# An Intelligent System For Diabet Diagnosis Based on Combined Intelligent Algorithm and Risk Factors in Patients

Mohammad Fiuzy [*,a,b] , Javad Haddadnia [a, b], Nasrin Mollania[c,] Maryam Hashemian [b], Kazem Hassan pour[d]

[a.] Department of Biomedical, Faculty of Electrical and Computer, Hakim Sabzevari University, Sabzevar, Iran
[b.] Research Center for Advanced Medical Technologies, Sabzevar University of Medical Sciences, Sabzevar, Iran
[c.] Departments of Biology, Faculty of Basic Sciences, Hakim Sabzevari University, Sabzevar, Iran
[d.]Department. of Clinical Sciences, Sabzevar University of Medical Sciences, Sabzevar, Iran
*Sabzevar, KHorasan Razavi, Iran, Postal Code : 65418-13187

**Abstract:** Diabetes occurs when the body is unable to produce or respond properly to insulin which is needed to regulate glucose. Besides contributing to heart disease, diabetes also increases the risks of developing kidney disease, blindness, nerve damage, and blood vessel damage. Diabetes disease diagnosis via proper interpretation of the diabetes data is an important (classification) problem. Diabet Diagnosis is a very problematic issue in medical diagnosis. Nowadays, many relatively complex clinical trials are carried out. Early diagnoses of diabetes dramatically reduce injuries and damage caused by the infection in community. In this study, a method for proper diagnosis based on the optimal features of the Risk Factors in patients is introduced. By Using a combined artificial intelligence methods, including search algorithms (BGA[1]) to explore or search and select the best features, Data mining methods (FCM[2]) got to classify and categorize data (patient characteristics led to the diagnosis of non-patient) Neural Network (NN[3]) for modeling or detection and identification of structural parameters of the disease, diabetic patient has been detected. Then, for better Comparison and show the Performance of the Proposed System, Patients tested based on Eight Factors of World Health Organization (WHO[4]) to Diabet Diagnosis by the same Intelligent System. The proposed system by using a combination of these methods was successful to achieve ***94.031*** % precision for diabetic patient identification. Accurate detection by combination and interaction of these methods based on the optimal appearance and Risk features, introduced by the proposed algorithm that Compared with the common methods of detection and diagnosis of patients with one side and artificial methods of the authorities on the other hand, its kind and even more accurate than other methods, the result is a smart combination. It's on operation kind has better than even more intelligent system that had been introduced, given in this document.
[M. Fiuzy, J. Haddadnia, N. Mollania, M. Hashemian, K. Hassan pour. **An Intelligent System For Diabet Diagnosis Based on Combined Intelligent Algorithm and Risk Factors in Patients**. *Life Sci J* 2013;10(4s):380-386] (ISSN:1097-8135). http://www.lifesciencesite.com. 57

**Keywords**: Diabet, Risk Factor, Artificial Intelligent Process

## 1. Introduction

Many people in the world suffer from diabetic diseases. There is no certain cure for diabetes. Diabetic diseases can be delayed or controlled with proper diet. In addition, diagnosis of diabetic disease is a difficult matter for physicians. Diabetic disease occurs due to not producing or properly using insulin in the body. Insulin is a hormone. The insulin hormone converts sugar, and (other food) into energy needed for daily life. The source of diabetes might be genetic and/or environmental factors. In the USA, there are 20.8 million children and adults that suffer from diabetes. 14.6 million Of these people have been diagnosed with diabetes. The rest of these people are unaware that they have the disease [1]. According to the reports, about 220 million people around the world suffer from this disease [2, 3, and 4]. That is mainly diagnosed by blood (Plasma, Glucose, and Fasting) and saliva tests. Diabetes is divided into three types; type I, type II and type III. Type 1 Diabetes or Insulin Dependent Diabetes Mellitus (IDDM) mostly are observed in children or at earlier age. In this type of disease, body is not able to produce insulin. Only 5% of people suffer from this type. Type II diabetes or Non Insulin Dependent Diabetes Mellitus (NIDDM) is observed in 90% to 95% of diabetic patients [3]. Insulin is secreted in this type, but body is resistance to use insulin. This type of diabetes is common in adults and obese, equally in both women and men. The third type of diabetes occurs in pregnant women and it mainly becomes type II diabetes after pregnancy [5]. In recent years, so much attention has been paid to medical data mining method. So if the information is in several categories, it would be better to combine them to gain higher accuracy [6]**.** There are many factors, which make it hard to diagnose diabetes. Moreover these factors make the physician's job difficult. There are two factors for determining the decisions of the physician: (1) evaluating the current test results of a patient, (2) comparing the patient with other patients with the same condition by referring to previous

