IIOAB JOURNAL

ARTICLE

MULTI-OBJECTIVE ARTIFICIAL BUTTERFLY OPTIMIZATION WITH LEVY DISTRIBUTION ALGORITHM BASED FEATURE SELECTION AND CLASSIFICATION MODEL FOR THYROID DISEASE DIAGNOSIS

Rajaram Pavithra[1*], Latha Parthiban[2]

[1]*Department of Computer Science, Bharathiyar University, TN, INDIA*
[2]*Department of Computer Science, Pondicherry University, Community College, Pondicherry, INDIA*

ISSN 0976-3104

ABSTRACT

Thyroid diseases are a commonly occurring endocrine illness over the globe, which affect the functions of thyroid gland resulting into abundant secretion of thyroid hormones. this paper presents a new feature selection (FS) based classification model for thyroid disease diagnosis using Multi-Objective Artificial Butterfly Optimization with Levy Distribution (MOABO-L) and deep neural network (DNN). The MOABOL-DNN algorithm is applied for both feature selection and tuning the training scheme of deep neural network (DNN). In order to improve the convergence rate of MOABO algorithm, Levy distribution concept is incorporated to it. The utilization of FS process helps to remove the unwanted features and increases the overall classification accuracy. For classification process, DNN with MOABO based fine-tuned training strategy is employed. To ensure the effectiveness of the MOABOL-DNN algorithm, a series of simulations takes place on two benchmark thyroid dataset. The obtained outcome indicated the goodness of the presented MOABOL-DNN technique with the maximum accuracy of 99.68% and 98.14% on the applied dataset 1 and dataset 2 respectively.

THE IIOAB JOURNAL

KEY WORDS
Feature selection, classification, thyroid disease, ABO algorithm, Levy flight

INTRODUCTION

The abnormal development in thyroid gland results in Congenital Hypothyroidism (CH) [1]. Globally, numerous people are infected by CH and especially in Europe and North America. Countries like Iran uses neonatal thyroid screening models for diagnosing the CH at earlier phase and treat accordingly. It is a cost-effective mechanism and Thyroid-Stimulating Hormone (TSH1) is verified within limited span; however distinct influential factors lead to cause CH even for small kids, adults, pregnant women, and genetic history of thyroid infection are also determined.

Authors [2] examined the efficiency of CH investigation and earlier remedy by correlating the Intelligence Quotient (IQ) value while examining the TSH level. Also, a statistical test with t-test and examination of covariance tests has implied the significance of earlier and massive doses of treatment in accomplishing a typical deployment in kids for CH diagnosis. Recently, Artificial Intelligence (AI) as well as Machine Learning (ML) methods was applied for achieving better interpretation of thyroid information, and medical investigation.

Diverse studies were carried out to examine the efficiency of the above mentioned models. For instance, [3] utilized the different classification models for thyroid nodule ultrasound photographs. [4] employed the two ML models for computing the structural classification of thyroid infection. Additionally, massive studies using different learning models to signify vital insights regarding thyroid disease. Under the application of ML as well as Computer Aided Diagnosis (CAD) methodologies, developers mitigate the feasible errors in diagnosing the biomedical data in both time as well as cost-efficient way. Traditionally, no classifications are deployed for CH analysis. Based on the survey, the newly developed models are highly efficient in diagnosing the thyroid disease with considerable performance.

Received: 22 Dec 2020
Accepted: 26 Feb 2021
Published: 9 Mar 2021

In [5], Random Forests (RF) as well as SVM methods are trained on the ultrasound images of thyroid lumps for cancer diagnosis as benign or malignant, and maximum accuracy is attained. In [6] used Radial Basis Function (RBF), Learning Vector Quantization (LVQ), MLP, Back Propagation Algorithm (BPA), AIRS, and Perceptron frameworks for UCI dataset. In this approach, MLP and BPA have attained supreme and inferior accuracy. In [7] related the function of decision tree (DT) models in thyroid disease prediction. Moreover, NBTree is a unification of NB and DT methods which has gained better accuracy, recall and precision.

