IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

A Critical Study on Data Mining Techniques in Healthcare Dataset

Prateek Pandey[1] and Dr. P. K. Roy[2]
Research Scholar, Department of Computer Science, A.P.S. University, Rewa, M.P. India.[1]
Head of Department of Computer Science, U.T.D. , A.P.S. University, Rewa, M.P., India.[2]

Abstract: *In the absence of medical diagnosis substantiations, it is complicated for experts to speak out about the grade of disease with affirmation. Generally, many tests are done that involve clustering or classification of large-scale data. However,, many tests could complicate the main diagnosis process and lead to difficulty obtaining the end results, particularly when many tests are performed. This kind of difficulty could be resolved*

Keywords: Medical diagnosis, Disease, Clustering Classification

I. INTRODUCTION

In the 21st century, humans use different technologies to improve society. Every day, human beings useuse vast amounts of data, and these data are in different fields. Among them diagnosis of disease is a very critical field where the human and computer incorporate to achieve best results by balancing the knowledge of human experts in relating problems and goals with the search potential of computers. This kind of difficulty could be resolved with the aid of machine learning techniques. In this paper survey on four different disease diagnoses are taken into the consideration. The heart Disease, Breast Cancer Disease, the Diabetes Disease and Liver Disease are analyzed and observed with existing works. This survey paper reveals various existing approaches that have processed for diagnosis these diseases using data mining techniques.

Comparison were targeted as a main objective in many studies that mainly aimed to develop a prediction model in a critical fields, like medicine, by investigating several data mining methods, intending to get the model that have the highest Prediction accuracy. The Aim of this survey is to analyze some of the famous earlier works in data mining techniques in the diagnosis of heart disease, breast cancer, diabetes and Liver disease.

II. SURVEY ON HEART DISEASE DIAGNOSIS USING DATA MINING TECHNIQUES

The numbers of medical decision support systems are employed using different approaches. George et al. have anticipated decision support system to classify and detect agitation transition. In this system support vector machines are used for detection. This system is for Dementia patients. This system presents a decision confidence measure and two new SVM architectures, which were useful to agitation detection and agitation transition detection. An accuracy of 91.4% was achieved; in assessment with 90.9% for the conventional SVM[1]. Haitham and Alan have projected automated recognition of obstructive sleep apnea syndrome using the SVM classifier. In this study, they evaluated features from the magnitude and phase of the thoracic and abdominal respiratory effort signals for OSA detection. This supports on the physiological fact that all through normal breathing the abdominal and thoracic efforts happen simultaneously. The aim of this study is to appraise classification of whole night regular and apneic epochs using mined features from the phase and magnitude of the respiratory effort signals, compared and combined with some other features from HRV and oxygen saturation signals [2] [3]. Support Vector machines have also been utilized in decision support systems such as [4].An intelligent system based support vector machine along with a radial basis function network is accessible for the diagnosis. The support vector machine with sequential minimal optimization algorithm is applied to India based patients' data set. Then, the Radial Basis Function (RBF) network structure qualified by Orthogonal Least Square (OLS) algorithm is functional to same data set for predictions [5]. Tsai and Watanabe proposed a genetic algorithm (GA) based method and implemented for influential the set of fuzzy membership