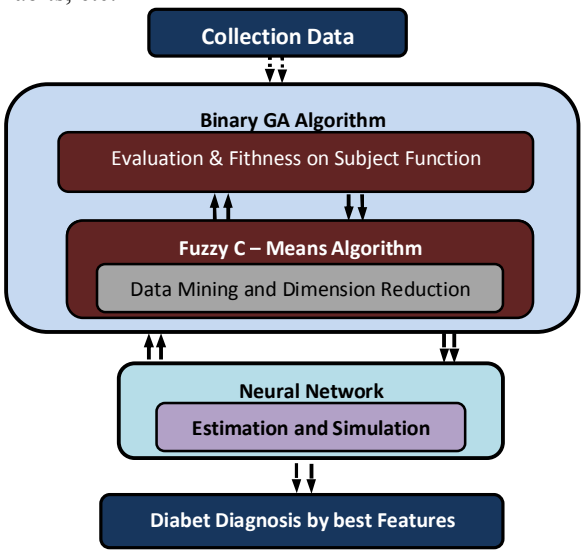
decisions [7]. For this reason, diagnosis of diabetes for a physician is very difficult [8]. So far all the researches that have been done, have used Eight World Health Organization diagnostic features that involve (**1-Number of pregnancy, 2-Plasma glucose concentration a 2 h in an oral glucose tolerance test, 3- Diastolic Blood Pressure (mm Hg,4-Triceps skin fold thickness (mm), 5- ,2-h serum insulin (lU/ml), 6- Body mass index (weight in kg/(height in m)2, 7- Diabetes pedigree function, 8- Age (years), and Class variable (0 or 1)**) has been conducted such as [10, 11, 12] World Original References and Data Bases and even in Iran such as [9] that indexed in SID[5] or major Academic References or Data bases. But the 8 features of patients would be time consuming and costly for them and physicians. However, medical studies have determined the diabetes prevalence in patients depending on several factors such as a history of other diseases (e.g. heart disease or nerve disease), family history of diabetes and many other parameters. So to achieve stable and favorable answer, has been provided composition of desire a smart way because diagnosis diabetes a big help on controlling in disease progress and faster treatment. Therefore we introduced an Automatic diagnosis System for Diabetes based on combined Intelligent Algorithm and Risk Factors by their Parameters that involves 23 Features such as: Demographic Characteristics, Health History, Physiological Characteristics and Parameters of the Patient's Family. As noted above, the major references have done several works or publication in the diabetes diagnosis field. But in this context, based on the authors' research and Indexing scientific references, only one study in [13] has been done by risk factors but it has afew features. Therefore, in order to select the best features and achieve the highest accuracy, a new and efficient algorithm was proposed in this study for feature selection, classification of samples, identification and modeling. Firstly, in this paper, the general method is described. Then the outline from Flowchart 1 is discussed. Therefore, it is hoped that the study results despite being short could be significant in improving methods, treatment and diagnosis of Diabet Diagnosis.

## 2. Importance of Pattern Recognition, Feature in Selection and Extraction

In fact, feature selection is the choice of features that have maximum power at output prediction [14]. To solve the problems which depend on optimal subset [15], feature selection algorithms are divided into two major categories. If features selection is done independent of any type of learning algorithm, it will be called Filter method, in which, the result of the selected features are determined before processing. If the assessment process is associated with a classified algorithm, the method of feature selection will be called Wrapper or Closed loop. This is a common pattern recognition system which consists of 4 sections: feature extraction and selection; designing and training classifier; and testing. In this study, feature extraction and testing by neural networks selected best features by GA algorithm, and FCM classifier was used for class and sample classification.

