Fuzzy Rule-based Expert System for Diagnosis of Thyroid Disease

S. Amrollahi Biyouki
Department of Industrial Engineering, Amirkabir University of Technology
Tehran, Iran
sajjadamrollahi@aut.ac.ir

M.H. Fazel Zarandi
Department of Industrial Engineering, Amirkabir University of Technology
Tehran, Iran
zarandi@aut.ac.ir

I.B. Turksen
Department of Mechanical and Industrial Engineering,
University of Toronto
Toronto, Canada
turksen@mie.utoronto.ca

Abstract—The diseases in glands of human bodies have been increased with high rate in the last decade. Thyroid is one of these glands that its disease has spread worldwide. The main function of thyroid gland is to balance the metabolism and cells' activity. Since it does its own task abnormally, thyroid disorders will occur and the negligence of them may cause irreparable events. Because of inaccessibility of Endocrinologist experts for most people, modeling and developing an expert system for diagnosis thyroid's disease that can be accessible in everywhere is vital.

This paper presents a fuzzy rule-based expert system for diagnosis thyroid's disease. This proposed system includes three steps: pre-processing (feature selection), neuro-fuzzy classification and system evaluating. In the proposed system, the process of diagnosis encounters with vagueness and uncertainty in final decision. So, we handled the imprecise knowledge by using fuzzy logic. In neuro-fuzzy classification step, we generated initial fuzzy rules by k-means algorithm and then scaled conjugate gradient algorithm (SCG) was used to determine the optimum values of parameters. In the last step, we used the generated fuzzy rules to model and evaluate the system.

This system can help non-experts who are suspicious of their thyroid function or it can be used as a diagnosis assistance system to help experts for assuring their diagnosis.

Keywords— medical expert system; fuzzy rule-based; thyroid disease; medical diagnosis; neuro-fuzzy classification

978-1-4799-6926-5/15/$31.00 ©2015 IEEE

I. INTRODUCTION

A. Thyroid Gland

The thyroid gland is one of the most important glands in the body that is located in the neck. The function of the thyroid gland is to take iodine, found in many foods, and convert it into thyroid hormones: thyroxin (T4) and triiodothyronine (T3). Thyroid cells are the only cells in the body which can absorb iodine.

The thyroid gland has many diseases. The most of these are goiters, thyroid cancer, solitary thyroid nodules, hyperthyroidism, hypothyroidism, thyroiditis, etc. [1]

In hyperthyroidism condition, thyroid produces hormones more than its normal condition and the cells of body consume more energy and faster in comparison with normal condition, but in the other condition hypothyroidism, the function of thyroid is on the contrary. In this condition, thyroid produces hormones less than its normal and the cells of body consume less energy [2]. Goiter is an abnormal enlargement of thyroid gland and the most common cause of Goiter is the lack of iodine in the diet. Thyroid cancer starts when the cells in the thyroid begin to change and grow uncontrollably, forming a mass called a tumor.

B. Fuzzy Expert System

The development of artificial intelligence (AI) methodology has been recognized as an important requirement in complex problem solving situations [3]. One of the branches of Artificial Intelligence is Expert Systems that provide expert advice by choosing the most related and expertise solution of the problem.

The use of artificial intelligence in medicine (AIM) started in the end of the 1960's and some experimental systems were introduced such as:

- INTERNIST was a rule-based expert system for the diagnosis of complex problems in general internal medicine. This system covered 80% of the knowledge of internal medicine, but was criticized for the shallowness of their knowledge.
- MYCIN was a rule-based expert system to diagnosis and recommend treatment for certain blood infections
- CASNET was an expert system for the diagnosis and treatment of glaucoma.
- EXPERT was an extension generalized of the CASNET formalism which was used in creating consultation systems in rheumatology and endocrinology.
- ONCOCIN was a rule-based medical expert system for oncology protocol management. It was designed to assist physicians in treating cancer patients receiving chemotherapy. [4]; [5]

Because of the difficulty of considering all of the measurements and symptoms of a disease, experts often have imprecise knowledge and this imprecise knowledge causes uncertainty and vagueness in diagnosis procedure. For handling imprecise knowledge, Fuzzy logic is often applied which is introduced by Zadeh [6]. By applying Fuzzy logic to Expert systems, Fuzzy expert systems (FES) were developed in medical diagnosis.

