International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

Diagnosis of Diabetes using Artificial Intelligence Techniques by using Bio Medical Signal Data

Saurabh Pandey
Department of Computer Science & Engg
Integral University, Lucknow
saurabhpdy16@gmail.com

Manish Madhava Tripathi
Department of Computer Science & Engg
Integral University, Lucknow
mmt@iul.ac.in

Abstract-- **Data Mining pursuits at discovering know-how out of information and supplying it in a shape that is effortlessly compressible to humans. It is a technique this is evolved to study huge quantities of data automatically gathered. Artificial intelligence approach like fuzzy, ANN and so forth are presently used for fixing a wide variety of issues in distinctive application area for choice based version designing. These systems lets in us to introduce the getting to know and variation talents subsequently such form of framework has been used in numerous unique manner of diagnosis of disease. It enables in developing computational paradigm that gives a mathematical tool for dealing with the uncertainty and the imprecision standard of human reasoning. Relational between symptoms and dangers elements for Diabetic based on the expert's medical know-how is taken and additionally associated complications or due to a few commonplace metabolic disease it can cause imaginative and prescient loss, coronary heart failure, stroke, foot ulcer, nerves. In this work overview is supplied on various strategies which are is taken into consideration in analysis where symptoms located inside the patient and relation representing the clinical information that relates the signs and symptoms in set S to the diseases in set D to diagnose the set B of the feasible sicknesses of the patients can be inferred by the compositional rule of inference. It has been found that Neural Networks are correctly used for getting to know membership capabilities, fuzzy inference rules and other context dependent patterns; fuzzification of neural networks extends their capabilities in applicability. The classifier is the heart of the automated prognosis system. The reliable classifier have to diagnose the ailment at as high accuracy as feasible even though there are many uncontrolled versions. In literature, special classifiers have been proposed for automatic analysis of PD. The NNs and adaptive neuro fuzzy classifier with linguistic hedges (ANFIS-LH) are investigated for automatic analysis of PD. The performance of probabilistic neural community (PNN) for automated diagnosis of PD is evaluated. SVM classifier is likewise investigated for the identical goal. NNs have some drawbacks consisting of need of long schooling time and uncertainties in activation function to be used in hidden layer, variety of cells in hidden layer, and the range of hidden layer. In case of SVM, type of kernel function and penalty consistent and so forth influences the category performance. If these parameters are not appropriately decided on, the class performance of SVM degrades. Similarly, the performance of ANFIS depends on type and parameters of membership characteristic and output linear parameters.**

Key words: ANN, ANFIS, ANFIS-LH, PNN

1. Introduction:

Diabetes mellitus is a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels. Several pathogenic processes are involved in the development of diabetes. These range from autoimmune destruction of the - cells of the pancreas with consequent insulin deficiency to abnormalities that result in resistance to insulin action. The basis of the abnormalities in carbohydrate, fat, and protein metabolism in diabetes is deficient action of insulin on target tissues. Deficient insulin action results from inadequate insulin secretion and/or diminished tissue responses to insulin at one or more points in the complex pathways of hormone action. Impairment of insulin secretion and defects in insulin action frequently coexist in the same patient, and it is often unclear which abnormality, if either alone, is the primary cause of the hyperglycemia. Symptoms of marked hyperglycemia include polyuria, polydipsia, weight loss, sometimes with polyphagia, and blurred vision. Impairment of growth and susceptibility to certain infections may also accompany chronic hyperglycemia. Acute, life-threatening consequences of uncontrolled diabetes are hyperglycemia with ketoacidosis or the nonketotic hyperosmolar syndrome.

Long-term complications of diabetes include retinopathy with potential loss of vision; nephropathy leading to renal failure; peripheral neuropathy with risk of foot ulcers, amputations, and Charcot joints; and autonomic neuropathy causing gastrointestinal, genitourinary, and cardiovascular symptoms and sexual dysfunction. Patients with diabetes have an increased incidence of atherosclerotic cardiovascular, peripheral arterial, and cerebrovascular disease. Hypertension and abnormalities of lipoprotein metabolism are often found in people with diabetes. The vast majority of cases of diabetes fall into two broad etiopathogenetic categories. In one category, type 1 diabetes, the cause is an absolute deficiency of insulin secretion. Individuals at increased risk of developing this type of diabetes can often be identified by serological evidence of an autoimmune pathologic process occurring in the pancreatic islets and by genetic markers. In the other, much more prevalent category, type 2 diabetes, the cause is a combination of resistance to insulin action and an inadequate compensatory insulin secretory response. In the latter category, a degree of hyperglycemia sufficient to cause pathologic and functional changes in various target tissues, but without clinical symptoms, may be present for a long period of time before diabetes is detected. During this asymptomatic period, it is possible to demonstrate an abnormality in carbohydrate metabolism by measurement of plasma glucose in the fasting state or after a challenge with an oral glucose load.