A study performed in [8] showcases that ML approaches with fundamental infrastructure for thyroid diagnosis. From the above mentioned approaches, artificial neural network (ANN) has implied optimal outcome. One of the major disadvantages of this model is that, Genetic Algorithm (GA) technique is not applicable and provides insignificant result. [9] deployed a scheme termed as Data Mining (DM) with the help of NN that is applied in early thyroid disease prediction. The system is trained under the application of BP and gradient method at the same time. However, the differences in layers of different network attributes are not applied in the training phase. [10] operated on the trial of 21 variables for training a method with the help of classification algorithm. It is trained by applying k-nearest neighbor (KNN),

***Corresponding Author**
Email:
pave0581@gmail.com

IIOAB JOURNAL ISSN: 0976-3104

artificial neural network (ANN) and fuzzy ANN models and compares the accuracy. It is clear that Fuzzy ANN outperforms than alternate classifiers.

This paper introduces a new feature selection based classification model for thyroid disease diagnosis using Multi-Objective Artificial Butterfly Optimization with Levy Distribution (MOABO-L) and deep neural network (DNN). The MOABOL-DNN algorithm is applied for both feature selection and tuning the training scheme of DNN. To increase the capability of the MOABO algorithm, Levy distribution concept is incorporated to it. The application of FS process helps to eliminate the unwanted features and raises the overall classification accuracy. To perform classification process, DNN with MOABO based training strategy is employed. For validating the performance of the MOABOL-DNN algorithm, a set of simulations takes place on two benchmark thyroid dataset.

METHODS

The proposed MOABOL-DNN algorithm for FS with classification processes for thyroid disease diagnosis involves three main processes namely preprocessing, feature selection, and classification. Here, MOABOL-DNN algorithm is employed for both FS and classification process. The entire workflow of the proposed MOABOL-DNN algorithm is illustrated in [Fig. 1].