Fuzzy expert systems have been already used in medical diagnosis area such as: diagnosis of MS [5], Asthma [7] and heart disease diagnosis [8] etc.

C. *Architecture of Fuzzy Expert Systems*

Expert system (ES) is a knowledge-based system that employs knowledge about its application domain and uses an inference reasoning procedure to solve problems that would otherwise require human expertise. Expert systems contain four main modules: knowledge-based, working memory, inference engine, graphical user interface [9].

In knowledge-based module, the knowledge can gather based on direct approach e.g. interview with experts, using questionnaire or indirect approach by using data [3]. The knowledge is in the form of IF-THEN rules. The facts of fired rules store in working memory module for reasoning [13].

Inference engine contains an inference reasoning procedure that could have the ability to conclude [13]. And finally; the system should design and develop for none-experts who have no knowledge in that specific domain. A well-graphical user interface allows the user to ask the system how the result is concluded.

The differences between Fuzzy Expert Systems (FES) and Expert Systems (ES) are in the architecture of system and the rules exist in knowledge-based module. In the architecture of FES, Fuzzification and Defuzzification modules are added to ES modules. The task of these new modules is to convert crisp data to fuzzy form of data that score in the interval [0, 1] and vice-versa. Also, the knowledge of FES is in fuzzy forms of IF-THEN.

Figure1 represents the structure of Expert Systems and figure 2 represents the structure of Fuzzy Expert Systems for realizing the differences of these systems.

.

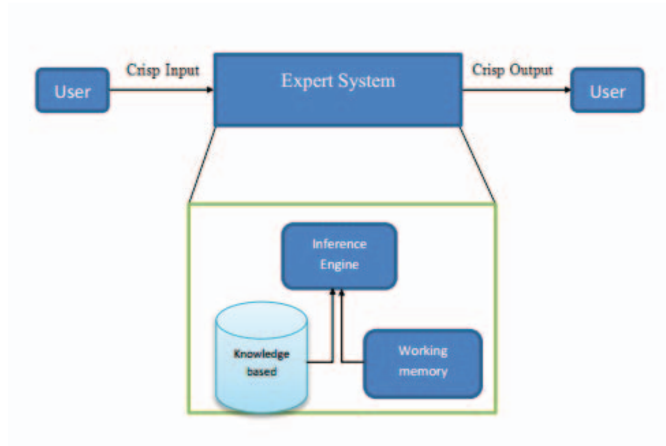

Fig. 1. Structre of Expert Systems

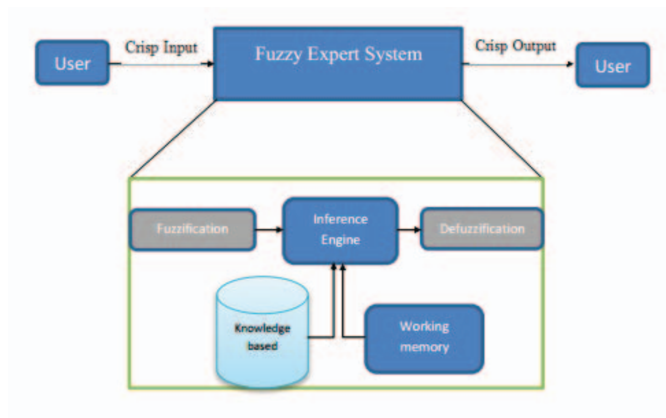

Fig. 2. Structure of Fuzzy Expert Systems

II. BACKGROUND

A. *Literature of thyroid diagnosis*

In the diagnosis of thyroid disease literature, different methods were applied. Some methods which were reviewed are given below:

