Date of publication JUN-30, 2024, date of current version MAY-19, 2024.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.23.2.3548

An Intelligent System for Thyroid Dysfunction Prediction

ELAF A. ABD AL-KAREEM[1], SAMAH J. SABA[2]
[1]Department of Al qura'n Science, College of Islamic sciences, University of Diyala, Iraq
[2]Department of computer science, College of Science, University of Diyala, Iraq

Corresponding author: Elaf A. Abd Al-Kareem (e-mail: elaaf.ali1989@gmail.com).

ABSTRACT The prediction of thyroid diseases has become increasingly important in recent years. The feature engineering part is generally less studied in existing approaches than model optimization. To circumvent these restrictions, feature engineering for machine learning (ML) and deep learning models are examined in our research. This paper proposes a system to predict thyroid dysfunction. A model comprises three stages: preprocessing of the data, selection of features, and classification. A standard scalar normalizes the data after filling missing values with mean values. A Deep Neural Network (DNN) is utilized for classification, and the Improved Salp Swarm Algorithm (ISSA) is used for feature selection. Finally, accuracy is utilized to estimate the system proposed. The UCI thyroid illness datasets are used in our investigation. In experiments with the DNN model, deep learning classifiers based on selected features produce the highest accuracy of 99.68%. The model proposed has superior performance compared to state-of-the-art methods.

KEYWORDS thyroid prediction; DNN; ISSA; SSA; feature selection.

I. INTRODUCTION

IN the healthcare industry, computational biology is being utilized to develop new treatments. By doing so, we could gather the patient information that had been kept to predict illness in the past. Various clever prediction algorithms are available for the early-stage detection of the illness. Many datasets exist in the medical information system, but no intelligent algorithms can quickly analyze diseases. ML algorithms are essential in the long run for resolving the challenging nonlinear issues in creating a prediction model. Any disease prediction model must prioritize the characteristics chosen from the various datasets and utilized to classify healthy patients as accurately as possible. In such a situation, a healthy patient could receive unnecessary treatment as a result of misclassification. Thus, it is extremely important to foresee any health problem associated with thyroid disease [1, 2].

The thyroid is a neck-based endocrine gland that produces hormones (FT3 and FT4) that are then released into the circulation. In addition to controlling heart rate and body temperature, thyroid hormones are responsible for regulating metabolism, which is how the body utilizes and absorbs nutrition [3]. Significant problems can arise from an overactive or underactive thyroid gland (hyperthyroidism with excessive hormones or hypothyroidism with decreased hormones) [4]. In addition, the thyroid gland may become inflamed (thyroiditis) or expand due to one or more tumors growing inside (multinodular goiter, nodules). Cancerous tumors can sometimes be found in these nodules. Due to this, treating thyroid issues is a critical issue [5].

This research presents a model for predicting thyroid dysfunction. The model is divided into three stages: feature selection, data preparation, and classification. The data is handled by substituting the mean value for any missing values before being normalized using a standard scalar. Both the deep neural network (DNN) and the improved Salp Swarm algorithm (ISSA) are used in the classification process. Finally, the model is assessed using F-measure, accuracy, and precision.

The paper is arranged as follows: the related work is demonstrated in section II, the theoretical background is discussed in section III, the proposed method is discussed in section IV, the results are discussed in section V, and the conclusion is outlined in section VI.

II. RELATED WORK

Numerous studies have examined how hormonal parameters and patient data, such as age and gender, can be used to identify thyroid diseases. The most notable difference is that some studies use machine learning models for prediction and classification, whereas others use deep neural networks.

K. Shankar [6], proposed an algorithm utilizing kernel-based classifiers and optimal feature selection to classify

thyroid diseases. Using multi-kernel support vector machines, the authors enhanced the classifying process by using improved gray wolf optimization to enhance feature selection. The suggested thyroid prediction results in specificity, sensitivity, and accuracy of 94.5%, 99.05%, and 97.49% compared to the existing methods.

