Research Article

Optimal Feature Subset Selection with Multi-Kernel Extreme Learning Machine for Medical Data Classification

[1]M. Y. MOHAMED PARVEES, [2]M. RAJA

[1]Department of Computer and Information science, Annamalai University, Annamalainagar – 608002, India. yparvees@gmail.com
[2]Research Scholar, Department of Computer and Information science, Annamalai University, Annamalainagar – 608002, India. mraja.cdm@gmail.com
Article History: *Do not touch during review process(xxxx)*

Abstract: Medical data classification is treated as a crucial process in the domain of medical informatics. The recently developed machine learning (ML) algorithms are found useful for the medical diagnosis. This paper presents a new ML based medical data classification model using the whale optimization algorithm (WOA) based on feature selection with Multi-kernel extreme learning machine (MKELM) model, called WOA-MKELM. The proposed model could perform the medical data classification using two processes, namely the WOA based FS and the MKELM based classification. In the first stage, the WOA-FS algorithm is executed on the medical data to generate the feature reduced subset. In the second stage, the MKELM algorithm is applied to allocate the appropriate class labels for the medical data. The incorporation of WOA-FS model, prior to classification, helps to increase the detection rate significantly. The performance of the WOA-MKELM model has been tested using three medical datasets, namely hepatitis, UCI-Indian Liver Patient (UCI-ILD), and thyroid. The obtained experimental could values verify the effectiveness of the WOA-MKELM model with the maximum accuracy of 98.36%, 98.72% and 98.93% for the applied hepatitis, UCI-ILD and thyroid dataset respectively.

Keywords: Medical data classification, Healthcare, Disease diagnosis, Feature selection, Machine learning

1. Introduction

Recently, Clinical Decision Support System (CDSS) has been considered a significant application in healthcare sector for guiding the physicians to make medical based decisions **Sumalatha G, Muniraj NJR (2013)**.Machine Learning (ML) and Data Mining (DM) models have been employed for different types of clinical data classification. Health data classification depends upon the learning mechanism where the records are acquired from various datasets and aim in enhancing the quality of DSS. The documents are comprised of greater association in eliminating the unwanted noises. It is significant to accomplish pre-processing phases for resolving the complexities involved in clinical data. However, a dataset is assumed to be non-identical, and so no pre-processing is applied. The pre-processing method for specific dataset is impossible with no maintenance and estimations. In this framework, generalized structure is composed of clinical data referred as input, data pre-processing steps, and the DM models for understanding the historical data to classify the unknown data. Eliminating the noise from data, as well as feature subset selection, is essential pre-processing phases to enhance the working function of a classifier. ML is a significant step in developing knowledge discovery. The major constraint of ML is to identify the new patterns for the individuals and to disturb the data patterns for providing significant and advantageous for enhancing the decision-making operation. It is employed in identifying the meaningful patterns to guide the significant operations of clinical diagnosis and treatments. The Classification, Clustering, and soft computing methods are employed in the clinical DM.

The count of soft computing relied classifiers, presented and examined, in this study, help in categorizing the medical data significantly. **Abbass HA. (2002)** presented a system along with pareto-differential estimation approach using local search method, namely the Memetic Pareto Artificial Neural Network (MPANN), for diagnosing breast cancer. It is followed by, **Kiyan T, Yildirim T. (2004)** which proposed a statistical NN relied method for diagnosing breast cancer. **Karabatak M, Ince MC. (2009)** deployed the professional mechanism helps detect breast cancer, in which, to limit the dimensions of a dataset. The association Rules (AR) have been applied. A hybrid Feature Selection (FS) model is employed for addressing the problems of maximum dimensionality of biomedical information and performed on breast cancer dataset. **Fana C-Y, Changb P-C, Linb J-J, Hsiehb J. (2011)** unified the case-based data clustering as well as Fuzzy Decision Tree (FDT) for developing a hybrid approach for clinical data classification. This method is implemented on 2 datasets, namely Wisconsin Breast Cancer WBC and liver infections. In 3 classifiers, such as Radial Basis Function (RBF), Multilayer Perceptron (MLP), as well as Probabilistic Neural Network (PNN), is processed on breast cancer dataset. Here, the PNN has exhibited better performance when compared with the MLP.

Then, **Anooj PK. (2012)** has applies weighted fuzzy rules for making the CDSS in order to predict Heart Disease (HD). Initially, fuzzy rules, relied on traditional data, have been generated to perform better learning, and to deploy the CDSS model. Additionally, the fuzzy rules are weighted on the basis of important parameters. Then **Samb ML, Camara F, Ndiaye S, Slimani Y, Esseghir MA. (2012)** presented an extended version of support vector machine with recursive feature elimination (SVM-RFE) and performed experiments on various clinical datasets (SPECT Heart Data). Moreover, the local search process is embedded within the method. The feature selection is done by applying Fuzzy Entropy (FE). The classical task performed feature selection concept related to Kernel F-Score. **Samb ML, Camara F, Ndiaye S, Slimani Y, Esseghir MA. (2012)** deploys a hybrid scheme by applying K-Nearest Neighbour (KNN) as well as Genetic Algorithm (GA). The intelligent CDSS has resulted in evolutionary principle where it has applies the NN models like GA, SVM, KNN, MLP, RBF, PNN, Self-Organizing Map (SOM), as well as the Naive Bayes (NB) as the classification models. **Jabbar MA, Deekshatulu B, Chandra P (2013)** establishes a CDSS and performs on clinical datasets which depends upon the experiments and indicate the SVM as the best classifier for making CDSS. An effective clinical data classification approach is relied on Adaptive Genetic Fuzzy System (AGFS). Here, the rules are generated from the data initially and optimize the rules selection which is processed by applying the GA. **Khanmohammadi S, RezaeiahariM. (2014)** develops a hybrid intelligent scheme and perform experiments on breast cancer, Diabetes, Liver cancer datasets, and so on. A breast cancer detection and developed weighted classification, based on association rules (WCBA), is a proficient Weighted Classification (Association rules) approach. Therefore, it is implied that WCBA, performs well when compared with the alternate classifiers and Association Classification (AC) methods. Many studies have defined with limited count of datasets.