## 3. Proposed Method

At first, data are collected from the Sabzevar Diabet Research Center in VASEI Hospital, Sabzevar, Iran. Around 375 cases (184 Diabetic samples and 191 Healthy samples) form Patients were collected in the survey for the proposed work. The inputs designed for the system are 23 that are shown in Table 1. The inputs values have been assigned based on the comparison between men who have diabetes and non-diabetic men. Based on this comparison, a probability value has been assigned to each of these inputs and these values are fed into the system. These input values range is its Experimental numerical value and in other between **[ -5 and 5 ]** based on the condition or physical appearance of that person according to Patients and Physicians Decisions. The inputs to the system have been designed on the basis of common symptoms of diabetes mellitus and some of the factors which could lead to diabetes like high blood pressure or food habits, etc.




**Flow Chart 1 : Proposed System**

After that the data were collected from People, they were delivered to Binary GA with neural network in its cost function. Binary GA algorithm, based on optimal features selection at database, generates optimal answers (0 or 1) due to the cost

function. The network output is then delivered to the classification algorithm "FCM" for reducing dimension, estimating accuracy of Binary GA algorithm and correcting the classification of patient from impatient. The output of this algorithm returns to the Neural Network again and this process is repeated in the cost function of genetic binary algorithm until the most optimal features for increasing maximum accuracy of feature selection are achieved. Flowchart 1 represents all stages in this study.

**Table 1: All Feature for Data Base**

| | | | |
|---|---|---|---|
| **Stress level** | **History of heart disease** | **History of Smoking** | **History of diabetes** |
| **Carbohydrate Consumption** | History of neurological illness | History of lung disease | The Thirst |
| **Blood urea** | Hyperpidy disease | History of kidney disease | Physical activity |
| **Initial blood glucose** | Weight | History of liver disease | Sex |
| **Systolic blood pressure** | Length | History of gout | History of gastrointestinal disease |
| **Diastolic blood pressure** | BMI | History of endocrine disease | |

### 3.1 Genetic Algorithm

Genetic algorithm was developed for the first time by John Holland [16]. During 1960-1970 [17], the genetic algorithm was used to find approximate answers in optimization and search problems derived from natural concepts like: heredity, mutation and recombination. The major objective in this algorithm is the chance to have good samples, to continue their life in the next generation, and to improve the answers. The procedure of algorithm device is determined in pseudo – code [17]:

**Procedure**: Applied Genetic Algorithm
**Initialization**
t = 0;
Set population size or pop size, number of generation or max_gen, probability of crossover or $p_c$ and probability of mutation or $p_m$
Initialize parent population ***P (t)****:*
Evaluate *P (t) and select the best solution* $\sigma^*$ *with the* Optimum objective function among ***P (t)***
**While** (no termination criteria) **do**
**Regenerate** C (t) from P (t) by applying the crossover And mutation operations:
**Evaluate** C (t) and select the current best solution $\sigma$ With the minimum objective value among C (t)
Update the best solution $\sigma^*$ solution, I, e, if $\sigma < \sigma^*$, **then** $\sigma^* = \sigma$;
**Select** P (t+1) from P (t) and C (t);
t = t+1;
**End while;**
**End procedure**

Here, the genetic algorithm was used from n-gene chromosomes to illustrate the problem space based on 23 features, for example, as a flowing vector. Many genes are "1" (effective in generation or effective features) and many are "0" (uneffective in generation or uneffective features).

| | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |

The rotating selection (roulette wheel) is used to select chromosomes to combine and generate [18]. The method applied for mutation is in the inverse way. In this method, a sub-collection of chromosome's gene is selected and then it is reversed in replacement. To combine chromosomes with the above mentioned structure, the Partially Mapped cross over (PMX) method is used [19]. The objective function is the equation that gets merits for each generation. It seems that accuracy in identifying is more important than the small selected subsets.
Although the two series have the same accuracy, the smaller set is preferred, so we suggest the following fitness function.

