Physician's Decision Process for Disease Diagnosis of Overlapping Syndrome in Liver Disease using Soft Computing Model

Mohammed Abdullah Alghamdi, Sunil G Bhirud, Afshar M. Alam

Abstract- An understanding towards the overlapping syndromes is essential to cope up the liver disease with the paradigm shift which is underway. Autoimmune hepatitis (AIH), primary biliary cirrhosis (PBC) and primary sclerosing cholangitis (PSC) all belong to the family of autoimmune liver diseases. This can result in a transient phenotypic overlap or a combined syndrome with characteristics of both diseases. It demonstrates mixed clinical presentations of immune-mediated liver injury. Due to overlapping features of AIH, some patients connote such phenotype that leads to a dubious diagnosis of any disease. In some cases, this type of diagnosis system causes the patient's death in the absence of well-validated diagnostic criteria. Some improvements in diagnosis instrumentation and automation that has become to capture experience of any veteran person, is needed. Simulation of such diagnosis experience is extremely important as it leads to knowledge discovery. In this paper, we gave a soft computing model based disease diagnosis system for overlapping syndrome of liver disease. It helped the in physician's decision process that comes after long experience in new individual.

Index Terms— AIH, PBC, PSC, physician's, diagnostic, phenotypic, cholangitis.

I. INTRODUCTION

The recent development in the computer aided medical diagnosis is increasing gradually. Although, this is doubtless that assessment of data taken from patient and experts decisions are the most significant factors in diagnosis. But, dealing with data for classification to help the expert requires different artificial intelligence techniques. Classification systems provide medical data to be examined in shorter time and more detailed. They also minimize the possible errors and make medical diagnosis systems more rigorous medical diagnosis systems. Identifying the liver disorder is too difficult in an early stage while it will be functioning normally in spite of partially damaged [1]. Undetectable this early stage of liver disorder increases the patient's diagnosis problems. Analyzing the level of enzymes in patients' blood helps to diagnose the liver disorder [2]. Moreover, body monitoring devices are widely used to administer the condition of any patient. These body sensor networks help to inform patient's present condition. Apart from these computer aided systems, we need automatic classification algorithms for evaluation of patient's data. Researchers have devoted a lot of effort to make the diagnosis systems more robust.

Manuscript Received on December 2014.

Mohammed Abdullah Alghamdi, College of Engineering, Bharti Vidyapeeth Bhavan, Pune, India.

Prof. Sunil G. Bhirud, Department of Computer Engineering & Information Technology, VJTI,Mumbai India.

Prof. M. Afshar Alam, Department of Computer Science, Jamia Hamdard, New Delhi, India.

Implementation of expert systems along with monitoring devices for liver disease and others has reduced the patients queues at the experts such as hepatologists, haematologist etc. Liver disorders are also an important disease in medicine. Levels of enzymes combined with blood are analyzed in Liver Disorder diagnosis. It can be a lot of possible errors in this diagnosis due to the number of enzymes to be many as well as the effects of taken different alcohol rates to be varied from one patient to the other. Liver disease is one of the major causes of all the natural mortalities, and still increasing year-on-year throughout the world. Most often, liver failure occurs gradually and over many years. However, a more rare condition known as acute liver failure occurs rapidly (in as little as 48 hours) and can be difficult to detect initially. Nearly 70 percent of patients with autoimmune hepatitis are women, most between the ages of 15 and 40. The disease may start at any age, but is most common in adolescence or early adulthood [3]. It is a disease getting constantly challenged by many eminent and premier researchers. Although, some advancements have been reported for its clinical prevention and cure and there has been a noticeable decline in the lives lost, but they are not quite adequate [4]. The lack of affordable treatment and early detection is the crux of this hostile situation. It is becoming harder for the perennial biomedical scientists and researchers to exorcize this demon. Figure 1 demonstrates the rising tide of liver disease deaths given by the British Liver Trust. This official statistic finds out some truths pertaining to liver disease. Liver disease is the 5th dangerous disease in England and Wales, after heart, cancer, stroke and respiratory disease. In 2008, 16,087 people (13,805 people in England and Wales, 1,903 in Scotland and 379 in Northern Ireland) died in the UK those were suffering from liver disease. This is a 4.5 % increase since 2007, and by 12% in just three years, since 2005, totaling 46,244 lives lost. Analysts predicted that deaths from liver disease will be doubled in 20 years if this continues. It means twice as many people now die from liver disease as in 1991 [3].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

