https://doi.org/10.46344/JBINO.2022.v11i03.09

KIDNEY AND LIVER DISEASE PREDICTION USING HYBRID NEURAL NETWORK ALGORITHM

Dr. S. Vijayarani[1], Dr. C.Sivamathi[2] , Mr.S.Dhayanand[3]

[1]Assistant Professor, [2] [1]Assistant Professor, [2] M.Phil Research Scholar,
[1,3] Department of Computer Science, Bharathiar University, Coimbatore, Tamilnadu
[2]Department of Computer Science, PSG College of Arts & Science, Coimbatore, Tamilnadu

ABSTRACT

Data mining applications in healthcare sector becomes very essential and significant and it provides benefit to all parties those who are involved in the healthcare domain. The massive amounts of data generated by healthcare transactions are too complex and voluminous and these are processed and analyzed by traditional methods. Data mining methodologies have transformed these mounds of data into useful information for accurate decision making. The primary objective of this research work is to predict kidney and liver diseases by using neural network algorithms. In liver disease prediction, the diseases like CBCL and acute hepatitis diseases are predicted. In kidney disease prediction, Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure and Chronic Glomerulonephritis diseases are predicted. In this research work, a new hybrid neural network algorithm ANFIS+RBF is proposed which is a combination of Adaptive Neuro-Fuzzy System (ANFIS) and RBF (Radial Basis Function). This proposed algorithm was compared with existing Artificial Neural Networks (ANN) and ANFIS algorithms. Experiments are conducted with existing and proposed algorithms. From the experimental results, we have observed that the performance of the proposed hybrid ANFIS+RBF algorithm is more efficient than existing algorithms.

Keywords: Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, Radial Basis Function, ANFIS+RBF, Kidney diseases, Liver diseases.

1 Introduction

Data mining can be defined as the process of discovering previously unknown patterns and trends in databases and using that information to build predictive models [5]. Otherwise, it can be defined as the process of data selection and exploration and building models using vast data stores to uncover previously unknown patterns [9]. Data mining is not new, it has been used intensively and extensively by financial institutions, retail and wholesale marketers, manufacturers. Data mining techniques gives some economical beneficial information, decision making information, predictions to the organization.

An Artificial Neural Network (ANN) is a computational model, based on biological neural networks in the human brain. Artificial Neural Networks have huge applications in Machine Learning research and industry, speech recognition, computer vision and text processing, Image Processing and Character recognition, Forecasting, classification of diseases etc. Basically, there are 3 different layers in an artificial neural networks. They are:

i. Input Layer - In this layer, all the inputs are fed in the model.
ii. Hidden Layers – In this layer, entire processing steps are carried out. There can be any number of hidden layers, based on the application.
iii. Output Layer – In this layer, the output are received, after processing.

An adaptive neuro-fuzzy inference system (ANFIS) is one of artificial neural network technology. It is based on Takagi–Sugeno fuzzy inference system. It integrates both artificial neural networks and fuzzy logic principles and thereby it combines the benefits of both techniques in a single framework. The inference system of ANFIS is similar to a set of fuzzy IF–THEN rules. Usually, ANFIS is referred as a universal estimator. ANFIS can be used in learning part of neural networks. It uses Fuzzy Logic to process the input through hidden layers. Thus ANFIS works by applying Neural Network learning methods and tune the parameters of a Fuzzy Inference System (FIS). To illustrate the ANFIS architecture, consider two fuzzy IF-THEN rules based on a first order Sugeno model is given in following rules.

Rule 1: If x is A_1 and y is B_1 , Then $f_1= p_1x + q_1y + r_1$.

Rule 2: If x is A_2 and y is B_2 , Then $f_2= p_2x + q_2y + r_2$.

Here x, y are inputs, A, B are fuzzy sets, f is fuzy region, p,q,r are parameters.

The main objective of this research work is to predict kidney and liver diseases using neural network algorithms. This work proposes a hybrid algorithm called ANFISRBF, based on Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). This work aims to predict kindey diseases like Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure and Chronic Glomerulonephritis. Also the work predicts liver diseases such as Cirrhosis, Bile Duct, Chronic Hepatitis, Liver Cancer and Acute Hepatitis.

The remaining portion of the paper is organized as follows. Related works are discussed in Section 2. The existing and

the proposed methodologies and their respective algorithms are given in Section 3. Section 4 analyzes the experimental results. Section 5 gives conclusion.

2 Literature Review

Antonio SalgadoCastillo et.al [3] evaluated an ANN and SVM designed using Matlab 7.9.0.529 for the classification and prediction of patients with diabetic neuropathy, using Pulse Waves Sequences of Blood Volume. Efficiency was evaluated taking into account the algorithms and training time as well as effectiveness in classification and prediction. From the experimental result the best classification results were obtained with the ANN with the Gradient descent learning algorithm with adaptive learning rate. Chaitrali S.

Dangare et.al [4] has analyzed prediction systems for Heart disease using more number of input attributes. The authors implemented Decision Trees, Naive Bayes, and Neural Networks classification techniques for analyzing Heart disease database. The performance of these techniques was compared, based on accuracy. Authors' analysis shows that out of these three classification models Neural Networks predicts Heart disease with highest accuracy.

Abeer Y. Al-Hyari et al [1] proposed in their research by using Artificial Neural Network, Decision Tree and Naïve Bayes to predict chronic kidney disease. From the experimental result authors concluded that the proposed ANN algorithm has produced good results. Krishnaiah. V et.al [6] discussed briefly and examined the potential use of classification based data mining techniques such as Rule based, Decision tree, Naïve Bayes and Artificial Neural Network to massive volume of healthcare data. Aim of the researchers was to propose a model for early detection and correct diagnosis of the disease which would help the doctor in saving the life of the patient.

Michael E Brier et.al [7] used an artificial neural network to predict the occurrence of DGF (Delayed graft function) and compared with traditional logistical regression models for prediction of DGF. Artificial neural networks are used for Covariate analysis and logistical regression was done in the prediction of number of occurrence of DGF. From the experimental results authors concluded that artificial neural networks may be used for prediction of DGF in cadaveric renal transplants

Ananda Kumar. K et.al [1] used a supervised machine learning algorithms to find the smallest set of genes that could confirm high accuracy in classification of cancer from micro array data. They proposed a method Fast Adaptive Neuro-Fuzzy Inference System (FANFIS) as a classification model and used Modified Levenberg-Marquardt algorithm for learning phase. From the experimental results they suggested that proposed method results with better classification accuracy and minimum time in execution when compared to existing technique Adaptive Neuro-Fuzzy Inference System (ANFIS).

Hyontai Sug [2] has suggested a method based on oversampling in minor

classes to compensate the insufficiency of data effectively in order to generate more accurate decision trees for liver disorder diseases. He has used two representative algorithms of decision trees; C4.5 and CART for predict liver disorder disease from BUPA liver disorder dataset. From the experimental results it is observed that both the algorithms has produced very good results so that they recommend oversampling for the data set to generate decision trees.

Karthik. S et.al, [3] were applied a soft computing technique for intelligent diagnosis of liver disease. They have implemented classification and its type detection in three phases. The authors classified liver diseases using Artificial Neural Network in first phase. Then in second phase, the authors implemented Learn by Example (LEM) algorithm to generate classification rules. In final phase, the types of liver diseases are identified using fuzzy rules.

Sadik Kara et.al [8] had concentrated on the diagnosis of optic nerve disease through the analysis of pattern electro retinography (PERG) signals with the help of artificial neural network (ANN). The authors used Levenberg Marquart (LM) back propagation algorithm to train Neural Network. In this work, a Multilayer feed forward Artificial Neural Network was used for classification. The classified the patients into healthy patients and diseased patients.

Rong-Ho Lin et.al [7] described a study and constructed an intelligent liver diagnosis model. Integrated artificial neural networks, analytic hierarchy process, and case-based reasoning methods and examined the patients suffered from liver disease and determined the liver disease types. Performance evaluation indicated that the AHP-weighted CBR outperformed the pure CBR in terms of accuracy cadaveric renal transplants.

3. Proposed ANFIS+RBF algorithm

ANFIS (Adaptive Neuro-Fuzzy Inference System) method is employed as a pedagogics for Sugeno-type fuzzy systems [10]. System parameters are known by the help of ANFIS. Typically the quantity and kind of fuzzy system membership functions are outlined by user once applying ANFIS. ANFIS –method could be a hybrid methodology that consists of two parts; gradient methodology and least square methodology. The gradient method performs the calculation of input membership function parameters and the least square methodology performs the calculation of output function parameters [14]. An ANFIS can be used as IF-THEN rules model, so that it can map inputs to outputs. Such prediction problem can be done using a function, in order to predict output for a given input vector. This output vector should be as close as possible to its actual output. Hence the problem is to determine which ANFIS function is effective, so that the difference between the actual output and the predicted one is minimized. ANFIS uses fuzzy rules to minimize the error rate in prediction. In this work, ANFIS was hybrid with RBF for better performance. The proposed algorithm is the technique for interpolation in multidimensional space. First the fuzzy system is initialized

into the neural networks using genfis commands. The parameters to measure the performance risk analysis are declared. Anfis command used to get the parameters set of the consequent of each rule. Before start learning, we are considering number of iterations and tolerance as learning parameters. Now train the data with ANFIS membership functions.

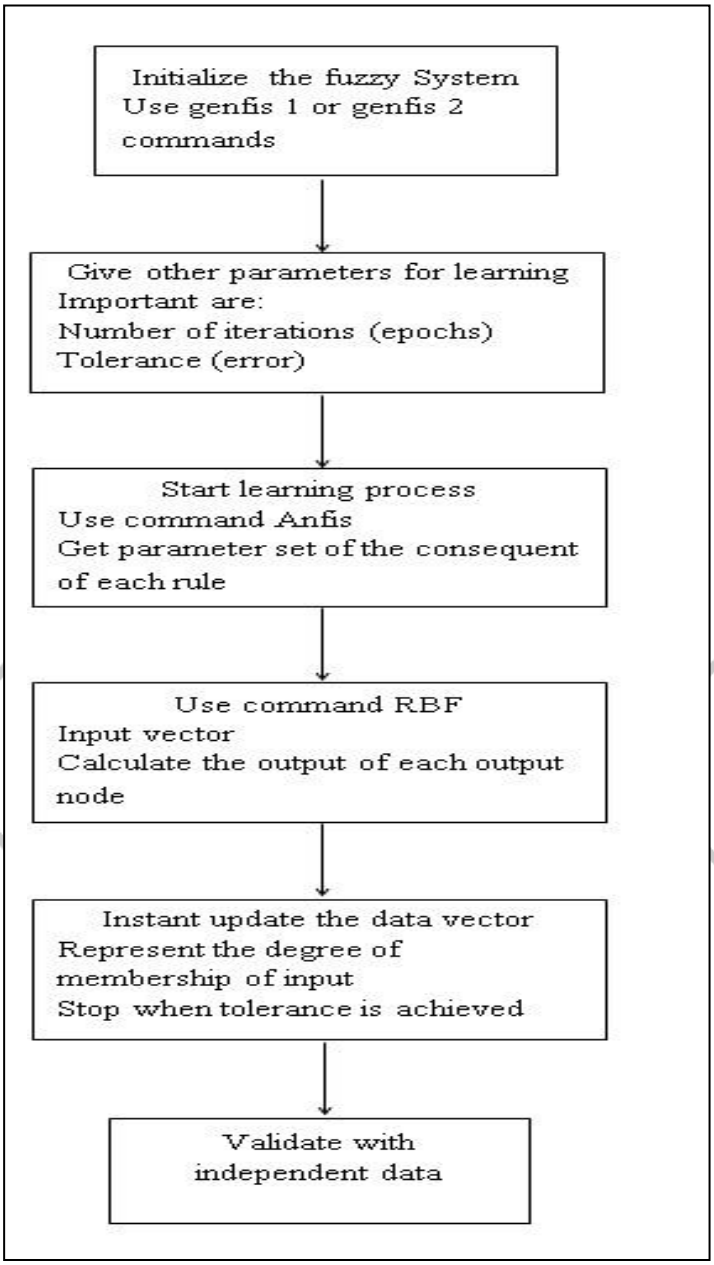

Figure 1: Flow Chart of ANFIS+RBF

Then Radial Basis Function has been applied in the area of neural networks where they may be used as a replacement for the sigmoidal hidden layer transfer characteristic in multilayer perceptron. Finally, test the input (weight) with RBF methodology to predict hidden values. For each RBF each input vector/pattern (x). Calculate the output (y) of each output node

$$y_o^p = w_{bo} + \sum_{h=1}^{m} w_{ho} \emptyset(x^p, R_h, \sigma_h)$$

Here w checks with the target value and predicts the disease type based on weight. Since the both ANFIS and RBF are combined together in hybrid format, the classification performances are highly improved.

4 Experimental Results

In this section, the algorithms ANN, ANFIS and proposed ANFIS+RBF algorithms are implemented and their performance are analyzed. This work is implemented in Matlab 2013 tool. MATLAB (MATrix LABoratory) is a multi-paradigm numerical computing environment. It is high performance language that integrates computation, visualization, and programming in a single framework. The performance analysis measures used are: Accuracy Measure, Correctly Classified Instances, Incorrectly Classified Instances, TP Rate, Accuracy, Precision, F-measure and Execution time.All the above performance measures are analyzed for both kidney and liver dataset.

Kidney Dataset: For predicting kidney disease, several Patient data are collected from various laboratories, medical centres and hospitals. Using these data the synthetic kidney function test (KFT) dataset have been created for analysis of kidney disease. The dataset contains five hundred and eighty four instances and six attributes namely Age, Gender, Urea, Creatinine and Glomerular Filtration Rate (GFR). Here age and gender are patient personal details. Creatinine is a: breakdown waste product of muscles. The kidneys filters these waste and pass out through urine. If kidney does not function properly then , these Creatinine remain in body and results in chronic kidney diseases. GFR is a best measure to test the function of kidney. It is also used to find the stage of kidney disease.

Liver Dataset: For predicting lever disease, Indian Liver Patient Dataset (ILPD) taken from the UCI Repository are used in the experiment. The data set was collected from north east of Andhra Pradesh, India. This dataset has five hundred and seventy six instances and ten attributes namely Age, Gender, Total Bilirubin (TB), Direct Bilirubin (DB), Alkaline Phosphotase (ALP), Sgpt Alamine Aminotransferase, Sgot Aspartate Aminotransferase, Total Protiens (TP), Albumin (ALB) and Albumin and Globulin Ratio (A/G Ratio). Age and gender are personal details of patients. Bilirubin is a yellowish orange pigment that arises normally when red blood cells break down. TB and DB is bilirubin measure to predict liver disease. ALP represents the amount of alkaline phosphatase enzyme in the bloodstream. If the level of ALP is high, then it means the there is some disease or fault in liver. Sgpt and sgot are enzymes mixed into blood when liver damages. Albumin and Globulin are protein levels, that are used to predict dehydration of patients. TP, ALB and A/G ration measures are used to predict dehydration in patients. Table 1 represents the performance measures of algorithms in both liver and kidney datasets.

Table 1. Performance measures of algorithms in both liver and kidney datasets.