2. Related Work:

Tuba kiyan [2] et al. 2004 has discussed that statistical neural networks can be used to perform breast cancer diagnosis effectively. The scholar has compared statistical neural network with Multi Layer Perceptron on WBCD database. Radial basis function(RBF), General Regression Neural Network(GRNN),Probabilistic Neural Network(PNN) were used for classification and their overall performance were 96.18% for Radial Basis Function (RBF), 97% PNN, 98.8% for GRNN and 95.74% for MLP. Hence it is proved that these statistical neural network structures can be applied to diagnose breast cancer. **Xin Yao** [24] et al. 1999 has attempted to implement neural network for breast cancer diagnosis. Negative correlation training algorithm was used to decompose a problem involuntarily and solve them. In this article the author has discussed two approaches such as evolutionary approach and ensemble approach, in which evolutionary approach can be used to design compact neural network automatically. The ensemble approach was aimed to tackle large problems but it was in progress. **Dr.S.Santhosh baboo and S.Sasikala** [27] have done a survey on data mining techniques for gene selection classification. This article dealt with most used data mining techniques for gene selection and cancer classification, particularly they have focused on four main emerging fields. They are neural network based algorithms, machine learning algorithms, genetic algorithm and cluster based algorithms and they have specified future improvement in this field Ilias Maglogiannis [9] et al. 2009 have presented an article on An intelligent system for automated breast cancer diagnosis & prognosis using SVM based classifiers with Bayesian classifiers and ANN for prognosis & diagnosis of breast cancer disease. Wisconsin diagnostic breast cancer datasets were used to implement SVM model to provide distinction between the malignant & benign breast masses. These datasets involve measurement taken according to Fine Needle Aspirates (FNA). The article provides the implementation details along with the corresponding results for all the assessed classifiers. Several comparative studies have been carried out concerning both the prognosis and diagnosis problem demonstrating the superiority of the proposed SVM algorithm in terms of sensitivity, specificity and accuracy Clinical diagnosis of breast cancer helps in predicting the malignant cases. A lump felt during the examination roughly give clues as to the size of tumour and its texture. The various common methods used for breast cancer diagnosis are Mammography, Biopsy, Positron Emission Tomography and Magnetic Resonance Imaging. The results obtained from these methods are used to recognise the patterns which are aiming to help the doctors for classifying the malignant and benign cases. There are various data mining techniques, statistical methods and machine learning algorithms that are applied for this purpose. This section consists of the review of various technical and review articles on data mining techniques applied in breast cancer diagnosis. In [14] A. Soltani Sarvestani, A. A. Safavi, N.M. Parandeh and M.Salehi provided a comparison among the capabilities of various neural networks such as Multilayer Perceptron (MLP), Self Organizing Map(SOM), Radial Basis Function (RBF) and Probabilistic Neural Network(PNN) which are used to classify WBC and NHBCD data. The performance of these neural network structures was investigated for breast cancer diagnosis problem. RBF and PNN were proved as the best classifiers in the training set. But the PNN gave the best classification accuracy when the test set is considered. This work showed that statistical neural networks can be effectively used for breast cancer diagnosis as by applying several neural network structures a diagnostic system was constructed that performed quite well. In [15] **Orlando Anunciacao, Bruno C**. Gomes, Susana Vinga, Jorge Gaspar, Arlindo L.Oliveira and Jose Rueff explored the applicability of decision trees for detection of high-risk breast cancer groups over the dataset produced by Department of Genetics of faculty of Medical Sciences of Universidade Nova de.