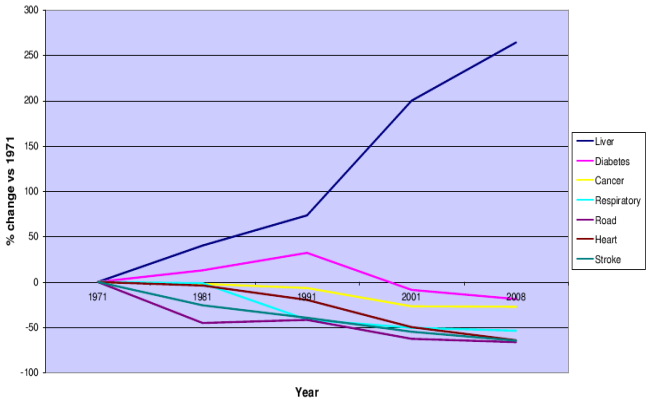

Figure 1: Mortality Statistics due to Liver Disorder

There are two types of liver disorder: Hepatitis and Cirrhosis. Basically, Hepatitis is an inflammation of the liver that can be caused by a virus, by inherited disorders, and sometimes by certain medications or toxins such as alcohol and drugs. Apart from these, Cirrhosis has some other causes, including [1], [4]:

1. It can be caused by excessive alcohol consumption.
2. chronic infection with hepatitis B virus
3. inherited disorders of iron and copper metabolism
4. severe reactions to certain medications
5. fatty liver caused by obesity
6. infections from bacteria and parasites usually found in the tropics
7. repeated episodes of heart failure with liver congestion and bile-duct obstruction

Moreover, this disease becomes more serious in case of overlapping syndrome that makes the prognosis too difficult. Figure 2 shows the overlapping syndrome of autoimmune liver disease. Medical diagnosis has become one of the most progressive disciplines in recent years. Some improvements in diagnosis instrumentation and automation are needed for this type of overlapping syndromes. This computer aided diagnosis process captures the experience of veteran persons [5]. Simulation of such diagnosis experience extremely important as it leads to knowledge discovery. This diagnosis system is very helpful for non-expertise to make the right prescription. On a physician's realistic practice, Medical diagnosis relies to a huge degree it may take years for a physician, particularly a novel or junior one, to build up enough experiences. The different disease diagnosis models have been improved to assist physicians deal with diagnostic problems, as they take care of medical diagnosis as a decision making process during which the physician induces the finding of a novel and strange case from disorder clinical experience [6]. Fast developments of information technology have set aside physicians the troubles of transliterating patient's medical histories literal in latest years. The data processing systems, medical histories and other associated medical information can be competently accumulated in great amount in databases for quick recovery, cross-reference and study. A huge majority of earlier studies on disease diagnosis employed statistical techniques for modeling mainly due to the complexity in accessing a substantial amount of medical data [2]. Liver disease is one of the disease diagnoses attainable. The liver is the biggest internal organs in the human body, taking part a major role in metabolism and serving numerous vital functions, e.g. decomposition of red blood cells, etc. Liver disease is generally caused by inflammation or injured hepatocytes, registers a tenacious presence on the list of top ten incurable diseases in World [7].

