Iranian Journal of Medical Informatics (2014) Vol 3, Issue 1

Diabetes Diagnosis via XCS Classifier System

Navid Moshtaghi, Arezoo Yazdani Seqerlou

Abstract — *This study aims to use novel concepts of artificial intelligence to design an expert clinical system. This system is able to diagnose the diabetes disease at the right time automatically. Target population in this research was existing information on the web site of California University including 768 patients. Volume of samples was taken 500 in this study. Selection of the patients was conducted randomly and the data required for designing this system were extracted accordingly. The expert system developed in this paper was a learning system as an improved version of eXtended Classifier Systems (XCS). Extended classifier systems are known as one of the most successful learning agents in this field of artificial intelligence. They are comprised of a set of simple rules with "if-then" format. Each rule predicts a particular reaction (i.e. type of disease) regarding the information received from the environment. This set of rules is "evolved" interacting with real data, while their prediction accuracy is gradually enhanced. This evolution is usually done using the patterns inspired from the nature such as genetic algorithm. In this research, the system started to learn by application of a real dataset collected. It performance was then examined on some 268 other patients, the results of which were compared with some conventional data mining methods. This comparison indicates preference of the proposed method with other techniques in terms of prediction accuracy. Installation of these systems in hospitals and application of them as a handy tool for physicians can improve decision-making process for diagnosis and provide more comfort for the patien[1]t.*

Keywords — expert clinical system, extended classifier system (XCS), diabetes disease.

I. Introduction

Diabetes mellitus, or simply diabetes, is a group of metabolic diseases in which a person has high blood sugar, either because the pancreas does not produce enough insulin, or because cells do not respond to the insulin that is produced. This high blood sugar produces the classical symptoms of polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger) [2].There are three main types of diabetes mellitus (DM). Type 1 DM results from the body's failure to produce insulin, and presently requires the person to inject insulin or wear an insulin pump. This form was previously referred to as "insulin-dependent diabetes mellitus" (IDDM) or "juvenile diabetes". Type 2 DM results from insulin resistance, a condition in which cells fail to use insulin properly, sometimes combined with an absolute insulin deficiency. This form was previously referred to as non-Insulin-dependent diabetes mellitus (NIDDM) or "adult-onset diabetes". The third main form, gestational diabetes occurs when pregnant women without a previous diagnosis of diabetes develop a high blood glucose level. It may precede development of type 2 DM 8.Other forms of diabetes mellitus include congenital diabetes, which is due to genetic defects of insulin secretion, cystic fibrosis-related diabetes, steroid diabetes induced by high doses of glucocorticoids, and several forms of monogenic diabetes. All forms of diabetes have been treatable since insulin became available in 1921, and Type 2 diabetes may be controlled with medications. Both types 1 and 2 are chronic conditions that cannot be cured. Pancreas transplants have been tried with limited success in type 1 DM; gastric bypass surgery has been successful in many with morbid obesity and type 2 DM. Gestational diabetes usually resolves after delivery. Diabetes without proper treatments can cause many complications. Acute complications include hypoglycemia, diabetic ketoacidosis, or nonketotic hyperosmolar coma. Serious long-term complications include cardiovascular disease, chronic renal failure, and diabetic retinopathy (retinal damage). Adequate treatment of diabetes is thus important, as well as blood pressure control and lifestyle factors such as smoking cessation and maintaining a healthy body weight. Globally, as of 2012, an estimated 346 million people have type 2 diabetes [3].The classic symptoms of untreated diabetes are loss of weight, polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger). Symptoms may develop rapidly (weeks or months) in type 1 diabetes, while they usually develop much more slowly and may be subtle or absent in Type 2 diabetes [7].Prolonged high blood glucose can cause glucose absorption in the lens of the eye, which leads to changes in its shape, resulting in vision changes. Blurred vision is a common complaint leading to a diabetes diagnosis; type 1 should always be suspected in cases of rapid vision change, whereas with Type 2 change is generally more gradual, but should still be suspected. A number of skin rashes that can occur in diabetes are collectively known as diabetic dermadromes.