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

functions that can provide an optimal classification of myocardial heart disease from ultrasonic images. In this method an average classification rate of 96% is achieved [6]. In a different advance genetic algorithm is used to determine the attributes which donate more towards the diagnosis of heart ailments which ultimately reduces the number of tests which are desirable to be taken by a patient [7]. Yang and Honavar have projected a feature subset algorithm using genetic algorithm. A genetic algorithm to decide on the optimal feature subset for use with back propagation artificial neural networks has been described [8].A genetic algorithm for feature selection as well as for optimization of Support Vector Machine (SVM) parameter has been proposed by Haung. The projected method performs feature selection and parameters setting in an evolutionary way [9]. Very recently, a real coded Genetic algorithm for critical feature analysis for heart disease diagnosis has been portrayed [10]. Diagnosis of Heart Disease using Data mining Algorithm proposed by Rajkumar and Sophia, In their approach the preliminary diagnosis of a heart attack is made by a combination of clinical symptoms and characteristic electrocardiogram (ECG) changes. The accuracy with this technique is 52.33% [11]. Palaniappan and Awang proposed Intelligent Heart Disease Prediction System Using Data Mining Techniques. This investigation has developed a prototype Intelligent Heart Disease Prediction System (IHDPS) using data mining techniques, namely, Decision Trees, Nave Bayes and Neural Network. Results show that each technique has its unique strength in realizing the objectives of the defined mining goals [12]. The features of the artificial neural network (ANN), high accuracy and learning rate, make it worth trying as an algorithm to the prediction of heart disease [13].

III. SURVEY ON BREAST CANCER DIAGNOSIS

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 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 Network(GRNN),Probabilistic 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 the tumor 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 recognize 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 .

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

The performance of these neural network structures was investigated for breast cancer diagnosis problems. 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 Lisboa with 164 controls and 94 cases in WEKA machine learning tool. To statistically validate the association found, permutation tests were used. They found a high-risk breast cancer group composed of 13 cases and only 1 control, with a Fisher Exact Test(for validation) value of 9.7*10-6 and a p-value of 0.017. These results showed that it is possible to find statistically significant associations with breast cancer by deriving a decision tree and selecting the best leaf. In [16] Dr. Medhat Mohamed Ahmed Abdelaal and Muhamed Farouq investigated the capability of the classification SVM with Tree Boost and Tree Forest in analyzing the DDSM dataset for the extraction of the mammographic mass features along with the age that discriminates true and false cases. Here, SVM techniques show promising results for increasing diagnostic accuracy of classifying the cases witnessed by the largest area under the ROC curve comparable to values for tree boost and tree forest. In [17] Wei-pin Chang, Der-Ming and Liou explored that the genetic algorithm model yielded better results than other data mining models for the analysis of the data of breast cancer patients in terms of the overall accuracy of the patient Certified Journal classification, the expression and complexity of the classification rule. The artificial neural network, decision tree, logistic regression, and genetic algorithm were used for the comparative studies and the accuracy and positive predictive value of each algorithm were used as the evaluation indicators. WBC database was incorporated for the data analysis followed by the 10-fold cross-validation. The results showed that the genetic algorithm described in the study was able to produce accurate results in the classification of breast cancer data and the classification rule identified was more acceptable and comprehensible. In [18] K. Rajiv Gandhi, Marcus Karnan and S. Kannan in their paper constructed classification rules using the Particle Swarm Optimization Algorithm for breast cancer datasets. In this study to cope with heavy computational efforts, the problem of feature subset selection as a pre-processing step was used which learns fuzzy rules bases using GA implementing the Pittsburgh approach. It was used to produce a smaller fuzzy rule bases system with higher accuracy. The resulted datasets after feature selection were used for classification using particle swarm optimization algorithm. The rules developed were with rate of accuracy defining the underlying attributes effectively. In [19] J. Padmavati performed a comparative study on WBC dataset for breast cancer prediction using RBF and MLP along with logistic regression. Logistic regression was performed using logistic regression in SPSS package and MLP and RBF were constructed using MATLAB. It was observed that neural networks took slightly higher time than logistic regression but the sensitivity and specificity of both neural network models had Shelly Gupta et al a better predictive power over logistic regression. When comparing RBF and MLP neural network models it was found that RBF had good predictive capabilities and also time taken by RBF was less than MLP. In [20] Chul-Heui Lee, Soen-Hak Soc and Sang-Chul Choi in their study proposed a new classification method based on the hierarchical © 2015, IRJET ISO 9001:2008 granulation structure using the rough set theory. The hierarchical granulation structure was adopted to find the classification rules effectively. The classification rules had minimal attributes and the knowledge reduction was and lower accomplished by using the was upper approximations of rough sets. A simulation performed on WBC dataset to show the effectiveness of the proposed method. The simulation result showed that the proposed classification method generated minimal classification rules and made the analysis of information system easy. In [21] Aboul Ella Hassanien, and Jafar M.H.Ali in their paper presented a rough set method for generating classification rules from a set of observed 360 samples of the WBC data. The attributes were selected, normalized and then the rough set dependency rules were generated directly from the real value attribute vector. Then the rough set reduction technique was applied to find all reducts of the data which contains the minimal subset of attributes that are associated with a class label for classification. They showed that the total number of generated rules was reduced from 472 to 30 rules after applying the proposed simplification algorithm. They also made a comparison between the obtained results of rough sets with the well known ID3 decision tree and concluded rough sets showed higher accuracy and generated more compact rules. In[22] Sudhir D. Sawarkar et al applied SVM and ANN on the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

