

A Novel Method of Hybrid Extreme Learning Machine for Diabetes Mellitus Diagnosis

C. Thiyagarajan¹, K. Anandha Kumar² and A. Bharathi³

¹Bharathiar University, Coimbatore – 641046, Tamil Nadu, India; tr061417@gmail.com

²Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam – 638401, Tamil Nadu, India; k448089@gmail.com

³Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam – 638401, Tamil Nadu, India; smresearch8@gmail.com

Abstract

Objectives: To design a classifier for the detection of Diabetes Mellitus with optimal cost and precise performance. **Method and Analysis:** The diagnosis and interpretation of the diabetes data is must because major problem occurs due to this data maintenance. Several researches are made with machine learning but still needs improvements. In this paper a new method is evaluated as hybrid Extreme Learning Machine (HELM) with African Buffalo Optimization (ABO). **Findings:** ELM is used to select the input data because of fast learning speed. Optimization technique is used for searching and classifying the good diabetic data. The ABO is a population based algorithm in which individual buffalos work together to identify the diabetics data by updating fitness value for best output solution. The proposed HELM technique is successfully implemented for diagnosing diabetes disease. By using this machine learning algorithm, the classification accuracy is achieved for classifying the diabetes patients by using much of the data set for training and few data sets for testing. In order to improve the quality as well as accuracy there is a need for algorithm. The combination of ELM-ABO classifier is applied in training dataset taken from PRIMA Indian dataset for classification and the experimental results are compared with SVM and other ELM classifiers applied on the same database. **Improvement:** It is observed that the HELM method obtained high accuracy in classification with less execution time along with performance evaluation of parameters such as recall, precision and F-Measure.

Keywords: African Buffalo Optimization (ABO), Diabetes Mellitus (DM), Extreme Learning Machine (ELM), Machine Learning Algorithms, Optimization

1. Introduction

Machine learning techniques suiting computer-aided medical diagnosis should have good comprehensibility. It has the transparency of diagnostic knowledge and the explanation ability. Machine learning methods for classification provide inexpensive means to perform diagnosis, prognosis, or detection of certain outcomes in health care research. Diabetes is disease in which the body does not properly produce insulin. To prevent or postpone such complications strict control over the diabetic blood glucose level is needed. Number of computer-based system is available to diagnose the diabetes¹. The goal of this study

is to propose various statistical normalization procedures to improve the classification accuracy.

Classification is one of the most important decision making techniques in many real world problem. In² proposed a new approach for diagnosis of diabetes based on the Small-World Feed Forward Artificial Neural Network (SW-FFANN). The classification performance of the SW-FFANN was better than that of the conventional FFANN. This work is high of cost and architecture. In³ proposed a method of data handling (GMDH), Genetic Algorithms (GA) and Probabilistic Neural Network (PNN) models the prediction of software maintainability and it was found that GMDH models predict more

*Author for correspondence

accurately than the other machine learning models. In⁴ proposed a novel approach for the diagnosis of diabetes with the help of neural networks and other computing technologies. The proposed work presented the initial results for a simple client and server two-tier architecture for healthcare. However it had higher computational complexity. In⁵ proposed a noninvasive method to detect DM and Non Proliferative Diabetic Retinopathy (NPDR). Initially three groups of features were extracted from the images of retina. A color gamut was established with 12 colors representing the features of the images. 13 features were extracted from tongue images based on measurements, distances, areas, and their ratios represent the geometry features. Applying a combination of the 34 features, the proposed method can separate Healthy/DM tongues as well as NPDR/DM-sans NPDR (DM samples without NPDR) tongues using features from each of the three groups with average accuracies of 80.52% and 80.33%, respectively, but has high resistance to noisy images. In⁶ proposed new learning algorithm for single-hidden layer feed forward neural networks which is called Extreme Learning Machine (ELM). Both in theory and experimental results, this learning algorithm gives better generalization performances and extremely faster learning speed than traditional popular gradient based learning algorithm. In⁷ proposed an African Buffalo Optimization: A Swarm-Intelligence Technique. BO was not only able to obtain better solutions but at a faster speed. In⁸ build classification models and risk assessment tools for diabetes, hypertension and comorbidity using machine-learning algorithms. In this work, the main objective is to classify the data as diabetic or non diabetic and improve the classification accuracy. For many classification problem, the higher number of samples chosen but it doesn't leads to higher classification accuracy. In many cases, the performance of algorithm is high in the context of speed but the accuracy of data classification is low. The main objective of the proposed model is to achieve high accuracy. Classification accuracy can be increase if we use much of the data set for training and few data sets for testing⁹. This survey has analyzed various classification techniques for classification of diabetic and non diabetic data. In¹⁰ proposed an improved learning algorithm for classification which is referred to as voting based ELM. It integrates the voting method into the ELM for webpage quality classification applications.