Figure 2: Overlapping symptoms of autoimmune liver disease

Specialist systems and machine learning methods are more and more introduced to assist the liver disease diagnosis. With the assist of diagnostic systems, the feasible faults experts made in the course of diagnosis can be avoided, and the medical information can be scrutinized in shorter time and more detailed too [8]. Earlier research on disease diagnosis employs statistical techniques for modeling. Statistical techniques, on the other hand, need assumptions and are generally implemented to examine linear information. Therefore, they are less competent of dealing with huge and difficult nonlinear and dependent information, such as the vast data gathered incessantly through health examination and medical treatment [6]. Hence, to examine huge and complex data, more efficient strategies are required. For instance, data mining and soft computing methods may offer a constructive result. Soft computing varies from conventional (hard) computing in that, (unlike hard computing), it calculates using the tolerance of indistinctness, uncertainty, partial truth, and approximation. Actually, the role model for soft computing is the human mind. The directing principle of soft computing is: use the tolerance for imprecision, uncertainty, partial truth, and approximation to attain tractability, robustness and low solution cost. Soft computing methods, such as genetic algorithms, fuzzy sets, artificial neural networks, rough sets, inductive logic programming, Chaos Theory, Fractal Theory, Possibility Theory, Probability Theory, and Learning Theory have been launched for application to medical related fields [9] over the past few decades. For medical diagnostic processes, artificial neural networks can be applied as one of the most famous techniques. Artificial neural networks have previously demonstrated its efficiency and popularity for the medical diagnostic processes with dissimilar existing applications worldwide [10]. This paper is an attempt to propose a soft computing model to diagnose the liver disease along with overlapping symptoms intelligently for the physician decision process. Primarily, artificial neural networks classification used for categorizing the liver disease. Rough set rule induction employed in Learn by Example (LEM) algorithm used to produce classification rules. This rule induction surmounts the disadvantages of Levenber-Marquardt algorithm (MLP) and enhanced the precision. Finally the fuzzy rules were employed to recognize the kinds of the liver disease.

II. Related Work

We wish always to have a such intelligence system that will work

on our behalf and learn at every stage adaptively. Simulated brain and its learning ability led to the invention of artificial neural network (ANN). Its beneficial application can be seen in the medical diagnosis systems, especially in the diagnosis and determine disease severity or predict disease [6], [11]. A good example is diagnosis system of liver disorders that gave better results. These results help physicians recognize liver disorder more accurately and provide better prescription. There is no doubt that ANN based solutions having some error, but this is natural. It is becoming easier for the perennial biomedical scientists and researchers to exorcize this daemon using these approaches. In [12], ASAMC (Stochastic Approximation Monte Carlo) was proposed. It was compared with SAMC, SA and BFGS using 250 samples of BUPA liver disorders. ASAMC outperforms the others in terms of training error and testing error rate. In [7], a new neural network based method for typical extraction was proposed. It reduced the dimensionality of input vectors use for supervised learning problems and made the classification more efficient. Topology-optimization Evolutionary Neural Network was proposed in [13] that gave the optimal structure of the MLP to a given problem. This was first evolutionary neural network technique. It was compared with the five data mining algorithm using classification and regression process for BUPA liver disorders. It outperformed the others with accuracy of 70.9%. Pattern classification technique was proposed in [14]. It was improved version of the NN model. Based on the basic concepts of MLP, RBF and SVM, a new neural network with delimited weights is suggested, and some empirical results were reported. Then, the relaxed conditions for RBF networks to be universal approximations are presented. They present that RBF networks can uniformly near any continuous function on a compress set supplied that the radial basis activation function is almost continuous in every location, in a particular place basically marked off, and not a polynomial. In [15], two diagnosis approaches for liver disorder were proposed. These were Hopfield Neural Network (HNN) and Fuzzy Hopfield Neural Network (FHNN), these two were more accurate, more economical and fast as compared to traditional diagnosis systems. HNN and FHNN gave pleasant improvement results for liver disorder with the accuracy of 88.2% and 92% respectively. Apart from the neural network, another concept was used to make diagnosis systems more accurate which was based fuzzy logic. Several fuzzy based methods were proposed like FES-LD [16], 3-phase diagnosis method in [17] and [18], and [19] etc. Due qualitative nature of medical data, it has some noise that requires a mapping function for changing to quantitative. Fuzzy logic is a very good source to make the medical data noiseless [20]. It makes doctors able to recognize the liver disease with higher accuracy [21], [5]. The liver disease may not cause any indications at an earlier stage or the symptoms may be vague, like feebleness and loss of energy. Symptoms partially depend on the type and the extent of liver disease. Liver diseases are diagnosed based on the liver function test. Anyway, this disease cannot be predicted at an earlier stage due to lack of symptoms and signs, in this research attempt to apply soft computing techniques for intelligent diagnosis of liver disease. Fuzzy based methods use ANN, LEM, SVM, FREx etc. for classification and rule induction in liver diagnosis systems. ANN is used for classifying the liver disorder, LEM to create classification rules [17]. FREx is fuzzy extraction algorithm which extracts explainable knowledge as rules from trained SVM. FREx was tested with BUBA liver disorder database. The result was quite good besides, they presented an ignorable error tax. The resultant coverage, in their tests with 3 fuzzy sets, was more than 94% with only 21 rules. This may be considered an excellent result. The FREx method is being extended to multiple classes SVM and it will be appraised in other benchmark applications [18]. Now the aim was to propose such a method that would learn from samples to diagnose the medical data. Such method was proposed in [19]. This method could diagnose with accuracy 63.26% of liver disorder [19]. . Furthermore, FES-LD was proposed for diagnosing of liver disorders. This fuzzy system was faster, cheaper, also more responsible and more accurate than other traditional systems. In addition, on time diagnosis of sickness and nominating the ratio of liver disorders development has been experienced and its confirmation 91% [16]. Moreover, particle swarm optimization based methods came into existence to make the diagnosis systems more rigorous. CBRPSO [22] and GA-CBRPSO [23] belong to this category. These were hybrid model of case based reasoning, particle swarm optimization and genetic algorithm. Initially a case-based reasoning method is applied to preprocess the data set; therefore, a weight vector for each feature is derived, while PSO is used for clustering to diminish the effects of initial conditions and then number of clusters was reduced to two [22]. A genetic algorithm is applied for medical data classification and to assemble a decision-making system for illness identified. The average performance for liver disorders of CBRPSO is 78.18% and average predicting accuracy for breast cancer of CBRPSO model is 97.4% and for liver disorders is 76.8% [23]. GA-CBRPSO is the best method with average accuracy 76.8%.