3. Methodology:

The workflow of AI Technique analysis arising from the outlined clinical situations which provides a brief overview of the fundamental steps that should be followed to apply AI Technique s for the purposes of medical diagnosis with sufficient confidence.

For the reasons discussed above, the network receives patient's data to predict the diagnosis of a certain disease. After the target disease is established, the next step is to properly select the features (e.g., symptoms, laboratory, and instrumental data) that provide the information needed to discriminate the different health conditions of the patient. This can be done in various ways. Tools used in chemometrics allow the elimination of factors that provide only redundant information or those that contribute only to the noise. Therefore, careful selection of suitable features must be carried out in the first stage. In the next step, the database is built, validated and "cleaned" of outliers. After training and verification, the network can be used in practice to predict the diagnosis. Finally, the predicted diagnosis is evaluated by a clinical specialist. The major steps can be summarized as:

Features selection

Building the database

Data cleaning and preprocessing

Data homoscedasticity

Training and verification of database using AI Technique
Network type and architecture
Training algorithm
Verification
Robustness of AI Technique based approaches
Testing in medical practice

The individual steps listed above will be shortly commented and some details given.

3.1 Features selection:
Correct diagnosis of any disease is based on various, and usually incoherent, data (features): for example, clinicopathologic evaluation, laboratory and instrumental data, subjective anamnesis of the patient, and considerations of the clinician. Clinicians are trained to extract the relevant information from each type of data to identify possible diagnoses. In artificial neural network application such data are called "features". Features can be symptoms, biochemical analysis data and/or whichever other relevant information helping in diagnosis. Therefore, the experience of the professional is closely related to the final diagnosis. The ability of AI Technique's to learn from examples makes them very flexible and powerful tools to accelerate medical diagnosis. Some types of neural networks are suitable for solving perceptual problems while others are more adapted for data modeling and functional approximation [34]. Regardless of the features selected, those chosen for training the neural network should be "robust" indicators for a given clinical situation or pathology. In general, feature selection relies upon previous clinical experience. Features that bring insufficient, redundant, non-specific, or noisy information about the investigated problem should be avoided. The selection/extraction of suitable features among all available ones is usually carried out using various approaches. The most important and best-known tools for variable selection are powerful mathematical means of data mining such as principal components analysis, genetic algorithm, or AI Technique's.

3.2 Building the database:
The neural network is trained using a suitable database of " example " cases. An " example " is provided by one patient whose values for the selected features have been collected and evaluated. The quality of training and the resultant generalization, and therefore the prediction ability of the network, strongly depend on the database used for the training. The database should contain a sufficient number of reliable " examples " (for which the diagnosis is known) to allow the network to learn by extracting the structure hidden in the dataset and then use this " knowledge " to " generalize " the rule to new cases. In addition, clinical laboratory data should be in a form that is readily transferable to programs for computer-aided diagnosis.

Data cleaning and preprocessing
Data in the training database must be preprocessed before evaluation by the neural network. Several approaches are available for this purpose. Data are normally scaled to lie within the interval [0, 1] because the most commonly used transference function is the socalled logistic one. In addition, it has been demonstrated that cases for which some data are missing should be removed from the database to improve the classification performance of the network. A decrease in the classification performance of the network is observed for imbalanced databases (those with a different number of cases for each class) [35].

Data homoscedasticity
Once the suitable features, database, data preprocessing method, training algorithm, and network architecture have been identified, data concerning " new " patients who are not included in the training database can be evaluated by the trained network. The question asked is whether the new data belong to the same population as those in the database (homoscedasticity). Failure at this step might lead the network to misclassify the new data. This problem can be solved by the use of an additional parameter that indicates the population to which a certain sample belongs.

3.3 Training and verification of database using AI Technique
Network type and architecture
Although multilayer feed-forward neural networks are most often used, there are a large variety of other networks including bayesian, stochastic, recurrent, or fuzzy. The optimal neural network architecture must be selected in the first stage. This is usually done testing networks with different number of hidden layers and nodes therein. The optimal architecture is that for which the minimum value of E for both training and verification is obtained.

Training algorithm
Various training algorithms are available. However, the most commonly used is back propagation [33]. As discussed in *"Network learning"* section, backpropagation algorithm requires the use of two training parameters: (*i*) learning rate and (*ii*) momentum. Usually, high values of such parameters lead to unstable learning, and therefore poor generalization ability of the network. The optimal values of the training parameters depend upon the complexity of the studied system. In general, the value of momentum is lower than that of learning rate. In addition, the sum of their values should be approximately equal to one.