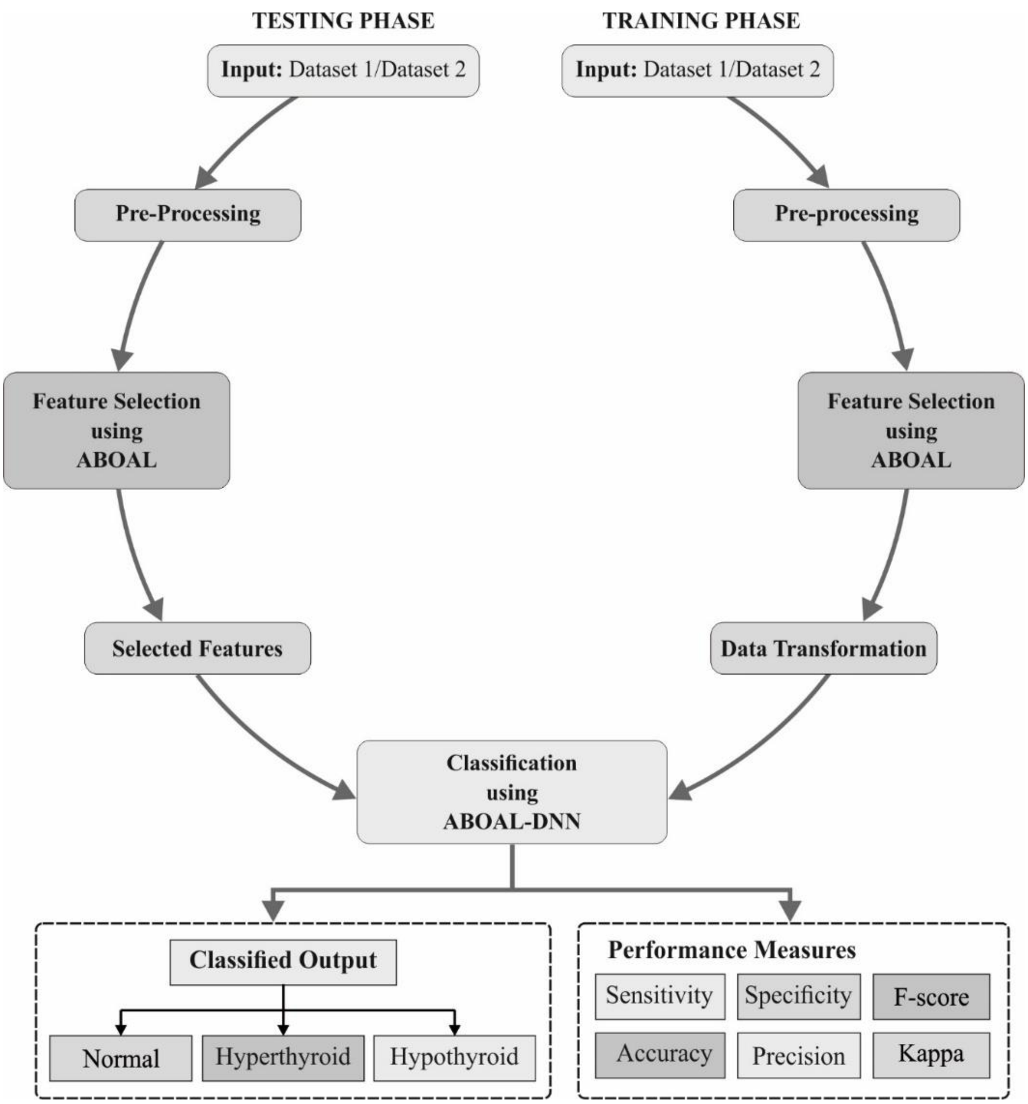

Fig. 1: Workflow of MOABOL-DNN model.

Preprocessing

In the beginning, the actual medical data is completely preprocessed to remove the unwanted noise and missing value replacement takes place via mean method. In addition, minimum-maximum (min-max) data normalization method is used to normalize the data values in the range of 0 to 1.

IIOAB JOURNAL ISSN: 0976-3104

ABOA with Levy Flight

Based on the mate-finding principle of speckled woods, Qi et al. [11] projected a novel meta-heuristic approach named ABO. In general, the butterfly population is classified as 2 groups named sunspot and canopy. A butterfly with optimal fitness makes sunspot butterflies and remaining is referred as canopy butterflies, and various flight principles have been employed for all applications. In this framework, three flight modes are composed of ABO such as sunspot, canopy, and free flight modes.
Only few rules are developed for normalizing the mate-finding principle of butterflies in ABO approach:

- The possibility of identifying female butterflies can be increased by male butterflies where it attempts to find an optimal location called as sunspot;
- A remarkable sunspot can be accomplished when a sunspot butterfly flies towards closer sunspot;
- The canopy butterfly moves to a sunspot butterfly.

Assume that $S = \{x_1, x_2, \dots, x_N\}$ is a search space where $x_i \in \mathfrak{R}^n$. A multi-objective optimization issues desires to identify the global minimum $x^* \in \mathcal{S}$ which reduces a function set implied by f, i.e.:

$$x^* = \arg \min_{\forall x \in \mathcal{S}} (f(x)) = \arg \min_{\forall x \in \mathcal{S}} (f_1(x), f_2(x), \dots, f_M(x)), \tag{1}$$

subject to:

$$g_i(x) = 0 \quad \forall i = 1,2, \dots, p, \tag{2}$$

$$h_i(x) \geq 0 \quad \forall i = 1,2, \dots, q. \tag{3}$$

Collection of values satisfies the given constraints that describe the possible region and a point in this region is assumed as possible solution. In multi-objective issues, no single solution is effective interms of objectives while assuming the conflicting objects. Hence, solution for multi-objective optimization issues are not considered as a scalar value, however a collection of solutions is named as Pareto-optimal set.

Initially, the Pareto Dominance is defined as a solution vector x^a is referred as a dominative solution vector x^b $(x^a \prec x^b)$ when $(x_i^a) \leq f(x_i^b), \forall i = \{1,2, \dots, N\}$, as well as $\exists i \in \{1,2, \dots, N\}$ where $f(x_i^a) < f(x_i^b)$. Regarding the Pareto Dominance, a solution vector x^a is assumed to be a Pareto-optimal when $x^b, f_j(x^a) \leq f_j(X^b), j = 1,2, \dots, M$, and $j \in \{1,2, \dots, M\}$ in which $f_j(x^a) < f_j(x^b)$. Hence, Pareto-optimal set P^* defines the multi-objective optimization issues $f(x)$ by means of Pareto-optimal solutions are depicted as:

$$P^* = \{x \in \mathcal{S} | f(x) \prec f(x'), \forall x' \in \mathcal{S}\}. \tag{4}$$

The Pareto-optimal front PF^* by means of multi-objective optimization issues $f(x)$ and Pareto-optimal set P^* is depicted as given in the following:

$$PF^* = \{f(x) | x \in P^*\}. \tag{5}$$

Moreover, Pareto-dominance models are classified as 2 familiar classes namely, (i) indicator-related model as well as (ii) decomposition-related model.