Seera M, Lim CP. (2014) recommends a data analysis pattern for dimensional reduction process. In order to verify the class level for the applied data, it depends upon the adaptive arrangement model. The applied system gains Eigenvector and Eigen matrix for high dimensional reduction by applying the principal component analysis (PCA) based on mass updating. **Kuncheva LI, Faithfull WJ (2014)** provides statistical principles for classifying the feature extracted by PCA. It follows the, dimensionality reduction method which limits the features by using the feature extraction and computes the data classification with maximum accuracy. **Jayanthi SK, Sasikala S (2014)** applies a model for clinical data categorization on the basis of Orthogonal Local Preserving Projection (OLPP). The OLPP with a classifier has been applied to enhance the results in clinical DM. In classification, the artificial bee colony (ABC) approach is repeated with the ANN. Thus, the dimension reduction, applied, incurs no loss of accuracy.

Tarle B, Jena S (2017) projects the integration of Bays theorem as well as the balanced unlikelihood relied on cross parameter selection approach. The present model has been used on hyperactive pressure analysis issues. Thus, in the entire function, the classification has been improvised. **Park HW, Li D, Piao Y, Ryu KH (2017)** projects major applications of FS such as filters and wrappers with Particle Swarm Optimization (PSO) for the medical records. The relationship of attribute selection models depends upon the function with GA. Hence, the 2 models are applied for medical records sets and showcase the classification accuracy (CA). **Harb HM, Desuky AS (2014)** presents the best cross attribute selection method to find the applicable Single Nucleotide Polymorphisms and select optimal SNP subclass. The developed approach depends upon the combination of filter and wrapper. Under the application of Conditional Mutual Information Maximization, the SVM attribute elimination is computed.

2. Significance of the Study

This paper presents a new machine learning (ML) based medical data classification model using the whale optimization algorithm (WOA) based feature selection (FS) with Multi-kernel extreme learning machine (MKELM) model, called WOA-MKELM. The presented model uses the WOA-FS algorithm on the medical data to generate the feature reduced subset. Then, the MKELM algorithm is applied to allocate the appropriate class labels for the medical data. The incorporation of WOA-FS model, prior to classification, helps to increase the detection rate significantly. The performance of the WOA-MKELM model has been tested using three medical datasets, namely Hepatitis, UCI-Indian Liver Patient (UCI-ILD), and thyroid.

3. Objectives of the Study

- To search and consider the required medical data set, namely Hepatitis, Indian Liver Patient, Thyroid to apply efficient medical data classification algorithm.
- To propose the exact machine learning based medical data classification model according the datasets considered.

- To propose medical data classification model using the whale optimization based feature selection with Multi-Kernel extreme learning machine, called WOA-MKELM.
- To find out the experimental values to verify the effectiveness of the WOA-MKELM.

4. Hypotheses of the Study

- The Medical datasets namely Hepatitis, Indian Liver Patient, Thyroid are considered to evaluate the proposed WOA-MKELM.
- There is no significant machine learning based medical data classification model using the whale optimization based feature selection with Multi-Kernel extreme learning machine called WOA-MKELM. Therefore, yet there is no significant the experimental values to verify the effectiveness of the WOA-MKELM.

5. Methodology

In this proposed model, the features are selected from the pre-processed medical data using WOA and the selected features are utilized for categorizing them into distinct class labels using MKELM. The complete workflow of the proposed model is given in Fig. 1.

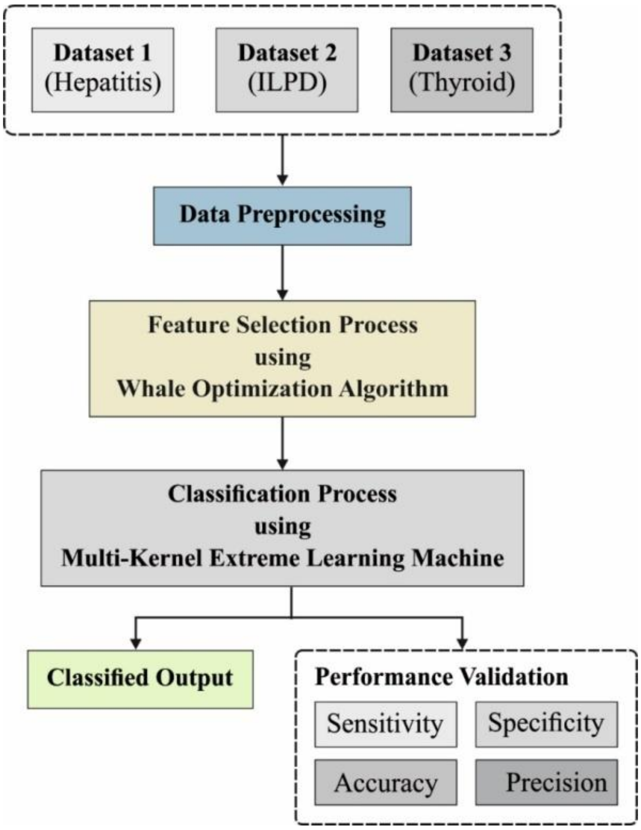

Fig. 1. The complete workflow of the proposed model

The WOA works based on the search for prey, encircle prey and bubble-net attack. The Fig. 2 depicts the complete WOA steps. The WOA is applied to identify the best feature subset that enhances the classification function. In this model, MKELM receives the input as selected features from WOA. In MKELM, rather than using a single kernel mapping, various kernels are used for combining multiple-kernel learning principles for gaining effective classification functions as given in Algorithm 1.

Algorithm 1: Multi-KELM algorithm
Input: A set of medical training samples (x_i, y_i), $i = 1, 2, N, a$ test sample x^, specified Q kernels $kq(\mathrm{x}_i, \mathrm{x}_j)$, parameters $\lambda 1, \ldots, \lambda Q$.
Output: Predicted label y.

1. Estimate the mixed kernel matrix K with(K)= $\sum_{q=1}^{Q} \lambda_q k_q (\mathrm{x}_i, \mathrm{x}_j)$ for $\forall i, j \in \{1, \ldots, N\}$
2. Estimate the mixed kernel $k(\hat{x}, x_i) = \sum_{q=1}^{Q} \lambda_q k_q (\hat{x}, \mathrm{x}_i)$ for $\forall i \in \{1, \ldots, N\}$
3. Estimate the resultant S and accomplish the predicted label$y = \mathrm{sign}(S)$.

