Journal of College of Education for pure sciences(JCEPS)
Web Site: http://eps.utq.edu.iq/ Email: com@eps.utq.edu.iq
Volume 7, Number 1, January 2017

Thyroid Disease Diagnosis using Genetic algorithm and Fuzzy Neural Network

Wijdan R .Abdulhussein

Wijdan_r_a@yahoo.com and wijdan.rashid@gmail.com

University of Thi-Qar, College of Computers and Mathematics Sciences

Abstract

The use of expert systems Techniques were still have a great importance in the diagnosis of diseases, thyroid disease diagnosis is one of the important classification problems because thyroid hormones help regulation of the body's metabolism. In this paper, hybrid intelligent model for diagnosis is proposed which is able to assist the physicians to take effective decisions. The behavior of two models has been described in classification of thyroid disease and diagnosis its state (normal, hyperthyroidism, and hypothyroidism). Firstly, training of fuzzy neural network have been used trial and error approach, then combining it with Genetic algorithm to obtain the best parameters using the proposed representation for the chromosomes and fitness function. Training the two models was on the same dataset that are available on UCI dataset.

The experimental results show high classification accuracy for both systems , but the best results have been gotten when combining with genetic algorithm obtaining smaller structure with less number of both hidden layer and neurons for each one , also much less error for training and testing . That is mean the proposed model can be used to facilitate thyroid disease diagnosis in more accuracy.

Key words: Thyroid disease, Diagnosis techniques, Fuzzy Neural network, genetic algorithm, classification.

Journal of College of Education for pure sciences(JCEPS)
Web Site: http://eps.utq.edu.iq/ Email: com@eps.utq.edu.iq
Volume 7, Number 1, January 2017

تشخيص مرض الغدة الدرقية باستخدام الخوارزمية الجينية والشبكة العصبية المضببة

وجدان رشيد عبد الحسين
Wijdan_r_a@yahoo.com and wijdan.rashid@gmail.com
جامعة ذي قار – كلية علوم الحاسوب والرياضيات

الخلاصة

استخدامتقنياتالنظمالخبيرةكانومايزاللهاالاهميةالكبريفيتشخيصمختلف الامراض , وان تشخيص مرض الغدة الدرقية واحدة من مشاكل التصنيف المهمة لان هرمونات الغدة الدرقية تساعد على تنظيم عملية التمثيل الغذائي في الجسم . تم في هذا البحث اقتراح نموذج ذكي هجينللتشخيص له القدرة على مساعدة الاطباء في إتخاذ القرارات المناسبة, و تم وصف سلوك نموذجين لتصنيف مرض الغدة الدرقية وتشخيص الحالة (طبيعي , فرط نشاط الغدة الدرقية , القصور الدرقي) , بدايةً, دُربت الشبكة العصبية المضببة باستخدام إسلوب المحاولة والخطأ, بعد ذلك تم دمج الخوارزمية الجينية للحصول على افضل براميترات للشبكة المضببة, و باستخدام التمثيل للكروموسوم ودالة الافضلية المقترحين. ان تدريب النظامين كان على نفس البيانات التي اخذت من قاعدة بيانات *جامعة كالوفورنيا - ارفن .*

اوضحت النتائج دقة عالية في التصنيف لكلا النظامين لكن النتائج الافضل كانت عند دمج الخوارزمية الجينية فتم الحصول على هيكل اصغر للشبكة باقل عدد للطبقات المخفية والخلايا العصبية المكونة لها , كذلك معدل مربع خطا اقل بكثير في مرحلتي التدريب والاختبار. ذلك يعني ان النظام المقترح من الممكن استخدامه لتسهيل تشخيص مرض الغدة الدرقية بدقة اكثر.

الكلمات المفتاحية : مرض الغدة الدرقية , تقنيات التشخيص, الشبكات العصبية المضببة,الخوارزمية الجينية , التصنيف.