In order to reduce the search space and avoid stagnation, levy flight concept is incorporated into the ABO algorithm. In this work, a Lévy flight plays an important role in accomplishing better results [12]. Subsequent to the massive rounds, the distance from actual walks to random walk intends to make stable distribution. Lévy flights are characterized using inverse square distribution of a step length that optimizes the random searching process if the targets become scarce in resources. Unlike, Brownian motion is more applicable when there is a requirement for placing numerous preys. The traits of 2 random walks resulted in enhancing the Swarm Intelligence (SI) optimization in which Lévy flights maximizes the eligibility of "exploration" whereas Brownian motion supports the "exploitation". Numerically, Lévy flights are considered as type of random walk with step lengths of heavy-tailed Lévy alpha-stable distribution with respect to power-law notion, $L(s) \sim |s|^{-1-\beta}$, where $0 < \beta \leq 2$ is referred as index. A common function of Lévy distribution is illustrated as

$$L(s, \gamma, \mu) = \begin{cases} \sqrt{\dfrac{\gamma}{2\pi}} exp\left[-\dfrac{\gamma}{2(s-\mu)}\right] \dfrac{1}{(s-\mu)^{3/2}}, & 0 < \mu < s < \infty; \\ 0, & s \leq 0. \end{cases} \tag{6}$$

Considering the $2D$-Lévy flights for sample applies a Lévy distribution whereas the directions satisfy a uniform distribution. It is pointed that, Lévy flights are highly significant in identifying unwanted and large-scale search space when compared with Brownian walk. The major cause for this discussion is that

IIOAB JOURNAL
ISSN: 0976-3104

variance of Lévy flight $\delta^2(t) \sim t^{3-\beta}, 1 \le \beta \le 2$ enhances robustly than Brownian random walks which is referred as, $\delta^2(t) \sim t$.

ABOA with Levy Flight based FS

In this section, the ABO-L algorithm is applied for the FS process to choose an optimal set of features. The FS problem can be considered as an optimization problem, and ABO-L algorithm can be used to resolve it. This work has presented a multi-objective FS mechanism under the application of weighted-sum technique with an aim of reducing the classification error. In mathematical format, fitness function helps the agent to accomplish optimal solution is equated as given below:

$$f = \arg \min_{\forall x \in \mathcal{S}} \left(\sum_{i=1}^{M} w_i f_i(x) \right), \tag{7}$$

where $\sum_{i=1}^{k} w_i = 1$ with $w \ge 0, f_i(x)$ denotes classification error of class i, and M implies the overall class values. In addition, the extension of next technique fits the first one with tiny difference where the feature set size and classification error has to be reduced. Therefore, fitness function is projected in the following:

$$f = \arg \min_{\forall x \in \mathcal{S}} \left(\sum_{i=1}^{M} w_i f_i(x) + w_{M+1} f_{M+1}(x) \right), \tag{8}$$

where $f_{M+1}(x)$ represents the count of features.

DNN based Classification

In general, DL approach is applicable for accomplishing higher level dimension features from input dataset. Next, the features are gathered from Deep Neural Network (DNN) and employed for improvising the performance efficiency [13]. Moreover, typically used DL approach is a DNN classification approach developed by the integration of stack of autoencoder (AE) with the help of SM classification technology.

In general, AE contains input, hidden and output layers. Hence, the resultant level is similar to the level of input. Furthermore, the AE is trained for embedding the input with feature spaces, where dimension is low than input space. Next, a dimension of a code space is decided as higher than input space to maximize the efficiency of classification to a greater extent. Hence, AE manages to provide best implication of input vector through the replacement of the proper code.

Hybridization of ABO-L with L-BFGS Model: Limited memory Broyden–Fletcher-Goldfarb-Shannon (L-BFGS) is referred as productive optimization method on the basis of BFGS obtained from Quasi-Newton family for high level optimizing problems. Quasi-Newton model has been developed extensively for identifying the desired models for making hessian or inverse hessian of function (f) to be limited.