Verification
AI Technique's based medical diagnosis should be verified by means of a dataset different from that one used for training.

Robustness of AI Technique based approaches
It is well known that AI Technique's are able to tolerate a certain level of noise in the data and consequently they

typically provide sufficient prediction accuracy. However, this noise might sometimes cause misleading results, especially when modeling very complex systems such as the health condition of a human body. Such noise would not only impact the normal uncertainty of the measured data but might also impact secondary factors, for example the coexistence of more than one disease. Crossed effects cannot be predicted unless they have been considered during building of the training database. Any factor that influences the symptoms of the disease under study must be taken into account by including such cases in the database. Only in this way can the network correctly classify the patient. Of course, one way to avoid this is to combine the experience of the clinical specialist with the discriminative power of AI Technique based approaches.

3.4 Testing in medical practice:

As the final step in AI Technique aided diagnosis should be testing in medical practice. For each new patient the network ’s outcome is to be carefully examined by a clinician. Medical data of patients for which the predicted diagnosis is correct can be eventually included in the training database.

4. Result and Discussion:

This chapter covers the simulation results of the algorithm developed using MATLAB for detection of diabetes using ANFIS approach from minimum number of data inputs selection using regression analysis. The database that is used is shown below.

dataname =
nop : no. of time pregnant
gconc : glucose concentrations
bp : blood pressure
skinthick : skin thickness
insulin :insulin level
bmi :bmi
pedigree :pedigree value
age
output :diabetes present or not

Thus the data consist of 8 types of input values and the one output. There are 768 input samples for all 8 input some of the data and respective output data are shown in table 1.

Table 1: Data record for diabetes symptoms and respective input.

S.No.	I/P 1	I/P 2	I/P 3	I/P 4	I/P 5	I/P 6	I/P 7	I/P 8	O/P
	nop	gconc	bp	skinthick	insulin	bmi	pedigree	age	output
1	6	148	72	35	0	33.6	0.627	50	1
2	1	85	66	29	0	26.6	0.351	31	0
3	8	183	64	0	0	23.3	0.672	32	1
4	1	89	66	23	94	28.1	0.167	21	0
5	0	137	40	35	168	43.1	2.288	33	1
6	5	116	74	0	0	25.6	0.201	30	0
7	3	78	50	32	88	31	0.248	26	1
8	10	115	0	0	0	35.3	0.134	29	0
9	2	197	70	45	543	30.5	0.158	53	1
10	8	125	96	0	0	0	0.232	54	1
11	4	110	92	0	0	37.6	0.191	30	0
12	10	168	74	0	0	38	0.537	34	1
13	10	139	80	0	0	27.1	1.441	57	0
14	1	189	60	23	846	30.1	0.398	59	1
15	5	166	72	19	175	25.8	0.587	51	1
16	7	100	0	0	0	30	0.484	32	1
17	0	118	84	47	230	45.8	0.551	31	1
18	7	107	74	0	0	29.6	0.254	31	1
19	1	103	30	38	83	43.3	0.183	33	0
20	1	115	70	30	96	34.6	0.529	32	1
21	3	126	88	41	235	39.3	0.704	27	0
22	8	99	84	0	0	35.4	0.388	50	0
23	7	196	90	0	0	39.8	0.451	41	1
24	9	119	80	35	0	29	0.263	29	1
25	11	143	94	33	146	36.6	0.254	51	1
26	10	125	70	26	115	31.1	0.205	41	1
27	7	147	76	0	0	39.4	0.257	43	1
28	1	97	66	15	140	23.2	0.487	22	0
29	13	145	82	19	110	22.2	0.245	57	0

30	5	117	92	0	0	34.1	0.337	38	0
31	5	109	75	26	0	36	0.546	60	0
32	3	158	76	36	245	31.6	0.851	28	1
33	3	88	58	11	54	24.8	0.267	22	0
34	6	92	92	0	0	19.9	0.188	28	0
35	10	122	78	31	0	27.6	0.512	45	0

All the input values are normalized by diving it with respective maximum value. By normalization all the input shifts to a similar range of 0 to 1 as shown in table 2.

