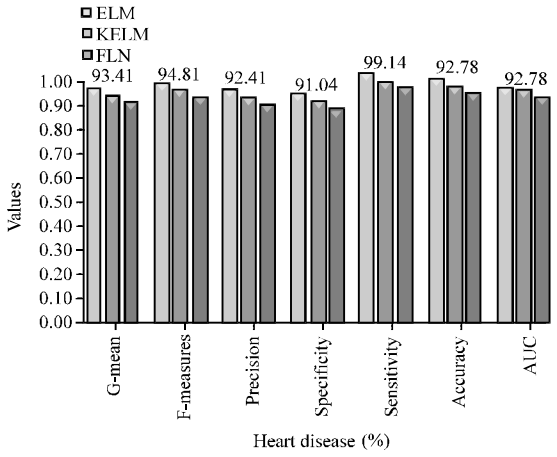

Fig. 7: Results for last feature subset on Parkinson (PD)

CONCLUSION

In this study, effective hybrid method (mRMR-FLN) has been created to support methods in diagnosing diseases. The proposed hybrid method contains with FLN classifier, whose parameters are explored in detail. With the aid of the feature selection techniques, especially the mRMR filter, the performance of FLN classifier was recorded when comparing it with the other classifiers. The promising performance obtained on five different diseases dataset has proven that the proposed hybrid method (mRMR-FLN) can distinguish well enough between infected patients and healthy patients. It is observed that (mRMR-FLN) achieves the better classification accuracy in heart disease 90.07%, breast cancer 93.02%, diabetes 90.16%, hepatitis disease 996.66% and Parkinson's (PD) 96.84% via 10-fold CV analysis. The future study takes a strong focus on investigation the developed hybrid method with other powerful feature selection algorithm, trying to achieve better performance in the diagnostic problem.

REFERENCES

Almouti, M.M. and P. Bayat, 2016. Presenting a hybrid model for early diagnosis of hepatitis by applying data mining techniques. Intl. J. Comput. Inf. Technol., 4: 1-10.

Ananthapadmanabhan, K.R. and G. Parthiban, 2014. Prediction of chances-diabetic retinopathy using data mining classification techniques. Indian J. Sci. Technol., 7: 1498-1503.

Chavanpatil, A.B. and P.S.S. Sonawane, 2017. To predict heart disease risk and medications using data mining techniques with an IoT based monitoring system for post operative heart disease patients. Procedia Intl. J. Emerg. Trends Technol., 4: 8274-8281.

Chen, H.L., G. Wang, C. Ma, Z.N. Cai and W.B. Liu *et al.*, 2016. An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease. Neurocomput., 184: 131-144.

Choubey, D.K. and P. Sanchita, 2016. GA_MLP NN: A hybrid intelligent system for diabetes disease diagnosis. Intl. J. Intell. Syst. Appl., 8: 49-59.

De Lannoy, G., D. Francois, J. Delbeke and M. Verleysen, 2012. Weighted conditional random fields for supervised interpatient heartbeat classification. IEEE Trans. Biomed. Eng., 59: 241-247.

Diz, J.M., 2014. [Application of data mining techniques to support the diagnosis of breast cancer]. Ph.D Thesis, Instituto Superior de Engenharia do Porto, Porto, Portugal. (In Portuguese)

Dominic, V., D. Gupta and S. Khare, 2015. An effective performance analysis of machine learning techniques for cardiovascular disease. Appl. Med. Inf., 36: 23-32.

Fawcett, T., 2003. ROC Graphs: Notes and practical considerations for data mining researchers. HP Laboratories Palo Alto, Palo Alto, CA. http://www.hpl.hp.com/techreports/2003/HPL-2003-4.pdf.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett., 27: 861-874.

Foster, K.R., R. Koprowski and J.D. Skufca, 2014. Machine learning, medical diagnosis and Biomedical Engineering research-commentary. Biomed. Eng. Online, 13: 94-102.

Hariharan, P. and K. Arulanandham, 2017. Design an disease predication application using data mining techniques for effective query processing results. Adv. Comput. Sci. Technol., 10: 353-361.

Hariraj, V., K.W. Vikneswaran and I. Zunaidi, 2017. An efficient data mining approaches for breast cancer detection and segmentation in mammogram. J. Adv. Res. Dyn. Control Syst., 9: 185-194.