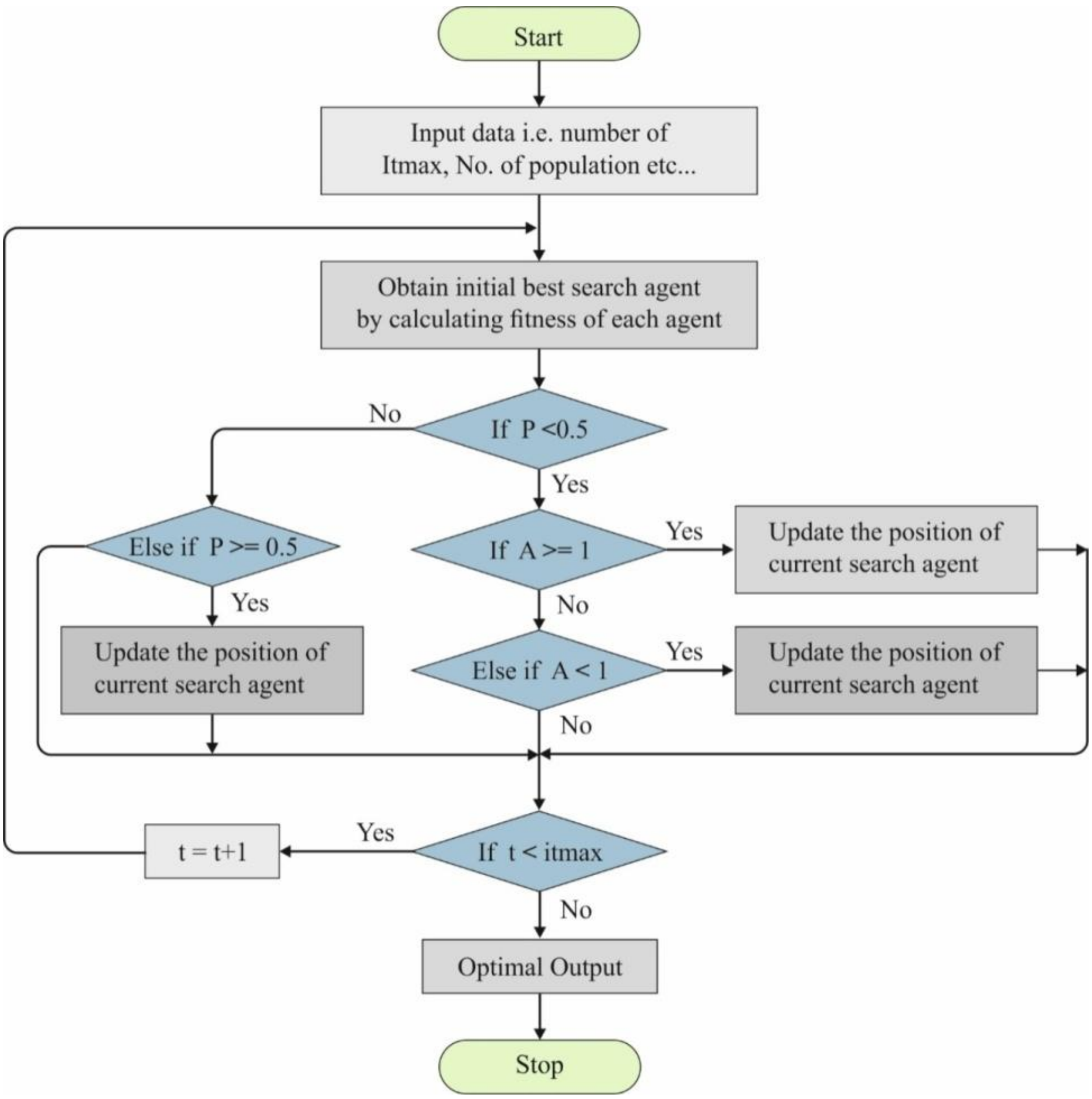

Fig. 2. The complete steps of WOA.

The experimental validation of the proposed WOA-MKELM model has been tested against three benchmark datasets, namely ILP dataset (https://archive.ics.uci.edu/ml/datasets/ILPD+(Indian+Liver+Patient+Dataset), Hepatitis dataset (https://archive.ics.uci.edu/ml/datasets/Hepatitis), and Hypothyroid dataset (http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease). The first ILP dataset includes a set of 416 instances with the presence of 11 attributes and the attributes are numbered as shown in Table 1. Similarly, the second hepatitis dataset comprises a total of 155 instances with the existence of 20 features given in Table 2. Finally, the third hypothyroid dataset includes 7200 instances with 22 attributes given in Table 3.

Table 1: Features Description of UCI-ILP Dataset

No. of Features	Feature name	Feature Description
1	Age	Age of the patient
2	Gender	Gender of the patient

3	TB	Total Bilirubin
4	OB	Direct Bilirubin
5	Alkphos	Alkaline Phosphatase
6	Sgpt	Aminotransferase
7	Sgot	Aminotransferase
8	TP	Total Proteins
9	ALB	Albumin
10	NC Ratio	Albumin and Globulin Ratio
11	Target Class	LD/NLD

Table 4 and Fig.3 investigate the FS outcome of the WOA-FS model when compared with the FS models in terms of the best cost involved. The WOA-FS model has found to be effective in the selection of optimal features with the least good cost on all the applied datasets. For instance, on the applied hepatitis dataset, the WOA-FS model has reached a minimum best cost of 0.643 whereas the GWO-FS and SA-FS models have obtained a higher best cost of 0.742 and 0.893 respectively. Similarly, on the applied UCI-ILP dataset, the WOA-FS model has reached a lower best cost of 0.598, whereas the GWO-FS and SA-FS models have gained a higher best cost of 0.869 and 0.874 respectively. At last, on the applied thyroid dataset, the WOA-FS model has accomplished the minimum best cost of 0.619, whereas the GWO-FS and SA-FS models have gained a higher best cost of 0.755 and 0.860 respectively. Table 5 shows the detailed comparative results analysis of the WOA-MKELM model on the applied hepatitis dataset.