In order to improve the quality as well as accuracy there is a need for algorithm. The ELM¹¹ is used for intelligent classification purpose. Recently, ELM has fascinated the attention of many researchers in different applications. ELM is an advancement of single layer feed forward neural network which is an improved version of standard feed forward neural network.

In this work Hybrid Extreme Learning Machine (HELM) is proposed for the classification of diabetes as diabetic and non-diabetic¹² by the combination of African Buffalo Optimization attempts to develop a totally-new algorithm that will demonstrate exceptional capacity in the exploitation and exploration of the search space. African Buffalo Optimization (ABO) works on updating the position of the best buffalo to avoid early convergence or stagnation and in the case where the best buffalo location is not improved in a number of iterations, the entire herd is re-initialized. Similarly ABO ensures fast convergence with its use of very few parameters. The advantage of ELM is obvious in shorter training time and in compact model size (i.e., computer memory to store the trained model) while the generalization of ELM is comparable to that of SVM. In this work, the performances of HELM (with or without prior duplication) in different aspects were evaluated by comparing the results with Support Vector Machine (SVM), ELM and Transductive Extreme Learning Machine (TELM).

This paper organized as follows: Section II describes about proposed methodology in combination of ABO and ELM, Section III deals about experimental results for data used from diabetes mellitus, and Section IV concludes the paper.

2. Proposed Methodology

The proposed new framework of HELM based ABO an attempt compared with the existing algorithms to solving limitations earlier algorithms, especially the problems of accuracy and inefficiency. The ABO was implemented to update the current positions of population in the discrete searching space, for the better classification purpose. In the second stage, the effective and efficient ELM classifier is conducted based on the optimal feature subset obtained in the first stage.

2.1 Extreme Learning Method

ELM mainly applied for Single Hidden Layer Feed forward Neural Networks (SLFNs) it is the process of randomly selecting the input weights and systematically determines the output weights of SLFNs. This algorithm tends to the best generalization performance at extremely fast learning speed⁴³. ELM contains the three layers they are input layer, hidden layer and an output layer. ELM has several significant features which are differ from traditional learning algorithms applied for feed forward neural networks. The learning speed of ELM could be completed in seconds or less than seconds for many traditional applications. In traditional algorithm there exists a virtual speed barrier in which the algorithms cannot process and it is not unusual way to take long time for train a feed-forward network using classic learning algorithms for uncomplicated applications. The ELM has better simplification performance compared with gradient based learning algorithms such as back propagation. The gradient based learning algorithms and some other learning algorithms may face many issues such as local minima, improper learning rate and over fitting, etc. The methods are implemented to overcome the above issues such as weight decay and stopping methods.

In real applications, the number of hidden N nodes will always be less than the number of training samples N and the training error cannot be made exactly zero but can be a nonzero training error ϵ . The hidden node parameters a_i and (input weights and biases or centers and impact factors) of ELM need not be tuned during training and may simply assigned with random values according to continuous sampling distribution. If the number of neurons in the hidden layer is equal to the number of samples, then H is square and invertible. Otherwise, the system of equations needs to be solved by numerical methods, concretely by solving

$$\|H(w_1, \dots, w_M, b_1, \dots, b_M)\hat{\beta} - T\| = \min_{\beta} \|\beta - T\|$$

The result that minimizes the norm of this least squares equation is

Where was the Moore-Penrose generalized inverse of matrix H ?

The three important properties are

- Minimum training error.

- Smallest norm of weights and best generalization performance.
- The minimum norm least-square solution of $H\beta = T$ is unique,

Give a training set

$N = \{X_1, \dots, X_N \mid X_1 \in \mathbb{R}^n, t_1 \in \mathbb{R}^m, t_1 = 1, \dots, N\}$ activation

function $g(x)$ and hidden neuron , do the following

- Assigning random value to the input weight w_i and the bias $b_i, i=1, \dots, N$
- Find the hidden layer output matrix H .
- Calculate the output weight β

Figure 1. Algorithm of ELM.