T. Alyas [7] used several algorithms of machine learning, such as random forests, decision trees, KNNs, and artificial neural networks to optimize the prediction of disease depending on the parameters of the dataset. For the classification, the dataset has also been manipulated to make accurate predictions. Several datasets were classified based on unsampled and sampled data for better comparisons. After dataset manipulation, the RF algorithm attained the highest accuracy, equal to 91% specificity and 94.8% accuracy.

G. Chaubey [8] predicted and assessed their accuracy using three different techniques: decision trees, logistic regression, and KNN algorithms. The thyroid data set from UC Irvin knowledge discovery in the databases archive was used by the machine learning repository for this. The highest accuracy achieved by KNN is 96.875%.

I. Ioniţă and L. Ioniţă [9] examined and contrasted the Radial Basis Function Network, Multilayer Perceptron, Decision Tree, and Naive Bayes classification models. All of the classification models mentioned above have substantial accuracy, according to the data; the decision tree model has the best classification rate. The UCI machine learning repository supplied the dataset to develop and evaluate the classifier. The highest accuracy achieved is 96.5% with the decision tree algorithm.

S. Sankar [10] considered the knowledge discovery dataset from UC Irvin. This research suggests using the XGBoost algorithm to forecast thyroid illness precisely. A comparison is made between the effectiveness of the proposed XGBoost algorithm and the logistic regression, decision tree, and k-nearest Neighbor (kNN) approaches. We compare and evaluate the performance of all four methods. The accuracy of the proposed XGBoost achieves 98.59%, KNN has an accuracy of 96.875%, a Decision tree – 87.5%, and logistic regression – 81.25%.

III. METHODOLOGY

A. DATASET

The UCI thyroid illness datasets are used in our investigation. The UCI machine learning repository has several datasets related to thyroid conditions [11]. There are 9172 sample observations in the dataset, and 31 characteristics are used to represent each sample.

Table 1. Class-wise target Description [11].

FEATURE	COUNT
Miscellaneous	281
Ant thyroid treatment	33
General health	436
Hypothyroid	601
Binding protein	376
Replacement therapy	136
Hyperthyroid	182
	6771

B. IMPROVED SALP ALGORITHM (ISSA)

Salps have a translucent barrel-shaped body and are members of the Salpidae family [12]. A salp's shape is seen in Figure 1(a). Salps frequently group to create salp chains in deep waters. Figure 1(b) shows an illustration of this chain. Although the primary cause of this behavior is still unclear, some scientists think it is done to enhance mobility through quickly synchronized adjustments and feeding [13].

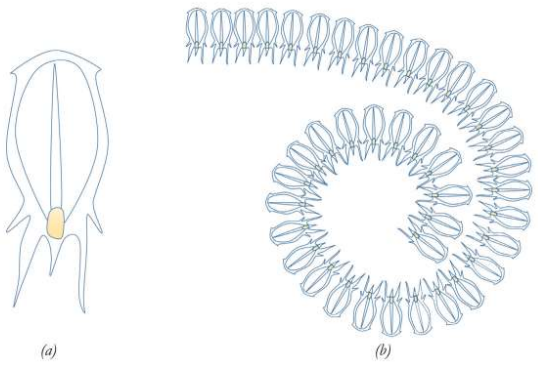

Figure 1. (a) Individual Salp, (b) Swarm of Salp (Salps Chain) [13].

A population is first split into two groups, leaders and followers, in order to mathematically describe the salp chains. Salps in the head of a chain are known as leaders, while the other salps are known as followers. These salps are called so because the leader leads the followers, and the swarm follows each other, either indirectly or directly following the leader. The salp's position is specified in an n-dimensional search space, where n is the number of variables in a given issue, just like in other swarm-based algorithms. Hence, a two-dimensional matrix named x contains the positions of every salp [14, 15].