Hauschild, V., K. Hauret, M. Richardson, B.H. Jones and T. Lee, 2017. A taxonomy of injuries for public health monitoring and reporting (No. APHC-PHIP-12-01-0717). Army Public Health Center, USA. http://www.dtic.mil/docs/citations/AD1039481.

Husain, N.P., N.N. Arisa, P.N. Rahayu, A.Z. Arifin and D. Herumurti, 2017. Least squares support vector machines parameter optimization based on improved ant colony algorithm for hepatitis diagnosis. J. Ilmu Komputer Informasi, 10: 43-49.

Kang, S., P. Kang, T. Ko, S. Cho and S.J. Rhee *et al.*, 2015. An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction. Expert Syst. Appl., 42: 4265-4273.

Kharya, S. and S. Soni, 2016. Weighted naive bayes classifier: A predictive model for breast cancer detection. Intl. J. Comput. Appl., 133: 32-37.

Li, F., J.M. Zurada, Y. Liu and W. Wu, 2017. Input layer regularization of multilayer feedforward neural networks. IEEE. Access, 5: 10979-10985.

Li, G., P. Niu, X. Duan and X. Zhang, 2014. Fast learning network: A novel artificial neural network with a fast learning speed. Neural Comput. Appl., 24: 1683-1695.

Lichman, M., 2013. Welcome to the UC Irvine machine learning repository. University of California Irvine, Irvine,California.http://archive.ics.uci.edu/ml/index.php

Linden, A. and P.R. Yarnold, 2016. Using data mining techniques to characterize participation in observational studies. J. Eval. Clin. Pract., 22: 839-847.

Long, N.C., P. Meesad and H. Unger, 2015. A highly accurate firefly based algorithm for heart disease prediction. Expert Syst. Appl., 42: 8221-8231.

Luukka, P., 2011. Feature selection using fuzzy entropy measures with similarity classifier. Expert Syst. Appl., 38: 4600-4607.

Muthukumar, P. and G.S.S. Krishnan, 2016. A similarity measure of intuitionistic fuzzy soft sets and its application in medical diagnosis. Appl. Soft Comput., 41: 148-156.

Ozcift, A. and A. Gulten, 2011. Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput. Methods Programs Biomed., 104: 443-451.

Padmanaban, K.A. and G. Parthiban, 2016. Applying machine learning techniques for predicting the risk of chronic kidney disease. Indian J. Sci. Technol., 9: 1-5.

Peng, H., F. Long and C. Ding, 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell., 27: 1226-1238.

Rajan, J.R., 2017. Prognostic system for early diagnosis of pediatric lung disease using artificial intelligence. Curr. Pediatr. Res., 21: 31-34.

Senturk, Z.K. and R. Kara, 2014. Breast cancer diagnosis via data mining: Performance analysis of seven different algorithms. Comput. Sci. Eng., 4: 35-46.

Shahbaba, B. and R. Neal, 2009. Nonlinear models using Dirichlet process mixtures. J. Mach. Learn. Res., 10: 1829-1850.

Suganya, P. and C.P. Sumathi, 2015. A novel metaheuristic data mining algorithm for the detection and classification of parkinson disease. Indian J. Sci. Technol., 8: 1-9.

Temurtas, H., N. Yumusak and F. Temurtas, 2009. A comparative study on diabetes disease diagnosis using neural networks. Expert Syst. Appl., 36: 8610-8615.

Thomas, J. and R.T. Princy, 2016. Human heart disease prediction system using data mining techniques. Proceedings of the 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT'16), March 18-19, 2016, IEEE, Nagercoil, India, ISBN:978-1-5090-1278-7, pp: 1-5.

Tomar, D. and S. Agarwal, 2013. A survey on data mining approaches for Healthcare. Intl. J. BioSci. BioTechnol., 5: 241-266.

Tomar, D. and S. Agarwal, 2015. Hybrid feature selection based weighted least squares twin support vector machine approach for diagnosing breast cancer, hepatitis and diabetes. Adv. Artificial Neural Syst., 2015: 1-10.

Verma, L. and S. Srivastava, 2016. A data mining model for coronary artery disease detection using noninvasive clinical parameters. Indian J. Sci. Technol., 9: 1-6.

Vijayarani, S. and S. Dhayanand, 2015. Liver disease prediction using SVM and Naive Bayes algorithms. Intl. J. Sci. Eng. Technol. Res., 4: 816-820.