Table 2: Data record for diabetes symptoms and respective input after normalization.

S.No.	I/P 1 nop	I/P 2 gconc	I/P 3 bp	I/P 4 skinthick	I/P 5 insulin	I/P 6 bmi	I/P 7 pedigree	I/P 8 age	O/P output
1	0.353	0.744	0.59	0.354	0	0.5	0.259	0.62	1
2	0.059	0.427	0.541	0.293	0	0.4	0.145	0.38	0
3	0.471	0.92	0.525	0	0	0.35	0.278	0.4	1
4	0.059	0.447	0.541	0.232	0.11	0.42	0.069	0.26	0
5	0	0.688	0.328	0.354	0.2	0.64	0.945	0.41	1
6	0.294	0.583	0.607	0	0	0.38	0.083	0.37	0
7	0.176	0.392	0.41	0.323	0.1	0.46	0.102	0.32	1
8	0.588	0.578	0	0	0	0.53	0.055	0.36	0
9	0.118	0.99	0.574	0.455	0.64	0.45	0.065	0.65	1
10	0.471	0.628	0.787	0	0	0	0.096	0.67	1
11	0.235	0.553	0.754	0	0	0.56	0.079	0.37	0
12	0.588	0.844	0.607	0	0	0.57	0.222	0.42	1
13	0.588	0.698	0.656	0	0	0.4	0.595	0.7	0
14	0.059	0.95	0.492	0.232	1	0.45	0.164	0.73	1
15	0.294	0.834	0.59	0.192	0.21	0.38	0.243	0.63	1
16	0.412	0.503	0	0	0	0.45	0.2	0.4	1
17	0	0.593	0.689	0.475	0.27	0.68	0.228	0.38	1
18	0.412	0.538	0.607	0	0	0.44	0.105	0.38	1
19	0.059	0.518	0.246	0.384	0.1	0.65	0.076	0.41	0
20	0.059	0.578	0.574	0.303	0.11	0.52	0.219	0.4	1
21	0.176	0.633	0.721	0.414	0.28	0.59	0.291	0.33	0
22	0.471	0.497	0.689	0	0	0.53	0.16	0.62	0
23	0.412	0.985	0.738	0	0	0.59	0.186	0.51	1
24	0.529	0.598	0.656	0.354	0	0.43	0.109	0.36	1
25	0.647	0.719	0.77	0.333	0.17	0.55	0.105	0.63	1
26	0.588	0.628	0.574	0.263	0.14	0.46	0.085	0.51	1
27	0.412	0.739	0.623	0	0	0.59	0.106	0.53	1
28	0.059	0.487	0.541	0.152	0.17	0.35	0.201	0.27	0
29	0.765	0.729	0.672	0.192	0.13	0.33	0.101	0.7	0
30	0.294	0.588	0.754	0	0	0.51	0.139	0.47	0
31	0.294	0.548	0.615	0.263	0	0.54	0.226	0.74	0
32	0.176	0.794	0.623	0.364	0.29	0.47	0.352	0.35	1
33	0.176	0.442	0.475	0.111	0.06	0.37	0.11	0.27	0
34	0.353	0.462	0.754	0	0	0.3	0.078	0.35	0
35	0.588	0.613	0.639	0.313	0	0.41	0.212	0.56	0

International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

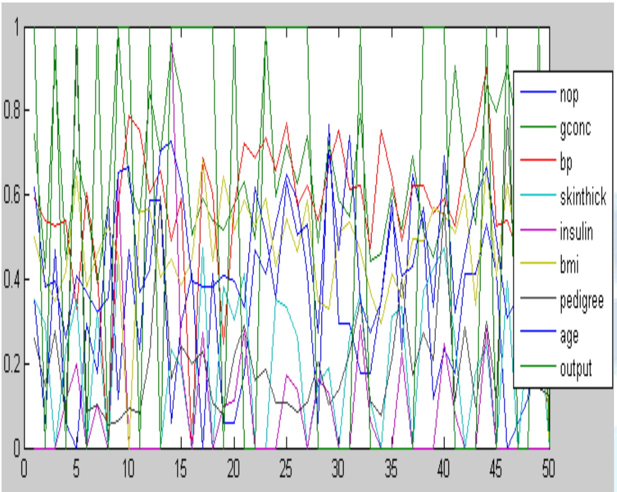

Fig 1: Plot showing 50 samples of all the normalized input data along with output.