- Adaptive conic section function neural network (CSFNN) [10]
- Multi-layer perceptron with fast back-propagation (MLP with fbp) [10]
- Multi-layer perceptron with back-propagation (MLP with bp) [10]
- Radial basis function (RBF) [10]
- Learning vector quantizer (LVQ) [11]
- Probabilistic potential function neural network (PPFNN) [11]
- Linear discriminant analysis (LDA) [12]
- C4.5 with default learning parameters (C4.5-1) [12]
- C4.5 with parameter c equal to 5 (C4.5-2) [12]
- C4.5 with parameter c equal to 95 (C4.5-3) [12]
- Neuro fuzzy classification (NEFCLASS -J) [2]
- Artificial immune recognition system (ARIS) [13]
- GDA – WSVM expert system [1]

In former studies, articles attempted to increase the classification accuracy. Although the classification accuracy is an important feature of a system but in this study, we focused on generated fuzzy-rules and the values of membership function's parameters. We used a new neuro-fuzzy classification based on k-means algorithm and scaled conjugate gradient (SCG) algorithm [14] for determining the values of parameters.

B. System Procedure

The procedure of developing the system organized as below:

- **Pre- processing phase**
  In this phase, we selected more significant features for diagnosis by consulting an Endocrinologist. This phase explained in section III.A.

- **Classification phase**
  In proposed Neuro-fuzzy Classification, K-means algorithm was used to generate initial fuzzy rules for determination the values of parameters. Finally, for obtaining the optimum values of nonlinear parameters, scaled conjugate gradient (SCG) algorithm [14] was applied. In section III.B, the proposed neuro-fuzzy classification and the structure of the system explained.

- **System modeling and evaluation**
  In section III.C, the Membership functions of features and generated rules represented and then the performance of the proposed system evaluated by calculating of classification accuracy.

Figure 3 represents the procedure of developing the proposed expert system for diagnosis of thyroid disease.