Table 2: Features Description of Hepatitis Dataset

No. of Features	Feature Description
1	Age
2	Sex
3	Steroid
4	Antivirals
5	Fatigue
6	Malaise
7	Anorexia
8	Liver Big
9	Liver Firm
10	Spleen Palpable
11	Spiders
12	Ascites
13	Varices
14	Bilirubin
15	Alk Phosphate
16	SGOT
17	Albumin
18	Protime
19	Histology
20	Target Class

Table 3: Features Description of Thyroid Dataset

No. of Features	Feature Description
1	Age
2	Sex
3	On thyroxine
4	Query on thyroxine
5	On anti-thyroid
6	Sick
7	Pregnant
8	Thyroid surgery
9	I131 treatment
10	Query hypothyroid
11	Query hyperthyroid
12	Lithium
13	Goiter

14	Tumor
15	Hypopituitary
16	Psych
17	TSH
18	T3
19	TT4
20	T4U
21	FTI
22	Target Class

Table 4: Result analysis of existing with proposed feature selection on applied dataset.

Methods	Dataset	Best Cost	Selected Features
WOA-FS	Hepatitis	0.643	2,3,4,6,7,8,9,13,15
	UCI-ILP	0.598	1,2,3,6,7
	Thyroid	0.619	1,2,3,5,8,9,10,12,13,19,20
GWO-FS	Hepatitis	0.742	1,2,3,11,12,14,15,16
	UCI-ILP	0.869	1,3,4,5,6,7,9
	Thyroid	0.755	1,2,4,5,6,7,12,14,15,19,20
SA-FS	Hepatitis	0.893	1,2,4,5,6,8,9,11,12
	UCI-ILP	0.874	2,3,4,5,9
	Thyroid	0.860	2,3,7,8,10,12,13,14,15,18,19

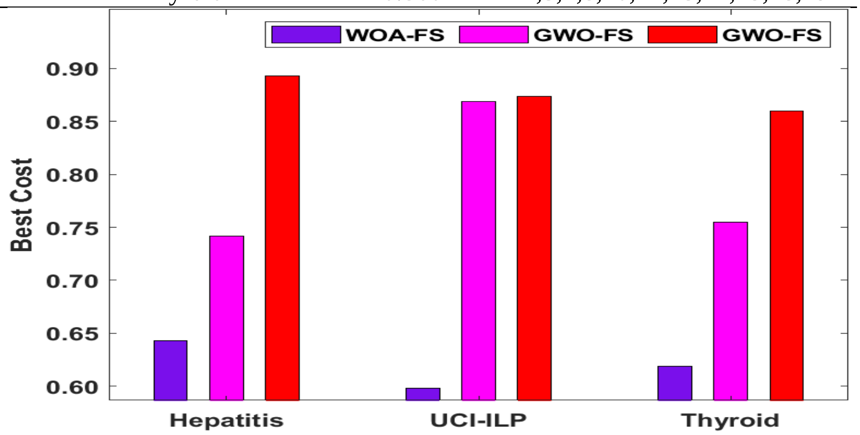

Fig. 3. Best cost analysis of the WOA-FS Model

Table 5 Result Analysis of Existing with Proposed WOA-MKELM on Hepatitis Dataset

Methods	Sensitivity	Specificity	Precision	Accuracy
WOA-MKELM	98.84	97.93	98.10	98.36
SVM-1	77.94	76.92	94.64	77.78
SVM-SA	98.50	84.61	97.05	96.25
K-NN	70.56	72.43	76.89	71.40
SVM-2	80.71	84.92	83.90	81.17
NN	76.46	80.64	79.07	78.31
ANFIS	78.34	81.54	80.15	79.67
NIPALS-SOM-ANFIS	91.09	94.65	94.65	93.06
PCA-LSSVM	93.75	95.76	95.87	95.00
PCA-AIRS	92.66	95.14	95.72	94.12

Fig.4 examines the classifier results analysis of the WOA-MKELM model, on the applied hepatitis dataset in terms of sensitivity and specificity. From the figure, it is noticed that the KNN model appear as a poor performer by offering a minimum sensitivity of 70.56% and specificity of 72.43%.Similarly, the NN model shows certainly higher performance over the KNN model by attaining a sensitivity of 76.46% and specificity of 80.64%. In line with the models cited, the SVM-1 model exhibits a slightly higher result with a sensitivity of 77.94% and specificity of 76.92%. Along with that, the ANFIS model demonstrates that even better result with a sensitivity of 78.34% and specificity of 81.54%. Eventually, the SVM-2 model accomplishes slightly higher sensitivity of 80.71% and specificity of 84.92%. Besides, the NIPALS-SOM-ANFIS model reaches a moderate classification outcome with a sensitivity of 91.09% and specificity of 94.65%. Concurrently, the PCA-AIRS model results in a manageable outcome with a sensitivity of 92.66% and specificity of 95.14%. At the same time, the PCA-LSSVM model depicts reasonable results with a sensitivity of 93.75% and specificity of 95.76%. Though the SVM-SA model accomplishes a near optimal sensitivity of 98.50% and specificity of 84.61%, the presented WOA-MKELM model outperforms the previous models with a sensitivity of 98.84% and specificity of 97.93%.