Fig 1 shows the plot of all the 8 input signals after normalization. We can observe that data input are ranging from 0 to 1.Hence all of them are in similar range now we can check the dependency of output using the regression analysis.

The regression analysis gives the value of regression coefficients which describes the output dependency on each individual input. The regression coefficients are shown below for 8 inputs:

(1) 0.024655 ,(2) 0.0038664 ,(3) -0.0049116 ,(4) $5.7516x10^{-005}$,(5) $4.7833x10^{-005}$,(6) 0.0041863,(7) 0.098013 ,(8) -0.0010427.
The regression coeff which are high shows higher significance in relation of output. Thus to determine significant inputs we have sorted the regression coeff to point out minimum important input data. Absolute value of regression coeff after sorting in ascending order is shown in table 3.

Table 3: Significance list after regression analysis

Input id	Regression coef.
5	$4.7833x10^{-005}$
4	$5.7516x10^{-005}$
8	0.0010427
2	0.0038664
6	0.0041863
3	0.0049116
1	0.024655
7	0.098013

From the table 3 above we can observe that input 5 and 4 are non significant and if we consider only 5 input out of 8 then the significant i/p ids are [2 6 3 1 7].Thus the significant i/p are significant input are [1nop ;2 gconc ;3 bp; 6 bmi ;7 pedigree].

5. Conclusion:

To accurately prediction of the disease a good model which can represent the diabetes presence by input characteristics is required. With a good prediction model and an accurate detection technique, diagnosis can be made more efficient for dynamic use of for disease detection tools. Based on the prediction methodology, medical practitioner can envision biomedical diagnosis by engineering tools which can adapt to any future unexpected conditions automatically. A long term prediction algorithm can definitely play a very important role in planning and provisioning. In most of the previous study done for more than two decades, the most likely choice of a model is a stochastic process. The stochastic processes which tend to be used in modeling are those which has less number of parameters. And, in most of the cases the parameters fitted with the measured statistics of an actual data tend to produce models which has similar statistical properties with the actual data. This kind of model will be able to achieve a forecasting technique comparable to the actual data stream. Thus, the behavior of a real time data can be predicted using stochastic processes. Ideally, such processes should be capable of accurately representing the statistical properties of the real data which is not always possible because of several complexity issues. But one main goal is to make sure that the first and second order statistics match that of the actual prediction of presence of diseases.

References:

[1] G. E. Sakr, I. H. Elhajj and H. A. Huijer, “Support vector machines to define and detect agitation transition," IEEE Transactions On Affective Computing, vol. 1, pp. 98-108, December 2010.
[2] M. Haitham, A. Angari and A. V. Sahakian, “Automated recognition of obstructive sleep apnea syndromeusing support vector machine classiifier," IEEE Transactions On Information Technology In Biomedicine,vol. 16, pp. 463-468, May 2012.
[3] F. Azuaje, W. Dubitzky, P. Lopes, N. Black and K. Adamsom, “Predicting coronary disease risk based on shortterm rr interval measurements: A neural network approach," Artificial Intelligence in Medicine, vol.15, pp. 275-297, March 1999.
[4] E. Comak, A. Arslan and T. Ibrahim, “A decision support system based on support vector machines fordiagnosis of the heart valve diseases," Computers in biology and Medicine, vol. 37, pp. 21-27, January 2007.
[5] S. Ghumbre, C. Patil, and A. Ghatol, “Heart disease diagnosis using support vector machine,"International Conference on Computer Science and Information Technology (ICCSIT), pp. 84-88,December 2011.
[6] D. Y. Tsai and S. Watanabe, “Method for optimization of fuzzy reasoning by genetic algorithms and is application to discrimination of myocardial heart disease," IEEE Nuclear Science Symposium and Medico1 Imaging Conference, pp. 2239-2246, December 1966.
[7] E. A. M. Anbarasi and N. Iyengar, “Enhanced prediction of heart disease with feature subset selection using genetic