Ref [14] suggests using the following equation to modify the leader's position:

$$x_j^i = \begin{cases} F_j + c_1\left((ub_j - lb_j)c_2 + lb_j\right) & c_3 \geq 0 \\ F_j - c_1\left((ub_j - lb_j)c_2 + lb_j\right) & c_3 < 0 \end{cases}, \quad (1)$$

where: F_j is a position of a food supply in a jth dimension, ub_j denotes an upper bound of a jth dimension, lb_j denotes a lower bound, c_3, c_1, and c_2 are random integers, and x_1 j presents a position of a first salp (leader) in a jth dimension. The leader merely modifies its location with relation to food supply, as demonstrated by Equation (1). Most crucial factor in SSA is the coefficient c_1, which balances the definitions of exploration and exploitation [14]:

$$c_1 = 2e^{-(\frac{Al}{L})^2}, \quad (2)$$

where: L is a maximum number of iterations, l is a current iteration.

Random integers evenly produced inside the interval [0,1] comprise the parameters c_2 and c_3. They specify the step size and whether the subsequent location in the jth dimension should be toward negative or positive infinity.

Using Newton's law of motion, the below equations are used to modify the followers' positions:

$$x_j^i = \frac{1}{2}at^2 + vot, \quad (3)$$

where: $i \geq 2$, x_j^i indicates ith follower salp position in jth dimension, t is the time, v_0 is the initial speed, and $a = v$ final v_0, where $v = x - x_0 t$.

Since optimization involves iterations, the difference between them equals 1. With $v_0 = 0$, this equation may be written as follows:

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}), \quad (4)$$

where: the position of the ith follower salp in the jth dimension is indicated by x_j^i and $i \geq 2$. It is possible to mimic the salp chains using Equations (1) and (4).

Figure 2 shows the block diagram of the SSA algorithm. The SSA method begins approximating the global optimum by starting several salps at random points, as this picture illustrates. Next, it determines which salp has the highest fitness by calculating each salp's fitness and designating the best salp's location as the source food that the salp chain will pursue by placing it in variable F. In the meanwhile, Equation (2) is used to modify the coefficient c1. Equation (1) is utilized to update the location of the leading salp for every dimension, while Equation (4) is used to update the position of the follower salps. Salp will be returned to the borders if any of it leaves the search area. With the exception of initialization, every step above is carried out repeatedly until a final requirement is satisfied [14].

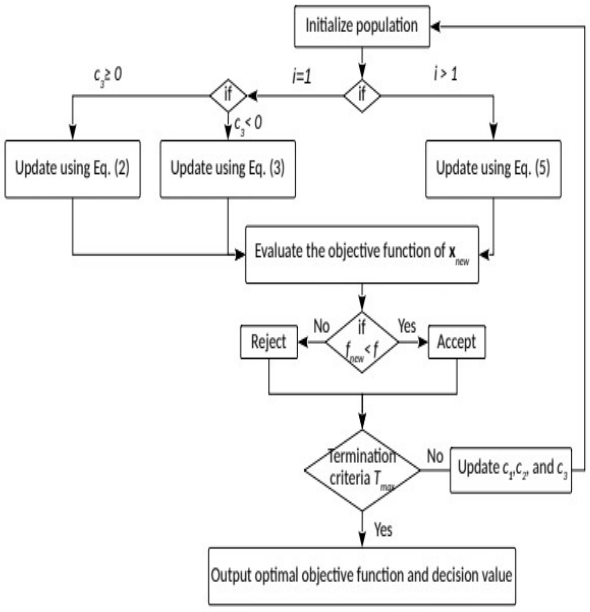

Figure 2. The block diagram of the SSA [14]

SSA keeps the optimal answer in the food source variable. As such, the process of utilizing search space is very competitive. Similar to several other optimization techniques, SSA continues to need help with poor convergence speed and local stagnation. To overcome these issues, Ref [16] suggested an improved SSA named ISSA. The ISSA improvements are as follows. Firstly, an inertia weight $w \in [0, 1]$ is presented in SSA. During the search, this parameter quickens the convergence pace. To get around local fixes, it also strikes a balance between exploration and exploitation capabilities. During the search, this parameter quickens the convergence pace. To get around local fixes, it also strikes a balance between exploration and exploitation capabilities. Equation (5) illustrates the update formula [15]:

$$x_j^i [k+1] = w\frac{1}{2}\left(x_j^i [k]\right) + c_1 x_j^i [k]) . \quad (5)$$