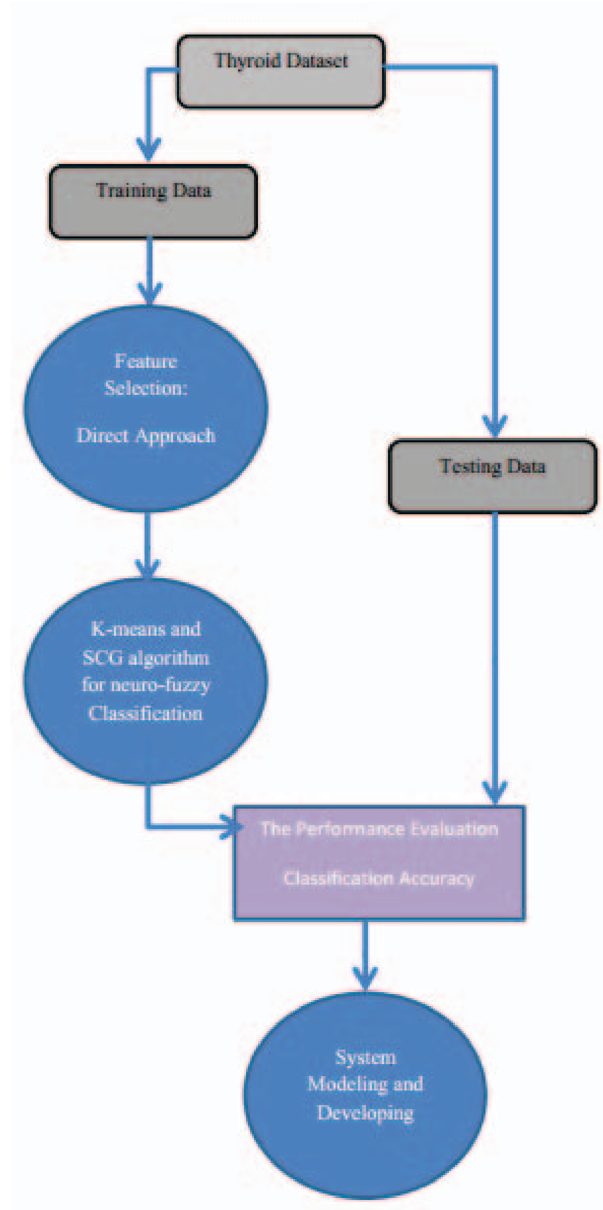

Fig. 3. the procedure of developing the proposed expert system for diagnosis of thyroid disease

III. THYROID DATASET

We used thyroid dataset that was taken from UCI machine learning repository. This dataset contains 215 samples, 3 classes and for each sample five features. These classes are assigned to the values that named as hyperthyroidism, hypothyroidism and Euthyroid (normal function).
The features of each sample are considered as below:

- Ftr_1 : T3-resin uptake test (percentage)

- Ftr_2 : Total serum thyroxine recognized as T4
- Ftr_3 : Total serum triiodothyronine recognized as T3
- Ftr_4 : Thyroid-stimulating hormone recognized as TSH
- Ftr_5 : Maximal absolute difference of TSH

These 215 samples include 150 instances of Euthyroid class, 35 instances of hyperthyroidism class and 30 instances of hypothyroidism class.

IV. Proposed Fuzzy Expert System

A. *Pre-Processing*

This phase was based on direct approach. By consulting with an expert Endocrinologist, four first features have been prescribed in the measurement of thyroid gland recently. So, we deleted Ftr_5 value and used the other features in our proposed system (feature selection). The performance of used features introduced in section C.

B. *The Proposed Neuro-fuzzy Classification*

Fuzzy classification is the task of partitioning a selected feature into fuzzy classes [15]. By employing adaptive networks to fuzzy classification, neuro-fuzzy classifier was introduced by Jang and his co-worker Sun in 1993 [15].In this study, the classifier used is based on Jang's neuro-fuzzy classifier.

In the proposed classifier, Gaussian membership function was used for fuzzy sets description and Sugeno inference method to get fuzzy results. At first, k-means algorithm was used to generate initial fuzzy rules for determination the values of parameters. Finally, for obtaining the optimum values of nonlinear parameters, scaled conjugate gradient (SCG) algorithm [14] was applied.

By applying proposed neuro-fuzzy classifier to thyroid dataset, three rules was generated that represent in next section.

Figure 4 demonstrates the structure of neuro-fuzzy classifier constructed for the proposed system.

By entering the features of thyroid's data as input data, this network will determine the class of thyroid function.

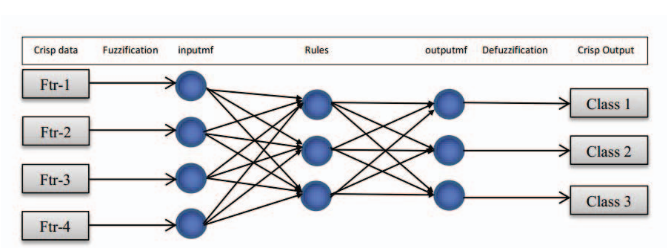

Fig. 4. Structure of the proposed neuro-fuzzy classifier

C. *System Modeling and Evaluation*

- Parameters of Proposed System

By applying the proposed neuro-fuzzy classifier, the parameters of the proposed diagnosis system was determined. These parameters of system are the membership function shapes and the parameters of each input and cluster. These parameters for each input represent as below:

- Feature 1: T-3 resin uptake test value (percentage value)

This feature is done to check thyroid function. Thyroid function depends on the action of many different hormones, including thyroid-stimulating hormone (TSH), T3, and T4. It helps see how much thyroxin binding globulin (TBG) is available. TBG is a protein that carries most of the T3 and T4 in the blood. Figure 5 demonstrates the membership function of feature 1 and table 1 represents the parameters of Gaussian membership function for each cluster of feature 1.

