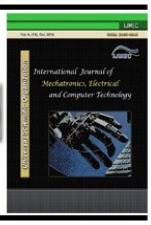
Vol. 6(21), Jul. 2016, PP. 2977-2986

Contents list available at IJMEC

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)

Journal Homepage: www.aeuso.org

PISSN: 2411-6173 - EISSN: 2305-0543



Hybrid algorithm based on K-nearest-neighbor algorithm and Adaboost with selection of feature by genetic algorithms for the diagnosis of diabetes

S. Bahramian and A. Nikravanshalmani *

[1] *Department of Computer, Karaj,Branch, Islamic Azad University, Karaj, Iran*

***Corresponding Author's E-mail:** nikravan@kiau.ac.ir*

Abstract

Diabetes is known to be one of the chronic and dangerous diseases the people in afflicted with which are increasing in number. On the other hand, early detection of this disease can be effective in treatment and control of disease progression for patients. Using the data mining techniques and machine learning, disease detection systems have been able to help patients and physicians to detect the disease in due time. In this paper, a hybrid algorithm based on the algorithm of K nearest neighbors and Adaboost has been proposed to improve the accuracy of detection of diabetes. In the first stage, the genetics algorithm is used to select the features and the dimensions of feature vector are reduced. Then, the combination of Adaboost algorithm and K nearest neighbor algorithm having been used, the samples are classified. To evaluate the algorithm, it was tested on PID dataset by 10-fold cross validation. The results obtained indicate the high accuracy and speed of the proposed method.

Keywords: *K-Nearest Neighbor algorithm, Adaboost algorithm, data mining, classification, diabetes, weak classifier composition, genetic algorithms, feature selection.*

1 Introduction

Diabetes is one of the most common and dangerous diseases and appearing when the body cannot produce the required insulin or take in the produced insulin [1]. The recent statistics of who indicate that diabetes is growing rapidly in the world, and especially in Asian countries [2]. Diabetes is a hidden disease because there may appear no sign in a lot of people in afflicted with it at the early stage and the disease the disease, at first its symptoms are not apparent until long after the disease may be detected only after a long time. At present, 400 million people in the world are inflicted with diabetes, half of whom are unaware of that diabetes. Also, diabetes causes deaths of 5 million people in the world annually, which is a person every six seconds. Most of the costs spent for diabetes are for treatment of its symptoms including cardiovascular, eye, and kidney and foot problems. While these complications can be prevented by due detection and paper treatment and late detection of diagnosis means that most people have suffered one of these complications at the disease detection time [3].Today, sensitivity of medical issues, the widespread useful information in a diagnostic area and different making by physician in there analysis have faced experts to a great challenge in the process of diagnosis detection. On the other hand, the emergences of decision - making support systems in medicine have become effective to overcome this challenge [4]. These systems are based on data mining and machine learning techniques. One of the applications in this field is detection and prognosis of different types of diseases performed through classification techniques [5]. In the classification process the classes are already predicted are the samples are assigned to one of these classes. In this process, each record of the available data includes features one of which is class feature and dividing of classes is determined through the values of these features. Classification belongs to the observation algorithms consisting two stages of learning and

Article History:
IJMEC DOI: 649123/10214
Received Date: Oct. 11, 2015
Accepted Date: Mar. 19, 2016
Available Online: Jul. 07, 2016