1. Introduction

The thyroid gland regulates a wide range of physiological functions in the body including growth, metabolism and energy homeostasis, via the secretion of thyroid hormone (TH) [1, 2, 3].

Two active thyroid hormones have been produced by the thyroid gland, triiodothyronine T3, and levothyroxine T4, which are essential in the body temperature regulation, to product of proteins, and in overall energy production and regulation [4-6].

There are two categories of disorders of the thyroid, hypothyroidism and hyperthyroidism, according to whether serum of thyroid hormone levels (T4 and T3) are decreased or increased [7].

Many basic organs of the body have been effected by the performance of thyroid. The patient have been suffered from coma or thyroid attack which might lead to the death, if the recognize of thyroid disorder was not on time. Therefore, there is quite vital in true diagnosis of thyroid disorder (low and high efficiency of Thyroid) based on laboratory and clinical tests (disease symptom)are quite vital .The most common Thyroid Testing are [8]:

- Blood Lab testing.
- • TSH test.
- Ultrasound of thyroid.
- Thyroid Scan.

- Thyroid fine needle biopsy.

There is rapidly growing for using of intelligent hybrid systems in many areas with successful applications including process control, credit evaluation, design of engineering, cognitive simulation , and medical diagnosis .In this regard, the research on diagnosis of thyroid disease by means of smart methods has been done by researchers.

Neural Networks are widely useful for solve problems with non-linear statistical modeling and modeling of complex databases of medical information. A popular approach is training algorithm for Back propagation network, which is used to adjust weights of an ANN to minimize a cost function [9].

Both fuzzy systems and neural networks are motivated by imitating human reasoning processes. In neural nets, the relations are coded in the network and its parameters, but are not explicitly given. In contrast to other systems that are knowledge- based, no clear knowledge is needed for the applications of neural networks. In fuzzy schemes, the representation for the relationships are clearly in if–then rules [10].

In fact, the neural network fuses well with fuzzy logic, which overcome some of the limitations for both techniques, some research tries to have given birth to the "fuzzy neural networks" or "fuzzy neural systems" .There are a lot of ideas and theories for combining neural networks and fuzzy logic applied in different fields [11,12].

In the same area, Genetic algorithm (GA) is subclass of an evolutionary algorithm that utilizes biology inspired mechanism, where search space consists of elements or (chromosomes) which correspond to a particular solution, and the best solution of a problem is reached through an iterative procedure [9]. This paper addressed the diagnosis of thyroid disease using mixed systems which is result from combine genetic algorithm and fuzzy neural system.

2. Thyroid Dataset

The thyroid dataset used in this work includes 215 instances are available on UCI machine learning as a thyroid gland database and using it as datasets for training / testing the both systems [13]. Five attributes are founded for each instance (5 features) with its class as shown in Table 1, the five features are: T3, Total serum triiodothyronine, Basal thyroid stimulating hormone (TSH), Total Serum thyroxin, and Maximal absolute difference of TSH value after injection thyrotropin compared to the base value.

3. Combining Genetic Algorithm with Fuzzy Neural network

Using the dataset that has been described in the previous section, taking their attributes as input data, and the class for it as output (we give 1, 2, 3 as output according to the class) as shown in Table 1.

Table 1: Classes of thyroid state and the used output

The class	Thyroid state	Output
First Class	150 instances (Normal)	1
Second Class	35 instances(hyperthyroidism)	2
Third Class	30 instances(hypothyroidism)	3

Firstly, the inputs patterns is transmitted from a range of values to new one, therefore the rescaling will be performed using Min-Max Normalization [14].

$$y = (x_{max} - x_{min}) * \frac{x_1 - x_{min}}{x_{max} - x_{min}} + x_{min} \quad (1)$$

Where:

x_{max} = maximum value in input data.

x_{min} = minimum value in input data.

x_1= current value want to normalize.