[1] N. Moshtaghi is with Mechatronic Department, Computer Engineering Faculty, Tehran University
(Corresponding Author: navid.moshtaghi@alumni.ut.ac.ir)
A. Yazdani Seqerlou is with Computer Engineering Faculty, Tehran University

II. CLASSIFICATION

Table 1. Comparison of type 1 and 2 diabetes [14].

Feature	Type 1 diabetes	Type 2 diabetes
Onset	Sudden	Gradual
Age at onset	Mostly in children	Mostly in adults
Body habitus	Thin or normal	Often obese
Ketoacidosis	Common	Rare
Autoantibodies	Usually present	Absent
Endogenous insulin	Low or absent	Normal, decreased or increased
Concordance in identical twins	50%	90%
Prevalence	~10%	~90%

Diabetes mellitus is classified into four broad categories: type 1, type 2, gestational diabetes and "other specific types". The "other specific types" are a collection of a few dozen individual causes [2]. The term "diabetes", without qualification, usually refers to diabetes mellitus [2]. The rare disease diabetes insipidus has similar symptoms as diabetes mellitus, but without disturbances in the sugar metabolism (*insipidus* means "without taste" in Latin).The term "type 1 diabetes" has replaced several former terms, including childhood-onset diabetes, juvenile diabetes, and insulin-dependent diabetes mellitus (IDDM). Likewise, the term "Type 2 diabetes" has replaced several former terms, including adult-onset diabetes, obesity-related diabetes, and noninsulin-dependent diabetes mellitus (NIDDM). Beyond these two types, there is no agreed-upon standard nomenclature.

A. Type 1 diabetes

Type 1 diabetes mellitus is characterized by loss of the insulin-producing beta cells of the islets of Langerhans in the pancreas, leading to insulin deficiency. This type can be further classified as immune-mediated or idiopathic. The majority of type 1 diabetes is of the immune-mediated nature, in which beta cell loss is a T-cell-mediated autoimmune attack. There is no known preventive measure against type 1 diabetes, which causes approximately 10% of diabetes mellitus cases in North America and Europe [6]. Most affected people are otherwise healthy and of a healthy weight when onset occurs. Sensitivity and responsiveness to insulin are usually normal, especially in the early stages. Type 1 diabetes can affect children or adults, but was traditionally termed "juvenile diabetes" because a majority of these diabetes cases were in children. "Brittle" diabetes, also known as unstable diabetes or labile diabetes is a term that was traditionally used to describe to dramatic and recurrent swings in glucose levels, often occurring for no apparent reason in insulin-dependent diabetes. This term, however, has no biologic basis and should not be used. There are many reasons for type 1 diabetes to be accompanied by irregular and unpredictable hyperglycemias, frequently with ketosis, and sometimes serious hypoglycemias, including an impaired counterregulatory response to hypoglycemia, occult infection, gastroparesis (which leads to erratic absorption of dietary carbohydrates), and endocrinopathies (e.g., Addison's disease)[7]. These phenomena are believed to occur no more frequently than in 1% to 2% of persons with type 1 diabetes [5].

B. Type 2 diabetes

Insulin resistance, which may be combined with relatively reduced insulin secretion. The defective responsiveness of body tissues to insulin is believed to involve the insulin receptor [2]. However, the specific defects are not known. Diabetes mellitus cases due to a known defect are classified separately. Type 2 diabetes is the most common type. In the early stage of type 2, the predominant abnormality is reduced insulin sensitivity. At this stage, hyperglycemia can be reversed by a variety of measures and medications that improve insulin sensitivity or reduce glucose production by the liver.