testing. Therefore, the data set are divided into training and testing sets. First in learning stage, the available data in training data set are lead to one of classification algorithms for model to be created. Then the model is assessed with the testing dataset [6]. With the output of these systems, the presence or absence of disease can be determined [7]. Selecting the input features properly is one of the significant issues in classification process as it directly the result obtained. Feature–selecting-based methods try to decrease the data dominions and increase the model accuracy through selecting a subset of the initial data, in that they remove those features making the classification algorithm deviated, having lower importance in classification or producing no useful information and then the classification is done with a fewer number of features learning to this increase of classification efficiency. Not removing these features provides any problems in information aspect but increases the computational load for the desired function. In addition, it causes the great unusual information to be stored with useful data. The selection of these features is of great importance. In another word, the proper significant factor in creating a model with high accuracy is the selection of features. If the total number of features is N, the number of possible subsets is 2^N, which is too high for the overage N. Complete heuristic search, Random Search and Search can be used to make a query among this numbering subsets from the features. But to find an optimal subset is found among the feature sets depends on evaluation function directly. If evaluation function assigns an unsuitable value to feature subset, this subset is not chosen as the optimal subset. There are found different evaluation functions such as Measure Distance, Information Measures, Dependence Measures, Consistency Measures and Classifier Error Rate Measures. Combination of classifiers leads to increase of classification efficiency [8, 9]. The Combination of base and weak classifiers and making a strong classifier is called the Ensemble learning [10, 11]. In a category, the combination of Classifiers is performed out in static and dynamic. Bagging and Boosting are two ways to this based on Filtering, Sub sampling and reweighting techniques. In Filtering method, it is assumed that the dataset is big and the sample chosen from it, are removed or returned to the dataset. In Subsampling, the datasets are used with fixed value and the data are resembled with a certain probability distribution. Reweighting is also used with datasets of fixed value but the data are valued with a week learning system and weighted. Bagging Method uses Sub sampling technique and the learners it produces are independent of each other. Boosting uses Reweighting technique and the learners are in series. Finally, the results of these week classification are combined. There are deferent methods to be the combination the most common of which averaging with maximum vote [12]. The Ensemble learning is better than the best basic classifier when the base classifier has an acceptable varied efficiency. This variation is created when the patterns classified improperly are different [13-16]. In this paper, using data mining and machine learning techniques, we tried to introduce an efficient and accurate algorithm to detect diabetes early to become a useful tool to help diabetes patients and medical society[17].

2 Relater work

Regarding the importance of the issue, various researches have focused on smart system of diabetes detection. These methods are based on learning and tested on PID dataset. In [18], there is a presented a system for classification of diabetic patients with the use of neural networks with complex values of (CVNN) and parametric modeling methods based on (RVNN). This system has gained an accuracy of 80.65% - 81.00%. The authors in [19], have focused on the use of fuzzy - neural classification model (Nefclass). They used from optimization algorithm of particles to correct the membership function. Their proposed system (pso-Nefclass) has the accuracy of 82.32%. In Ref [20], based on the Gini index and decision tree model, GG-FSDT model is proposed with the accuracy of 75.8%. Also in [21], with general regression neural network (GRNN), there is a proposed model to detect diabetes of 80.21% accuracy. The authors of [22], have used the algorithm of k-fuzzy for classification of diabetic patients, which has accuracy of 72.95%. The authors of [23] used the modified detection system of artificial immunity to classify the diabetes data, with the accuracy of 89.10%. In [24], genetic programming was used to detect diabetes. Of course, the algorithm they presented differs from standard genetic in manner of parents intersection and offspring production at the accuracy of 84.4%. The authors of [25], using MLCPN technique, focused on classification of PID data. The function MLCPN is created from the combination of two methods of Kohonen self-organized function (KSOM) and Grossberg at the accuracy of 97.0%.

3 Theory

3.1 Feature selection

Feature subset selection is very important in the field of data mining. Also, it is an essential pre-processing method to remove irrelevant and redundant data. It can be applied in both unsupervised and supervised learning. In supervised learning, feature selection aims to maximize classification accuracy. Features or characteristics used to describe a template explicitly define a correlation. On other hand, the increased dimensionality of data makes testing and training of general classification method difficult. So, development of a model the feature selection for classifying is useful. The relevant features are provided as input to classification algorithm.

Feature selection methods can be categorized into filter, wrapper, and embedded or hybrid. The filter approach selects features without involving any data-mining algorithm. The filter algorithms are evaluated based on four different evaluation criteria namely, distance, information, dependency and consistency. The wrapper approach selects feature subset based on the classifier and ranks feature subset using predictive accuracy or cluster goodness. It is more computationally expensive than the filter model. Entropy measure has been used as filter method[26-28].