Lastly, the algorithm convergence rate could be more stable but often slow. Consequently, the OBL method is used [17, 18]. By using this method, the algorithm may be brought closer to the global optima, giving it greater flexibility to explore the search space and accelerating its convergence to the ideal value. Mathematically, OBL can be indicated as in (6):

$$x_j^i [k+1] = (u_i - l_i) + w c_1 x_j^i [k]) . \quad (6)$$

The Improved SSA is illustrated in Figure 3.

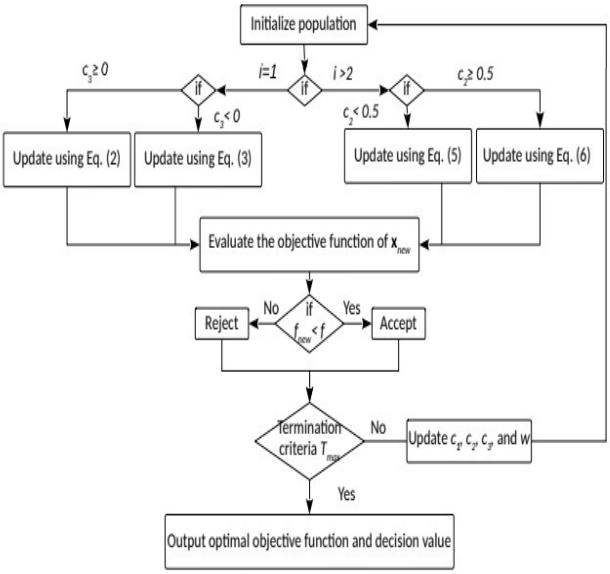

Figure 3. Block Diagram of ISSA [16]

C. DEEP NEURAL NETWORK (DNN)

DNNs are a subset of AI, also known as deep learning, which, according to John McCarthy, the computer scientist who invented the phrase in the 1950s, is "the engineering and science of creating intelligent machines with the ability to attain goals like humans" [19, 20].

Deep learning is a subdomain of neural networks in which neural networks include more than three layers, i.e., more than one hidden layer. The normal number of network layers utilized in deep learning nowadays ranges from five to over a thousand. Deep neural networks (DNNs) are used in this article to indicate the neural networks utilized in deep learning. Deep neural networks (DNNs) may learn higher-level characteristics with greater abstraction and complexity than shallower neural networks. Utilizing DNNs to process visual data is an illustration of this principle [21].

D. EVALUATIONJ METRICS

As part of the evaluation of any deep learning model, certain key classification metrics such as Accuracy, Recall, F1-score, and precision are computed. This is calculated by adding up the True Positives, False Positives, True Negatives, and False Negatives [22, 23].

In math, accuracy refers to the percentage of accurate

predictions made by a model. Among all metrics, it is the most important one because it indicates whether the model is successful. It is computed as follows [24, 25]:

$$\text{accuracy} = \frac{(TP+T\ )}{(TP+TN+FP+F\ )}. \quad (7)$$

IV. PROPOSED METHOD

In the first step, the dataset is processed. In the preprocessing module, the dataset is checked for missing values and then replaced with the mean value, and normalization is performed in this module. After preprocessing, the dataset is partitioned into two parts: the testing and training dataset.

In the second step of the framework, we performed feature selection using the improved Salp algorithm (ISSA). The DNN is used to predict thyroid dysfunction at the classification step. Then, the model is evaluated using accuracy, precision, and recall measures. Figure 4 illustrates the proposed method.

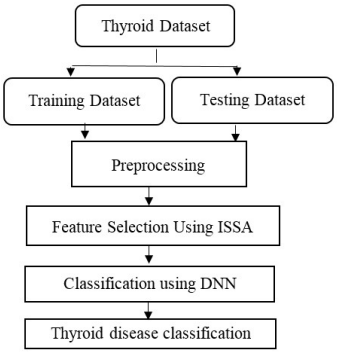

Figure 4. Proposed method design

Algorithm 1 illustrates the proposed ISSA-DNN algorithm.