C. Gestational diabetes

Gestational diabetes mellitus (GDM) resembles type 2 diabetes in several respects, involving a combination of relatively inadequate insulin secretion and responsiveness. It occurs in about 2%–5% of all pregnancies and may improve or disappear after delivery. Gestational diabetes is fully treatable, but requires careful medical supervision throughout the pregnancy. About 20%–50% of affected women develop type 2 diabetes later in life. Though it may be transient, untreated gestational diabetes can damage the health of the fetus or mother. Risks to the baby include macrosomia (high birth weight), congenital cardiac and central nervous system anomalies, and skeletal muscle malformations. Increased fetal insulin may inhibit fetal surfactant production and cause respiratory distress syndrome. Hyperbilirubinemia may result from red blood cell destruction. In severe cases, perinatal death may occur, most commonly as a result of poor placental perfusion due to vascular impairment. Labor induction may be indicated with decreased placental function. A Caesarean section may be performed if there is marked fetal distress or an increased risk of injury associated with macrosomia, such as shoulder dystocia. A 2008 study completed in the U.S. found the number of American women entering pregnancy with pre-existing diabetes is increasing. In fact, the rate of diabetes in expectant mothers has more than doubled in the past six years. This is particularly problematic as diabetes raises the risk of complications during pregnancy, as well as increasing the potential for the children of diabetic mothers to become diabetic in the future[14].

D. *Other types*

Prediabetes indicates a condition that occurs when a person's blood glucose levels are higher than normal but not high enough for a diagnosis of type 2 DM. Many people destined to develop type 2 DM spend many years in a state of prediabetes which has been termed "America's largest healthcare epidemic"[15]. Latent autoimmune diabetes of adults (LADA) is a condition in which type 1 DM develops in adults. Adults with LADA are frequently initially misdiagnosed as having type 2 DM, based on age rather than etiology. Some cases of diabetes are caused by the body's tissue receptors not responding to insulin (even when insulin levels are normal, which is what separates it from type 2 diabetes); this form is very uncommon. Genetic mutations (autosomal or mitochondrial) can lead to defects in beta cell function. Abnormal insulin action may also have been genetically determined in some cases. Any disease that causes extensive damage to the pancreas may lead to diabetes (for example, chronic pancreatitis and cystic fibrosis). Diseases associated with excessive secretion of insulin-antagonistic hormones can cause diabetes (which is typically resolved once the hormone excess is removed). Many drugs impair insulin secretion and some toxins damage pancreatic beta cells. The ICD-10 (1992) diagnostic entity, *malnutrition-related diabetes mellitus* (MRDM or MMDM, ICD-10 code E12), was deprecated by the World Health Organization when the current taxonomy was introduced in 1999 [6].

III. Signs and Symptoms

The classic symptoms of untreated diabetes are loss of weight, polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger). Symptoms may develop rapidly (weeks or months) in type 1 diabetes, while they usually develop much more slowly and may be subtle or absent in Type 2 diabetes [7]. Prolonged high blood glucose can cause glucose absorption in the lens of the eye, which leads to changes in its shape, resulting in vision changes. Blurred vision is a common complaint leading to a diabetes diagnosis; type 1 should always be suspected in cases of rapid vision change, whereas with Type 2 change is generally more gradual, but should still be suspected. A number of skin rashes that can occur in diabetes are collectively known as diabetic dermadromes.

A. *Diabetic emergencies*

People (usually with type 1 diabetes) may also present with diabetic ketoacidosis, a state of metabolic dysregulation characterized by the smell of acetone, a rapid, deep breathing known as Kussmaul breathing, nausea, vomiting and abdominal pain, and altered states of consciousness. A rare but equally severe possibility is hyperosmolar nonketotic state, which is more common in type 2 diabetes and is mainly the result of dehydration.

IV. Complications

All forms of diabetes increase the risk of long-term complications. These typically develop after many years (10–20), but may be the first symptom in those who have otherwise not received a diagnosis before that time. The major long-term complications relate to damage to blood vessels. Diabetes doubles the risk of cardiovascular disease. The main "macrovascular" diseases (related to atherosclerosis of larger arteries) are ischemic heart disease (angina and myocardial infarction), stroke and peripheral vascular disease [8].Diabetes also damages the capillaries (causes microangiopathy). Diabetic retinopathy, which affects blood vessel formation in the retina of the eye, can lead to visual symptoms, reduced vision, and potentially blindness [9]. Diabetic nephropathy, the impact of diabetes on the kidneys, can lead to scarring changes in the kidney tissue, loss of small or progressively larger amounts of protein in the urine, and eventually chronic kidney disease requiring dialysis. Diabetic neuropathy is the impact of diabetes on the nervous system, most commonly causing numbness, tingling and pain in the feet and also increasing the risk of skin damage due to altered sensation. Together with vascular disease in the legs, neuropathy contributes to the risk of diabetes-related foot problems (such as diabetic foot ulcers) that can be difficult to treat and occasionally require amputation.