3.2 Genetic Algorithm

Genetic algorithm is a subset of evolutionary algorithm developed from Darwin's theory of gradual evolution and fundamental ideas. Process of optimization is based on a random trend in genetic algorithm. Before the genetic algorithm can be implemented, we must first find encoding system for the intended problem. The most common way to show chromosomes in the genetic algorithms is in binary form. In this case, chromosome is a bit string, the length of which is determined by some existing parameters. In other words, each parameter is related to a bit in a string. In this algorithm, for a fixed number called population, a set of target parameters is produced randomly. The genetic algorithm applies the rule of surviving the best to get the better solutions and then it assigns the number representing the fitting of that set to the member of the population. This process is repeated for every single member. With the retrieval of genetic algorithm operators such as selection, Mutation and crossover imitated from natural genetics, better approximations can be obtained from final solution and this procedure continue to get the convergence criterion. A selection operator chooses some chromosomes among the available chromosomes in a population for reproduction. The methods of selection are selection of the elite, the roulette wheel, tournament, Boltzmann, ranking, etc. The crossover operator is a random merging which some parts of chromosomes are exchanged. This issue causes that the children are not exactly like their parents and have had a combination of characteristics of their parents. After the merging, Mutation operator is applied on chromosome. This operator chooses a gene from a chromosome randomly and changes the content of that gene. There are three criteria for algorithmtermination:1-The number of generations in algorithm,2-the population does not become better,3-classification accuracy of element with the best fitness does not exceed the threshold level.

3.3 Nearest Neighbor algorithm

K-nearest neighbor algorithm is one of the learning algorithms which is commonly used although it is very simple and gives good results in classification of linear data. This algorithm is resistant to noise and has good performance when the sample are great in number .The algorithm process is such that a sample x of test data sets is compared with all educational samples in terms of Euclidian distance. Of the educational samples calculated, k samples of less distance to sample x are chosen as neighbors of sample x and finally, sample x belongs to the class which has the greatest votes in the number of nearest neighbors. The most important in this algorithm is to choose the value of k which must be very low or very high. Therefore, k is tested in different values so that the best value is obtained. This value must be odd so that the similar cases in final voting from k closer neighbors are removed [29].

3.4 Adaboost algorithm

This is an adaptive reinforcement algorithm performing classification based on a basic algorithm and each time, the classifier is regulated in favor of misclassified samples .This algorithm uses the database repeatedly and it is not necessary to have big database [30]. In this method , the probability of selecting a sample of x_t to be located in educational data set of classifier j+1 is determined based on the error probability of classifier *cj* in that if sample x_t is not classified properly, the probability of being selected for the next classifier is increased and if the sample x_t is classified properly ,it is less likely to be selected for the next classifier. All the learners are simple and weak and must have error less than ½. Otherwise, the education is stopped since its continuation makes the learning become difficult for the next classifier. Also, the initial probability of selecting educational samples is considered to be uniform. In fact, the weight of sample shows the importance of the sample. The final hypothesis is obtained through weighted voting of T number of weak hypotheses [31]. The pseudo-code of this algorithm is shown below.

AdaBoost

Input: N labelled examples $< (x_1,v_1),...(x_n,v_n) >$, $x_i \in X$, $v_i \in \{-1,1\}$

Weak learning algorithm **KNN** lable with k=1

Integer T specifies the iteration number

Initialize: initial weights w , $w_i = 1/N$, for i=1, ... ,N.

T= 0 , $err' = 0$

Do while $t \leq T$ and $err^t < 0,5$

1. Normalize w^t, so that $\sum_{i=1}^{N} w_i^t = 1$
2. Call Knn , providing with the weight w^t , get hypothesis $h^t : X \rightarrow \{-1,1\}$.
3. Compute $err^t = \sum_{i=1}^{N} w_i^t e_i^t$, where $e_i^t = 1$, if $h^t(x_i) \neq v_i$, and 0 otherwise.
4. Set $\alpha^t = 0,5\ log\ [(1-err^t)/err^t]$.
5. Update the weights to be: $w_i^{t+1} = w_i^t exp(2\alpha^t e_i^t)$.
6. T=t+1.