The L-BFGS and BFGS use the identical approaches for generalizing a function of hessian matrix. At the initial stage, a positive definite as well as sparse symmetric matrix H_0 is obtained from f function, which is normalized. Next, H_k can be accomplished using m BFGS update to H_0 by utilizing data collected from m prior iterations when k is supreme than m. Therefore, process in L-BFGS is depicted as u_k and GD of a function is referred as gk.

$$\mathrm{H}_{\mathrm{k}+1} = \mathrm{V}_\mathrm{k}^\mathrm{Z} \mathrm{H}_\mathrm{k} \mathrm{V}_\mathrm{k} + \rho_\mathrm{k} \mathrm{s}_\mathrm{k} \mathrm{s}_\mathrm{k}^\mathrm{Z} \tag{9}$$

where$\rho_\mathrm{k} = 1/\mathrm{v}_\mathrm{k}^{\tau_{\mathrm{s}_\mathrm{k}}}, \mathrm{Y}_\mathrm{k} = \mathrm{I} - \rho_{\mathrm{k}\mathrm{s}_\mathrm{k}\mathrm{s}_\mathrm{k}^\mathrm{Z}}, \mathrm{s}_\mathrm{k} = \mathrm{u}_{\mathrm{k}+1} - \mathrm{u}_\mathrm{k}$ and $\mathrm{y}_\mathrm{k} = \mathrm{g}_{\mathrm{k}+1} - \mathrm{g}_\mathrm{k}$. The L-BFGS is suitable in resolving the computational demands because of numerous scale issues like DNN training. Additionally, L-BFGS is robust and needs lower storage for large-scale problems. To improve the exploration process of the L-BFGS, ABO-L algorithm is employed. Once the exploration is completed, L-BFGS generates better solutions and vector with iterative development where it is applied for computing general verification of proposed framework. Therefore, the inner variables of DNN architecture has been optimized using presented scheme to eliminate the local optimal issues in AEs and SM to obtain closer-optimal DNN and L-BFGS is used for local searching parameter vectors.

SM Classifier: It is a classification approach employed for multi-label classification problems. Here, the mapping is performed among input vector c with K class labels, as depicted below:

$$v_q = \frac{\exp\,(\theta_q^Z \mathrm{c})}{\sum_{\mathrm{k}=1}^{\mathrm{K}} \exp\,(\theta_\mathrm{k}^Z \mathrm{c})} (q = 1,2, \ldots \mathrm{K}) \tag{10}$$

where $\theta_\mathrm{k} = [\theta_{\mathrm{k}1} \theta_{\mathrm{k}2} \ldots \theta_{\mathrm{kN}}]^Z$ refers the weights to be tuned using effective optimization approach.

RESULTS

For validating the performance of the MOABO-DNN model, a detailed experimentation was carried out on two benchmark two thyroid dataset [14]. A set of measures used to investigate the classification performance interms of sensitivity, specificity, precision, accuracy, F-score, and Kappa. A brief discussion of the dataset is provided in the subsequent section.

Dataset Description

The Thyroid dataset 1 includes a 3772 training and 3428 testing samples. A set of 6667 samples comes under Normal (class 3), 166 samples under hypothyroidism (Class 2), and 367 samples under hyperthyroidism (Class 1). This dataset includes a total of 21 features. Similarly, the thyroid Dataset-2 has a total of 215 Samples with 5 features and 1 target class (1 = normal, 2 = hyperthyroid, 3 = hypothyroid).

[Table 1] provides the FS results attained by the ABOL-FS with WOA-FS and SA-FS models on the applied two datasets. From the table, it is evident that, on the applied dataset 1, the ABOL-FS model has achieved excellent results with the minimum best cost of 0.567 and 0.623 on the applied dataset 1 and 2 respectively. At the same time, the SA-FS model has led to worse performance with the maximum best cost of 0.809 and 0.958 on the applied dataset 1 and 2 respectively. Though the WOA-FS model has achieved moderate results with the best cost of 0.752 and 0.914, it fails to outperform the ABOL-FS model.

Table 1: Selected Features of Existing with Proposed ABOL-FS Algorithm on Dataset-1 and Dataset-2

Methods	Dataset	Best Cost	Selected Features
ABOL-FS	Dataset-1	0.567	1,3,4,6,7,8,9,12,15,18,20
	Dataset-2	0.623	2,3,5
WOA-FS	Dataset-1	0.752	1,2,3,4,6,7,8,10,11,12,14,15,16
	Dataset-2	0.914	1,3,4
SA-FS	Dataset-1	0.809	3,5,6,7,9,10,12,13,14,15,16,18,19,20
	Dataset-2	0.958	1,2,3,5