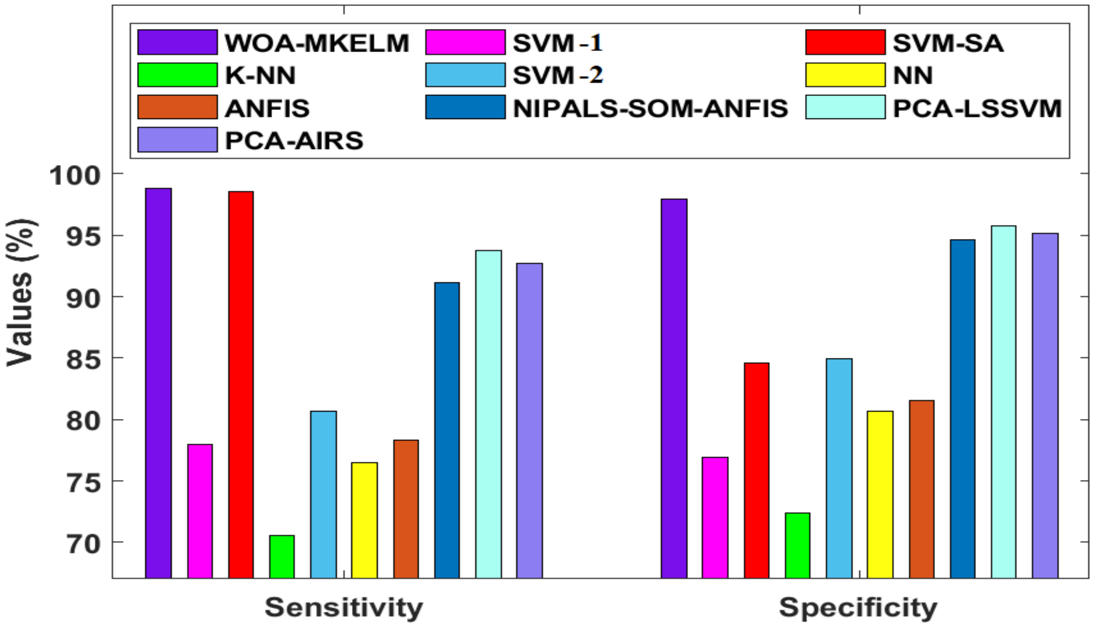

Fig.4. Sensitivity and specificity analysis of WOA-MKELM model on Hepatitis dataset

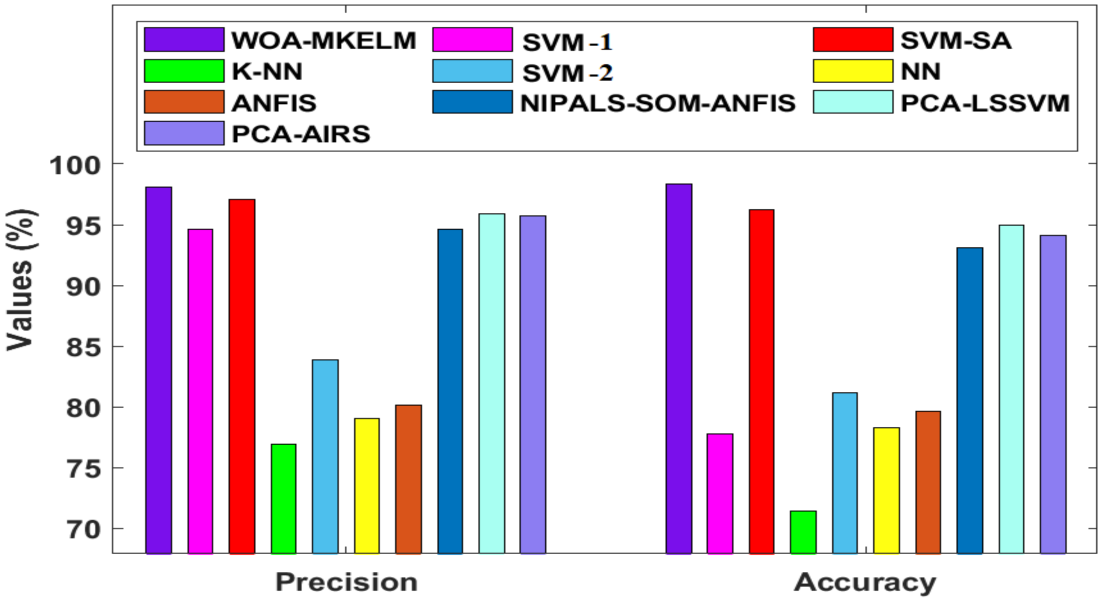

Fig. 5. Precision and accuracy analysis of WOA-MKELM model on Hepatitis dataset

Fig.5 predicts the classifier outcomes analysis of the WOA-MKELM method, on the applied hepatitis dataset, by employing precision and accuracy. From the figure, it is clear that the KNN approach exhibits inferior performers by giving the least precision of 76.89% and accuracy of 71.4%. At the same time, the NN approach has managed to imply certainly maximum performance over the KNN model by gaining a precision of 79.07% and accuracy of 78.31%.It is Follows by, the ANFIS framework which reaches moderate precision of 80.15% and accuracy of 79.67%. In line with this, the SVM-II technology has attained a considerable classification result with a precision of 83.9% and accuracy of 81.17%. Likewise, the SVM-I model implies an acceptable outcome with the precision of 94.64% and accuracy of 77.78%. At the same time, the NIPALS-SOM-ANFIS framework showcases better results with the precision of 94.65% and accuracy of 93.06%.In the meantime, the PCA-AIRS model demonstrates moderate results with the precision of 95.72% and accuracy of 94.12%. Also, the PCA-LSSVM model implements an even better result with the precision of 95.87% and accuracy of 95%. Even though the SVM-SA model gains closer optimal precision of 97.05% and accuracy of 96.25%, the proposed WOA-MKELM model surpasses the traditional methods with the precision of 98.1% and accuracy of 8.36%.

Table 6 provides comprehensive and comparative results analysis of the WOA-MKELM model on the given UCI-ILP dataset.

Table 6 Result Analysis of Existing with Proposed WOA-MKELM on UCI-ILP Dataset

Methods	Sensitivity	Specificity	Precision	Accuracy
WOA-MKELM	98.89	94.18	98.63	98.72
Random Forest	87.65	83.09	92.20	86.26
SVM	81.59	60.00	82.60	75.10
Naive Bayes	75.46	44.28	75.92	66.09
MLP Neural Network	82.53	67.16	86.16	78.11
PSO-SVM	94.93	93.33	96.77	94.42
Boosted C5.0	94.40	91.42	97.52	93.75
CHAID	75.59	24.24	79.33	65.00

Fig. 6 investigates the classifier result analysis of the WOA-MKELM model, on the applied UCI-ILP dataset, with respect to sensitivity and specificity. From the figure, it is portrayed that the NB method showcases as insignificant performer by offering a minimum sensitivity of 75.46% and specificity of 44.28%. At the same time, the CHAID method attempts to show moderate performance over the NB model by attaining a sensitivity of 75.59% and specificity of 24.24%. Similarly, the SVM technology implies the reasonable result with the sensitivity of 81.59% and specificity of 60%. In line with this, the MLPNN technique depicts manageable results with the sensitivity of 82.53% and specificity of 67.16%. Similarly, the RF model has gained better classification results with the sensitivity of 87.65% and specificity of 83.09%.In the meantime, the Boosted C5.0 model projects acceptable results with sensitivity of 94.4% and specificity of 91.42%. Although the PSO-SVM model reaches closer optimal sensitivity of 94.93% and specificity of 91.42%, the present WOA-MKELM model performs well the classical methods with the sensitivity of 98.89% and specificity of 94.18%.