Output: the hypothesis $h(x_i) = sign[\sum_{j=1}^{t} \alpha_j h_j(x_i)]$.

Figure 1: Pseudo code of Adaboost algorithm [30]

4 The proposed algorithm

In this paper, from PID dataset, using genetic algorithm based on entropy criterion, some more important features in detection of diabetes are selected and based on 10-fold cross validation model, they are divided into educational and test data. The learning is assessed with educational data and the system is assessed with test data. In creating learning model with k-nearest neighbor algorithm, the samples are classified into patient and non-patient groups and with Adaboost algorithm, the k-nearest neighbor's algorithm is reinforced as a weak classifier and the samples reclassified. The implementation was done in Matlab 2014 and classification assessment parameters including accuracy and error mean are compared.

4.1 Individual representation and population initialization and Feature selection

To select a subset of features, a string of binary chromosomes is considered at a now of the set given. Each binary figure show a feature with 0 or 1.A schematics of these chromosomes is shown in Fig 2. The initial population is also a set of chromosomes the number of which is fixed during the implementation. To evaluate the initial population, the competency function based on information measures is used. This criterion takes in to consideration the information obtained by a feature. Feature x, outperforms feature y in this method. If information obtained from feature x is greater than that of feature y, the entropy is a sample of Measures Information. Entropy is a parameter related to disorder

of a system and when disturbances increase and extraction of suitable features becomes complex maybe it is successful to only partially. In this paper, the entropy criterion was used to select the suitable features in that the entropy value is calculated from the following formula for each array of feature vector. The features with less entropy have greater effect in class separation. To quantify entropy, we used the following equation (Shannon & Weaver, 1949), where *H(A)* is the entropy of categorical variable *A*, *a* is a category of *A*, and p_a is the proportion of observed values within that category [32].

$$H(A)=-\sum_{a\in A}(p_a\, log_2(p_a)) \quad (1)$$

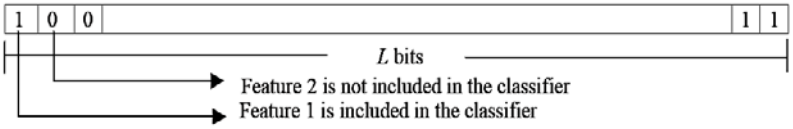


Figure 2. A feature vector with the length of L

In selecting selection operator, for parents, we used an elite-based method. The parents are chosen regarding the evaluation function elite idea adds a new feature to selection process. The best member of population survives and is present in the next population, and the member with the highest adaptation transferred to the new population automatically. The elitism applications in genetic algorithm usually improve the efficiency. As elitism causes the probability of selecting best chromosomes to increase, it can increase the genetic algorithm speed. In this paper, for crossover operator, single-point method was used. In this method, a point of chromosome is chosen randomly is based on production and test randomly the solution is not clear and we do not know which solution is optimal to define the termination condition as the solution the population. For this reason, certain number of generation, numbers of generations in the article 80 have been considered, we consider for termination condition. Genetic algorithm parameters are shown in table 1.

Tabel 1- Parameters specification in proposed algorithm

Population size	50
Number of iteration	80

4.2 Dividing database using k-folds cross validation method

In cross validation, the main data set is divided into k subsets of equal size, each time one is chosen as validity and the other k-1 subsets are used as education. This continues until each subset is educated once and k-1 times of validations are done. Finally, classification is obtained from mean of accuracy in each stage. In 10-fold cross validation, the training data are subdivided randomly into 10 disjoint equal-sized sets. The learning algorithm is applied 10 times, each time on a training set containing all but one of these subsets. The resulting classifier is tested on the one-tenth of the data that was held out. The performance of the 10 classifiers is averaged to provide an estimate of the overall performance of the learning algorithm when trained with the given features.