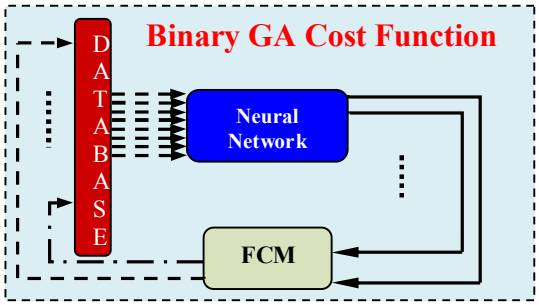
$$Fithness = \alpha.Accuracy + \beta.\frac{|n|-|s|}{|n|} \qquad \text{(Eq.1)}$$

In which |n| is the number of the total features and |s| is the number of selective features. First sentence is the carefully identified coefficient and the second sentence is the rate coefficient of reduced features. We supposed that the sum of α and β coefficient is fixed and equal to 100. Now, regarding the importance of the identified accuracy and due to the fact that fewer features are used in the diagnosis, α and β coefficients are set. Undoubtedly, detection accuracy is more important in this problem, and therefore, α (here 99) will be greater than β (here 1) [15]. Genetic algorithm parameters are set according to Table 2. Clearly, GA introduces the number of the subset which has the most and useful information. Respectively, GA searches in the 2-power 23 subsets to find the optimal solution. The

suggested algorithm located in the genetic algorithm cost function is observed in Flowchart 2.

**Table 2. GA Properties**

| ALG | **Binary GA** | POPsize | **200** |
|---|---|---|---|
| Maxiter | **250** | Varlow | **-10** |
| Var high | **+10** | Max generation | **600** |
| Recom Percent | **0.1** | Mut Percent | **0.4** |
| Cross Percent | **0.5** | Selection | **Roulette** |
| ALG | **Binary GA** | POPsize | **200** |
| Maxiter | **250** | Varlow | **-10** |
| Var high | **+10** | Max generation | **600** |
| Recom Percent | **0.1** | Mut Percent | **0.4** |
| Cross Percent | **0.5** | Selection | **Roulette** |




**Flow Chart 2: GA Cost Function**

### 3.2 Multi layer Progressive Algorithm located in Genetic Algorithm Cost Function

Progressive (feed forward) multi-layer neural networks are composed of many processing elements that are connected to each other, and are called neurons [20]. Generally the input neurons (P) are added and their weights are multiplied (w), and applied by bias (b) (width from sources), and result in activation function (f), and then the results (a) will be transferred to the next layer (Figure 1). Feed Forward Multi-Layer Neural Network is one of the most important types of static neural networks, and is used in many applications such as system identification problems, control, etc [20]. Error Back Propagation algorithm or BP model was introduced as a training algorithm for network by this model. This algorithm provides a solution for reducing gradient to minimize errors in the network. In this project, we used 3 layers of feed forward network with sigmoid activation function for input layer (i) and output layer (k) and linear function for input layer (i)( Figure 1 with adoption of [20]). First, we selected 23 neurons for input in [15]. Then, we selected the middle layer of network for the best selection of the structure. Afterwards to achieve the best result, 12 neurons were selected in the middle layer. The number of selections in the middle layer of neurons is very important. If their number is low, Network will face a shortage of learning resources for solving complex and nonlinear problems. If their number is high, the two following problems are caused: first, learning takes much more time, second, unimportant training data may be learned and the system network would be weak to solve the problems [21, 22, and 23]. In this method, the error is passed to the back. In each layer, necessary reforms are carried out on weights. In this process, the quite long error coverage reaches to minimum amount or Error Global coverage becomes equal to times of maximum repetition, which is pre-determined. Error Global should be set so that it avoids the Network from learning. In this project, 250 have been considered as the number of repetition. Error global value has also been determined to be 0.02. Properties of the suggested neural network are expressed in Table 3.

**Table 3. Neural Network (NN) Properties**

| Neurons | **23/ 12 /1** |
|---|---|
| Input / Output | **12/1** |
| Learning Algorithm | **BP** |
| MSE | **00.02** |
| Error Global | **00.02** |
| Maximum iteration | **1000** |
| 1 th Activation Function | **Sigmoid** |

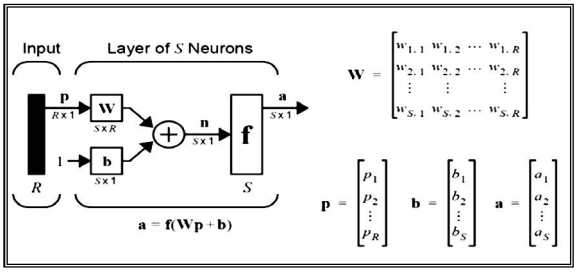



**Fig. 1) NN Structure**

### 3.3 FCM or fuzzy c-means algorithm

In 1969, Baraldi, proposed the first clustering model with fuzzy idea [24]. In this method, the amount of any membership or data belonging to each data to any cluster in matrix member is determined.