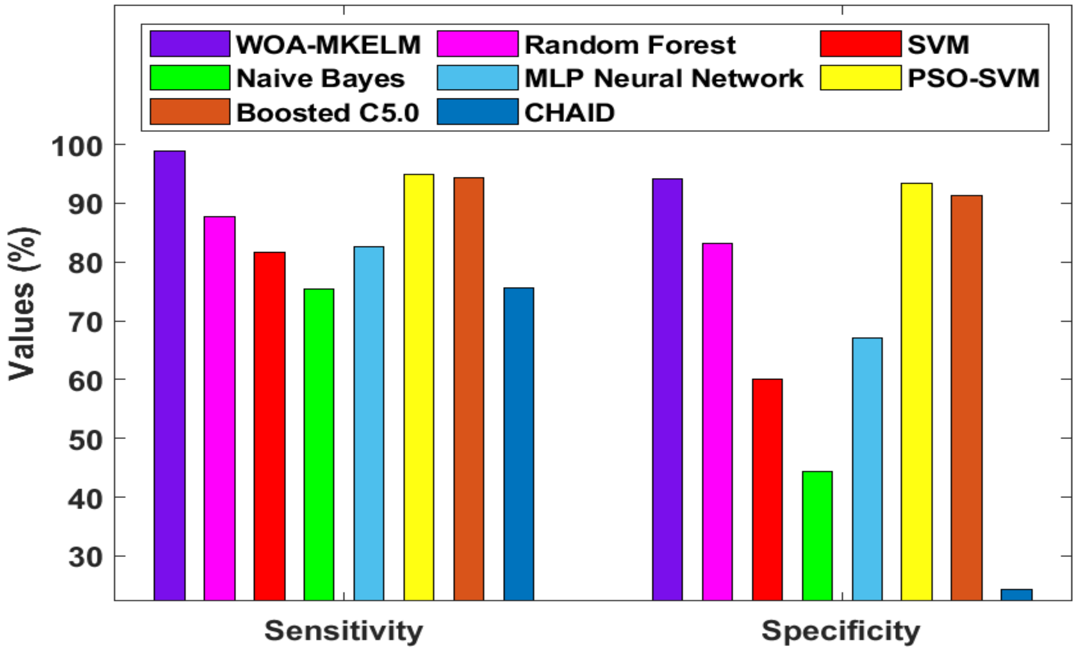

Fig.6. Sensitivity and specificity analysis of WOA-MKELM model on UCI-ILP dataset

Fig. 7 predicts the classifier outcome analysis of the WOA-MKELM method, on the given UCI-ILP dataset, by employing precision and accuracy. From the figure, it is pointed that the NB model appears as ineffective performer by offering a minimum precision of 75.92% and accuracy of 66.09%. Whereas, the CHAID model manages to display moderate performance over the NB model by reaching a precision of 79.33% and accuracy of 65%. Likewise, the SVM scheme showcases a reasonable precision of 82.6% and accuracy of 75.1%. Similarly, the MLPNN model achieves even better result with a precision of 86.16% and accuracy of 78.11%. In line with this, the RF method accomplishes a considerable classification result with a precision of 92.2% and accuracy of 86.26%. Similarly, the PSO-SVM approach offers better outcome with a precision of 96.77% and accuracy of 94.42%. Though the Boosted C5.frameworkreaches a near optimal precision of 97.52% and accuracy of 93.75%, the present WOA-MKELM model surpasses the previous models with a precision of 98.63% and accuracy of 98.72%.

Table 7 shows brief comparative results analysis of the WOA-MKELM method on the thyroid dataset. Fig.8presents the classifier results analysis of the WOA-MKELM model on the applied thyroid dataset in terms of sensitivity and specificity. From the figure, it is noticed that the IGWO+RBF-SVM model exhibits ineffective performer by offering a minimum sensitivity of 78.9% and specificity of 81.17%. At the same time, the IGWO+ANN model tries to show certainly higher performance over the IGWO+RBF-SVM model by gaining a sensitivity of 81.17% and specificity of 75.18%. Similarly, the IGWO+MKSVM model reaches a better classification outcome with a sensitivity of 90.05% and specificity of 94.5%. In the meantime, the IGWO+Linear-SVM model results in a manageable result with a sensitivity of 94.58% and specificity of 90.46%. The presented WOA-MKELM model performs well compared to the other approaches which achieve sensitivity of 94.87% and specificity of 99.12%.