Initially, the data was divided into (70%) samples for training and 30% samples for testing. the fuzzy neural network (FNN) with multilayer Backpropagation learning has been applied depending on trial and error approach for initialize the parameters using Gaussian membership function with 100 epochs.

This approach take a lot of time beside to some problems that appear to determine the best initialize for training the system, therefore, combining it with genetic algorithm can solve these problems, like how many number of fuzzy sets that are used for each hidden layer and the initial weights, number of hidden layers, beside the initial parameters that used for fuzzy sets or membership functions, we used the proposed encoding and initial population as shown in the next section. Algorithm (1) explains the overall operation.

Algorithm 1. Steps of an algorithm that used in this work

Step1: Generate the initial population G (1) using the proposed representation for the chromosomes that is shown in figure (1), with determine pop_ size of individuals and the maximum number of generations (Max _gen).

Step2: t=1.
Step3: While (Max _gen not reached) do {
Begin
- Inputs: - five features for all input patterns (70 % samples) .
- For all inputs, training the FNN using all individuals (chromosomes) in G (t).
- Select two chromosomes (g1 and g2) from G (t), which has the best fitness function as in Equation (2).
- Perform crossover and mutation to generate new offspring.
- Evaluate offspring.
- Replace an individual from G (t) with one of offspring only if it is better obtaining G(t+1).
- t=t+1 }
End while

Step4: The result is the best chromosome which has the best parameters for training the FNN.
Step5: Testing the system using another data (30 % samples).

3.1 Proposed Encoding and Initial Population

The following considerations for encoding and initial population used for genetic algorithm.

- The representation of chromosomes that is used as explain in Figure (1).
- The initial population of individuals are randomly generated with the following :
- - The number of hidden layers takes a random value between [1,7]
  - The number of fuzzy sets takes a random value between [1, 7]
  - The output weights also take random value between [-0.5, 0.5].
  - The population size is (10).

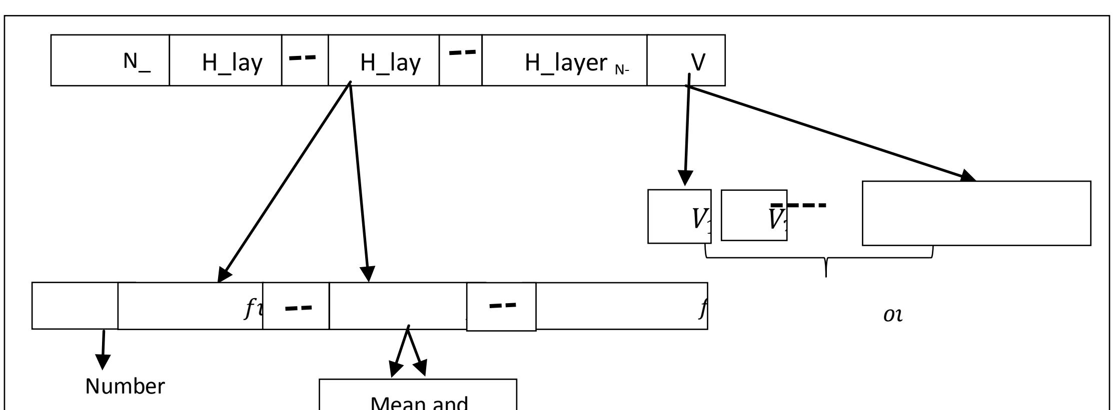

Figure 1. The used representation for the chromosomes

Where

N_ Hidden Number of hidden layers.

F_j Number of neuron (fuzzy sets) for hidden layer j.

V_{11} ----- $V_{F_{N_{hidden}}1}$ Initial weights for output layer.

Mean and variance Initial parameters for membership functions of fuzzy sets.