An extensive comparative results analysis of the MABOL-DNN model with previous models is carried out with respect to distinct measures, as depicted in [Table 2] and [Figs. 2-3] [15, 16]. The resultant values verified that the MABOL-DNN model has achieved effective classifier results on the applied dataset. On comparing with existing models, the CART method has accomplished least outcome with the sensitivity of 58.8%, specificity of 59.88%, precision of 57.7%, accuracy of 58.75%, F-score of 58.32%, and kappa value of 47%. It is observed that the RT model has portrayed somewhat higher results over CART with the sensitivity of 62.5%, specificity of 65.31%, precision of 62.7%, accuracy of 62.5%, F-score of 63.01%, and kappa value of 52.3%. It is noticeable that the J48 model has realized somewhat practicable outcome with the sensitivity of 63.8%, specificity of 68.51%, precision of 58.1%, accuracy of 66.25%, F-score of 69.02%, and kappa value of 56.8%. In line with, the NBTree model has reached to reasonable outcome with the sensitivity of 75%, specificity of 76.43%, precision of 77%, accuracy of 75%, F-score of 70.12%, and kappa value of 68%. At the same time, the IGWO-RBF-SVM model has led to moderate results with the sensitivity of 78.90%, specificity of 81.17%, precision of 68.79%, accuracy of 78.49%, F-score of 65.90%, and kappa value of 61.98%. On continuing with, even better performance with the sensitivity of 81.17%, specificity of 75.18%, precision of 72.71%, accuracy of 79.11%, F-score of 67.23%, and kappa value of 62.65% has been attained by the IGWO-ANN model. Moreover, the IGWO-Linear-SVM model has obtained a sensitivity of 82.58%, specificity of 90.46%, precision of 70.69%, accuracy of 93.96%, F-score of 71.83%, and kappa value of 64.32%. Besides, somewhat acceptable outcome is provided by the IGWO-MSVM model with the sensitivity of 94.65%, specificity of 94.5%, precision of 91.83%, accuracy of 97.49%, F-score of 91.76%, and kappa value of 90.65%.

Table 2: Result analysis of existing with proposed MOABOL-DNN method on thyroid dataset

Methods	Sensitivity	Specificity	Precision	Accuracy	F-score	Kappa
MOABOL-DNN (Dataset-1)	**97.75**	**99.54**	**95.39**	**99.68**	**96.50**	**95.73**
MOABOL-DNN (Dataset-2)	**96.16**	**98.90**	**94.27**	**98.14**	**95.16**	**93.58**
DNN (Dataset-1)	96.54	98.32	94.22	98.04	94.92	93.84
DNN (Dataset-2)	95.13	96.48	92.89	97.09	93.12	91.90
IGWO+MSVM	94.65	94.50	91.83	97.49	91.76	90.65
IGWO+Linear-SVM	82.58	90.46	70.69	93.96	71.83	64.32
IGWO+RBF-SVM	78.90	81.17	68.79	78.49	65.90	61.98
IGWO+ANN	81.17	75.18	72.71	79.11	67.23	62.65
NBTree	75.00	76.43	77.00	75.00	70.12	68.00
J48	63.80	68.51	58.10	66.25	69.02	56.80
Rand. Forest	65.00	69.06	66.00	65.00	66.87	55.10
Rand. Tree	62.50	65.31	62.70	62.50	63.01	52.30
CART	58.80	59.88	57.70	58.75	58.32	47.00

IIOAB JOURNAL ISSN: 0976-3104

Similarly, the D-KELM model has obtained moderate results on the dataset 2 with the sensitivity of 84.21%, specificity of 94.9%, precision of 78.31%, accuracy of 94.04%, F-score of 75.17%, and kappa value of 70.68%. Eventually, the D-KELM model on the dataset 1 has resulted to a somewhat sensible outcome with the sensitivity of 87.53%, specificity of 95.56%, precision of 79.08%, accuracy of 94.06%, F-score of 76.5%, and kappa value of 72.89%. Instantaneously, the OD-KELM model has showcased somewhat better results on the dataset 2 over all the other methods by realizing sensitivity of 86.79%, specificity of 96.19%, precision of 83.1%, accuracy of 94.11%, F-score of 84.78%, and kappa value of 79.93%. Next to that, the OD-KELM model has established reasonable results on the dataset 1 with the sensitivity of 92.67%, specificity of 97.88%, precision of 79.37%, accuracy of 98.01%, F-score of 83.98%, and kappa value of 78.29%.