Table 2. Diabetes diagnostic criteria[10,11]

Condition	2 hour glucose mmol/l(mg/dl)	Fasting glucose mmol/l(mg/dl)	HbA_{1c} %
Normal	<7.8 (<140)	<6.1 (<110)	<6.0
Impaired fasting	<7.8 (<140)	≥ 6.1(≥110) & <7.0(<126)	6.0–6.4

glycaemia			
Impaired glucose tolerance	≥7.8 (≥140)	<7.0 (<126)	6.0–6.4
Diabetes mellitus	≥11.1 (≥200)	≥7.0 (≥126)	≥6.5

Diabetes mellitus is characterized by recurrent or persistent hyperglycemia, and is diagnosed by demonstrating any one of the following:[6]

- Fasting plasma glucose level ≥ 7.0 mmol/l (126 mg/dl)
- Plasma glucose ≥ 11.1 mmol/l (200 mg/dL) two hours after a 75 g oral glucose load as in a glucose tolerance test
- Symptoms of hyperglycemia and casual plasma glucose ≥ 11.1 mmol/l (200 mg/dl)
- Glycated hemoglobin (Hb A1C) ≥ 6.5% [20].

A positive result, in the absence of unequivocal hyperglycemia, should be confirmed by a repeat of any of the above methods on a different day. It is preferable to measure a fasting glucose level because of the ease of measurement and the considerable time commitment of formal glucose tolerance testing, which takes two hours to complete and offers no prognostic advantage over the fasting test. According to the current definition, two fasting glucose measurements above 126 mg/dl (7.0 mmol/l) is considered diagnostic for diabetes mellitus [16].People with fasting glucose levels from 110 to 125 mg/dl (6.1 to 6.9 mmol/l) are considered to have impaired fasting glucose. Patients with plasma glucose at or above 140 mg/dL (7.8 mmol/L), but not over 200 mg/dL (11.1 mmol/L), two hours after a 75 g oral glucose load are considered to have impaired glucose tolerance [12]. Of these two prediabetic states, the latter in particular is a major risk factor for progression to full-blown diabetes mellitus, as well as cardiovascular disease [14]. Glycated hemoglobin is better than fasting glucose for determining risks of cardiovascular disease and death from any cause[14].

V. XCS

Machine learning is said to a wide range of supervised and unsupervised learning algorithms which are aiming at preventing from exhaustive search of data in the field of data mining and replacing this type of time consuming search is accomplished by intelligent methods which make it possible to simply classify or model their behavior via finding existing patterns among data. In the last two decades there are so many methods presented in the field of data mining in which different types of supervised, unsupervised or reinforcement learning algorithm are used for such goals as recognition and allocation of pattern. From among the most successful of these methods we can point out the classifier systems. In the general state, the classifier systems include a set of rules with "if-then" format, each of which presents a potential solution for the goal problem. This set of law is gradually evaluated by using a reinforced learning mechanism and updated in a specified intervals aided by a genetic algorithm. In the process of this gradual evolution, the system learns environmental behavior and then in application phase, presents suitable answers to the queries propounded by the user.

The first classifier system was suggested by Halland under the title of Learning Classifier System (LCS) in 1976. In this system, the value of each law was evaluated with an index called strength. The strength of a law in proportion with the amount of correct response to learning examples was increased within the framework of reinforced learning criteria and in specified intervals, an algorithm of evolution search (generally genetics algorithm) was responsible for producing new rules and omitting inefficient rules. At the end of training phase, this set of rules had the relative ability to present acceptable solutions when encountering new queries. Yet the successful performance of LCs was subject to selection of suitable amounts for controlling parameters of the system that mainly depends on the experience of designer of this system. When LCS was generated, the other types of classifier systems were recommended among which we can mention Extended Classifier Systems: XCS. Before introducing XCS in 1995, the ability of these systems was very limited in achieving suitable answers. But from that time on these systems were gradually changed to more intelligent and accurate agents and it is now believed that XCS and its improved versions are able to solve complicated problems with no need to adjust the parameters. Having introduced the classifier system with continuous variables (XCSR), some inherent weaknesses of binary classifier systems such as inability in introducing specified intervals of variable amounts were largely resolved and nowadays, these systems are recognized as one of the most successful Learning Agents for solving data mining problems in semi-observable environments.