WBC data .The results of SVM and ANN prediction models were found comparatively more accurate than the human being. The 97% high accuracy of these prediction models can be used to take decision to avoid biopsy. In [23] Sepehr M. H. Jamarani et al presented an approach for early breast cancer diagnosis by applying combination of ANN and multiwavelet based sub band image decomposition. The proposed approach was tested using the MIAS mammographic databases and images collected from local hospitals. The best performance was achieved by BiGHM2 multiwavelet with areas ranging around 0.96 under ROC curve. The Certified proposed approach could assist the radiologists in mammogram analysis and diagnostic decision making.

IV. SURVEY ON DIABETES DISEASE DIAGNOSIS USING DATA MINING TECHNIQUE

On A good number of researches have been reported in literature on diagnosis of different deceases. Sapna and Tamilarasi [24] proposed a technique based neuropathy diabetics. Nerve disorder is caused by diabetic mellitus. Long term diabetic patients could have diabetic neuropathies very easily. There is fifty (50%) percent probability to have such diseases which affect many nerves system of the body. For example, body wall, limbs (which is called as somatic nerves) could be affected. On the other hand, internal organ like heart, stomach, etc., are known as automatic nerves. In this paper the risk factors and symptoms of diabetic neuropathy are used to make the fuzzy relation equation. Fuzzy relation equation is linked with the perception of composition of binary relations that means they used Multilayer Perceptron NN using Fuzzy Inference System. Leonarda and Antonio [25] proposed automatic detection of diabetic symptoms in retinal images by using a multilevel perceptron neural network. The network is trained using algorithms for evaluating the optimal global threshold which can minimize pixel classification errors. System performances are evaluated by means of an adequate index to provide percentage measure in the detection of eye suspect regions based on neurofuzzy subsystem. Radha and Rajagopalan [26] introduced an application of fuzzy logic to diagnosis of diabetes. It describes the fuzzy sets and linguistic variables that contribute to the diagnosis of disease particularly diabetes. As we all know fuzzy logic is a computational paradigm that pro- vides a tool based on mathematics which deals with un- certainty. At the same time this paper also presents a computer- based Fuzzy Logic with maximum and mini- mum relationship, membership values consisting of the components, specifying fuzzy set frame work. Forty patients' data have been collected to make this relationship more strong. Jeatrakul and Wong [27] presented a comparison of neural network techniques for binary classification problems. The classification performance obtained by five different types of neural networks, i.e., Back Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), General Regression Neural Network (GRNN), Probabilistic Neural Network (PNN), and Complementary Neural Network (CMTNN). The comparison is done based on three benchmark data sets obtained from UCI machine learning repository. Zhou, Purvis and Kasabov [28] described a general method based on the statistical analysis of training data to select fuzzy membership functions to be used in connection with fuzzy neural networks. The technique is first described and then illustrated by means of two experimental examinations for medical data. Ten-Ho Lin and Von-Wun Soo [29] proposed alternative pruning method based on the Minimal Description Length (MDL) principle. The MDL principle is a tradeoff between theory complexity and data prediction accuracy given the theory. A greedy search algorithm of the minimum description length to prune the fuzzy ARTMAP categories one by one was proposed. The experiments showed that fuzzy ARTMAP pruned with the MDL principle gave better performance with fewer categories created compared to original fuzzy ARTMAP and other machine learning systems. They tested those techniques on a number of benchmark clinical databases such as heart disease, breast cancer and diabetes databases. Faezeh, Hossien, Ebrahim [30] proposed a fuzzy clustering technique (FACT) which determined the number of appropriate clusters based on the pattern essence. Different experiments for algorithm evaluation were per- formed which showed a better performance compared to the typical widely used K-means clustering algorithm. Data was taken from the UCI Machine Learning Repository [31]. Santi Waulan et al. [32] proposed a new SSVM for classification problems. It is called Multiple Knot Spline SSVM (MKS-SSVM). To evaluate the effectiveness of their method, they carried out an experiment on Pima Indian diabetes dataset. First, theoretical of MKS-SSVM was MKS-SSVM and presented. Then, application of comparison with SSVM in diabetes disease diagnosis was given. The results of this study showed that the MKS- SSVM was effective to detect diabetes disease diagnosis.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