Algorithm 1, The Proposed ISSA-DNN algorithm

Input: Thyroid disease dataset,
Output: Thyroid disease classification

Steps
1:xtrain, ytrain, xtest, ytest=split dataset randomly into(75:25).
2: Considering lb and ub initialize the Population of salp xi (i = 1, 2, ..., n)
3:WHILE (stopping condition is not met)
3.1:Compute the fitness of every search Agent (salp)
3.2:Specify nondominated Salps
3.3:Modify the selected features considering the obtained nondominated Salps
3.4:if the selected features becomes full Call the remove one resident
add the nondominated Salp
end
3.5:Select a food source from selected features: =SelectFood(selected features)
3.6:Modify c1 by Eq. (2)
3.7:for each salp(xi) if (i==1)
Modify leading salp position using Eq. (1)
ELSE
modify follower salp position using Eq. (4)
End
End
3.7: selected features=Amend the salps depend on the variables lower and upper bounds.
3.8:end while
4: m = Sequential()
m.add(Dense(256, input_shape=[x.shape[1]], activation='relu'))
m.add(Dropout(0.4))
m.add(Dense(128, activation='relu'))
m.add(Dropout(0.3))
m.add(Dense(63, activation='relu'))
m.add(Dropout(0.2))
m.add(Dense(1, activation='sigmoid'))
5: m.compile(optimizer=Adam(), loss='binarycrossentropy', metrics=['accuracy'])
6: m.evaluate(x_test, y_test)

Summary of the proposed DNN model is illustrated in Figure 5.

```
Model: "sequential_6"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_9 (Dense)             (None, 256)               1536

 dropout_6 (Dropout)         (None, 256)               0

 dense_10 (Dense)            (None, 128)               32896

 dropout_7 (Dropout)         (None, 128)               0

 dense_11 (Dense)            (None, 63)                8127

 dropout_8 (Dropout)         (None, 63)                0

 dense_12 (Dense)            (None, 1)                 64

=================================================================
Total params: 42,623
Trainable params: 42,623
Non-trainable params: 0
_________________________________________________________________
```

Figure 5. Summary of DNN

V. RESULTS AND DISCUSSION

This section describes the research results on thyroid illness prediction using deep learning and an optimization technique. We describe the outcomes of the feature selection strategy using ISSA and deep learning models. We divided the dataset into testing and training sets in a 75:25 ratio, using 75% of the data for model training and 25% for model testing.

After data splitting, we processed the data by replacing the missing values with the mean value. Then, the data is normalized. We use ISSA for feature selection, then, deep learning model with their best hyperparameter settings. Models are trained using significant characteristics chosen using feature selection methods and then tested with 25% of the data. We evaluate models in terms of accuracy. Table 2 shows the results of the proposed ISSA-DNN model. Figure 8 demonstrates the results compared with the related work.