According to the common approach for training XCSR, only the fitness of a rule is increased that responds a positive answer to training data. It means that the chance of each rule for not being omitted and participating in process of new rules production directly depends on the way of responding training data and to determine this chance realistically requires a large number of training data. Since the number of training data is limited in real applications and increasing the number of data are simply possible, using XCSR in such applications are not usually explainable in terms of time and computational expenses.

In the continue of this study a new method is presented for improving performance and increasing convergence rank of XCSR by means of limited training data.[1]

VI. Introducing Suggested Method

In the suggested method, firstly the limited set of training data is commonly applied for amending characteristics of rules consists of prediction, prediction error and fitness. This is done by means of the following relations:

Updating prediction and prediction error (1)

If expi < 1/β then Pi =Pi + (R-Pi) / expi, εi= εi+(|R-Pi|- εi) / expi

If expi ≥ 1/β then Pi =Pi + β (R-Pi), εi= εi+ β (|R-Pi|- εi)

Updating fitness (2)

If εi < ε0 then ki=1

If εi ≥ ε0 then ki= β (εi/ε0) –γ

Fi = fi+ β [(ki/∑kj) – fi]

In these relations, β is learning rank, γ is power of law accuracy, ε is prediction error, exp is law experiment, P is law prediction, R is reward received from environment, k is law accuracy and f is fitness. i index also indicates number of law in set of rules.

In the next phase for developing variety in set of data, several couples were selected as parents from among the fields that display the part of existing data condition using the method of "Stochastic selection with remainder"², and new data condition section is created using intermediate crossover method which are applied on the fields of parents. In this method, the quantity of each of the conditional variables is obtained from the following relation:

$$a_i= \alpha(a_i^{F})+(1-\alpha)(a_i^{M}) \quad (3)$$

in which a_i is the quantity of conditional variable of i in new data, a_i^{F} is the quantity of conditional variable i in the first parent (father), a_i^{M} is the quantity of conditional variable of i in the second parent (mother) and α is the coefficient of parents partnership which are determined in adaptive form. New data section performance is also produced using a non-linear mapping of conditional variables area to area of performance which are created by using the existing data.

Diversifying the existing data continues up to where learning stop condition (for example, when percent of system correct answers to the test data reach to a pre-determined threshold) is satisfied aided by completed dataIn the next chapter, some of the common algorithms are summarily defined for supervised learning and the results obtained from suggested method are studied with the results of these methods for Diabetes Disease Diagnosis Problem.

VII. Some Common methods for supervised learning

C4.5 algorithm is one of the most popular algorithms of making decision tree [3]. This algorithm was presented in 1993 and is an extended version of ID3 algorithm. Each intermediate node in the tree indicates a test on scales of an attribute and each branch also shows one of the authorized scales of attributes. The criteria used for selecting suitable attribute for a node is Information Gain which leads to a bios in favor of attributes with several values. To eliminate this problem Gairn Ratio criterion is also applied. ID3 algorithm only supports discrete attributes, while C4.5 algorithm in addition to attributes with discrete values also manages continuous attributes. Moreover, managing attributes with unspecified values is also one of the other advantages of C4.5 over ID3. To avoid the over fitness phenomenon in the generated classification model, pruning techniques of tree are used. The over-fitness phenomenon is happened when the generated classification accuracy of model is very high on the educational data, but not reaches a very high accuracy on the set of test data. In other words, over classification model is generated in proportion with educational data and this high proportion does not necessarily lead to higher classification accuracy on the set of test data. There are two main techniques for pruning a tree. In the first technique which is called Pre-Pruning, growth of tree in some paths is stopped before completion of tree, while in other technique under title of Post-Pruning, first the tree is grown completely and then, some of the sub-trees are replaced with a leaf node. The generated tree can be changed to a set of equal classification rule and then accomplishes pruning of rules by omitting some of its preconditions [1]. K-NN algorithm or the nearest neighbor is one of the learning algorithms based on samples. This algorithm in learning phase only restores the educational samples. For specifying class of one sample of data, the said algorithm calculates the space of this sample with other educational samples. The most common criterion for calculating such space is Euclidean norm. Although such criteria like Manhattan Minofski are also used for this purpose. After calculating the space, a majority voting is held between the nearest educational sample (k) to the current test sample and the majority label of this sample is allocated to the test sample. K is a parameter which is determined by the user. Such algorithms are called Lazy Algorithms because they do no special works in learning phase and merely restore test samples [1].