2.2 African Buffalo Optimization

ABO is a simulation of the alert ('*maaa*') and alarm ('*waaa*') calls of African buffalos in their foraging assignments. The *waaa* calls is used to warn the buffalos about the presence of predators, ward off an approaching inferior, assert dominance or express the lack of pastures in a particular area and therefore urge the herd to move on to safer or more rewarding areas (exploration). Whenever this call is made, the animals are asked to be alert and to seek a safer or better grazing field. The *maaa* calls is used to encourage the buffalos to be relaxed as there are good grazing fields around, reassure an inferior and to express satisfaction about the amount of pastures cum favorable grazing atmosphere at a particular location (exploitation). The buffalos are able to optimize their search for food source. The ABO is a population based algorithm in which individual buffalos work together to solve a given problem. Using the *waaa* (move on) signal or the *maaa* (hang around) signal, the animals are able to obtain amazing solutions in their exploration and exploitation of the search space.

The algorithm starts by initializing the population of buffalos with the function $f(x)$. The location allocation is random within the N -dimensional space for each buffalo. After allocating, it updates the buffalo's fitness separately within the search space. The following two factors vary based on the fitness value, if the fitness is better than the individual buffalo's maximum fitness (bpmax); it saves the location vector for the particular buffalo. In another case, if the fitness is better than the herd's maximum, then it saves it as the herd's maximum (bgmax). After completing all process the algorithm m updating the best buffalo

after that it moves on to validate the stopping criteria. Finally, if our global best fitness meets termination criteria, it gives the location vector as the solution to the above issues.

- Step1.** Objective function $f(x)$ $x=(x_1, x_2, \dots, x_n)^T$
- Step2.** Initialization: randomly place buffalos to nodes at the solution space;
- Step3.** Update the buffalos fitness values by following equation

$$W_{k+1} = w_k + lp_1 (bg_{max,k} - m_k) + lp_2 (bp_{max,k} - m_k)$$
 Where w_k and m_k represents the exploration and exploitation moves respectively of the k^{th} buffalo ($k=1, 2, \dots, N$); lp_1 and lp_2 are learning factors; r_1 and r_2 are random numbers between $[0, 1]$;
 bg_{max} is the herd's best fitness and bp_{max} , the individual buffalo's best
- Step3.** Update the location of buffalo k in relation to $bp_{max,k}$ and $bg_{max,k}$ using $m_{k+1} = \lambda (w_k + m_k)$. Where ' λ ' is a unit of time
- Step5.** Check bg_{max} is updating or not. If yes, go to 6 else, go to 2
- Step6.** If the stopping criteria is not met, go back to algorithm step 3
- Step7.** Output best solution.

Figure 2. African Buffalo Optimization (ABO) algorithm.

It is observed that the algorithm's movement has three parts as shown in Figure 4. Initially ' w_k ' represents the memory of the buffalos past location. A list of solutions represents the memory of each buffalo that can be used as an alternative for the current local maximum location. There is a probability of choosing one of the target lists of solutions of the buffalo's memory instead of the present herd's maximum point. Secondly $lp_1r_1 (bg_{max,k} - m_k)$ is concerned with the Cooperative part of the animals buffalos and is a pointer to the buffalo's social and information-sharing experience. Finally the third part $lp_2r_2 (bp_{max,k} - m_k)$ indicates the intelligence part of the buffalos. Hence the ABO exploits the memory and efficient caring capabilities of the buffalos in arriving at solutions.