$$U = [u_{i,j}]_{c \times n} = (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n) \quad \text{(Eq.2)}$$

In which, c is number of clusters and n is the data number. This method suffers from two main limitations: the first limitation was that no cluster should be empty.

$$(\sum_{j=1}^{u} u_{ij} > 0 \quad \forall_i \in \{1, \ldots\ldots, c\}) \quad \text{(Eq.3)}$$

The second limitation, called normalization constraints, was that the total membership of all clusters must be equal to "1" in each data class.

$$(\sum_{j=1}^{u} u_{ij} = 1 \quad \forall j \in \{1, \ldots\ldots, n\}) \qquad \text{(Eq.4)}$$

FCM tries for any data set, finds the parts that minimize the following cost Equation 1 or objective function.

$$J_f(X, U_f, C) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2} \qquad \text{(Eq.5)}$$

Where $d_{ij}$ is data distance between $X_j$ and the $i^{th}$ cluster center $m \in [1, \infty)$is degree of fuzziness. If "M" goes to one, clustering will be more difficult or crisp. On the contrary, if "M" goes to infinity, clustering will be fuzzier. Supposedly as demonstrated in Figure 4 we classified 500 random data in 2 classes.

Fuzzy –C- Means or "FCM" methods consist of 4 stages:

1. If function cannot be minimized directly, repeating algorithm can be used. To solve this problem, the optimal replacement scheme was used as follows: select the proper values for m, C and small positive number for "ε". The matrix C is randomly filled (middle or center of clusters) finally, set t = 0.
2. In (t = 0) membership matrix is calculated, and in (t > 0) updated membership matrix is determined. This means that the degree of membership for fixed parameters of clusters is optimized such as Equation 5 in the flowing:

$$u_{ij}^{(t+1)} = \frac{d_{ij}^{-2/(m-1)}}{\sum_{i=1}^{c} d_{ij}^{-2/(m-1)}} = \frac{1}{\sum_{l=1}^{c} (\frac{d_{ij}}{dij})^{1/(1-m)}} \qquad \text{(Eq .6)}$$

$$for \quad i = 1, \ldots c \quad and \quad j = 1, \ldots\ldots N$$

3. The final step is updating the center of the clusters with optimized membership matrix. In addition to these parameters, how the distance is measured is of the prime importance.
4. Repeating Steps 2 and 3 until

$$\left\| C^{(t+1)} - C^{(t)} \right\| < \varepsilon \quad \text{or} \quad \left\| U^{(t+1)} - U^{(t)} \right\| < \varepsilon \qquad \text{(Eq. 7)}$$

all above steps are applied on the random test data (Figure 2) and the second cluster is created. The result can be observed in Figure 3.

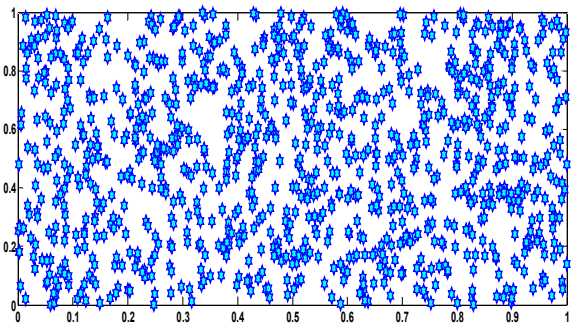


**Figure 1 : 500 rand data**

As shown in practice researches, this method does not get stuck on local peaks, and today FCM probable method is used as an inceptor in many clustering ways. More details exist on this method [24 - 28].