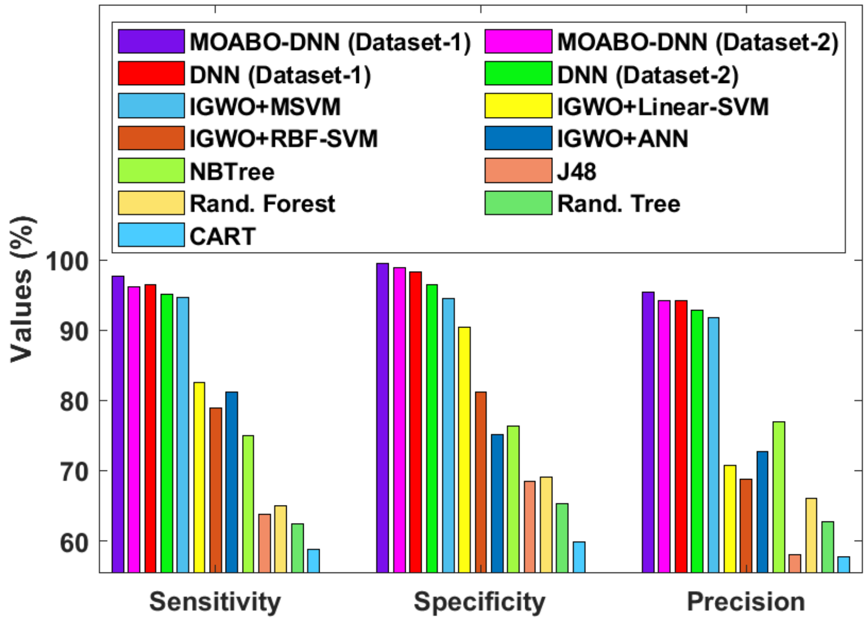

Fig. 2: Comparative result analysis of MOABO-DNN with other Models-I.

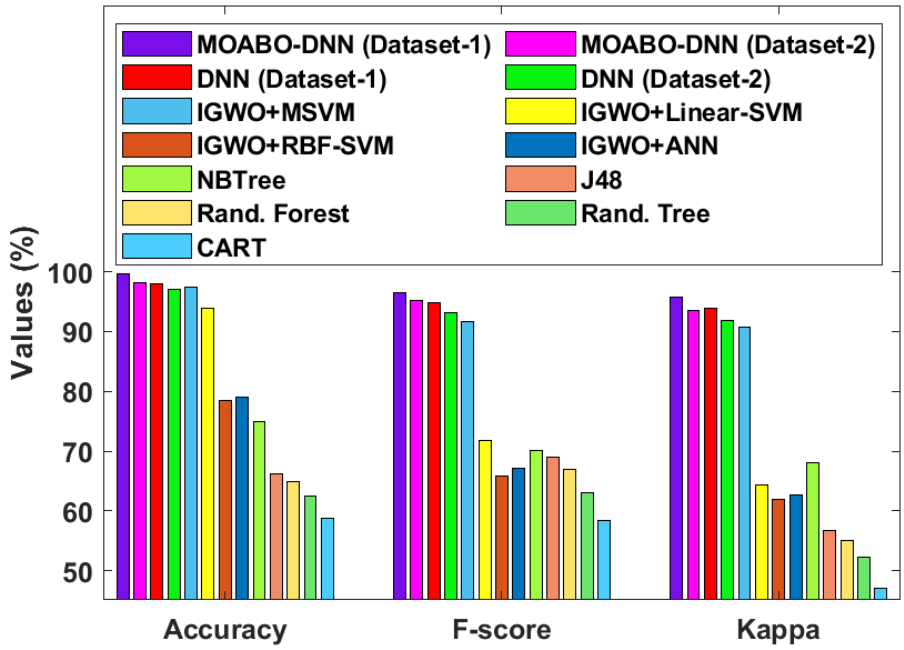

Fig. 3: Comparative result analysis of MOABO-DNN with other models-II.

However, the MOABOL-DNN model has accomplished superior results over the compared methods with the maximum sensitivity of 96.16%, specificity of 98.90%, precision of 94.27%, accuracy of 98.14%, F-score of 95.16%, and kappa value of 93.58%. On continuing with, the MOABOL-DNN model has resulted to

IIOAB JOURNAL ISSN: 0976-3104

effective classification on dataset 2 with the highest sensitivity of 96.16%, specificity of 98.90%, precision of 94.27%, accuracy of 98.14%, F-score of 95.16%, and kappa value of 93.58%. After observing the tables and resultant figures, it is apparent that the MOABOL-DNN model has achieved better results on the diagnosis of thyroid with the maximum accuracy of 99.68% and 98.14% on the applied dataset 1 and dataset 2 respectively. Therefore, it can be applied as a novel diagnostic tool for thyroid diseases.