III. PROPOSED METHODOLOGY

In our proposed diagnosis system, we used artificial neural network classification technique to categorize the overlapping and non-overlapping symptoms for pre-processing of the dataset. This classification helped to diagnose the liver disease more precisely. One of the most famous ANN classification techniques is back-propagation algorithm that gives high tolerance to noisy data, as well as their ability to classify patterns on which they have not been trained. The iterative learning process is the main characteristics of back-propagation algorithm in which weights along with their input values are adjusted each time. Structured network for a particular application is ready to be trained. To initiate this process, weights are assigned randomly, and then process is continued along with training followed by learning. In this process, structured network is trained by training data one at a time using a weight function in the hidden layers, and compares the obtained results against a want results. This training process is done again and again until unless connection weights are continually refined. Furthermore, in second stage we apply clustering technique to categorize the dataset into overlapping and non-overlapping symptoms. More formally, a support vector machine (SVM) [4] builds a hyperplane or set of hyperplanes in a high- or

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is accomplished by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier. After classification, support vector machines (SVMs) [4] are used to create classification rules in the second phase of the proposed method. In machine learning, SVMs examine the data and identify the patterns. Each instance of given training set is marked as belonging to one of the 2 categories. Trained SVMs are mapped with new test data set into the same space. It forecasts to belong to a category based on which side of the gap they fall on. Finally, SVMs surmounts the limitation of the Levenber-Marquardt algorithm (MLP), and we acquire the precision enhancement in the proposed method. Steps taken to diagnose the liver disorder are shown in Figure 3