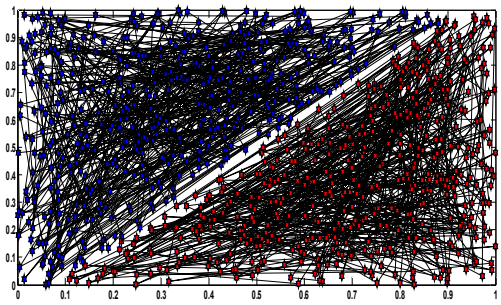


**Figure 2 : 2 Clusters Random data**

## 4. Result

MATLAB® and Origin® software is used for simulation. In this implementation, trained neural network is performed at one time. All weights will change for each input or for one row of data collection. The gained results are presented in Table 4 and figure (4). Approximately 171 from 174 patients were identified in the proposed method

**Table 4. Compare and Result**

| Patient identify | Best answer | Accuracy test | Method |
|---|---|---|---|
| ~ 172 Patient | 92.947 % | 92 % | Proposed algorithm |

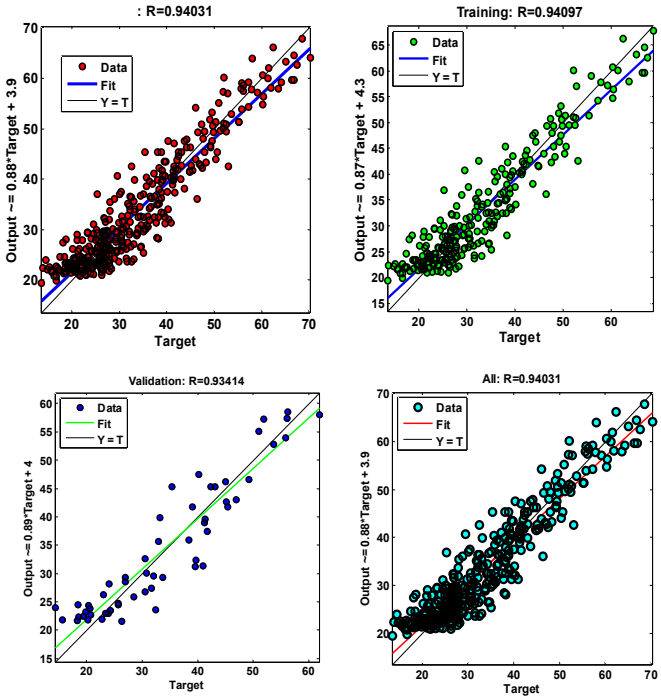



**Figure 3 :Final result Regression plot for output and target in validation and final result (train, test, valid) or all**

Therefore, the suggested method is precise and optimal. The best and most effective selected features in this suggested method are obtained in the following vector: this vector represents seven effective features (2, 3, 7, 9, 14, 15, 16, 19 and 21) from all the features involved in the diagnosis of Diabet by risk factor.

Figure 6 includes the obtained results in the final stage drawn in terms of factual result.

To review the performance of the proposed system, the below parameters are defined which have an important role in the final decision:

Tp = True positive = Cytological and suspicious positive diagnosis, which are positive in pathological tests.

Fp = False positive = Cytological and suspicious positive diagnosis, which are negative in pathological tests.

Tn = True negative = Cytological and suspicious negative diagnosis, which are negative in pathological tests.

Tp = True positive = Cytological and suspicious positive diagnosis, which are positive in pathological tests.

$$\text{Sensitivity} = \frac{T_p}{T_p + F_n} = 91.87\% \quad \text{(Eq.7)},$$

$$\text{Specialty} = \frac{T_n}{T_n + F_p} = 61\% \quad \text{(Eq.8)}$$

$$\text{PositivePredict value} = \frac{T_p}{T_p + F_p} = 73.27\% \quad \text{(Eq.9)},$$

$$\text{NegativePredictive value} = \frac{T_n}{T_n + F_n} = 92\% \quad \text{(Eq.10)}$$

$$\text{Accuracy} = \frac{T_{p+} T_n}{total} = 92\% \quad \text{(Eq.11)}$$

Number of true diagnosis For better comparison of the proposed algorithm and the medical method, their results are expressed in Table 5.