V. SURVEY ON LIVER DISEASE DIAGNOSIS USING DATA MINING TECHNIQUE

Model selection Paul Mangiameli et al., [2] proposed affects the decision support systems accurately. In their model selection, how to affects the accuracy of decision support system hydrides by single model and ensembles. They proposed single model is not more accurate than ensembles. Ahmed M. Hashem et al.,[18] proposed to Cirrhosis or fibrosis single stage predict Liver classification model and multistage classification model. In their model based on Decision Tree, Neural Network, Nearest Neighborhood clustering and Logistic Regression.

for Ziol.M et al., [3] proposed to evaluated liver fibrosis with chronic hepatitis C for patients using liver stiffness proposed (LSM).Z. Jiang.Z.,[4] discovering the corresponding degree of fibrosis by support vector machine (SVM). measurement

Kemal Polat et al.,[5] proposed resource allocation mechanism of AIRS was changed with a new one decided by Fuzzy-Logic. This approach called as Fuzzy- AIRS was used as a classifier in the diagnosis of Liver Disorders. In this Classification accuracies were evaluated by comparing them with reported classifier's accuracy, time and number of resources.

Piscaglia et al.,[6] proposed to predict Liver cirrhosis and other liver-related diseases used by Artificial neural network.

Dong-Hoi Kim et al.,[7] proposed machine learning technique and decision tree (C4.5).In this method is used for to predict the susceptibility to two liver diseases such as chronic hepatitis and cirrhosis from single nucleotide polymorphism (SNP) data. They also used to identify a set of SNPs relevant to those diseases.

Anh Pham,[8] developed optimizing the classification accuracy when analyzing some medical datasets. This proposed work done by new meta-heuristic approach, called the Homogeneity-Based Algorithm (or HBA).This approach used to predict error rates and associated penalty costs. These costs may be dramatically different in medical applications as the implications of having a false- positive and a false-negative case may be tremendously different.

and Rong-Ho Lin [9] proposed to predict accuracy of liver disease using case-based reasoning (CBR) classification and regression tree (CART) approach. He also integrates CART and CBR for the diagnosis of liver diseases. In this model included two major steps. (1) CART To diagnose whether a patient suffers from liver disease using CART. (2) To predict which types of Liver disease affected for patients using CBR.He also [12], proposed to determine whether patients suffer from liver disease or not using case-based reasoning, artificial neural networks and analytic hierarchy methods. They also predict which types of liver disease suffered human body.