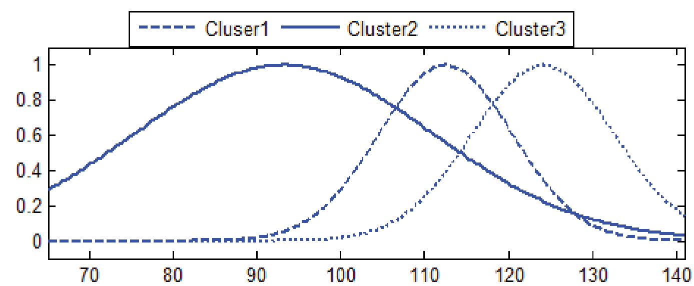

Fig. 5. Memnership functions of feature 1

TABLE I. MFs PARAMETERS OF FEATURE 1

Ftr_1	Parameters	
	Parameter 1	*Parameter 2*
Cluster 1	112.5	8.006
Cluster 2	93.17	17.96
Cluster 3	124.1	8.516

- Feature 2: total serum thyroxine (T4)

Thyroid produces a hormone, thyroxine, which is known as T4. This hormone plays a role in several of body's functions, including growth and metabolism. Some of T4 is called free T4. This means that it has not bonded to protein in your blood.

Total serum thyroxine test known as a total T4 test, measures both kinds of T4 (the T4 that has bonded to protein and the free T4) in blood. Figure 6 demonstrates the membership function of feature 2 and table 2 represents the parameters of Gaussian membership function for each cluster of feature 2.

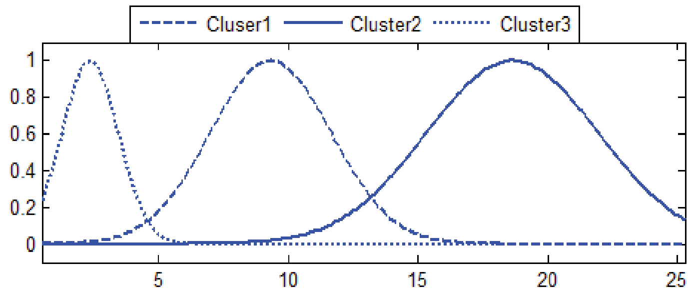

Fig. 6. Memnership functions of feature 2

TABLE II. MFS PARAMETERS OF FEATURE 2

Ftr_2	Parameters	
	Parameter 1	Parameter 2
Cluster 1	9.315	2.339
Cluster 2	18.64	3.311
Cluster 3	2.356	1.092

- Feature 3: total serum triiodothyronine (T3)

Thyroid produces triiodothyronine, a hormone known as T3. T3, along with T4 (which is also produced in the thyroid), regulates body's temperature, metabolism, and heart rate. Most of the T3 in body binds to protein. The T3 that does not do so is called free T3, and circulates unbound in blood. The most common kind of T3 test, known as the T3 total test, measures both kinds of T3 in blood.

Figure 7 demonstrates the membership function of feature 3 and table 3 represents the parameters of Gaussian membership function for each cluster of feature 3.

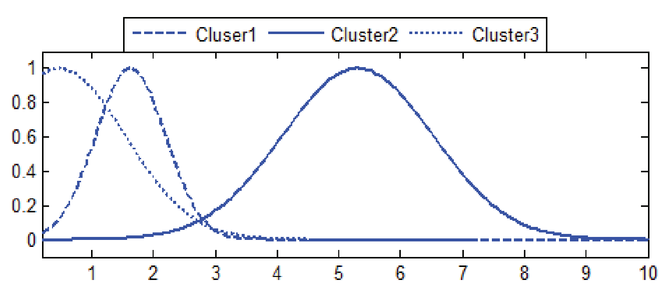

Fig. 7. Memnership functions of feature 3

TABLE III. MFS PARAMETERS OF FEATURE 3

Ftr_3	Parameters	
	Parameter 1	Parameter 2
Cluster 1	1.633	0.5669
Cluster 2	5.306	1.224
Cluster 3	0.4799	1.069