The Pima Indian Diabetes Data (PIDD) set is collected and used in HELM as data set. The labeled dataset is and divided into training and training sets to train the HELM with overall test performance. The ELM models for classification have been developed for the classifica-

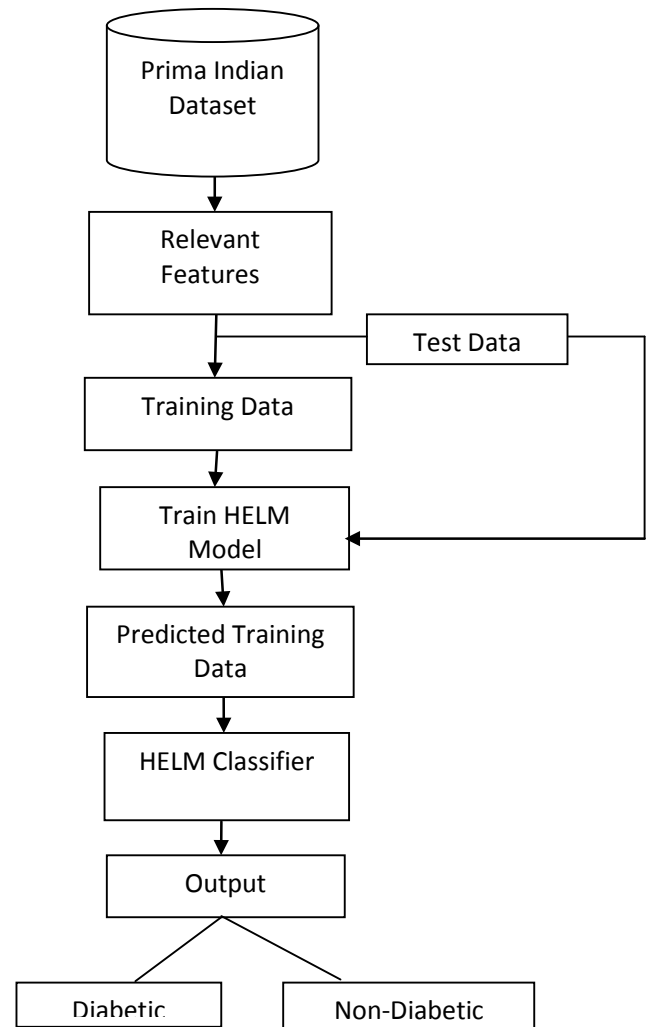


Figure 3. Proposed HELM selection model.

tion of diabetes dataset¹⁵. At last stage the classifier result as diabetic and Non-diabetic dataset is obtained with best solution. Here the training set is a straight forward technique which involves the input layer and random weight. The output weight is computed by updating of fitness value.

3. Experimental Results

It is important to conduct a set of experiments to set parameters and examine the effectiveness of our proposed user identification system based on clustering. This process is made in terms of recommendation accuracy and quality. In order to check the performance of the proposed algorithm ABO, a real-time dataset is applied in our simulations. The experimental results for proposed

HELM are carried out using PIMA dataset. PIDD set is available publicly from the machine learning database at UCI repository which is classified under two method. This dataset consist only females at the age of 21 of Pima Indian heritage living near Phoenix, Arizona². This data set is extracted from a larger database originally owned by the National Institute of Diabetes and Digestive and Kidney Diseases. The model is analyzed in two steps: at first model the datasets are trained and tested and the classifier is used to identify the diabetic and non-diabetic dataset. The proposed method HELM is executed using accuracy and execution time in PIMA dataset. Table 1 shows the performance evaluation of parameters among classifiers for the SVM, ELM and the proposed TELM and HELM classifiers.

Table 1. Performance evaluation of classifiers

Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F-Measure	Execution time (seconds)
SVM	79	85	71	0.74	45
ELM	94	79	82	0.79	21
TELM	96.5	87	74	0.85	19
HELM	98	94	93	0.92	12

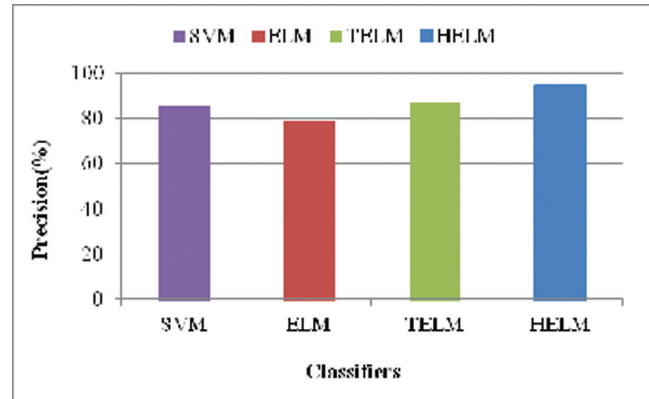


Figure 5. Percentage of precision in classifiers.