Chun-Ling Chuang et al., [15] proposed to diagnosis early Liver disease and predict classification accuracy by integrated case-based reasoning into classification and regression tree, back-propagation neural network (BPN), discriminatory analysis and logistic regression of classification methods in data mining techniques. In their methods used a ten-fold cross-validation to select a best

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 IRJET Volume: 02 Issue: 05 | Aug-2015 www.irjet.net p-ISSN: 2395-0072

model for more accurate diagnosis. Chin-Yuan Fan et al., [16] proposed to diagnose Liver disease used by hybrid model.In Their model, integrating case based reasoning and fuzzy decision tree for preprocessing and predict disease.

Pei-Chann Chang et al., [17] proposed to predict Liver disease used by integrating a case-based reasoning and a particle swarm optimization model.

Ihsan Ömür Bucak et al., [10] proposed to diagnose the liver disease using CMAC (Cerebellar Model Articulation Controller) neural network.

Ruxandra Stoean et al., [11] proposed to find five different possible level of liver fibrosis and chronic hepatitis C using two novel evolutionary Techniques. In these evolutionary techniques, hill climbing algorithm used for selection of attributes and support vector machine used for evaluation.

Catalin Stoean et al., [13] proposed to diagnosis Liver fibrosis using hill climbing algorithm for every new patients. They also proposed extracts a set of attributes (most meaningful) from the available attributes. He also [14] proposed to find five different possible levels of liver fibrosis and chronic hepatitis C using two novel Techniques. evolutionary In these evolutionary techniques, optimization engine used for selection of attributes and support vector machine used for evaluation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

VI. CONCLUSION

In this survey paper the problem of summarizing the different algorithm of data mining are used in the field of medical prediction are discussed. The main focus is on using different algorithm and combination of several targets attributes for different types of disease prediction using data mining. A foremost class of problems in medical science absorbs the diagnosis of disease, based upon an assortment of tests carried out upon the patient. When several tests are involved, the ultimate diagnosis may be difficult to obtain, yet for a medical expert. This has given rise, over the past few decades, to automated problem-solving tools, intended to assist the physician in making sense out of the welter of data. In healthcare, data mining is becoming increasingly more essential.

The selection of data mining approaches depends on the nature of the dataset if the dataset consist of the labeled features then the classification techniques can be suggested for best prediction. If the dataset is with unlabelled features then the clustering techniques are best suited for pattern recognition. If the optimization of the results needs to be improvised means then bio inspirational based techniques are best suited. Keeping in consideration with these existing problems this paper aims to survey the existing approaches in the field of medical sciences and the importance of data mining techniques used by various authors. The study reveals the importance of life threading disease should be diagnosed in the early stages. Future work is only concentrating on live disease diagnosis due to more number of death rate occurs because of it. Diagnosing in early stage suffers from high false alarm and detection rate is low.

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 using data prediction system 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

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

[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 high-risk 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
[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/compa r_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 Neural Network 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," International 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 Performance of Different Neural Networks for Binary Classifica tural Language Processing, Bangkok, 20-22 2009, doi:10.1109/SNLP.2009.5340935 October pp. 111-115.
[28] Q. Q. Zhou, M. Purvis and N. Kasabov, "Membership Function Selection Method for Fuzzy Neural Networks," University of Otago, Dunedin, http://otago.ourarchive.ac.nz/handle/10523/1027 2007.
[29] T.-H. Lin and V.-W. Soo, "Pruning Fuzzy ARTMAP Using the Minimum Description Length Principle in Learning from Clinical Databases," Proceedings of the 9th International Conference on Tools with Artificial Intelligence, Newport Beach, 3-8 November 1997, pp. 396- 403.
[30] F. Ensan, M. H. Yaghmaee and E. Bagheri, "Fact: A New Fuzzy Adaptive Clustering Technique," The 11th IEEE Symposium on Computers and Communications, Sardinia, 26-29 June 2006, pp. 442-447. doi:10.1109/ISCC.2006.73
[31] UCI Machine Learning Repository.
http://www.ics.uci.edu/mlearn/MLRepository.html [32] S. W. Purnami, A. Embong, J. M. Zain and S. P. Rahayu, "A New Smooth Support Vector Machine and Its Applications in Diabetes Disease Diagnosis," Journal of Computer Science, Vol. 5, No. 12, pp. 1006-1011.
[33] Wasyluk H, Onisko, Druzdzel M.J "Support of diagnosis of liver disorders based on a causal bayesian network model", In Medical Science Monitor, 7(Suppl. 1):327-332, May 2001.
[34] Paul Mangiameli, David West, Rohit Rampal, "Model selection for medical diagnosis decision support systems", Decision Support Systems 36 (2004) 247-259.
[35] M. Ziol, A. Handra-Luca, A. Kettaneh,"Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C", Hepatology 41 (2005) 48-54.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, March 2025
Impact Factor: 7.67