- Feature 4: Thyroid-stimulating hormone (TSH)

A TSH test measures the amount of thyroid stimulating hormone (TSH) in blood. TSH is produced by the pituitary gland. It tells the thyroid gland to make and release thyroid hormones into the blood.

Figure 8 demonstrates the membership function of feature 4 and table 4 represents the parameters of Gaussian membership function for each cluster of feature 4.

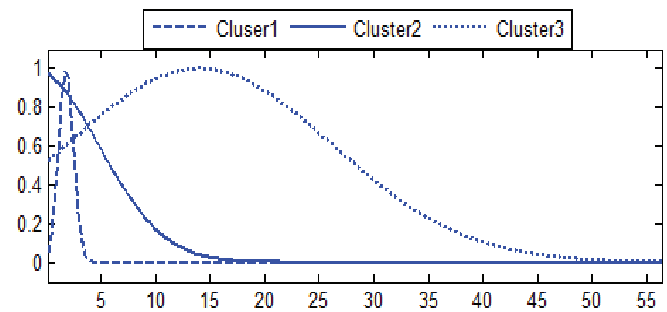

Fig. 8. Memnership functions of feature 4

TABLE IV. MFS PARAMETERS OF FEATURE 4

Ftr_4	Parameters	
	Parameter 1	Parameter 2
Cluster 1	1.778	0.6858
Cluster 2	-1.241	5.998
Cluster 3	14.01	12.24

- Rules of Proposed System

Rule base is the main part in fuzzy inference system and quality of results in a fuzzy system depends on the fuzzy rules [8]. The rule-based of the proposed system consists of three general rules. Antecedent part of all rules has three sections and consequent part of all rules has one section. The rules of the proposed system are as follows:

- Rule 1

If (in1 is in1cluster1) and (in2 is in2cluster1) and (in3 is in3cluster1) and (in4 is in4cluster1) then (out1 is out1cluster1)

- Rule 2

If (in1 is in1cluster2) and (in2 is in2cluster2) and (in3 is in3cluster2) and (in4 is in4cluster2) then (out1 is out2cluster1)

- Rule 3

If (in1 is in1cluster3) and (in2 is in2cluster3) and (in3 is in3cluster3) and (in4 is in4cluster3) then (out1 is out3cluster1)

For better view of the rule-based, Figure 9 represents the fuzzy rules of the proposed system.

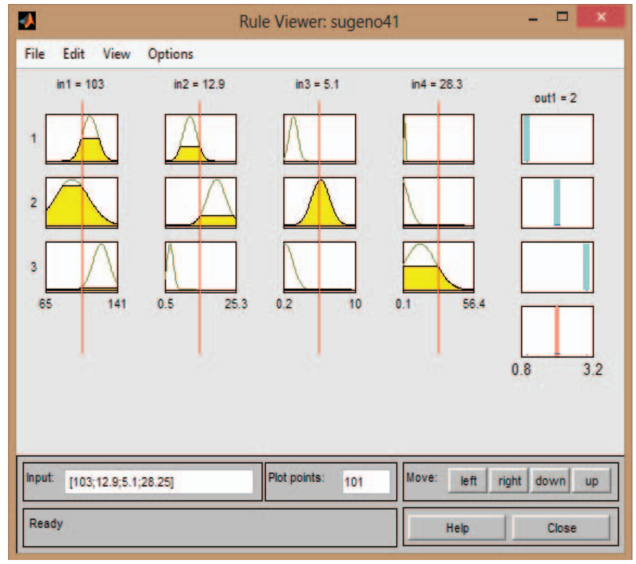

Fig. 9. Fuzzy Rule-based

- Performance Evaluation

In this study, we used classification accuracy as criteria for evaluating the performance of the proposed system. For this purpose, we divided the thyroid dataset to Training Data and Testing Data. Training Data consists of 160 sample data for Modeling and developing the system and 55 sample data as Testing Data for evaluating the proposed system. Table 5 represents each dataset with details (number of each class in Training Dataset and Testing Dataset) and Table 6 represents the test results of 55 testing data.