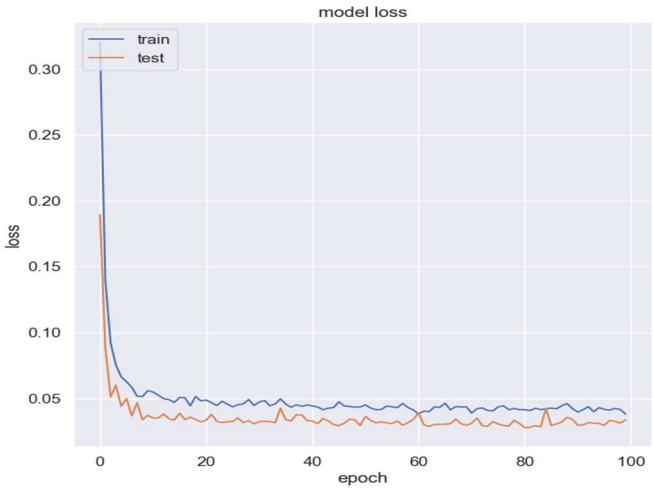

Figure 6. The loss results of the DNN model

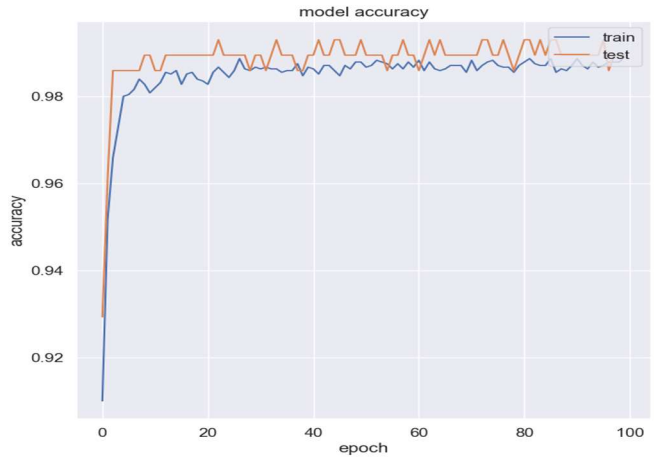

Figure 7. The accuracy of the DNN model

Table 2. Comparison of the proposed model with the related work

Reference	Model	Accuracy
K. Shankar [6]	optimal feature selection and KNN	97.49%
T. Alyas [7]	random forest	94.8%
G. Chaubey [8]	KNN	96.875%
I. Ioniţă and L. Ioniţă [9]	Decision tree	96.5%
S. Sankar [10]	XGBoost	98.59%
Ours	ISSA-DNN	99.68%

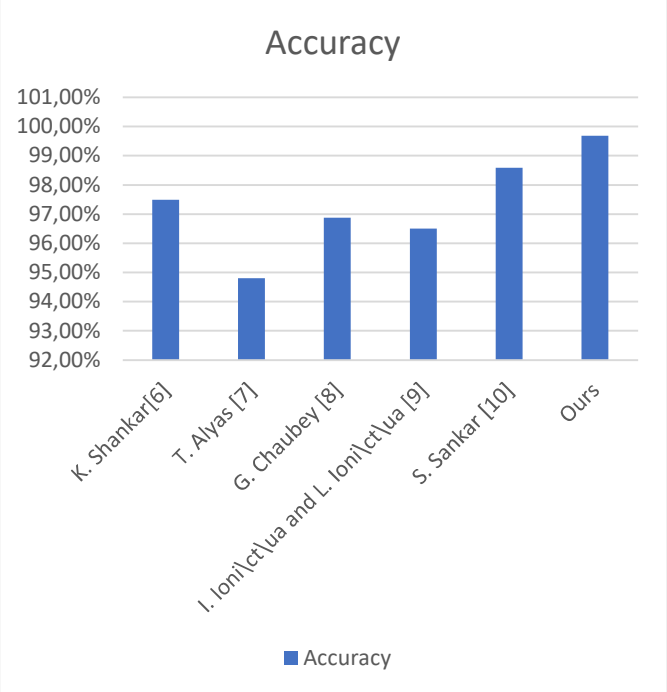

Figure 8. Comparison of the proposed model with the related work

K. Shankar [6] used optimal feature selection and KNN for classification and achieved an accuracy of 97.49%. T. Alyas [7] utilized RF for classification and attained an accuracy of 94.8%, I. Ioniţă and L. Ioniţă [9] achieved 96.5% with the decision tree, G. Chaubey [8] achieved 96.875% with the KNN algorithm, whereas S. Sankar [10] achieved 98.59% with XGBoost. The proposed ISSA-DNN outperformed the previous work with an accuracy score of 99.68%.