[36] Z. Jiang, K. Yamauchi, K. Yoshioka, et al., "Support vector machine-based feature selection for classification of liver Fibrosis grade in chronic hepatitis C, J. Med. Syst. 20 (2006) 389-394.
[37]. Kemal Polat,Seral Seralsahan, Halife Kodaz, Salih Gunes,"Breast cancer and liver disorders classification using artificial immune recognition system (AIRS) with performance evaluation by fuzzy resource allocation mechanism",Expert Systems with Applications 32 (2007) 172-183.
[38] Piscaglia F, Cucchetti A, Orlandini A, Sagrini E, Gianstefani A, Crespi C,"Prediction of significant fibrosis in chronic hepatitis C patients by artificial neural network analysis of clinical factors" Digestive and Liver Disease 2007;39:A25
[38] Dong-Hoi Kim, Saangyong Uhmn, Young-Woong Ko, Sung Won Cho, Jae Youn Cheong, Jin Kim, "Chronic Hepatitis and Cirrhosis Classification Using SNP Data, Decision Tree and Decision Rule" Computational Science and Its Applications ICCSA 2007 Lecture Notes in Computer Science Volume 47 Issue 07, 2007, pp 585- 596.
[39] Huy Nguyen Anh Pham, Evangelos Triantaphyllou "An application of a new meta-heuristic for optimizing the classification accuracy when analyzing some medical datasets" Expert Systems with Applications 36 (2009) 9240-9249
[40]. Rong-Ho Lin,"An intelligent model for liver disease diagnosis", Artificial Intelligence in Medicine (2009) 47, 53-62.
[41]. Ihsan Ömür Bucak, Semra Baki,"Diagnosis of liver disease by using CMAC neural network approach",Expert Systems with Applications 37 (2010) 6157-6164.
[42]. Ruxandra Stoean, Catalin Stoean, Monica Lupsor, Horia Stefanescu, and Radu Badea, "Evolutionary Conditional Rules versus Support Vector Machines Weighted Formulas for Liver Fibrosis Degree Prediction" Annals of the University of Craiova, Mathematics and Computer Science Series Volume 37(1), 2010, Pages 43-54.
[43] Z.Rong-Ho Lina, Chun-Ling Chuang b,n A hybrid diagnosis model for determining the types of the liver disease Computers in Biology and Medicine 40 (2010) 665-670.
[44].Catalin Stoean, RuxandraStoean, MonicaLupsor, HoriaStefanescu, RaduBadea,"Feature selection for a cooperative coevolutionar y classifi…classifications",computer methods and programs biomedicine 107(2012)382-392.
M.Rasmy,Khaled [49] Ahmed M.Hashem,M.Emad M.Washba Olfat G.Shaker "Single stage and multistage classification Models for the prediction of liver fibrosis degree in patients with chronic hepatitis C infection" Computer Methods and Programs in Biomedicine, Volume 105, Issue 3, March 2012, Pages 194-209.
[50] Polat K. "Application of Attribute Weighting Method Based on Clustering Centers to Discrimination of Linearly Non-Separable Medical Datasets" Journal of Medical Systems, August 2012, Volume 36, Issue 4, pp 2657-2673