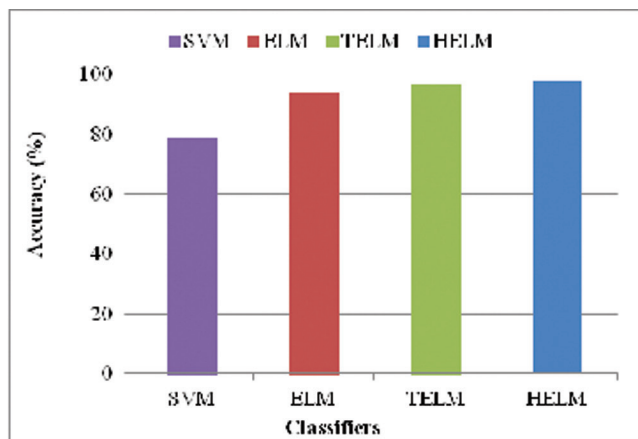


Figure 4. Percentage of accuracy in classifiers.

Figure 4 and Figure 5 shows the accuracy and precision percentage of SVM, ELM and the proposed TELM and HELM classifiers for our datasets. It shows that proposed HELM have high accuracy and the result shows that HELM is better than other classifiers in classification.

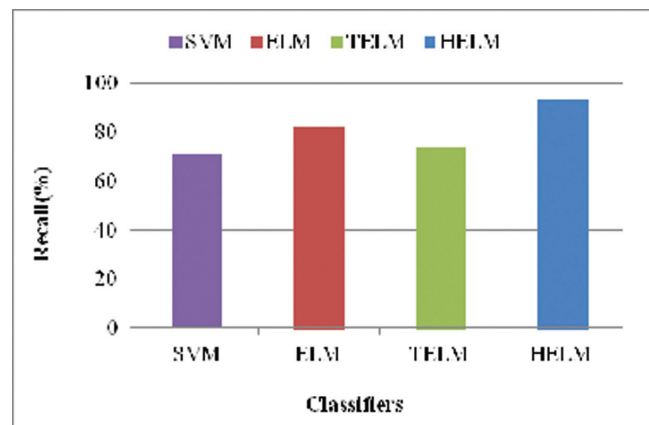


Figure 6. Percentage of recall in classifiers.

Figure 6 and Figure 7 shows the Recall and F-Measure percentage calculated for SVM, ELM and the proposed TELM and HELM classifiers. The execution time taken by the classifiers for classification for our dataset is compared and shown in Figure 8. It is observed that the execution time is gradually decreases for current HELM model and it performs than other classifiers.

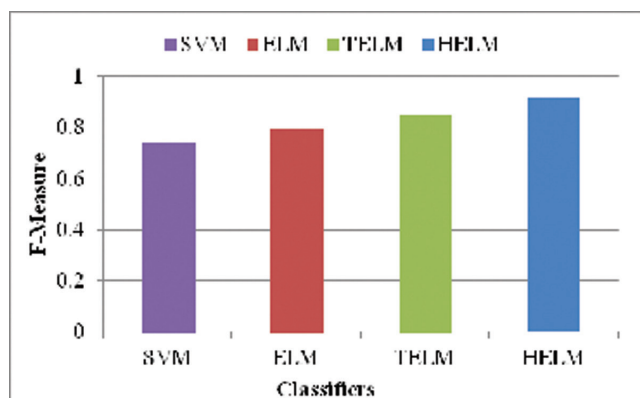


Figure 7. Percentage of F-measure in classifiers.

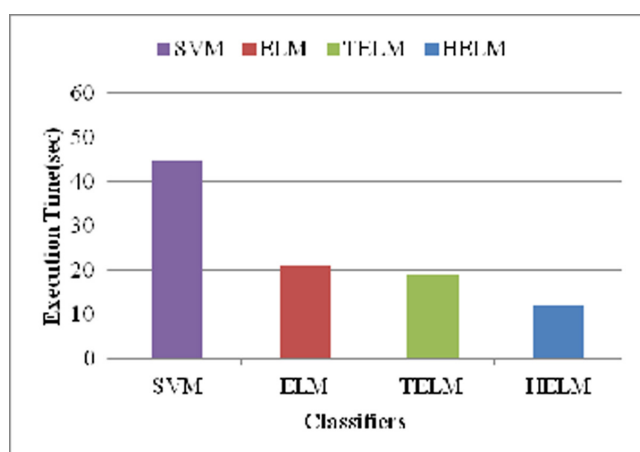


Figure 8. Comparison of execution time in classifiers.