**Table 5. Table for Comparing MD (Medical result) method with proposed ALG**

| | Proposed ALG | Medical result (MD) |
|---|---|---|
| **SEN (%)** | **100** | **91.87** |
| **SPE (%)** | **100** | **61** |
| **ACC (%)** | **100** | **73.27** |
| **PPv (%)** | **100** | **92** |
| **NPv (%)** | **100** | **92** |

The value of the diagnostic methods is in their ability to diagnose the disease. Two important parameters are defined: Sensitivity or cytological sensitivity in malignancy diagnosis, and Specificity or cytological specificity in malignancy diagnosis. Given the sensitivity and specificity values presented in this review, the obtained cases are reliable. Finally, for comparing the proposed and actual values diagnosed by a physician, the Receiver Operating Characteristics (ROC) Chart may be used. In this chart, the chart with more areas indicates that the system has better performance. In each of the 2 charts of the proposed system and real results (MD system), we used 2 new parameters: True Positive Fraction: Expresses the proportion of false positives to benign or

$$TPF = \frac{F_p}{T_p + F_p}, \quad \text{(Eq.12)}$$

and False Positives Fraction: Expresses whether the ratio of false negative is a true malignancy or

$$FPF = \frac{Fn}{Tn + Fn}. \quad \text{(Eq.13)}$$

This chart is displayed in Figure 5

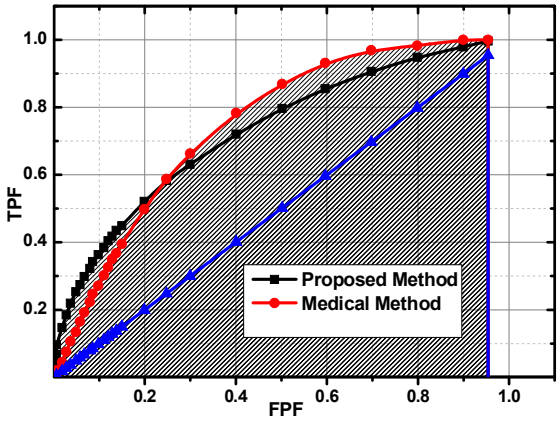



**Figure 5. Receiver Operating Characteristics (ROC) chart between proposed algorithm and medical system**

As the Figures 4 and 5 showed that the proposed system has a good performance. The results of the system are close to the real results.

## 5. Conclusion

In this opportunity, also by comparing relatively similar systems performance (not found similar system as far as authors know) determined that the proposed hybrid system can act more to detect or diagnose. Finally, according to the results in Table 4 Figure 5, With more accuracy in the diagnosis and also by less Patterns which applies binary optimization Algorithm and supports Vector Machine we seek to do more practice because by more accuracy (Due to reducing 23 into 9 features only those features Which have more impact in medical science and speed, it can be used as a tool for the diabetes diagnosis. Of course main goal of this research is to achieve 100% accuracy, to ensure that artificial intelligence systems give them practical aspects. To conclude, we propose a new method of

diagnosing diabetes by using newly designed inputs parameters. Proposed System has been used for diagnosis with a newly designed set of inputs, i.e. based on the basic symptoms which appear in diabetic patients and the physical conditions. The proposed work also reduces the cost for different medical tests and helps patients to take precautionary measures well in advance. With the successful implementation of this proposed work, the same method could also be applied in diagnosing other diseases like, coronary artery disease, hypertension etc. If you have a high knowledge, adequate control and understanding on intelligent system and data mining and disease then achieving such a goal is possible, In the not too distant future That God would want.

**Acknowledgment**

We would like to thank Sabzevar Diabet Research Center in VASEI Hospital for their help and support.