VI. CONCLUSIONS

Thyroid disease detection has become a major medical concern with an alarming rise in recent years, necessitating the use of effective computerized prediction algorithms. Model optimization and feature engineering are the main topics of current research; feature selection needs to be better explored. Feature selection and machine learning models are combined in this study to overcome these limitations. In this paper, a thyroid dysfunction prediction model is presented. The system is composed of three stages: data preprocessing, feature selection, and classification. The data is processed by filling missing values with mean values and then normalizing them using a standard scalar. ISSA is utilized for feature selection, and deep neural networks (DNN) are used for the classification process. Finally, the system is evaluated with accuracy. Results show that deep learning classifier-based selected attributes tend to achieve the highest accuracy of 99.68% when utilized with the DNN model. According to performance comparisons with the state-of-the-art approach, the proposed approach performs better than existing approaches.

References

[1] A. Tyagi, R. Mehra, and A. Saxena, "Interactive thyroid disease prediction system using machine learning technique," *Proceedings of the 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2018, pp. 689–693. https://doi.org/10.1109/PDGC.2018.8745910.

[2] F. Saiti, A. A. Naini, M. A. Shoorehdeli, and M. Teshnehlab, "Thyroid disease diagnosis based on genetic algorithms using PNN and SVM," *Proceedings of the 2009 3rd International Conference on Bioinformatics and Biomedical Engineering*, 2009, pp. 1–4. https://doi.org/10.1109/ICBBE.2009.5163689.

[3] S. Nussey and Saffron Whitehead., "Endocrinology: An integrated approach," *Oxford: BIOS Scientific Publishers*; 2001. PMID: 20821847. https://doi.org/10.1201/b15306.

[4] J. H. Romaldini, J. A. Sgarbi, and C. S. Farah, "Subclinical thyroid disease: subclinical hypothyroidism and hyperthyroidism," *Arq. Bras. Endocrinol. & Metabol.*, vol. 48, pp. 147–158, 2004. https://doi.org/10.1590/S0004-27302004000100016.

[5] L. Aversano *et al.*, "Thyroid disease treatment prediction with machine learning approaches," *Procedia Comput. Sci.*, vol. 192, pp. 1031–1040, 2021. https://doi.org/10.1016/j.procs.2021.08.106.

[6] K. Shankar, S. K. Lakshmanaprabu, D. Gupta, A. Maseleno, V. H. C. de Albuquerque, "Optimal feature-based multi-kernel SVM approach," *J. Supercomput.*, vol. 76, pp. 1128–1143, 2020. https://doi.org/10.1007/s11227-018-2469-4.

[7] T. Alyas, M. Hamid, K. Alissa, T. Faiz, N. Tabassum, and A. Ahmad, "Empirical method for thyroid disease classification using a machine learning approach," *BioMed Research Inernational*, vol. 2022, Article ID 9809932, pp. 1-10, 2022. https://doi.org/10.1155/2022/9809932.

[8] G. Chaubey, D. Bisen, S. Arjaria, and V. Yadav, "Thyroid disease prediction using machine learning approaches," *Natl. Acad. Sci. Lett.*, vol. 44, no. 3, pp. 233–238, 2021. https://doi.org/10.1007/s40009-020-00979-z.

[9] I. Ioniţă and L. Ioniţă, "Prediction of thyroid disease using data mining techniques," *BRAIN. Broad Res. Artif. Intell. Neurosci.*, vol. 7, no. 3, pp. 115–124, 2016. https://doi.org/10.24846/v25i3y201612.

[10] S. Sankar, A. Potti, G. N. Chandrika, and S. Ramasubbareddy, "Thyroid disease prediction using XGBoost algorithms," *J. Mob. Multimed.*, vol. 18, no. 3, pp. 1–18, 2022. https://doi.org/10.13052/jmm1550-4646.18322.

[11] R. Quinlan, "Thyroid disease," 1987 UCI Machine Learning Repository. https://doi.org/10.24432/C5D010.

[12] L. P. Madin, "Aspects of jet propulsion in salps," *Can. J. Zool.*, vol. 68, no. 4, pp. 765–777, 1990. https://doi.org/10.1139/z90-111.

[13] P. A. V. Anderson and Q. Bone, "Communication between individuals in salp chains. II. Physiology," *Proc. R. Soc. London. Ser. B. Biol. Sci.*, vol.