By using these results and confusion matrix method, the classification accuracy of the proposed system for diagnosis of thyroid disease was obtained about **89.09%.**

TABLE V. DETAILS OF DATASETS

	Thyroid dataset		**Total**
	Training Data	***Testing Data***	
Class 1 - hyper function	25	10	35
Class 2 - hypo function	20	10	30
Class 3-normal function	115	35	150
Total	160	55	215

TABLE VI. DETAILS OF TEST RESULTS

	Class 1 - hyper function	Class 2 - hypo function	Class 3-normal function
Class 1 - hyper function	35	-	-
Class 2 - hypo function	2	8	-
Class 3-normal function	4	-	6

- System Modeling And developing

User interface is one of the most important Modules of an expert system. It should be designed and developed for non-experts who have no knowledge in that specific domain. In this study, we focused on designing an appealing user interface for the patients so that it is easy to use.
Figure 10 and 11 represents the initial user interface of this proposed expert system.

V. DISCUSSION AND CONCLUSION

This paper represents a fuzzy rule-based expert system as an assistance system for diagnosing thyroid's function disease. This system uses the results of the prescribed measurement of thyroid gland as input data and by entering the input data, the output of the system is a crisp value which determine the class of thyroid function like an Endocrinologist expert.

In this study, we focused on identifying the rules and the parameters of the fuzzy system. Although the classification accuracy is a feature of a system, concentrating on rules and determination of the parameters values of the system is another important feature of a system. So, we used a new neuro-fuzzy classification based on k-means algorithm and scaled conjugate gradient (SCG) algorithm for determining the values of parameters.

For future study, we suggest using the other effective factors in diagnosis such as: sex, age, physical conditions and etc. In generating rules and determining the parameters values, type-2 fuzzy can be used for handling imprecise diagnosis.

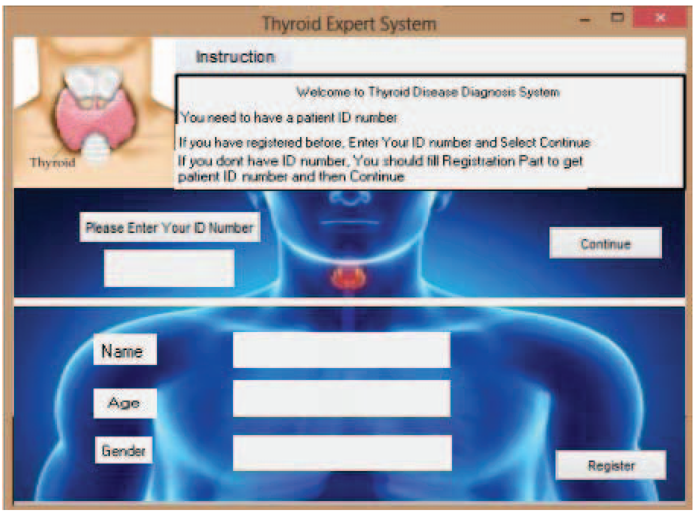

Fig. 10. User Interface (screen 1)