**Refrence**

1. http://www.diabetes.org/about-diabe-tes.jsp, 2007; http://www.fi.edu/biosci/healthy/diabetes.html, 2007; Polat, Gunes, & Arslan, 2008
2. Centers for Disease Control and Prevention, National Diabetes, Fact Sheet 2011 www.cdc.gov/diabetes
3. World Health Organization, Diabetes Center, Fact SheetN°312, www.who.int/mediacentre/factsheets / fs312/en
4. Na'im abdi Mohammad reza, Ahmady A.N, Tahami Ch, "Dabet Diagnosis by SVM", 14th ISCEE, SUMMER 2011
5. American Diabetes Association, Diabetes Basics www.diabetes.org/diabetes-basics
6. A. Al-Ani; M. Deriche, "A New Technique for Combining Multiple Classifiers using The Dempster-Shafer Theory of Evidence", Journal of Artificial Intelligence Research 17:333-361, 2002.
7. Duygu Calisir, Esin Dogantekin, "An automatic diabetes diagnosis system based on LOA-Wavelet Support Vector Machine Classifier ",Expert Systems with Applications 38 (2011) 8311-8315
8. Polat, K., Gunes, S., & Arslan, A. (2008). A cascade learning system for classification of diabetes disease: generalized discriminant analysis and least square support vector machine. Expert Systems with Applications, 34 , 482–487.
9. Datasets used for classification comparison of results, http://www.phys.uni.torun.pl/kmk/projects/datasets.html.
10. Tahami Ehsan, M. A. KHalil Zadeh, "Intelligent Diagnosis by MLP", 12th ICBME Iranian Conference of Bioelectric, Sahand University, Tabriz, Authumn 2005.
11. Y. Jiang and Z. Zhou, "Editing Training Data for kNN Classifiers with Neural Network Ensemble", in Proc. ISNN (1) , 2004, pp.356-361.
12. Pan Hongxia, Ma Qingfeng, Wei Xiuye, "Research on Fault Diagnosis of Gearbox Based on Particle Swarm Optimization Algorithm" ICM 2006, IEEE 3rd International Conference on Mechatronics.
13. Sumathy, Prof.Mythil, Praveen Kumar, Jishnujit T M, K Ranjith Kumar, "Diagnosis of Diabetes Mellitus based on Risk Factors", International Journal of Computer Applications (0975 – 8887), Volume 10– No.4, November 2010
14. Jensen R. Combining rough and fuzzy sets for feature selection. Ph.D Thesis, School of informatics, University Edinburgh, 2005.
15. Alipoor M, Haddadnia J. Introduction of an intelligent system for accurate diagnosis of breast cancer. IJBD. 2009; 2 (2): 33-40.
16. Joaquen Abelln, Andrés R, Masegosa. An ensemble method using credal decision trees. European Journal of Operational Research. 2010; 96 (1): 218-26.
17. Chiung Moon, Jongsoo Kim, Gyunghyun Choi, Yoonho Seo. An efficient genetic algorithm for the traveling salesman, problem with precedence constraints. European Journal of Operational Research. 2002; 140 (3): 606-617.
18. Ron Shelly. Roulette Wheel Study (1998). http://en.wikipedia.org/wiki/Roulette#History
19. Galiasso, Pablo , Wainwright, Roger L. A Hybrid Genetic Algorithm for the Point to Multipoint Routing Problem White Single Splite Paths. Proceedings of ACM/SIGAPP Symposium on Applied Computing. (SAC'01). 11-14 March 2001: 327-32.
20. Madan M, Gupta, Liang Jin, Noriyasu Homma. Static and Dynamic Neural Networks from fundamental to advanced theory. Forwarded by Lotfi A, Zadeh. IEEE Press. 2003
21. Nirooe Mehyar, Parviz Ab, Giti M. Simulating a Hybrid Model By GA and NN for Segregation patterns of benign and malignant breast cancer in mammography. Bio Phsics Iranian Journal. 2006; 13 (3): 25- 33.
22. Abdi N, Amirahmadi N, Tahami E, Rbbani. Diabet Diagnisis By SVM. 14th ISCEE Iran. summer 2011: 25-31.
23. Zhaohui L, Xiaoming W, Shengwen G, Binggang Y. Diagnosis of breast cancer tumor based on manifold learning and support vector machine. IEEE Internation Conf of Information and Automation. 2008; 703-707.
24. Baraldi A, Blonda P. A Survey of Fuzzy Clustering, Algorithms for Pattern Recognition, Part I and II. Ieee Transactions On Systems. Main, And Cybernetics Part B: Cybernetics. 1999; 29 (6): 786-801.
25. Jain AK, Murty MN, Flynn PJ. Data Clustering. A Review. ACM Computing Surveys. 1999; 31 (3): 264-323.
26. Rui Xu , Donald Wunsch. Survey of Clustering Algorithm. IEEE Transactions On Neural Networks. 2005; 30 (14): 2826-41.
27. Dring C, Lesot MJ, Kruse R. Data analysis with fuzzy clustering methods. Computational Statistics & Data Analysis. 2006; 51 (1): 192-214.
28. Asgarian E, Memarzadeh H, Saryani M, Jafar H. A new View On Clustering By GA. 14th CSICC 2009.

2/25/2013