210, no. 1181, pp. 559–574, 1980. https://doi.org/10.1098/rspb.1980.0153.
[14] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm algorithm: A bio-inspired optimizer for engineering design problems," *Adv. Eng. Softw.*, vol. 114, pp. 163–191, 2017. https://doi.org/10.1016/j.advengsoft.2017.07.002.
[15] L. Abualigah, M. Shehab, M. Alshinwan, and H. Alabool, "Salp swarm algorithm: A comprehensive survey," *Neural Comput. Appl.*, vol. 32, pp. 11195–11215, 2020. https://doi.org/10.1007/s00521-019-04629-4.
[16] S. Ben Chaabane, A. Belazi, S. Kharbech, A. Bouallegue, and L. Clavier, "Improved salp swarm optimization algorithm: Application in feature weighting for blind modulation identification," *Electronics*, vol. 10, no. 16, p. 2002, 2021. https://doi.org/10.3390/electronics10162002.
[17] S. Mahdavi, S. Rahnamayan, and K. Deb, "Opposition based learning: A literature review," *Swarm Evol. Comput.*, vol. 39, pp. 1–23, 2018. https://doi.org/10.1016/j.swevo.2017.09.010.
[18] A. A. Ewees, M. Abd Elaziz, and E. H. Houssein, "Improved grasshopper optimization algorithm using opposition-based learning," *Expert Syst. Appl.*, vol. 112, pp. 156–172, 2018. https://doi.org/10.1016/j.eswa.2018.06.023.
[19] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, 2017. https://doi.org/10.1109/JPROC.2017.2761740.
[20] R. Jha, V. Bhattacharjee, and A. Mustafi, "Increasing the prediction accuracy for thyroid disease: a step towards better health for society," *Wirel. Pers. Commun.*, vol. 122, no. 2, pp. 1921–1938, 2022. https://doi.org/10.1007/s11277-021-08974-3.
[21] A. Selwal and I. Raoof, "A Multi-layer perceptron based intelligent thyroid disease prediction system," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 17, no. 1, pp. 524–533, 2020. https://doi.org/10.11591/ijeecs.v17.i1.pp524-532.
[22] N. A. Saeed and Z. T. M. Al-Ta'i, "Heart disease prediction system using optimization techniques," in *New Trends in Information and Communications Technology Applications*, 2020, pp. 167–177. https://doi.org/10.1007/978-3-030-55340-1_12.
[23] N. A. Saeed and Z. T. M. Al-Ta'i, "Feature selection using hybrid dragonfly algorithm in a heart disease predication system," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 2862–2867, 2019. https://doi.org/10.35940/ijeat.F8786.088619.
[24] S. Raisinghani, R. Shamdasani, M. Motwani, A. Bahreja, and P. Raghavan Nair Lalitha, "Thyroid prediction using machine learning techniques," *Proceedings of the Advances in Computing and Data Sciences: Third International Conference, ICACDS 2019, Ghaziabad, India, April 12-13, 2019, Revised Selected Papers, Part 1 3*, 2019, pp. 140–150. https://doi.org/10.1007/978-981-13-9939-8_13.
[25] P. S. Prerana and K. Taneja, "Predictive data mining for diagnosis of thyroid disease using neural network," *Int. J. Res. Manag. Sci. & Technol.*, vol. 3, no. 2, pp. 75–80, 2015.

***ELAF ALI ABDULKAREEM**, an Assistant Lecturer at Department of Al qura'n Science, Computer education 2011, Computer Science 2021, she works at University of Diyala/College of Islamic sciences. Research interests are data mining and security.*
Email: elaaf.ali1989@gmail.com

***SAMAH J. SABA**, Assistant Instructor at Department of Computer Science, she got BSc from Department of Computer Science in 2008 and received Master's Degree in Computer Science in 2021, she works at University of Diyala/College of Science. Research interests are Data Security, AI.*
Email: Samahjalil@uodiyala.edu.iq