VIII. Dempster-Shafer Theory

Take θ be a set of all possible outputs in a test. The θ set is called the framework of observations [17,18]. In the Dempster-Shafer, belief is the amount used to express certainty of a condition and is usually calculated based on a function called Base Probability Allocation (BPA). BPA contains two conditions:

$$m(\phi)=0 \quad (4)$$

$$\sum_{A\subseteq\Theta} m(A)=1 \quad (4)$$

M(a) shows the amount of partial belief [20] which is exactly assigned to A, rather than total belief to A. All of the total belief allocated to A must be summed together:

$$Bel(A)=\sum_{\forall B\subseteq A} m(B) \quad (6)$$

Since the belief function cannot be indicative of contradict A (A'), Shafer has defined the uncertainty amount of A as A'.

The probability function calculates the complete amount of belief which can be assigned to A:

$$Pl(A) = 1 - Bel(A') = \sum_{B \subseteq \Theta} m(B) - \sum_{B \subset A'} m(B) = \sum_{B \cap A \neq \phi} m(B) \quad (7)$$

In comparing Bel(A) which gives a brief about all reasons for believing in A, PI(A) determines if anything that we are unaware about it supports A, how much belief must be made on A. Therefore, the amount of correct belief to A is in the range of [Bel(A),PI(A)].

Mixing rule of Dempster's evidences merges two independent evidences in one framework of observations and changes them to a body of evidences. This rule is in fact a method for mixing evidences from various sources. Thereby, the evidences are stacked together and yield a rule function which represents the correspondence of these rules. Assume that there are two bodies of observations as shown in Figure 3.

The value of BPA for the new body of evidences obtained from these two bodies of evidences would be equal to:

$$m(C) = (m_1 \oplus m_2)(C) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1 - N} \quad (8)$$

Thus, the rules of mixing Dempster's evidences calculate the amount of consistency between the two defined bodies of evidences in one framework of observations. In equation (8) which is called dempster's mixing rule, denominator is a normalizing agent which causes m(C) to be a BPA. N is the difference factor with the following value of

$$N = \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j) \quad (9)$$

Fig.1. Two bodies of evidences A and B

$$\{A_1, A_2, \ldots, A_l\}, \{m_1(A_1), m_1(A_2), \ldots, m_1(A_l)\}$$
$$\{B_1, B_2, \ldots, B_m\}, \{m_2(B_1), m_2(B_2), \ldots, m_2(B_m)\}$$

IX. Bayesian Method

In Bayesian method, confusion matrix is used to address the error due to incorrect classification by each classifier. The confusion matrix is depicted in Figure 1 for classifier et. In this matrix, the number of classes is equal to M. nt(ij) shows the number of samples with the real class of Ci. However, the classifiers define no class for them. BY using the confusion matrix, the values of correct assignment belief could be extracted as below:

$$Bel(X \in C_i \mid e_t(X) = j_t) = P(X \in C_i \mid e_t(X) = j_t) = \frac{n_{ij}^{(t)}}{\sum_{i=1}^{M} n_{ij}^{(t)}} \quad (10)$$