5 Experimental results

5.1 Diabetes disease database

In this paper, data available in PID dataset downloaded from UCI were used [33]. These data contains 8 registered features related to 768 women aged 21 at least, of whom 500 were years of age. The 8 features are: Number of times pregnant, Plasma glucose concentration, Diastolic blood pressure, Triceps

skin fold thickness, Insulin, Body mass index, Diabetes pedigree function, Age. From this number 500 people are healthy and 268 were diabetes. The characteristic of desired class was determined to be 0 and 1, meaning each individual with the set of features is inflicted with diabetes or not. The characteristic of class 1 means being patient and that of zero means lacked of disease.

5.2 Performance measures

The performance of the implemented classifier was evaluated by computing accuracy and classification error rate. The respective definitions of Accuracy and Error are as follows [34]:

(a) Accuracy %:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

(b) Error rate %:

$$Error = \frac{TP+TN}{TP+TN+FN+FP} - 1 \quad (3)$$

Where TP, TN, FP and FN denote respectively:

- True positives: classifies Diabetic as Diabetic.
- True negatives: classifies No Diabetic as No Diabetic.
- False positives: classifies No Diabetic as Diabetic.
- False negatives: classifies Diabetic as No Diabetic.

Accuracy parameter indicates that the classifier designed the percentage of correctly classified test has set records. So TP and TN are the values that should be maximum. Error parameter, exactly the opposite is the accuracy. When the best performance is established, the lowest value is zero and maximum amount is equal to one when the highest performance is achieved.

(c) Confusion matrix:

A confusion matrix contains information about actual and predicted classifications done by a classification system (Table 2).

Table 2 – Representation of confusion matrix [34]

Actual Predicted	Predicted	
	Negative	Positive
Negative	TN	FN
Positive	FP	TP

5.3 Results and discussion

We proposed a combination method for classifier, we test proposed method on diabetes dataset and compared the result. We implemented Adaboost algorithm based k-nearest neighbor. It leads to increase of classification accuracy. Results of proposed algorithm vs. knn-based classification by using 10fold

cross validation without using reduction algorithm are shown in table 3. According to the results can be explicitly expression that combination classification is a useful method. The accuracy of Adaboost algorithm has considerable increased.

Table 3- Compare the results of classification with basic algorithm VS classification by adaptive combination

Combined algorithm (Adaboost+KNN)	K-Nearest Neighbor algorithm	Parameters evaluated
0.9739	0.7188	Average accuracy
0.0261	02812	Average error
09893	0.0312	standard deviation
10.135552	1.604541	Elapsed time (s)

Also, according to what was said, reduction feature can be effective. Therefore using a genetic algorithm tried to select effective features. In this way we produce a population of candidate subsets, in other words, a set of features displayed by a binary sequence of length n. Those shown with 0 or 1 in place i, the presence or absence of feature set. *N* is the number of features available. The population of chromosomes is maintained. In each iteration of the algorithm, using an evaluation function, determined the merit of the current population and we choose the best elements of a next-generation population. New chromosome series of old chromosomes according to the following process:

- Combination: combine parts of two parents together to create new child.
- Mutation: changes the random bits of a parent to create a new child.

The initial population size is 50 and the number of iteration is 80 and Combination rate is 50 percent. In each round of implementation of the proposed algorithm the numbers of features selected are shown in table 4. The fourth feature (Triceps skin fold thickness) is not important, because in none of stages is not selected. As well as, second and sixth characteristics (Body mass index) are very effective. The eighth feature (age) is removed at a later stage which shows the importance of not much in the diagnosis of diabetes. The best case is choose6Attributes. So, accuracy propose algorithm by using feature selection method is shown in table 5.As well as, In compared with other methods presented in the diagnosis of diabetes on the dataset PID, our proposed method has gained higher accuracy. This claiming is shown in the diagram.