3.2 The Fitness Function

Selecting the individual for crossover and mutations are determined using the best fitness value. The following equation used for computing the fitness function.

$$fitness = \beta e + \alpha \left[\frac{\sum_{j=1}^{H} F\_fuzzy\ \ sets_j}{Max\_N\_fset * H} \right] \quad (2)$$

Where β , α are a constant in the range [0, 1], e is the output error, H is the number of hidden layers, $F\_fuzzy\ sets_j$ is Number of neuron for hidden layer j. and $Max_N_fset$is maximum number of fuzzy sets.

When GA has been stopped, the best chromosome which has the best fitness (a minimum number of Hidden layer and neuron at each one, lesstraining error) will be selected.

4. ExperimentalResults

The chromosome that resulted from GA will contain parameters of fuzzy neural network which are (best number for hidden layers and neurons at each one, membership functions' parameters of fuzzy sets and initial weights for output layer) ,which are used for training the fuzzy neural network without trial and error approach. Table 2 shows the best chromosome (best parameters) obtained from GA, thus Figure 2 illustrates the final structure using the best parameters.

Table (2): The best chromosome results from GA

No.Hidden layers	No. Fuzzy sets				Initial weights for output layer		
4	F_ H_layer 1	F_H_layer 2	F_H_layer 3	F_H_layer 4	V_{11}	V_{21}	V_{31}
	5	2	5	3	0.2453	-0.4728	0.3668

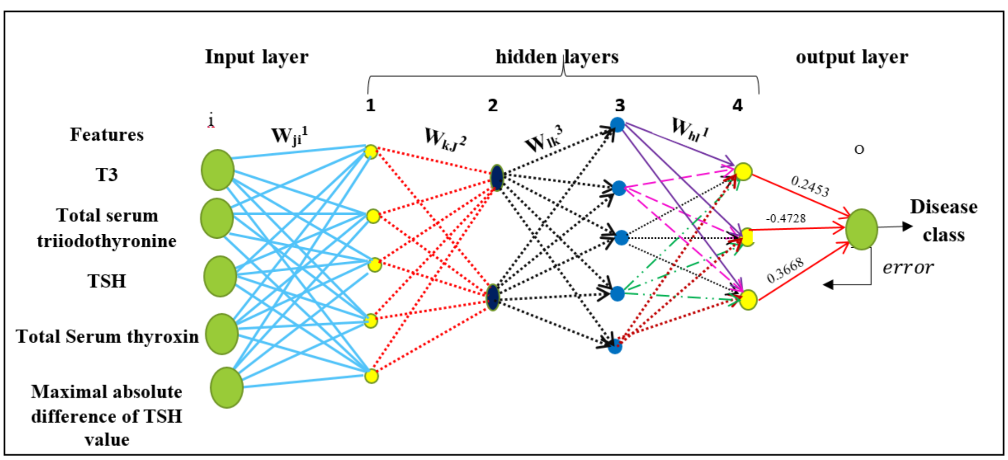

Figure 2. Structure for the network using best chromosome

Table(3): shows two methods with best parameters and its training and testing error.

the method	**Number of Hidden layers**	**Range Number of neurons**	**Training Error**	**Testing error**
FNN	**6**	**5-10**	0.0009	0.01
Genetic FNN	**4**	**2-5**	0.00005	0.003

As we see that combining Genetic with FNN reduce the number of hidden layers and neurons with minimum error. As shown in Table 2 and Table 3 ,that the proposed system with genetic has smaller structure with less number of hidden layers ,less fuzzy sets or neurons , and with more less error compared to system without using genetic algorithm .This mean reduced time for training and testing.