CONCLUSION

This paper has developed a FS based classification model using MOABOL-DNN and DNN for thyroid disease diagnosis. The proposed MOABOL-DNN algorithm for FS with classification processes for thyroid disease diagnosis involves three main processes namely preprocessing, feature selection, and classification. Here, MOABOL-DNN algorithm is employed for both FS and training scheme tuning of DNN. Initially, data preprocessing takes place to transform the raw data into a useful format. Then, MOABOL based FS process is executed to select an optimal number of features. Finally, DNN model is employed for classification and MOABOL-DNN algorithm is applied to fine tune the training process. For assuring the superior performance of the MOABOL-DNN algorithm, a set of simulations takes place on two benchmark thyroid dataset. The simulation outcomes ensured the effective diagnosis results with the maximum accuracy of 99.68% and 98.14% on the applied dataset 1 and dataset 2 respectively. In future, the performance of the MOABOL-DNN algorithm can be improved by clustering techniques.

REFERENCES

[1] Hashemipour M, Soheilipour F, Karimizare S, Khanahmad H, Karimipour M, Aminzadeh S, Kokabee L, Amini M, Hovsepian S, Hadian R. [2012] Thyroid peroxidase gene mutation in patients with congenital hypothyroidism in Isfahan, Iran. International journal of endocrinology.

[2] Najmi SB, Hashemipour M, Maracy MR, Hovsepian S, Ghasemi M. [2013] Intelligence quotient in children with congenital hypothyroidism: The effect of diagnostic and treatment variables. Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences. 18(5):395.

[3] Singh N, Jindal A. [2012] A segmentation method and comparison of classification methods for thyroid ultrasound images. International Journal of Computer Applications. (11):43-9.

[4] Erol R, Oğulata SN, Şahin C, Alparslan ZN. [2008] A radial basis function neural network (RBFNN) approach for structural classification of thyroid diseases. Journal of medical systems. 32(3):215-20.

[5] Prochazka A, Gulati S, Holinka S, Smutek D. [2019] Classification of thyroid nodules in ultrasound images using direction-independent features extracted by two-threshold binary decomposition. Technology in cancer research & treatment. 18:1533033819830748.

[6] Maysanjaya IM, Nugroho HA, Setiawan NA. [2015] A comparison of classification methods on diagnosis of thyroid diseases. In2015 International seminar on intelligent technology and its applications (ISITIA). 89-92.

[7] Turanoglu-Bekar E, Ulutagay G, Kantarcı-Savas S. [2016] Classification of thyroid disease by using data mining models: a comparison of decision tree algorithms. Oxford Journal of Intelligent Decision and Data Sciences. 2:13-28.

[8] Razia S, Rao MN. [2016] Machine learning techniques for thyroid disease diagnosis-a review. Indian J Sci Technol. 9(28):10-7485.

[9] Prerana PS, Taneja K. [2015] Predictive data mining for diagnosis of thyroid disease using neural network. International Journal of Research in Management, Science & Technology. 3(2):75-80.

[10] Umadevi S, JeenMarseline KS. [2017] Applying classification algorithms to predict thyroid disease. Int J Eng Sci Comput. 7(10):15118-20.

[11] Qi X, Zhu Y, Zhang H. [2017] A new meta-heuristic butterfly-inspired algorithm. Journal of computational science. 23:226-39.

[12] Li D. [2014] Cooperative quantum-behaved particle swarm optimization with dynamic varying search areas and Lévy flight disturbance. The scientific world journal.

[13] Pustokhina IV, Pustokhin DA, Gupta D, Khanna A, Shankar K, Nguyen GN. [2020] An effective training scheme for deep neural network in edge computing enabled Internet of medical things (IoMT) systems. IEEE Access. 8:107112-23.

[14] Dataset Source: https://archive.ics.uci.edu/ml/datasets/Thyroid+Disease

[15] Shankar K, Lakshmanaprabu SK, Gupta D, Maseleno A, De Albuquerque VH. [2020] Optimal feature-based multi-kernel SVM approach for thyroid disease classification. The journal of supercomputing. 76(2):1128-43.

[16] Turanoglu-Bekar E, Ulutagay G, Kantarcı-Savas S. [2016] Classification of thyroid disease by using data mining models: a comparison of decision tree algorithms. Oxford Journal of Intelligent Decision and Data Sciences. 2:13-28.