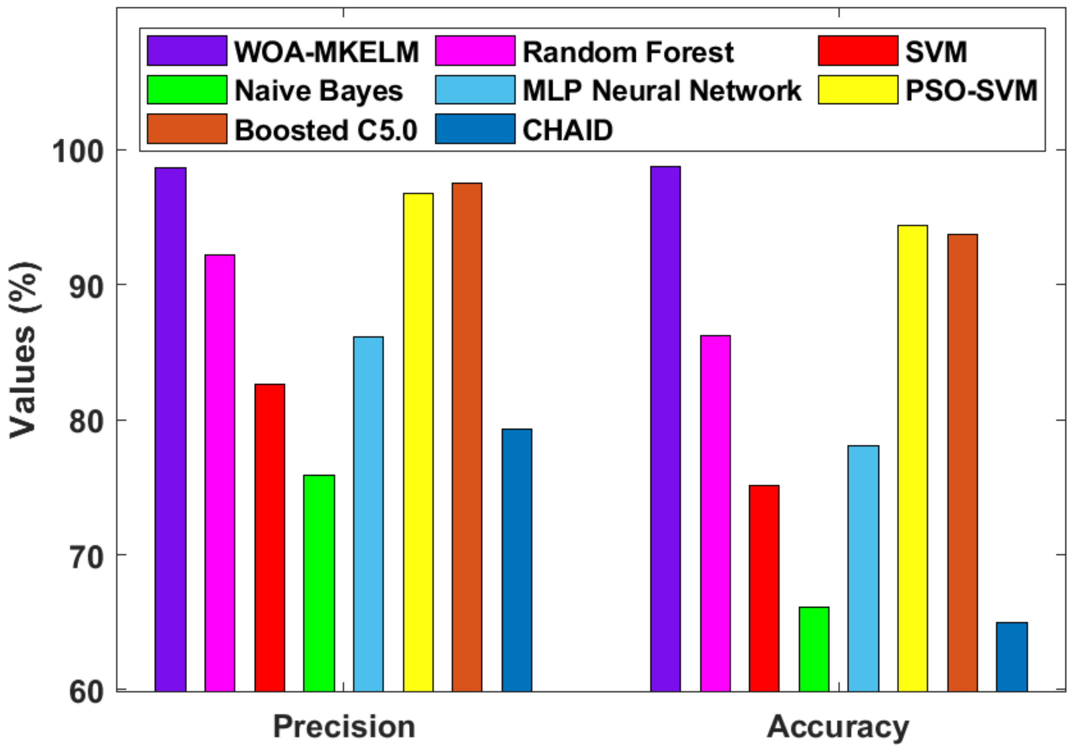

Fig. 7. Precision and accuracy analysis of WOA-MKELM model on UCI-ILP dataset

Table 7 Result Analysis of Existing with Proposed WOA-MKELM on Thyroid Dataset

Methods	Sensitivity	Specificity	Precision	Accuracy
WOA-MKELM	94.87	99.12	94.65	98.93
IGWO+MKSVM	90.05	94.50	79.11	97.49
IGWO+Linear-SVM	94.58	90.46	78.69	93.96
IGWO+RBF-SVM	78.90	81.17	68.79	78.49
IGWO+ANN	81.17	75.18	72.71	79.11

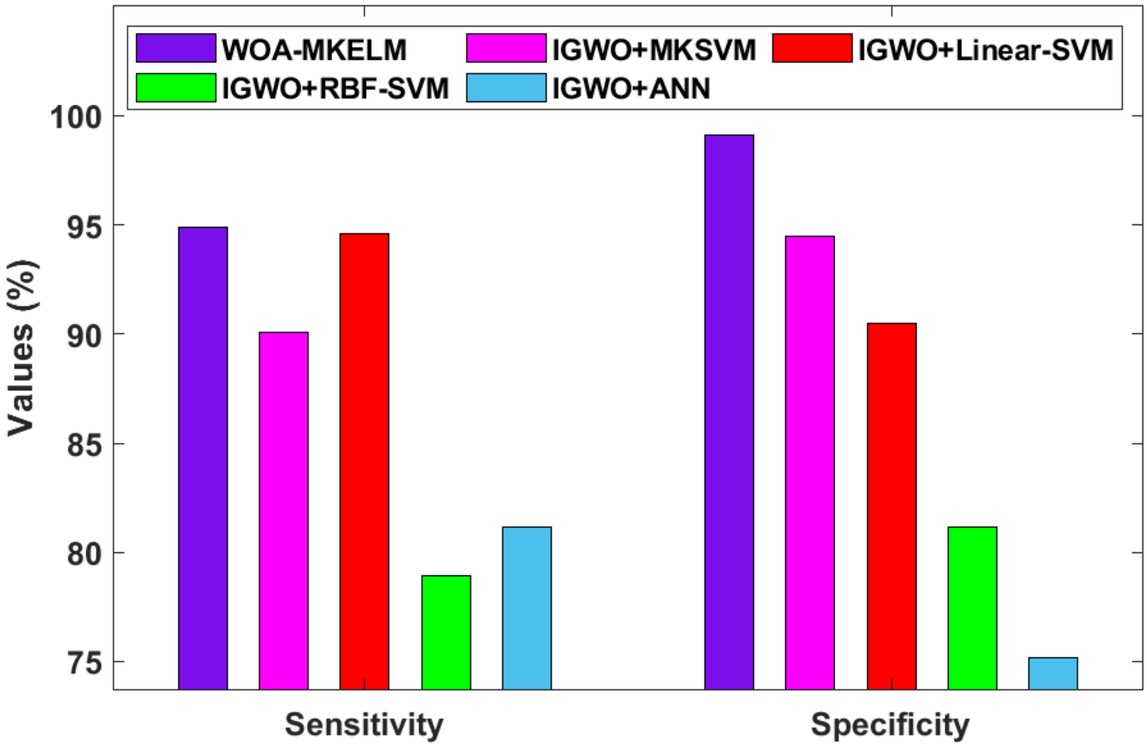

Fig. 8. Sensitivity and Specificity analysis of WOA-MKELM model on Thyroid dataset

Fig. 9 projects the classifier results analysis of the WOA-MKELM method on the applied thyroid dataset by employing precision and accuracy. From the figure, it is clear that the IGWO+RBF-SVM model appears as a poor performer by offering a minimum precision of 68.79% and accuracy of 78.49%. Concurrently, the IGWO+ANN model tries to show certainly moderate performance over the IGWO+RBF-SVM model by accomplishing a precision of 72.71% and accuracy of 79.11%.