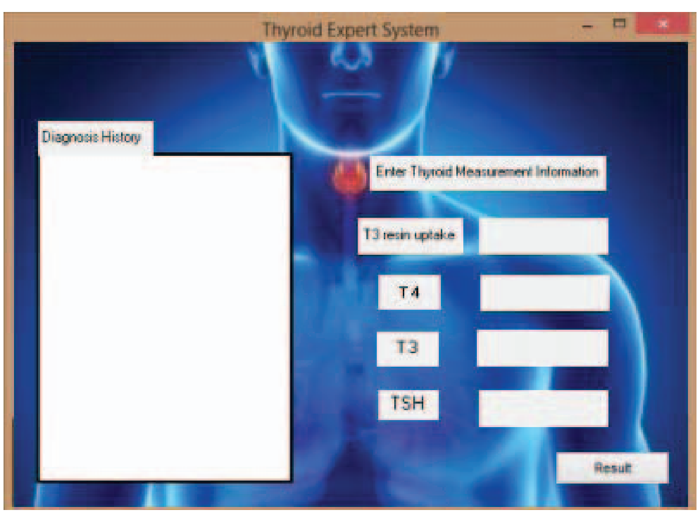

Fig. 11. User Interface (screen 2)

References

[1] Dogantekin. E , Dogantekin. A & Avci. D, "An expert system based on Generalized Discriminant Analysis and Wavelet Support Vector Machine for diagnosis of thyroid diseases", *Expert Systems with Applications*, January 2011, vol. 38, no. 1, pp. 146-150.

[2] keleş. A & keleş. A, "ESTDD:Expert system for thyroid diseases diagnosis", *Expert Systems with Applications*, January 2008, vol. 34, no. 1, pp. 242-246.

[3] Innocent. PR , John. RI & Garibaldi. JM., "Fuzzy Methods for Medical Diagnosis", *Applied Artificial Intelligence*, 2004, vol. 19, no. 1, pp. 69-98.

[4] Schatz. CV and Schneider. FK, "Intelligent and Expert Systems in Medicine–A Review".

[5] Fazel Zarandi. M. H , Harirchian, M. H , Damirchi-Darasi. S.R & Ghahazi. M. A, "Fuzzy rule based expert system for diagnosis of multiple sclerosis", in *2014 IEEE Conference on Norbert Wiener in the 21st Century (21CW)*, Boston, MA, 2014, pp. 1-5.

[6] Zadeh. L.A., "Fuzzy Logic and approximate reasoning", *Synthese*, 1975, vol. 30, no. 3-4, pp. 407-428.

[7] Fazel Zarandi. M. H , Zolnoori. M , Moin. M & Heidarnejad. H, "A fuzzy rule-based expert system for diagnosing asthma", 2010, pp. 129-142.

[8] Adeli. A & Neshat. M, "A Fuzzy Expert System for Heart Disease Diagnosis", in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Hong Kong, 2010.

[9] Liebowitz. L," *Introduction to expert system"*, 1988.

[10] Ozyilmaz. L & Yildirim. T, "Diagnosis of thyroid disease using artificial neural network methods", in *Proceedings of ICONIP'02 nineth international conference on neural information processing*, Orchid Country Club, Singapore, 2002, pp. 2033-2036.

[11] Serpen. G , Jiang. H & Allred. L., "Performance analysis of probabilistic potential function neural network classifier", in *Proceedings of artificial neural networks in engineering conference*, St.Louis, 1997, pp. 471-476.

[12] Pasi. L, "Similarity classifier applied to medical data sets", in *International conference on soft computing*, Helsinki, 2004.

[13] Polat. k , Sahan. S & Güneş. S, "A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted pre-processing for thyroid disease diagnosis", *Exper Systems with Applications*, May 2007, vol. 32, no. 4, pp. 1141-1147.

[14] Martin Fodslette Møller, "A scaled conjugate gradient algorithm for fast supervised learning", *Neural Networks*, 1993, vol. 6, no. 4, pp. 525-533.

[15] Jang. J-S & Sun. C-T., "A Neuro-Fuzzy Classifier and Its Applications", in *Second IEEE International Conference on Fuzzy Systems*, San Francisco, CA, 1993, pp. 94-98.