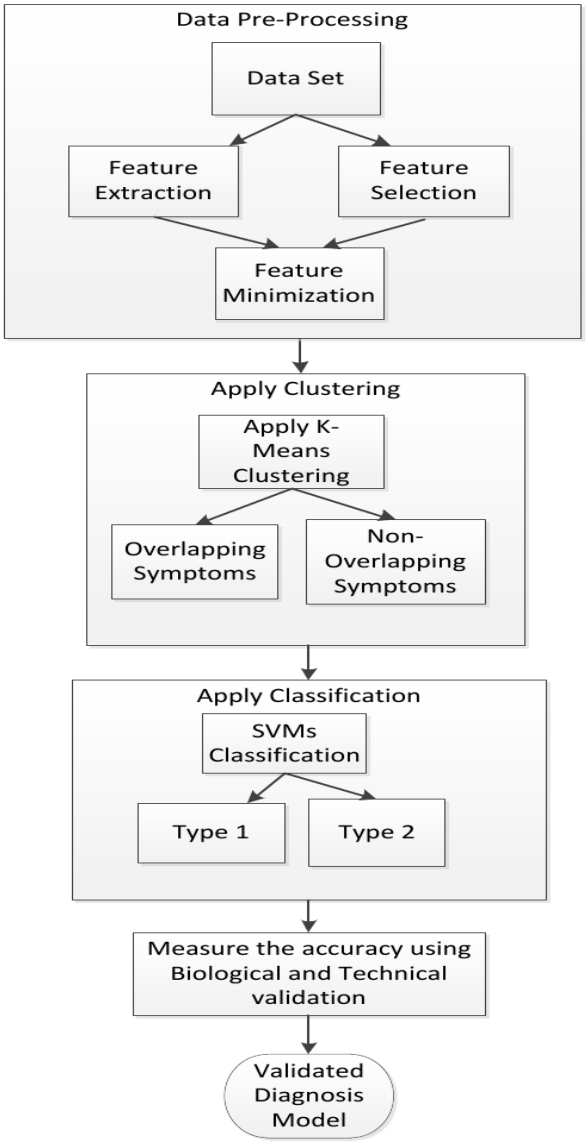

Figure 3: Steps taken to diagnose the Liver Disorder

Dataset:

In our experiment, we used 2 datasets of UCI machine learning storeroom for benchmark test: one of liver disorder, and other of hepatitis. These two datasets have overlapping symptoms. A dataset of liver disorder, namely BUPA dataset contains 345 records that are taken on seven attribute. Last attribute, namely selector is used to classify the whole data into two sets. Hepatitis is one of the types of liver disease. We used data of hepatitis for this study that is available on UCI machine learning repository [23]. The Liver Disorders data is named as Hepatitis Domain. This data is contributed by Carnegie-Mellon University that includes 155 records consisting of 18 fields and two classes. These are classified as DIE and LIVE. Each sample is taken from a single man. 85 of these samples are of DIE class with remaining 70 are possessed by to the LIVE class. First five attributes of the collected data samples are the outcomes of blood test while the last attribute includes daily alcohol consuming [18]. Liver disease data set is given in Table 1.

Table 1: Attributes of Dataset for Liver Diseases

S.No.	Hepatitis		BUPA liver disorders	
1	Class	DIE, Live	Mcv	mean corpuscular volume
2	AGE	10,20,30,40, 50,60,70,80	alkphos	alkaline phosphotase
3	SEX	male, female	Sgpt	alamine aminotransferase
4	STEROID	no, yes	Sgot	aspartate aminotransferase
5	ANTIVIRALS	no, yes	gammagt	gamma-glutamyl transpeptidase
6	FATIGUE	no, yes	drinks	Number of half-pint equivalents of alcoholic beverages drunk per day
7	MALAISE	no, yes	selector	field used to split data into two sets
8	ANOREXIA	no, yes		
9	LIVER BIG	no, yes		
10	LIVER FIRM	no, yes		
11	SPLEEN PALPABLE	no, yes		
12	SPIDERS	no, yes		
13	ASCITES	no, yes		
14	VARICES	no, yes		
15	BILIRUBIN	0.39, 0.80, 1.20, 2.00, 3.00, 4.00		
16	ALK PHOSPHATE	33, 80, 120, 160, 200, 250		
17	SGOT	13, 100, 200, 300, 400, 500		
18	ALBUMIN	2.1, 3.0, 3.8, 4.5, 5.0, 6.0		

19	*PROTIME*	*10, 20, 30, 40, 50, 60, 70, 80, 90*		
20	*HISTOLOGY*	*no, yes*		

Note: Common Attributes are alkphos and sgot that are showing the overlapping symptoms of both diseases.

IV. Result Analysis and Discussion

The proposed system was implemented using data mining techniques for pre-processing and Data prediction of liver disease. In this preprocessing, we used the ANN classification technique. We used the confusion matrix to examine the obtained experimental results. The experimental results show that ANN classification can achieve more accuracy when using more than 25 attributes in our datasets which preprocess the given dataset with overlapping symptoms. A confusion matrix (Kohavi and Provost, 1998) contains information about actual and predicted classifications done by a classification system. The performance of such systems is commonly evaluated using the data in the matrix. Figure 4 shows the confusion matrix for two class classifier. Following this matrix, table 2-5 were drawn that depict the Confusion matrix of Liver Disorder Prediction for 10, 15, 20 and 25 attributes respectively.