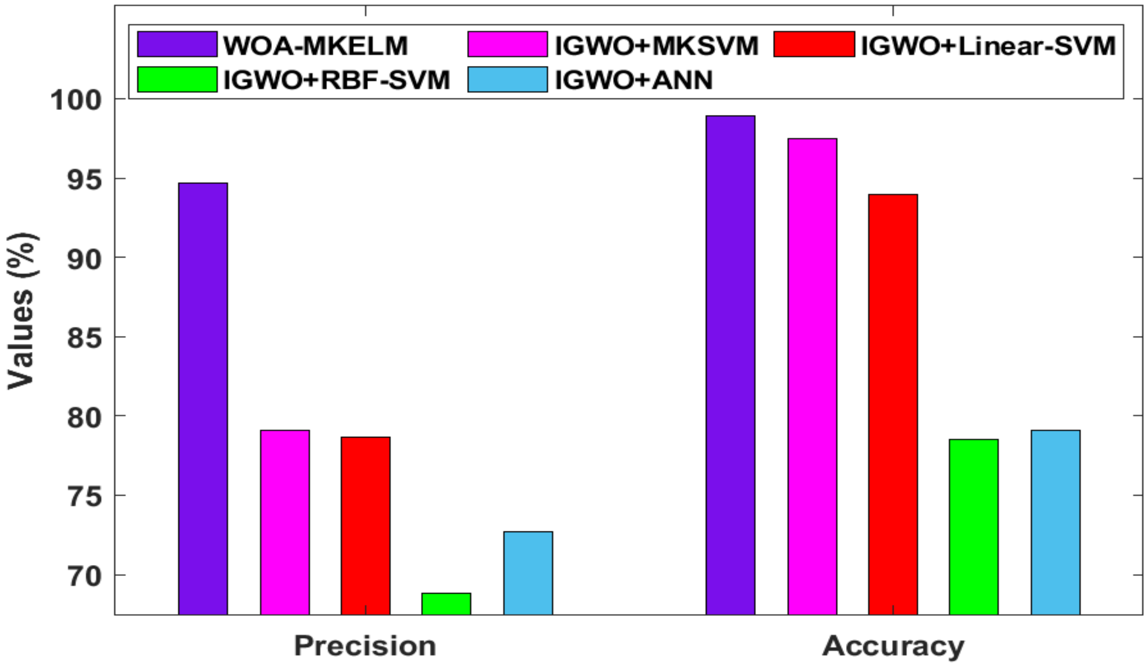

Fig. 9. Precision and accuracy analysis of WOA-MKELM model on Thyroid dataset

The IGWO+Linear-SVM model reaches to considerable classification outcome with a precision of 78.69% and accuracy of 93.96%. Meanwhile, the IGWO+MKSVM model attains a reasonable outcome with the precision of 79.11% and accuracy of 97.49%. The present WOA-MKELM model surpasses the classical methods with a precision of 94.65% and accuracy of 98.93%.

6. Conclusion

This paper has presented a ML based medical data classification model using the WOA-MKELM. The input medical data is primarily pre-processed to improve the data quality which is followed by WOA based FS model applied to choose the optimum subset of features. At last, the MKELM model is applied to categorize the feature subset to distinct class labels. The incorporation of WOA-FS model, prior to classification, helps to increase the detection rate significantly. The performance of the WOA-MKELM model has been tested using three medical dataset, namely hepatitis, UCI-ILD, and thyroid. The obtained experimental values verify the effectiveness of the WOA-MKELM model with the maximum accuracy of 98.36%, 98.72% and 98.93% on the applied hepatitis, UCI-ILD and thyroid dataset respectively. As the part of future scope, the classifier results can be enhanced using deep learning (DL) models.

References

1. Sumalatha, G., & Muniraj N.J.R. (2013). Survey on Medical Diagnosis Using Data Mining Techniques. *In: IEEE proceedings of international conference on optical imaging sensor and security*, 2 - 3 July 2013, Coimbatore, Tamil Nadu, India.
2. Abbass, H. A. (2002). An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artif Intell Med.,* 25(3):265-281.
3. Kiyan, T., & Yildirim, T. (2004). Breast cancer diagnosis using statistical neural networks. *J Electr Electron Eng.*, 4(2):1149-1153.
4. Karabatak, M. & Ince, M.C. (2009). An expert system for detection of breast cancer based on association rules and neural network. *Expert Syst Appl.*, 36(2):3465-3469.
5. Fana, C.Y., Changb, P.C., Linb, J.J., & Hsiehb, J. (2011). A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. *Appl Soft Comput*., 24: 632-644.
6. Anooj, P.K. (2012). Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules. *J. King Saud Univ. Comput. Inf. Sci.*, 2012; 11(1): 27-40.
7. Samb, M.L., Camara, F., Ndiaye, S., Slimani, Y., Esseghir, M.A. (2012) A novel RFE-SVM-based feature selection approach for classification, *Int J Adv Sci Technol.*, 43. doi. 10.1.1.641.826
8. Jabbar, M.A., Deekshatulu, B., Chandra, P. (2013). Classification of heart disease using k-nearest neighbor and genetic algorithm. *Elsevier Procedia Technology*, 10: 85–94.
9. Khanmohammadi, S., & Rezaeiahari, M. (2014). AHP based classification algorithm selection for clinical decision support development. *Elsevier Procedia Computer Science*, 36: 328–34.
10. Seera, M., Lim, C.P. (2014). A hybrid intelligent system for medical data classification. *Expert Syst Appl.*, 41: 2239-2249.
11. Kuncheva, L.I., & Faithfull, W.J. (2014). PCA feature extraction for change detection in multidimensional unlabeled data. *IEEE Trans Neural Netw Learn Syst.*, 25(1): 69–80.
12. Jayanthi, S.K., & Sasikala, S. (2014). Naive bayesian classifier and PCA for web link spam detection. *Comput Sci Telecommun.*, 41(1): 3–15
13. Tarle, B., & Jena, S. (2017). An artificial neural network based pattern classification algorithm for diagnosis of heart disease. *In: IEEE proceedings of international conference on computing, communication, control and automation (ICCUBEA) on 17–18 Aug 2017*, Pune. pp. 1–4.
14. Park, H.W., Li, D., Piao, Y., & Ryu, K.H. (2017). A hybrid feature selection method to classification and its application in hypertension diagnosis. *In: Bursa M, Holzinger A, Renda M, Khuri S (eds.) Information technology in bio- and medical Informatics. ITBAM 2017,* vol. 10443. Lecture notes in computer science. Springer, Cham.
15. Harb, H.M., & Desuky, A.S. (2014). Feature selection on classification of medical datasets based on particle swarm optimization. *Int J Comput Appl.*, 104(5):14–17
16. Alzubi, R., Ramzan, N., Alzoubi, H., & Amira, A. (2018). A hybrid feature selection method for complex diseases SNPs, *IEEE Access*, 6: 1292–1301.
17. Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*: 51-67.

18. Zhai, Q.H., Ye, T., Huang, M.X., Feng, S.L., & Li, H. (2020). Whale Optimization Algorithm for Multiconstraint Second-Order Stochastic Dominance Portfolio Optimization. *Computational Intelligence and Neuroscience*, *2020*.
19. Sharawi, M., Zawbaa, H.M., & Emary, E. (2017), February. Feature selection approach based on whale optimization algorithm. In *2017 IEEE Ninth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 163-168.
20. Zhang, Y., Wang, Y., Zhou, G., Jin, J., Wang, B., Wang, X., & Cichocki, A. (2018). Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. *Expert Systems with Applications*, *96*, 302-310.
21. https://archive.ics.uci.edu/ml/datasets/ILPD+(Indian+Liver+Patient+Dataset)
22. https://archive.ics.uci.edu/ml/datasets/Hepatitis
23. http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease