Comparison of the Particle Swarm Optimization with the Genetic Algorithms as a Training for Multilayer Perceptron Technique to Diagnose Thyroid Functional Disease

Ali Sharifi and Kamal Alizadeh

1Department of Chemistry, Lorestan University, Khorramabad, Iran

*Corresponding author: Department of Chemistry, Lorestan University, Postal Code: 6813717133, Khorramabad, Iran. Email: alizadehkam@yahoo.com

Received 2019 December 28; Revised 2020 February 26; Accepted 2020 February 28.

Abstract

Background: Thyroid is a gland that controls body key functions. Since thyroid hormones are responsible for controlling metabolism, the thyroid glands are one of the major organs in the body. Disruption of the glands causes thyroid disease, which is one of the most common endocrine diseases worldwide. Early diagnosis of thyroid disease is difficult since early symptoms are easily confused with those of other illnesses. In recent years, the employment of computer techniques to utilize data mining and intelligent algorithms accelerates the early diagnosis of the disease.

Objectives: The current study aimed at evaluating the role of the new automatic method according to a multilayer perceptron (MLP) in the diagnosis of thyroid disease.

Methods: The study aimed at comparing the particle swarm optimization (PSO) with the genetic algorithms (GA) as training for MLP technique used to diagnose thyroid functional disease. The data were collected for three classes: 150 cases of euthyroidism, 30 cases of hypothyroidism, and 35 cases of hyperthyroidism. MLP was used to elucidate the pattern in the data and species responsible for separating the classes. Furthermore, improved PSO and GAs were used to train the system, and the sensitivity and specificity of the model were studied in terms of accuracy, sensitivity, and specificity. All analyses were performed using MATLAB software.

Results: For the proposed model, the simulations results showed that the GA algorithm had a higher performance than the PSO algorithm in the diagnosis of functional thyroid disease, and the means of classification accuracy, sensitivity, and specificity were 95%, 96%, and 96%, respectively.

Conclusions: The results of real data indicated that the GA-MLP can be used with high diagnostic accuracy as an effective tool to clinically diagnose thyroid functional disease. The current study was a step towards prototype system development of the classification of knowledge in this area with a much lower computational cost.

Keywords: Thyroid Functional Disease, Artificial Neural Networks, Chemometrics, Multi-Layer Perceptron, Genetic Algorithms, Particle Swarm Optimization

1. Background

Thyroid diseases are among the most common endocrine disorders worldwide. The thyroid is a small, butterfly-shaped gland that lies in the middle of the lower neck. Thyroid disorders include thyroid benign and cancerous nodules. Deficiency and abnormal hypersecretion of thyroid hormones are termed hypothyroidism and hyperthyroidism, respectively. Levothyroxine is a drug of choice in hypothyroidism, whereas hyperthyroidism is controlled by antithyroid drugs, which reduce the synthesis of thyroid hormones, or decreasing the thyroid tissue amount with radioiodine (131I) treatment or thyroidectomy (1). The exact description of thyroid hormones is not found in the ancient literature of Unani medicine. However, Jalinoos was the first physician to describe the anatomy of the thyroid gland (2). Ancient Unani physicians treated thyroid disorders with seaweed, which contains a high amount of iodine and advised heated sponge for goiter (3).

Symptoms of thyroid disease are easily confused with those of the other illnesses that can make the diagnosis difficult. Once thyroid disease is gotten initially, the disorder can be monitored even before starting the treatment. Early and cost-effective diagnosis and treatment of thyroid disorders is a challenge in both Unani and modern medicine.

Physicians’ diagnosis dataset and disease grouping methods are valuable in medical diagnosis for experts. Grouping methods can minimize fault diagnosis rates that
may happen to low-experience physicians, and the results can be achieved in the short-run (4).

Artificial neural networks (ANNs) (5) are applied in many researches in various medical fields. ANNs are also used to classify thyroid diseases. Keleş and Keleş (6) studied the thyroid disease diagnosis for the first time and introduced an expert system according to the neuro-fuzzy classifier with 95.33% accuracy. Temurtas (7) conducted a study on designing a model for thyroid disease diagnosis with 93.19% accuracy using the multilayer perceptron (MLP) and Levenberg-Marquardt (LM). The study by Chen et al. (8) on thyroid disease suggested a particle swarm optimization (PSO) that optimized support vector machines with the fisher score (FS-PSO-SVM) CAD system and an average accuracy of 97.49%. Razia et al. (9) used the UCI dataset to compare MLP and radial basis function performances. Ma et al., (10, 11) proposed a convolutional neural network according to the computer-aided diagnostic method for thyroid diseases using the single-photon emission computerized tomography (SPECT) images.

Based on the stated researches, it is assumed that the unique capability of ANNs can be efficient and useful to categorize, differentiate, and diagnose diseases. In the current study, it was attempted to diagnose the disease with simple parameters through chemometric methods using MLP. The PSO and genetic algorithm (GA) techniques optimizing models were used to diagnose thyroid functional disease. The overall suggested framework of algorithms is illustrated in Figure 1.

2. Objectives

The current study aimed at investigating the early diagnosis of acute kidney disease by increasing the accuracy of diagnosis using the MLP method to reduce the duration and cost of treatment.

3. Methods

3.1. Data

The current study used the dataset including the results of five laboratory tests regarding the thyroid gland from the UCI machine learning repository as follows: total serum thyroxin (T₄), the maximal absolute difference (∆TSH), T₃-resin uptake test (RT₃U), total serum triiodothyronine (T₃RIA), and basal thyroid-stimulating hormone (TSH) (12). The data were collected for three classes: 150 cases of euthyroidism, 30 cases of hypothyroidism, and 35 cases of hyperthyroidism. The diagnoses were based on a complete medical record, including scan images, anamnesis, etc. The input, hidden, and output layers of NN to diagnose thyroid functional disease are demonstrated in Table 1.

| Layers          | Details                  |
|-----------------|--------------------------|
| **Input layer** | Factors                  |
|                 | T₄                       |
|                 | T₃RIA                    |
|                 | RT₃U                     |
|                 | TSH                      |
|                 | TRH                      |
| **Hidden layer(s)** | Number of hidden layers |
|                 | 1                        |
| **Output layer** | Dependent variables      |
|                 | Displacement             |
|                 | Number of units          |
|                 | 1                        |
|                 | Rescaling method for scale dependents |
|                 | Standardized             |
|                 | Activation function      |
|                 | Identity                 |
|                 | Error function           |
|                 | Sum of squares           |

3.2. Multilayer Perceptron

ANNs are utilized in innovation and science in biology, chemistry, and physics. An ANN is the human neural architecture mathematical representation reflecting its learning and generalization abilities (13). The ANN basic architecture applied to implement this classification task is a supervised network of multilayer feed-forward with the algorithm of backpropagation learning (14). The training and testing algorithm was coded using the basic MATLAB in order to analyze the data.

The MLP includes interconnected neurons or nodes in a simple system (Figure 2), which is a model among an input and an output vector representing a nonlinear mapping. The nodes can be viewed as weighted directed graphs and generate an output signal, which is a function of the sum of the inputs to the modified node by a simple nonlinear transfer or activation function.

A typical MLP network of processing nodes comprises at least three layers: an input layer that gets external inputs, at least one output layer, and a hidden layer that creates the classification. MLP is a prominent NN model owing to its comparably simple algorithm and the clear architecture. Due to the variety of NN training methods, choosing the most optimal training method can increase the accuracy of NN predictions. The current study employed three algorithms to train NN: the Levenberg-Marquardt algorithm, PSO, and GAs.
3.3. Particle Swarm Optimization
PSO is a populace-based stochastic streamlining procedure introduced by Kennedy and Eberhart (15), propelled by social conduct of bird flocking or fish schooling, in which every person is treated as a minuscule particle in the
n-dimensional space, with the velocity vector of particle $i$ and position vector demonstrated as $X_i(t) = (X_{i1}(t), X_{i2}(t), ..., X_{in}(t))$ and $V_i(t) = (V_{i1}(t), V_{i2}(t), ..., V_{in}(t))$. The particles travel as indicated by the following formula:

$$V_{id}(t + 1) = W \times V_{id}(t) + c_1 r_1 (P_{id}(t) - X_{id}(t)) + c_2 r_2 (P_{gd}(t) - X_{id}(t))$$

$$X_{id}(t + 1) = X_{id}(t) + V_{id}(t + 1), \quad i = 1, 2, \ldots, M; \quad d = 1, 2, \ldots, n$$

where acceleration coefficients including $c_1$ and $c_2$ and particle $i$ are recognized as the personal best position (pbest), which is the best previous position (the position providing the best value of fitness) shown as vector $P_i = (P_{i1}, P_{i2}, ..., P_{in})$, and the global best position (gbest) is the best position between the particles individual best positions in the populace demonstrated as vector $P_g = (P_{g1}, P_{g2}, ..., P_{gn})$.

Two random numbers uniformly dispersed in $(0, 1)$ are $r_1$ and $r_2$. Commonly, the $V_{id}$ value is constrained in the interval $[-V_{max}, V_{max}]$. Shi and Eberhart introduced inertia weight $w$, generally applied to accelerate the algorithm convergence speed (16, 17).

In PSO, in the swarm, every particle demonstrates a point in the solution space. The particles move around the space to find the optimum solution while thinking about the best solution (point) attended using the entire swarm and using the individual. Inertia weight is another significant parameter that adjusts the ability of algorithm local search and the global search. The large inertia weight values increase the global search algorithm. Once the algorithm converges to the optimum solution, it can be considered as a disadvantage to select a large value for the inertia weight. Hence, strategies are offered to alter the inertia weight (17, 18) adaptively. Regarding the classification and performance, a constant inertia weight was utilized in the current study. The PSO algorithm diagram is demonstrated in Figure 3.

3.4. Genetic Algorithms

GA is a technique of global search applied in the computational approach classified as a heuristic one (19). GA is a search technique to find approximate solutions to optimization problems. Therefore, it is not ensured to get the optimal solution. It is a global search procedure obtained and meant for proficient search and plan techniques. Three essential genetic operators manage the current exploration: mutation, selection, and crossover. GAs utilize these systems to locate an ideal explication for any multidimensional issue (20, 21). GA is a particular class of evolutionary algorithms (also recognized as evolutionary computation), which utilizes methods enlivened by evolutionary biology, including crossover, selection, mutation, and inheritance. Lately, the inclination to hybridize ANN and GA is common among authors (22-25). The preferences suggested by the current procedures shape a preferable GA-ANN hybrid intelligence framework, which could promote speculation and, in the meantime, facilitate the ANN configuration procedure.

3.5. Data Cleaning and Preprocessing

The database should be preprocessed before assessment by the NN data in training. For this purpose, a few methodologies are available. To lie within the interval $(0, 1)$, data are commonly scaled according to the most usually utilized transference function, which is the so-called logistic one. Furthermore, it is exhibited that instances of missing information ought to be expelled from the database to enhance the characterization accomplishment of the system (26). A reduction in the network performance classification is observed for the database that is imbalanced, the ones with various number cases for any class (27).

3.6. Software

The PSO and GAs, as training for the MLP models, were calculated by classification toolbox 4.0 in MATLAB environment (MATLAB R2018a, The Mathworks Inc., Natick).

4. Results

The designed network has just one output neuron, similar to the current study, figured as a binary categorizing...
issue once the three classes are 2, 1, and 0. In the current study, five inputs were selected and viewed as elements to predict thyroid functional disease. The inputs were based on thyroid functions: the isotopic displacement method by which $T_4$ was measured within 5.2 - 11.6 μg/dL as the normal range; radioimmunoassay by which $T_3$RIA was measured within 0.8 - 2.5 μg/L as the normal range, radioimmunoassay by which RT$_3$U was measured within 94% - 118% as the normal range, TSH calculated within 0.3 - 2.5 μIU/mL as the normal range; and $\Delta$TSH was the 5th variable in which the TSH value maximal absolute difference after the injection of 200 μg TRH was compared with that of the basal value.

The study employed a feed-forward MLP with one hidden layer. Then the Levenberg-Marquardt algorithm was used to train the MLP. Furthermore, the NN initial weights and bias were defined by PSO and GA. Last but not least, the weights and bias in each layer were coded in sequence in PSO and GA. After designing the network data processing by the ANN technique, training and testing were the two key phases. To implement the algorithm, no standard NN software was applied; rather, the basic MATLAB was employed. The normalized inputs were fed into the network to train the NN until the network was given adequately small goal mean squared error (MSE). In the current study, a fully-connected feed-forward NN with one hidden layer was applied.

5. Discussion

All 215 individuals with the five features mentioned in these algorithms participated in the study, and data of 70% of them were selected as training, 10% as validation, and 20% as test results. To achieve the best classification, different architectures for NN were used; 4 - 20 neurons in the middle layer and weight optimization of the features through all three algorithms in 400 Epoch. The MSE index was used as the performance indicator calculation, and the best model for each algorithm was reported. In case of no improvement in fitness values after 40 Epoch, the algorithm was stopped.

An expert diagnostic system based on the comparison of PSO with GA as training for MLP technique classifier was introduced to diagnose thyroid functional disease. The results mentioned above for the problems of the thyroid functional disease classification confirmed that the network diagnoses thyroid functional disease with the 95% performance of the network, as demonstrated in Table 2. The significance of each term was determined by $P$- and $R\%$ values. The smaller the $P$ value and the larger the $R\%$ value, the more significant is the corresponding term. The results of the analysis of variance (ANOVA) in the diagnosis of thyroid disease after applying GA as NN training, validation, and testing showed a significant difference at the 0.05 level and the better performance of the GA-MLP model compared with the PSO-MLP and BP-ANN models.

The performance of the suggested method was contrasted with other techniques in the literature employing the current datasets. The results outline is demonstrated in Table 3. The examination demonstrated that the employed strategy had the second most astounding diagnosis accuracy, despite the fact that its accuracy was somewhat lower than those of some of the best previous methods. Due to the advantages of the mentioned method with respect to reported ones, it may be used as an alternative method in most of the laboratories.

5.1. Conclusions

The current study results recommended that ANNs may likewise be utilized to develop a precise and successful model to predict thyroid functional disease. The results for the classification of the thyroid functional disease indicated that the system characterizes thyroid functional disease with the NN performance of 95% for GA-MLP and performance of 85.5% for PSO-MLP. The NN performance can further be upgraded by defining a higher goal error and training the ANN to progress the number of epochs to accomplish the predetermined objective.

The current study aimed at formulating diagnostic models of functional thyroid disease with suitable analytical characteristics by MLP ANNs. The results showed that GA is more successful than PSO in the diagnosis of functional thyroid disease and that the proposed GA-MLP can achieve very high diagnosis accuracy (95%). To streamline the diagnostic procedure in daily routine, prevent misdiagnosis, and reduce the cost of diagnosis in the early stages of the disease without using invasive procedures, the GA-MLP can be employed.

Footnotes

**Authors’ Contribution:** Study design: Kamal Alizadeh; analysis and interpretation of data and editing of the manuscript: Ali Sharifi and Kamal Alizadeh.

**Conflict of Interests:** The authors declared no conflict of interest.

**Ethical Approval:** The study did not include any human participants.

**Funding/Support:** The study was supported by the School of Science, Lorestan University.
Table 2. The Statistical Parameters of the Classifier Methods

| Classifier | Training | Validation | Testing | Mean |
|------------|----------|------------|---------|------|
| BP-ANN     | 71       | 70         | 67      | 68   | 70   | 68   |
| PSO-MLP    | 90       | 89         | 85      | 97   | 88   | 85.5 |
| GA-MLP     | 97       | 96         | 95      | 99   | 97   | 95   |

Table 3. The Diagnostic Accuracy of Approaches Available in the Literature and the Current Study for Thyroid Functional Disease

| Method | Method Information | Classification Accuracy, % | Reference |
|--------|--------------------|---------------------------|-----------|
| GDA, WSVM | The feature extraction - feature reduction phase, classification phase, and GDA, WSVM test for thyroid disease phase determination, individually. | 91.86 | (28) |
| AIRS | Thyroid disease diagnosis with a novel hybrid machine learning technique comprising this grouping system. By hybridizing AIRS with a progressed fuzzy weighted preprocessing, a system was achieved to explain this diagnosis issue via classifying. | 81.00 | (29) |
| AIRS with fuzzy weighted preprocessing | | 85.00 | |
| MLPNN with LMe | A comparative thyroid disease diagnosis was comprehended by the probabilistic, multilayer, learning vector quantization NNs. | 93.19 | (7) |
| PNN | By LDA, an experimental study was performed to reach more reliable accuracy. | 94.62 | |
| LDA | The research examined RBFNN and MLPNN for the structural classification of thyroid diseases. | 94.8 | (30) |
| RBFNN | Different characterization models were utilized to order T4, T3, TSH, and goiter. Many classification methods, such as K nearest neighbor, Naive Bayes, and support vector machine, were applied. | 94.44 | (31) |
| Naive Bayes classifier | Application of medical information gained based on AIRS. | 95.9 | |
| AIRS | The features, such as variance, mean, histogram feature, coefficient of local variation feature, homogeneity, and NMSID, were extracted and applied to train the classifiers such as SVM and ELM. | 93.56 | (32) |
| SVM | To diagnose thyroid disorders, a fuzzy rule-based classifier was designed. | 93.56 | (33) |
| ELM | A three-stage expert system based on support vector machines for thyroid disease diagnosis. | 93.56 | (34) |
| Radon-based approach | | 90.9 | |
| Fuzzy rule-based expert system | To diagnose thyroid functional disease. | 97 | (35) |
| The current study method (GA-MLP) | A comparison of PSO with GAs as training for the MLP method to diagnose thyroid functional disease. | 95 | This work |
| The current study method (PSO-MLP) | | 85.5 | |

References

1. Fauci AS, Braunwald E, Kasper DL, Longo DL, Loscalzo J. Harrison’s Principles of Internal Medicine. 17th ed. New York: Mc Graw Hill Press; 2008.
2. Jameson JL, De Groot LJ. Endocrinology Adult and Paediatrics: The thyroid gland. 6th ed. Philadelphia: Elsevier Saunders; 2010.
3. Harding AS. Milestone in Health and Medicine. 263. Oryx Press; 2000. 103 p.
4. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artif Intell Med. 2001;23(1):39–109. doi: 10.1016/s0933-3657(01)00077-x. [PubMed: 11470218].
5. Nageshwar Rao B, Lakzi Sriniwas Reddy D, Bhaskar G, editors. Thyroid Diagnosis Using Multiplayer Perceptron. International Conference on E-Business and Telecommunications. Springer Cham; 2019. p. 452–9.
6. Keleş A, Keleş A. ESTDD: Expert system for thyroid diseases diagnosis. Expert Systems with Applications. 2008;34(1):242–6. doi: 10.1016/j.eswa.2006.09.028.
7. Temurtas F. A comparative study on thyroid disease diagnosis using neural networks. Expert Systems with Applications. 2009;36(1):944–9. doi: 10.1016/j.eswa.2007.10.010.
8. Chen HL, Yang B, Wang G, Liu J, Chen YD, Liu DY. A three-stage expert system based on support vector machines for thyroid disease diagnosis. J Med Syst. 2012;36(3):1953–63. doi: 10.1007/s10916-011-9565-8. [PubMed: 21286792].
9. Razia S, Narasingarao MR, Sridhar GR. A decision support system for prediction of thyroid disease-a comparison of multilayer perceptron neural network and radial basis function neural network. Journal of Theoretical & Applied Information Technology. 2015;80(3).
10. Ma L, Ma C, Liu Y, Wang X. Thyroid Diagnosis from SPECT Images Using Convolutional Neural Network with Optimization. Comput Intell Neurosci. 2019;2019:6212759. doi: 10.1155/2019/6212759. [PubMed: 3076599]. [PubMed Central: PMC6350547].
24. Ahmad F, Mat Isa NA, Hussain Z, Osman MK, Sulaiman SN. A GA-based
      feature selection and parameter optimization of an ANN in diagnosing
      breast cancer. Pattern Analysis and Applications. 2014;18(4):861-70.
doi:10.1007/s10044-014-0375-9.
25. Xiao L, Wang Z, Peng X, editors. Research on Congestion Control
      Model and Algorithm for High-Speed Network Based on Genetic Neu-
      ral Network and Intelligent PID. International Conference on Wireless
      Communications, Networking and Mobile Computing. 2009. p. 1-6.
26. Amato F, López A, Peña-Méndez EM, Vanhara P, Hampl A, Havel J.
      Artificial neural networks in medical diagnosis. Journal of Applied
      Biomedicine. 2013;11(2):47-58. doi:10.2478/v10136-012-0031-x.
27. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD.
      Training neural network classifiers for medical decision making: the
      effects of imbalanced datasets on classification performance. Neural
      Netw. 2008;21(2-3):427-36. doi:10.1016/j.neunet.2007.12.031. [PubMed:
      18272329]. [PubMed Central: PMCID:PMC2344433].
28. Dogantekin E, Dogantekin A, Avci D. An expert system based on Gen-
      eralized Discriminant Analysis and Wavelet Support Vector Machine
      for diagnosis of thyroid diseases. Expert Systems with Applications.
      2013;40(1):146-50. doi:10.1016/j.eswa.2010.06.029.
29. Polat K, Sahan S, Gunes S. A novel hybrid method based on arti-
      ficial immune recognition system (AIRS) with fuzzy weighted pre-
      processing for thyroid disease diagnosis. Expert Systems with Applica-
      tions. 2007;32(4):414-7. doi:10.1016/j.eswa.2006.02.007.
30. Rasitha G. Predicting Thyroid Disease using Linear Discriminant
      Analysis (LDA) Data Mining Technique. Communications on Applied
      Electronics. 2016;4(1):4-6. doi:10.5201/cae201665990.
31. Erol R, Ogulata SN, Sahin C, Alparslan ZN. A radial basis function neu-
      ral network (RBFNN) approach for structural classification of thyroid
      diseases. J Med Syst. 2008;32(3):215-20. doi:10.1007/s10916-007-9125-5.
      [PubMed: 18444358].
32. Chandel K, Kunwar V, Sabitha S, Choudhury T, Mukherjee S. A com-
      parative study on thyroid disease detection using K-nearest neigh-
      bor and Naïve Bayes classification techniques. CSI Transactions on ICT.
      2017;4(2-4):313-9. doi:10.5120/cae201665990.
33. Kodaz H, Orsen S, Arslan A, Günes S. Medical application of infor-
      mation gain based artificial immune recognition system (AIRS): Diagno-
      sis of thyroid disease. Expert Systems with Applications. 2009;36(2):3086-92.
      doi:10.1016/j.eswa.2008.01.026.
34. Selvathi D, Sharnitha VS, editors. Thyroid classification and segmen-
      tation in ultrasound images using machine learning algorithms. IEEE
      International Conference on Signal Processing, Communication, Com-
      pting and Networking Technologies (ICSCCN), 2011. p. 836-41.
35. Asaad Sajadi N, Borouzet S, Mahjub H, Farhadian M. Diagnosis
      of hypothyroidism using a fuzzy rule-based expert system. Clinical
      Epidemiology and Global Health. 2019;7(4):519-24. doi:10.1016/j.jcegh.2018.11.007.