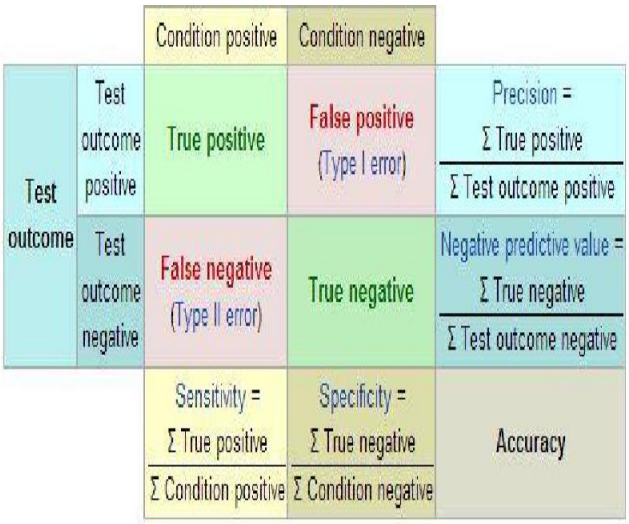

Figure 4: Confusion Matrix

Table 2: Confusion Matrix for Liver Disorder Prediction for 10 attributes

Confusion Matrix for 10 attribute of liver disorder		Predicted	
		Negative	Positive
Actual	Negative	93	0
	Positive	17	149

Table 3: Confusion Matrix for Liver Disorder Prediction for 15 attributes

Confusion Matrix for 15 attribute of liver disorder		Predicted	
		Negative	Positive
Actual	Negative	157	0
	Positive	13	266

Table 4: Confusion Matrix for Liver Disorder Prediction for 20 attributes

Confusion Matrix for 20 attribute of liver disorder		Predicted	
		Negative	Positive
Actual	Negative	184	0
	Positive	11	345

Table 5: Confusion Matrix for Liver Disorder Prediction for 25 attributes

Confusion Matrix for 25 attribute of liver disorder		Predicted	
		Negative	Positive
Actual	Negative	343	0
	Positive	5	536

We obtained accuracy 93.43 for 10 attributes. Maximum accuracy of the system is 99.43 at 25 attributes. Figure 5 shows graphical representation of accuracy against number of attributes. It has been observed that accuracy increases as number of attributes increases.

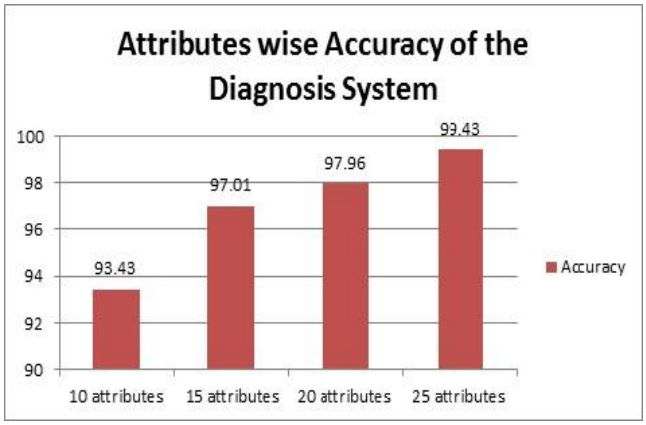

Figure 5: Attribute Wise Accuracy For Liver Disease Prediction System

V. Conclusion

Overlapping symptoms of liver disease makes the suspicious prescription for an expert that increases chances of other chronic diseases. Due to lack of standardized diagnostic measures, a suitable diagnosis remains hindered. In particular, it enables us to propose a diagnostic system that detects overlapping symptoms of the diseases. The outcome of the proposed work will be a classifier which takes in a set of symptoms to be analysed for a particular disease and the outcome would be to successfully conclude whether the sample is overlapping or non-overlapping symptoms, if diseased then the type of the disease (if any). The purpose of this work is pre analysis of a patient's symptoms for the pronouncement of a disease and its type, prior to adoption of medication under a physician's supervision. This work will certainly save precious time and effort of physicians and will bring to an end sufferings of diseased people waiting for diagnosis.