In order to achieve total belief values for each class in a mixed classification system, the belief values associated with that class are mixed with each other in the classifiers:

$$Bel(i) = (X \in C_i) = \frac{\prod_{t=1}^{T} P(X \in C_i \mid e_t(X) = j_t)}{\prod_{t=1}^{T} P(X \in C_i)} \quad (11)$$

Class Ci with its maximum value, is selected as the final classification decision. If the chosen class value is smaller than the threshold value, the mixed classification system will specify no class for the sample x.

$$PT_t = \begin{bmatrix} n_{11}^{(t)} & \cdots & n_{1j}^{(t)} & \cdots & n_{1(M+1)}^{(t)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ n_{i1}^{(t)} & \cdots & n_{ij}^{(t)} & \cdots & n_{i(M+1)}^{(t)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ n_{M1}^{(t)} & \cdots & n_{Mj}^{(t)} & \cdots & n_{M(M+1)}^{(t)} \end{bmatrix}$$

Fig.2. Confusion matrix

X. BORDA COUNT METHODE

Borda count method is in fact an extended version of the majority rule. In this method, each classifier generates a sorted list of classes for the input sample X. The first class and the last class of this list are introduced as the best and the worst selection by that classifier. Assume that there are T classifiers and M classes. The classifier t assigns rtk rank (score) to class Ck with k=1, M and t=1,…, T. These scores are given by the classifier such that the first element of this list is scored as M-1, the second element of this list is scored as M-i and the last element is scored zero. Afterwards, the total score for each class Ck is obtained from the following equation:

$$r_k = \sum_{t=1}^{T} r_k^t \quad (12)$$

With respect to the total score obtained for each class, the sample X belongs to a class which has the maximum total score.

Implementation of Borda Count is very simple and has easy calculations so it does not need training. However, this method considers the same efficiency for all classifiers and ignores capabilities and accuracy of classifications for all of them. This method can be modified to solve this deficiency and improve its efficiency such that it would become trainable and each classifier will be assigned a weight. In this case, total score of each class Ck is calculated as follows:

$$r_k = \sum_{t=1}^{T} w_t r_k^t \qquad (13)$$

Where, wt is the weight allocated to the classifier t.

XI. Comparing Results of Suggested Algorithm with other Supervised Learning Methods

Diabetes Disease is of such complications the diagnosis of which is hardly possible. So, some systems are needed to help the correct diagnosis of diseases in this regard. Several systems were recommended in this regard the basis of most of which is to use techniques that can discover and generalize a relation between data,correctly.
To apply the improved XCS classifier method data is needed firstly. Hospital systems are of those systems that deal with a large amount of data and can provide us information. Machine learning storage UCI provide such facilities that data base in each research field are accessible and applicable for intelligent systems. The data base of diabetes disease diagnosis is in Diabetes & Kidney and Digestive Problems Institute. We will explain information related to 768 patients each of them has 9 characteristics.

Table 3. Diabetes Data Set

Number of pregnancies
Plasma glucose concentration
Diastolic blood pressure (MMHG)
Triceps skin fold thickness
2 hours serum insulin
Body mass index
Age
Class variable
Hemoglobin A1c

After applying suggested method, the results obtained from performing improved XCSR algorithm was compared with four other algorithms in the following table in test phase.

Table 4. result

Suggested method	91.32
XCS	87.19
AD TREE	73.18
SVM	77.84
C4.5	71.32
K STAR	70.24
DEMPSTER-SAFER	81.69
THEORY	
BAYESIAN METHOD	85.91
BORDA COUNT METHODE	84.77

XII. Conclusion

In this paper, extended Classifier Systems (XCS) have been proposed for initial diagnosis of diseases based on clinical symptoms and patient's history. A new edition of these systems in which their performance has been improved trough binary use of the finite data was introduced. Meanwhile, disease diagnosis problem for diabetes was investigated with the aid of existing real data in databases of California University, Irvine to show efficiency of the proposed method. The results of this study demonstrated that the suggested method has greater accuracy and convergence rate in comparison with other conventional data mining techniques. Successful application of such methods makes this idea promising that initial and early diagnosis of a significant number of common diseases in the society could become possible in near future at low cost and acceptable accuracy with no need to the presence of specialist physician.