Table 4-Results for k-Nearest Neighbor algorithm and proposed algorithm with the 10- folds cross validation model

The number of selected feature	No. feature							
	1	2	3	4	5	6	7	8
7	*	*	*		*	*	*	*
6	*	*	*		*	*	*	
5	*	*			*	*	*	
4	*	*				*	*	
3		*				*	*	
2		*				*		

Table 5-Results for proposed algorithm by 6 features with the 10- folds cross validation model

The number of repeat Adaboost	Average accuracy	Average error	standard deviation	Elapsed time (seconds)
10Times	0.9766	0.0234	0.0082	8.818487
15 Times	0.9818	0.0182	0.0125	13.217965
20Times	0.9896	0.0104	0.0134	17.519928

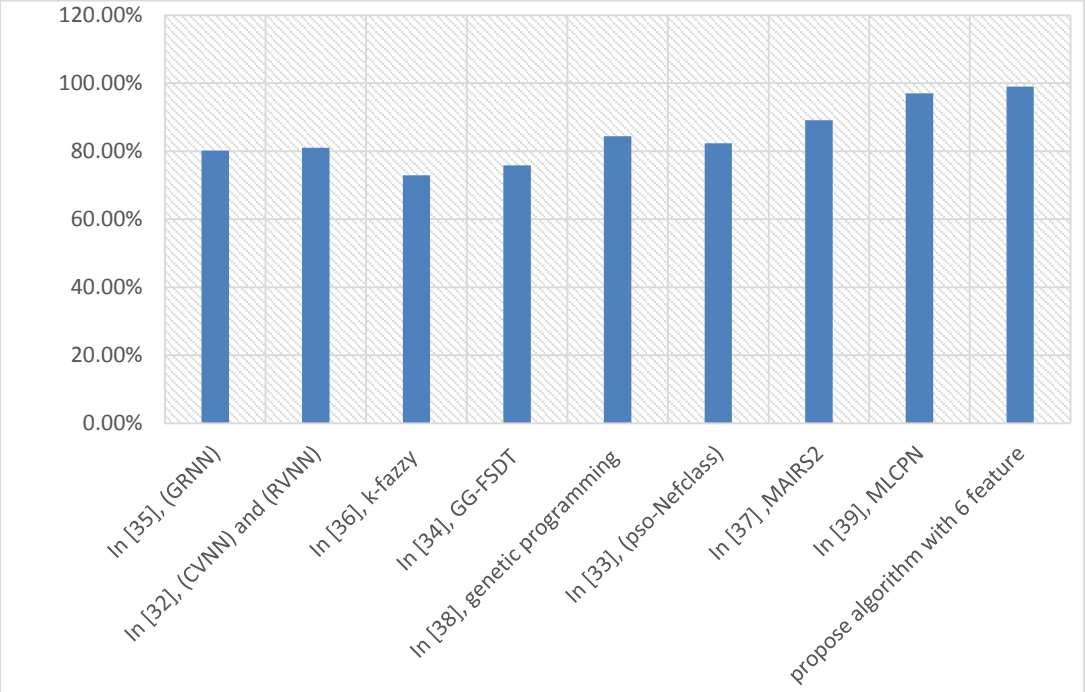


Figure 3. Increasing the accuracy of the proposed method compared to other methods

Conclusion

In this paper, we proposed a new algorithm of classification based on feature selection with genetic algorithm and combination of k-nearest neighbor and Adaboost algorithms to increase the accuracy of classifying PID dataset. Genetic algorithm with entropy criterion focuses on the reduction of non-necessary features in main datasets and after selection of optimal features, the data set with 10-fold cross validation are divided into educational and test data and used in learning model. Algorithm KNN as a basic and weak algorithm is reinforced by Adaboost algorithm through weighting samples and during implementation; classifier is regulated in favor of wrong samples to classify more samples. The practical tests showed that the proposed algorithm has greater accuracy than basic algorithm KNN and other methods.