References

[1] Lin, R.H.: An intelligent model for liver disease diagnosis. Artificial

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Intelligence in Medicine 47(1), 53–62 (2009)
[2] Lippincott Williams & Wilkins by Schiff Eugene R.; Sorrell, M.F.M.W.C.: Schiff's Diseases of the Liver, 11th edition edn. Wiley-Blackwell (2011)
[3] British liver trust, (2009) [online]. available:http://www.britishlivertrust.org.uk/media-centre/facts-about- liver-disease/
[4] Trivedi, P.J., Hirschfield., G.M.: Review article: overlap syndromes and autoimmune liver disease. Alimentary pharmacology & therapeutics 36(6), 517–533 (2012)
[5] De, S.K., Biswas, R., Roy, A.R.: An application of intuitionistic fuzzy sets in medical diagnosis. Fuzzy Sets and Systems 117(2), 209–213 (2001)
[6] Nayak, R., Jain, L., Ting, B.: Artificial neural networks in biomedical engineering: a review. In: Asia-Pacific Conference on Advance Computation (2001). URL http://eprints.qut.edu.au/1480/
[7] Perantonis, S., Virvilis, V.: Dimensionality reduction using a novel neural network based feature extraction method. In: Neural Networks, 1999. IJCNN'99. International Joint Conference on, vol. 2, pp. 1195–1198. IEEE (1999)
[8] Tarr, G.L.: Multi-layered feedforward neural networks for image segmentation. Tech. rep., DTIC Document (1991)
[9] Machine-learning-databases (last accessed: 7April 2005)
[10] Forsyth, R.S.: Bupa liver disorders dataset:UCI machine learning repository (last accessed: 10 july 2014).URL https://archive.ics.uci.edu/ml/datasets/Liver+Disorders
[11] Lin, R.H.: An intelligent model for liver disease diagnosis. Artificial Intelligence in Medicine 47, 53–62 (2009)
[12] Liang, F.: Annealing stochastic approximation monte carlo algorithm for neural network training. Machine Learning 68(3), 201–233 (2007)
[13] Rocha, M., Cortez, P., Neves, J.: Evolution of neural networks for classification and regression. Neurocom-puting 70(16), 2809–2816 (2007)
[14] Liao, Y.: Neural networks for pattern classification and universal approximation (2002)
[15] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20.3 (1995): 273-297.
[16] Neshat, M., Yaghobi, M., Naghibi, M., Esmaelzadeh, A.: Fuzzy expert system design for diagnosis of liver disorders. In: Knowledge Acquisition and Modeling, 2008. KAM'08. International Symposium on, pp. 252–256. IEEE (2008)
[17] Karthik, S., Priyadarishini, A., Anuradha, J., Tripathy, B.: Classification and rule extraction using rough set for diagnosis of liver disease and its types. Adv. Appl. Sci. Res 2(3), 334–345 (2011)
[18] Chaves, A., Vellasco, M.M.B.R., Tanscheit, R.: Fuzzy rule extraction from support vector machines. In: Hybrid Intelligent Systems, 2005. HIS '05. Fifth International Conference on (2005)
[19] Shevked, Z., Dakovski, L.: Learning and classification with prime implicants applied to medical data diagnosis. In: Proceedings of the 2007 international conference on Computer systems and technologies, p. 103. ACM (2007)
[20] Steimann, F.: On the use and usefulness of fuzzy sets in medical ai. Artificial Intelligence in Medicine 21(1), 131–137 (2001)
[21] Friedrich Steimann, K.P.A.: A fuzzy medical data model. In: Proceedings of the 12th European Meeting on Cybernetics and Systems Research(World Scientific, Singapore,, vol. 1, p. 271278 (1994)
[22] Lin, J.J., Chang, P.C.: A particle swarm optimization based classifier for liver disorders classification. In: Computational Problem-Solving (ICCP), 2010 International Conference on, pp. 63–65. IEEE (2010)
[23] Chang, P.C., Lin, J.J., Liu, C.H.: An attribute weight assignment and particle swarm optimization algorithm for medical database classifications. Computer methods and programs in biomedicine 107(3), 382–392 (2012)

Prof. First Author P**rofile** which contains their education details, their publications, research work, membership, achievements, with photo that will be maximum 200-400 words.

Dr. Second Author P**rofile** which contains their education details, their publications, research work, membership, achievements, with photo that will be maximum 200-400 words.

Prof. Third Author P**rofile** which contains their education details, their publications, research work, membership, achievements, with photo that will be maximum 200-400 words.