References

1. M.Shariat Panahi, N.Moshtaghi Yazdani, An Improved XCSR Classifier System for Data Mining with Limited Training Samples, Global Journal of Science, Engineering and Technology, (ISSN : 2322-2441) Issue 2, 2012 , pp. 52-57.
2. http://www.npr.org/blogs/health/2012/06/21/155505445/how-to-spot-a-neglected-tropical-disease.
3. Williams textbook of endocrinology (12th ed.). Philadelphia: Elsevier/Saunders. pp. 1371–1435. ISBN 978-1-4377-0324-5.
4. Dorner M, Pinget M, Brogard JM (May 1977). "Essential labile diabetes" (in German). *MMW Munch Med Wochenschr* 119 (19): 671–4. PMID 406527.
5. Lawrence JM, Contreras R, Chen W, Sacks DA (May 2008). "Trends in the prevalence of preexisting diabetes and gestational diabetes mellitus among a racially/ethnically diverse population of pregnant women, 1999–2005". *Diabetes Care* 31 (5): 899–904.
6. Cooke DW, Plotnick L (November 2008). "Type 1 diabetes mellitus in pediatrics". Pediatr Rev 29 (11): 374–84; quiz 385. doi:10.1542/pir.29-11-374. PMID 18977856.
7. Emerging Risk Factors Collaboration (2010). "Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies". The Lancet 375 (9733): 2215–22. doi:10.1016/S0140-6736(10)60484-9.
8. Boussageon R, Bejan-Angoulvant T, Saadatian-Elahi M, Lafont S, Bergeonneau C, Kassaï B, Erpeldinger S, Wright JM, Gueyffier F, Cornu C (2011). "Effect of intensive glucose lowering treatment on all cause mortality, cardiovascular death, and microvascular events in type 2 diabetes: meta-analysis of randomised controlled trials".
9. Risérus U, Willet W (January 2009). "Dietary fats and prevention of type 2 diabetes". Progress in Lipid Research 48 (1): 44–51.
10. Vijan, S (March 2010). "Type 2 diabetes". Annals of Internal Medicine 152 (5): ITC31-15. doi:10.1059/0003-4819-152-5-201003020-01003. PMID 20194231.

11. "Diabetes Care" January 2010". American Diabetes Association.http://care.diabetesjournals.org/content/33/Supplement_1/S3.full. Retrieved 2010-01-29.
12. Santaguida PL, Balion C, Hunt D, Morrison K, Gerstein H, Raina P, Booker L, Yazdi H. "Diagnosis, Prognosis, and Treatment of Impaired Glucose Tolerance and Impaired Fasting Glucose". Summary of Evidence Report/Technology Assessment, No. 128. Agency for Healthcare Research and Quality. http://www.ahrq.gov/clinic/epcsums/impglusum.htm. Retrieved 2008-07-20.
13. Selvin E, Steffes MW, Zhu H, Matsushita K, Wagenknecht L, Pankow J, Coresh J, Brancati FL (2010). "Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults". N. Engl. J. Med. 362 (9): 800–11. doi:10.1056/NEJMoa0908359.
14. Shoback, edited by David G. Gardner, Dolores (2011). Greenspan's basic & clinical endocrinology (9th ed.). New York: McGraw-Hill Medical. pp. Chapter 17. ISBN 0-07-162243-8.
15. "Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications" (PDF). World HealthOrganisation..
16. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia : report of a WHO/IDF consultation. World Health Organization. 2006. p. 21. ISBN978-92-4-159493-6.
17. Huang, Y. S., Suen, C. Y., "*A Method of Combining Multiple Experts for the Recognation of Unconstrained Handwritten Numerals*", IEEE Transaction on Pattern analysis & Machine Intelligence, vpl. 17, pt. 1, pp.90-94, 1995.
18. Al-Ani, A, Deriche, M, "*A New Technique for Combining Multiple Classifiers using The Dempster-*
19. *Shafer Theory of Evidence"*, Journal of Artificial Intelligence Research 17:333-361, 2002