References

[1] National Diabetes Information Clearinghouse (NDIC): 2014; Available from: http://diabetes.niddk.nih.gov/dm/ pubs/d-iagnosis.

[2] Guariguata, L., et al., GLOBAL ESTIMATES OF DIABETES PREVALENCE FOR 2013 AND PROJECTIONS FOR 2035. Diabetes research clinical practice, 2014. Volume 103(Issue 2,): p. Pages 137–149.

[3] Williams, R., J. Tuomilehto, and S. Bjork, THE ECONOMICS OF DIABETES CARE: an international perspective. London: Blackwell Science; 2000.

[4] Velu, C.M. and K.R. Kashwan, VISUAL DATA MINING TECHNIQUES FOR CLASSIFICATION OF DIABETIC PATIENTS, in IEEE 3rd International Conference on Advance Computing Conference (IACC). 2012. p. 1070 - 1075.

[5] Lavrac, N., SELECTED TECHNIQUES FOR DATA MINING IN MEDICINE. Artificial Intelligence in Medicine, 1999. 16(1): 3-23.

[6] Kotsiantis, S.B., SUPERVISED MACHINE LEARNING: A REVIEW OF CLASSIFICATION TECHNIQUES. Department of computer science and technology, 2007.

[7] Miller, R.A., MEDICAL DIAGNOSTIC DECISION SUPPORT SYSTEMS--PAST, PRESENT, AND FUTURE: A THREADED BIBLIOGRAPHY AND BRIEF COMMENTARY. JAm Med Inform Assoc, 1994. v.1 (1).

[8] Schapire, R.E. and Y. Singer, IMPROVED BOOSTING ALGORITHMS USING CONFIDENCE-RATED PREDICTIONS. Machine Learning 1999. 37(3): p. 297-336.

[9] Gomez-Verdejo, V., J. Arenas-Garcia, and A.R. Figueiras-Vidal, FIGUEIRAS-VIDAL, BOOSTING BY WEIGHTING CRITICAL AND ERRONEOUS SAMPLES. Neurocomputing 69 2006: p. 679-685.

[10] Dietterich, T.G., MACHINE LEARNING RESEARCH: FOUR CURRENT DIRECTIONS. Artificial Intell, 1997. 18: p. 97-136.

[11] Duda, R.O., P.E. Hart, and D.G. Stork, PATTERN CLASSIFICATION. 2001, New York: Wiley.

[12] Galar, M., et al., A REVIEW ON ENSEMBLES FOR THE CLASS IMBALANCE PROBLEM: BAGGING-, BOOSTING-, AND HYBRID-BASED APPROACHES IEEE Systems and Cybernetics Society, 2012 42(4): p. 463 - 484

[13] Kuncheva, L.I., M. Skurichina, and R.P.W. Duin, AN EXPERIMENTAL STUDY ON DIVERSITY FOR BAGGING AND BOOSTING WITH LINEAR CLASSIFIERS. Information Fusion, 2000. 3: p. 245-258.

[14] Krogh, A. and J. Vedelsby, NEURAL NETWORK ENSEMBLES, CROSS VALIDATION, AND ACTIVE LEARNING. Advances in Neural Information Processing Systems, 1995. 7: p. 231-238.

[15] Hansen, L. and P. Salamon, NEURAL NETWORK ENSEMBLES. IEEE, Pattern Analysis and Machine Intelligence, 1990. 12: p. 993-1001.

[16] Hashem, S., OPTIMAL LINEAR COMBINATIONS OF NEURAL NETWORKS: AN OVERVIEW. Neural Networks, 1997. 10: p. 599-614.

[17] Nooraliei, A. and B. Masoumi, Persian Handwritten Digits Recognition Using Zoning and Histogram Projection with Different Dimension of Feature Vector. International Journal of Mechatronics, Electrical and Computer Technology. 4(10): p. 1133-1151.

[18] Aibinu, A.M., M.J.E. Salami, and A.A. Shafie, APPLICATION OF MODELING TECHNIQUES TO DIABETES DIAGNOSIS, in IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010),. 2010.