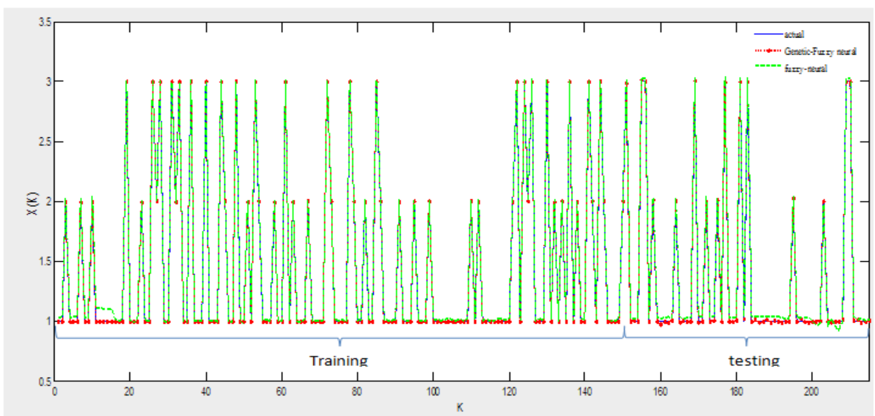

Figure (3): Desired and output during training and testing for fuzzy neural network with and without Genetic algorithm