4. Conclusion

Diabetes Mellitus is diagnosed by implementing the machine learning classification as hybrid ELM-ABO is proposed. The patients' data are classified by HELM as diabetic data and non-diabetic data. The performance parameters such as the classification accuracy, precision, recall, F-measure and execution time show high performance for proposed HELM. From the obtained results it is concluded that proposed HELM shows better results when compared with traditional techniques with high classification accuracy and less execution time.

5. References

1. Noronha K, Acharya UR, Nayak KP, Kamath S, Bhandary SV. Decision support system for diabetic retinopathy using discrete wavelet transforms. In the Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2013 Mar; 227(3):251–61. Crossref.
2. Erkeymaz O, Ozer M. Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. Chaos, Solitons and Fractals, Elsevier, ScienceDirect. 2016 Feb 29; 83:178–85. Crossref.
3. Torkaman M, Bagheri G, Ahmadi A, Saburi A. A comparative study of blood glucose measurements using glucometer readings and the standard method in the diagnosis of neonatal hypoglycemia. Iranian Journal of Neonatology (IJN). 2016 Mar 20; 7(1):41–6.
4. Karan O, Bayraktar C, Gümüşkaya H, Karlık B. Diagnosing diabetes using neural networks on small mobile devices. Expert Systems with Applications, Elsevier, ScienceDirect. 2012 Jan 31; 39(1):54–60. Crossref.
5. Zhang B, Kumar BV, Zhang D. Detecting diabetes mellitus and non proliferative diabetic retinopathy using tongue color, texture, and geometry features. Institute of Electrical and Electronics Engineers (IEEE) Transactions on Biomedical Engineering. 2014 Feb; 61(2):491–501. Crossref.
6. Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feedforward neural networks. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) International Joint Conference, Neural Networks. 2004 Jul 25–29; 2:985–90.
7. Odili JB, Kahar MNM, Anwar S. African Buffalo Optimization (ABO): a swarm-intelligence technique. Procedia Computer Science, Elsevier, ScienceDirect. 2015 Jan 1; 76:443–8. Crossref.
8. Olatunji SO, Selamat A, Raheem AAA. Modeling permeability prediction using extreme learning machines. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, Bornea, Malaysia; 2010 May 26–28. p. 29–33. Crossref.
9. Bin-Sabbar MS, Al-Rodhaan MA. Diabetes monitoring system using mobile computing technologies. International Journal of Advanced Computer Science and Applications (IJACSA). 2013 Feb; 4(2):23–31.
10. Cao J, Lin Z, Huang GB, Liu N. Voting based extreme learning machine. Information Sciences, Elsevier, ScienceDirect. 2012 Feb 15; 185(1):66–77. Crossref.

11. Liang Y, Niu D, Cao Y, Hong WC. Analysis and modeling for china's electricity demand forecasting using a hybrid method based on multiple regression and extreme learning machine: a view from carbon emission. *Energies*. 2016 Nov 11; 9(11):941. Crossref.
12. Faust O, Acharya R, Ng EYK, Ng KH, Suri JS. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *Journal of Medical Systems*, Springer. 2012 Feb 1; 36(1):145–57. Crossref.
13. Chen HL, Wang G, Ma C, Cai ZN, Liu WB, Wang SJ. An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease. *Neurocomputing*, Elsevier, ScienceDirect. 2016 Apr 5; 184:131–44. Crossref.
14. Emary E, Zawbaa HM, Grosan C, Hassenian AE. Feature subset selection approach by gray-wolf optimization. In *Afro-European Conference for Industrial Advancement, Advances in Intelligent Systems and Computing*, Springer. 2015; 334:1–13. Crossref.
15. Leung RK, Wang Y, Ma RC, Luk AO, Lam V, Ng M, So WY, Tsui SK, Chan JC. Using a multi-staged strategy based on machine learning and mathematical modeling to predict genotype-phenotype risk patterns in diabetic kidney disease: a prospective case-control cohort analysis. *Bio Med Central (BMC) Nephrology*. 2013 Jul 23; 14(1):162. Crossref.