[19] el habib daho, M., et al., RECOGNITION OF DIABETES DISEASE USING A NEW HYBRID LEARNING ALGORITHM FOR NEFCLASS, in 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA). 2013.

[20] Kamadi, V.S.R.P., et al., A COMPUTATIONAL INTELLIGENCE APPROACH FOR A BETTER DIAGNOSIS OF DIABETIC PATIENTS. Computers and Electrical Engineering, Contents lists available at SciVerse ScienceDirect., 2013.

[21] Kayaer, K. and T. Yildirim, MEDICAL DIAGNOSIS ON PIMA INDIAN DIABETES USING GENERAL REGRESSION NEURAL NETWORKS, in Yildiz Technical University , Department of Electronics and Comm. Eng. Besiktas, Istanbul 34349 TURKEY. 2013.

[22] Madhavi Pradhan, M., et al., DESIGN OF CLASSIFIER FOR DETECTION OF DIABETES USING NEURAL NETWORK AND FUZZY K-NEAREST NEIGHBOR ALGORITHM. International Journal of Computational Engineering Research (ijceronline.com) 2013. 2(5).

[23] Saidi, M., M.A. chikh, and N. settouti, AUTOMATIC IDENTIfiCATION OF DIABETES DISEASES USING A MODIfiED ARTIfiCIAL IMMUNE RECOGNITION SYSTEM2 (MAIRS2). Journal of Medical Systems, 2013. 36(5): p. 2721-2729.

[24] Waqar Aslam, M. and A.K. Nandi, DETECTION OF DIABETES USING GENETIC PROGRAMMING, in 18th European Signal Processing Conference (EUSIPCO-2010) Aalborg. 2010. p. 23-27.

[25] Velu, C.M. and K.R. Kashwan, MULTI-LEVEL COUNTER PROPAGATION NETWORK FOR DIABETES CLASSIFICATION in International Conference on Signal Processing, Image Processing and Pattern Recognition [ICSIPRI. 2013. p. 190 - 194

[26] Jacek, J. and C. Susan, GENETIC ALGORITHMS FOR FEATURE SELECTION AND WEIGHTING, in IN PROCEEDINGS OF THE IJCAI'99 WORKSHOP ON AUTOMATING THE CONSTRUCTION OF CASE BASED REASONERS. 1999, citeseerx.ist.psu.edu.

[27] Mousavirad, S. and H. Ebrahimpour-Komleh, Wrapper Feature Selection using Discrete Cuckoo Optimization Algorithm.

[28] Mousavirad, S. and H. Ebrahimpour-Komleh, Wrapper Feature Selection using Discrete Cuckoo Optimization Algorithm. International Journal of Mechatronics, Electrical and Computer Technology. 4(11): p. 709-721.

[29] Cover, T. and P. Hart, NEAREST NEIGHBOR PATTERN CLASSIFICATION. IEEE Information Theory Society, 2013. 13(1): p. 21 - 27

[30] Freund, Y. and R. Schapire, EXPERIMENTS WITH A NEW BOOSTING ALGORITHM. In Proceedings of the thirteenth international conference on machine learning, 1996: p. 148–156.

[31] Xuchun, L., W. Lei, and S. Eric, ADABOOST WITH SVM-BASED COMPONENT CLASSIFIERS,. Engineering Applications of Artificial Intelligence, 2008. 21: p. 785-795.

[32] Ghahramani, Z., INFORMATION THEORY, M.R. Ltd, Editor. 2007, Encyclopedia of Cognitive Science: University College London United Kingdom. p. 1-7.

[33] Pima Indians Diabetes Data set. Available from: https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes.

[34] Fayssal, B. and M.A. CHIKH, DESIGN OF FUZZY CLASSIFIER FOR DIABETES DISEASE USING MODIFIED ARTIFICIAL BEE COLONY ALGORITHM. Computer Methods and Programs in Biomedicine 2013. 112(1): p. 92-103.