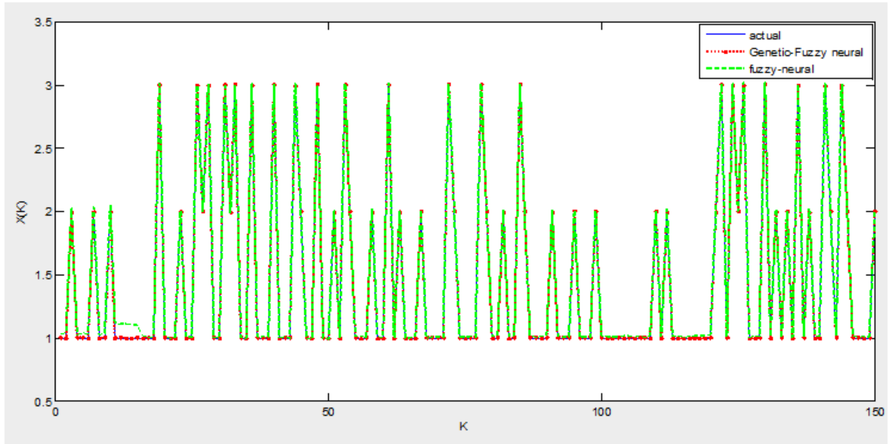

Figure (4): Desired and output during training for both systems

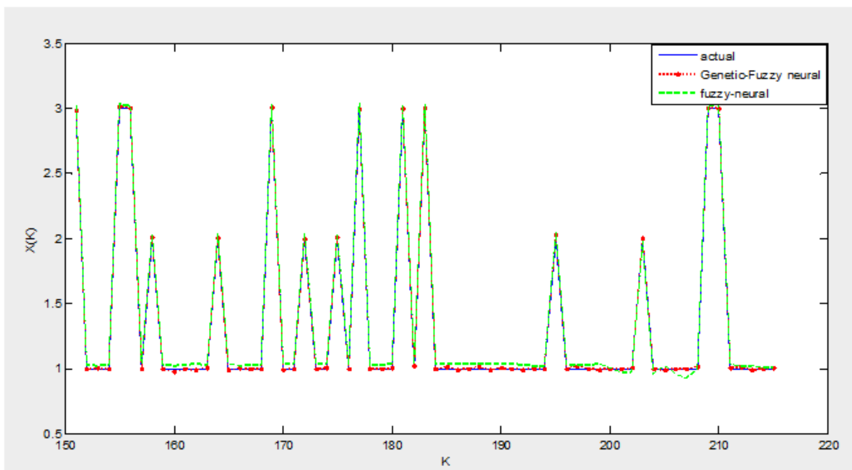

Figure (5): Desired and output during testing for both systems

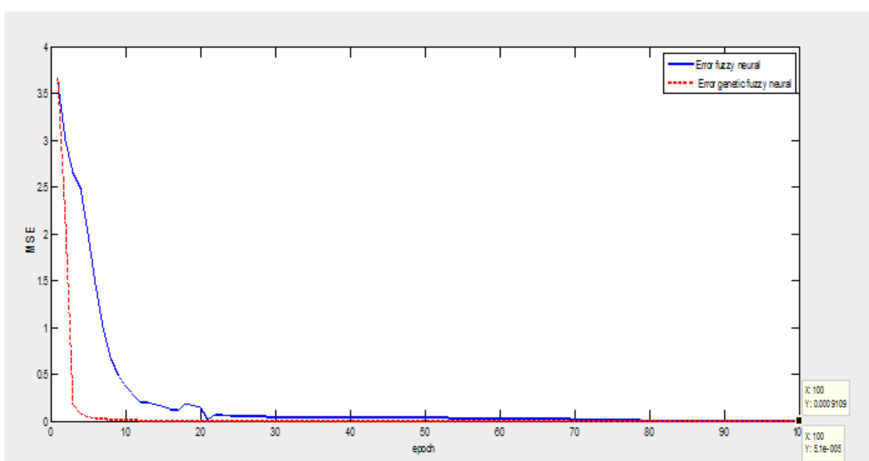

Figure (6): The error during training phase

5. Conclusions

Diagnosis of thyroid disease is one among the necessary problems .To develop a medical decision support system, in this work hybrid intelligent scheme which is result from combining three methods, fuzzy logic, neural networks and genetic algorithm has been proposed. Using new representation for the chromosomes to obtain the best parameters for the network, which is useful to make it more accurate for the classifications of the disease.

As shown in experimental results section, we conclude that Genetic Fuzzy Neural System give good result for training and testing that mean it is benefit in thyroid disease diagnosis. It was confirmed that proposed model reach classification performances near optimum, because less time in training and testing as result from smaller structure and its needed processing.

References

[1] Garg ,A. and Vanderpump, K.(2013).Subclinical thyroid disease. Br Med Bull, 107 ,101-16.

[2] Yen, P. (2005).Genomic and nongemomic actions of thyroid hormones in Werner &Ingbar's The Thyroid: a fundamental and clinical text. Ninth Edition, Eds: L Braverman and R Utiger. Pus: Lippincott Williams & Wilkins.. Chapter 8 , 135-150.

[3] James,St.(2014). Update on Thyroid Disorders, National Medicines Information Centre, Dublin 8, 20(1).

[4] Ozyılmaz, L. and Yıldırım, T.(2012).Diagnosis of thyroid disease using artificial neural network methods. In Proceedings of ICONIP'02 .9th international conference on neural information processing. Singapore: Orchid Country Club, 2033– 2036.

[5] Zhang, G.and Berardi, L. V.(1998). An investigation of neural networks in thyroid function diagnosis. Health Care Management Science. 29–37.

[6] Polat et al. (2007). A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted preprocessing for thyroid disease diagnosis. Expert Systems with Applications, 32, 1141–1147.

[7] Jack, D. (2002). Thyroid Hormone Tutorial: Thyroid Pathology, Endocrine Module (PYPP 5260), Thyroid Section.

[8] Deaf, H. (2010). Thyroid Disorders, Rochester Recreation Club for the Deaf.

[9] Navneet, W.et al.(2015). A Survey on Applications of Adaptive Neuro Fuzzy Inference System, International Journal of Hybrid Information Technology, 8(11) , 343-350 .

[10] Robert, B.(2002). Neuro-Fuzzy Methods for Modeling and Identification, Springer-Verlag, Heidelberg.161–186.

[11] Hui, Z.(2015).A New Fuzzy Neural Network Algorithm for Rural Public Service Performance Evaluation, International Journal of Smart Home, 9(10), 153-160.

[12] Yuanyuan, C.et al (2015). A New Fuzzy Neural System with Applications, ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24.

[13] UCI repository of machine learning databases.(2015). University of California atrvine, Department of Computer Science, http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease, accessed.

[14] Jayalakshmi ,T. and Santhakumaran ,A.(2011).Statistical normalization and back propagation for classification.International Journal of Computer Theory and Engineering ,IACSIT.3(1), PP. 89-93.