algorithm," International Journal of Engineering Science and Technology, vol. 2, pp. 5370-5376, November 2010.
[8] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," IEEE Intelligent Systems,pp. 44-49, March 1998.
[9] C. L. Huang and C. J.Wang, "A ga-based feature selection and parameters optimization for support vector machines," Expert Systems with applications, vol. 31, pp. 231-240, October 2006.
[10] J. Z. H. Yan and C. Xiao, "Selecting critical clinical features for heart diseases diagnosis with a realcoded genetic algorithm," Applied Soft Computing, vol. 8, pp. 1105-1111, March 2008.
[11] A. Rajkumar and G. S. Reena, "Diagnosis of heart disease using datamining algorithm," Global Journal of Computer Science and Technology, vol. 10, pp. 38-43, December 2010.
[12] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques,"International Journal of Computer Science and Network Security, pp. 343-350, January 2008.
[13] W. G. Baxt, "Application of artificial neural networks to clinical medicine," Lancet, vol.346, pp. 1135-1138, October 1995
[14] Sarvestan Soltani A. , Safavi A. A., Parandeh M N. and Salehi M., "Predicting Breast Cancer Survivability using data mining techniques," Software Technology and Engineering (ICSTE), 2nd International Conference, 2010, vol.2, pp.227-231.
[15] Anunciacao Orlando, Gomes C. Bruno, Vinga Susana, Gaspar Jorge, Oliveira L. Arlindo and Rueff Jose, "A Data Mining approach for detection of highrisk Breast Cancer groups," Advances in Soft Computing, vol. 74, pp. 43-51, 2010.
[16] Abdelaal Ahmed Mohamed Medhat and Farouq Wael Muhamed, "Using data mining for assessing diagnosis of breast cnacer," in Proc. International multiconfrence on computer science and information Technology, 2010, pp. 11-17.
[17] Chang Pin Wei and Liou Ming Der, "Comparision of three Data Mining techniques with Ginetic Algorithm in analysis of Breast Cancer data".[Online]. Available:http://www.ym.edu.tw/~dmliou/Paper/compar_threedata.pdf
[18] Gandhi Rajiv K., Karnan Marcus and Kannan S., "Classification rule construction using particle swarm optimization algorithm for breast cancer datasets," Signal Acquisition and Processing. ICSAP, International Conference, 2010, pp. 233 – 237.
[19] Padmavati J., "A Comparative study on Breast Cancer Prediction Using RBF and MLP," International Journal of Scientific & Engineering Research, vol. 2, Jan. 2011.
[20] Lee Heui Chul, Seo Hak Seon and Choi Chul Sang, "Rule discovery using hierarchial classification structure with rough sets," IFSA World Congress and 20th NAFIPS International Conference, 2001, vol.1 , pp. 447-452.
[21] Hassanien Ella Aboul and Ali H.M. Jafar, "Rough set approach for generation of classsification rules of Breast cnacer data," Journal Informatica, 2004, vol. 15, pp. 23–38.
[22] Sudhir D., Ghatol Ashok A., Pande Amol P., "Neural Network aided Breast Cancer Detection and Diagnosis",7th WSEAS International Conference on Neural Networks, 2006.
[23] Jamarani S. M. h., Behnam H. and Rezairad G. A., "Multiwavelet Based NeuralNetwork for Breast Cancer Diagnosis", GVIP 05 Conference, 2005, pp. 19-21
[24] M. S. Sapna and D. A. Tamilarasi, "Fuzzy Relational Equation in Preventing Neuropathy Diabetic," Internati- onal Journal of Recent Trends in Engineering, Vol. 2, No. 4, 2009, p. 126.
[25] L. Carnimeo and A. Giaquinto, "An Intelligent System for Improving Detection of Diabetic Symptoms in Retinal Images," IEEE International Conference on Information Technology in Biomedicine, Ioannina, 26-28 October 2006.
[26] R. Radha and S. P. Rajagopalan, "Fuzzy Logic Approach for Diagnosis of Diabetes," Information Technology Jour-nal, Vol. 6, No. 1, pp. 96-102. doi:10.3923/itj.2007.96.102
[27] P. Jeatrakul and K. W. Wong, "Comparing the Perform- ance of Different Neural Networks for Binary Classifica- tural Language Processing, Bangkok, 20-22 October 2009, pp. 111-115. doi:10.1109/SNLP.2009.5340935