Dataset	Algorithms	TP Rate	Precision	F Measure	Correctly Classified Instances (%)	In Correctly Classified Instances (%)	Execution time in MS
Kidney Dataset	**ANN**	0.877	0.9	0.295	87.7	12.3	918
	ANFIS	0.908	0.942	0.374	90.82	9.18	486
	ANFIS+RBF	0.95	0.894	0.502	95	5	310
Liver Dataset	**ANN**	0.83	0.799	0.432	83	17	719
	ANFIS	0.91	0.855	0.491	91	9	474
	ANFIS+RBF	0.95	0.894	0.502	95	5	310

Based on liver disease measures, TB, DB, ALP, Sgpt, Sgot, TP, ALB and A/G Ratio, liver diseases like Child Behaviour Checklist (CBCL), Acute Hepatitis are predicted. This is shown in table 2. From the results shown in figure 3, it was found that the proposed ANFIS +RBF algorithm could predict with same number of patients as in existing algorithm. The analyses is further extended, based on gender. Table 4 shows the prediction of disease based on gender i.e for male and female.

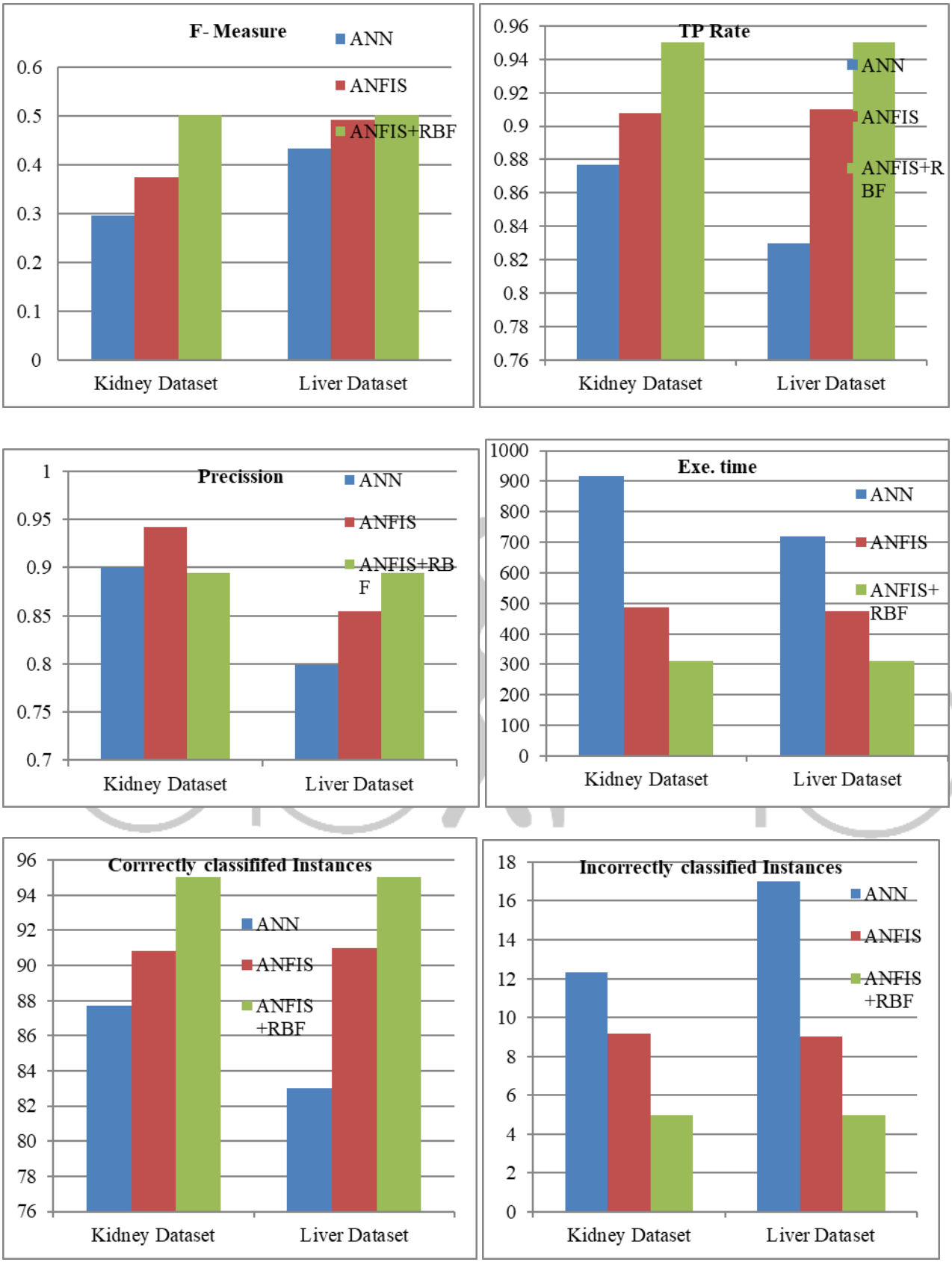

Figure 2. Performance measures of algorithms

Table 2. Prediction of Liver Diseases

Liver diseases	ANN	ANFIS	ANFIS +RBF
Normal	363	361	361
CBCL	171	173	173
Acute Hepatitis	24	24	24
Outliers	18	18	18

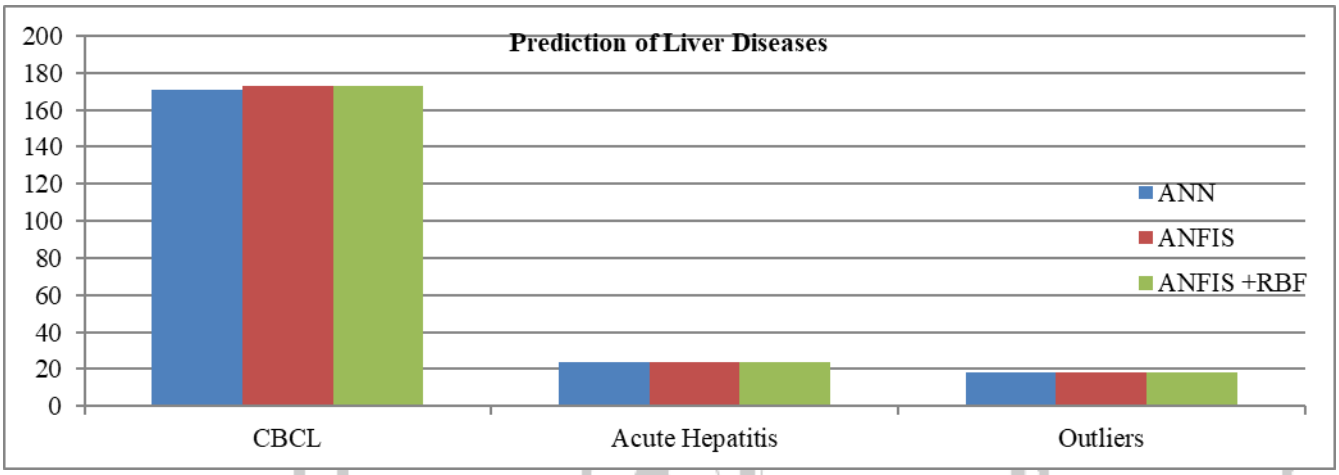

Figure 3. Prediction of Liver Diseases

Table 3. Prediction of Liver Patients in Gender wise

Algorithms	Normal		CBCL		Acute Hepatitis		Outliers	
	Male	Female	Male	Female	Male	Female	Male	Female
ANN	271	90	143	30	21	3	15	3
ANFIS	271	92	138	33	19	5	15	3
ANFIS+RBF	265	96	137	36	19	5	15	3

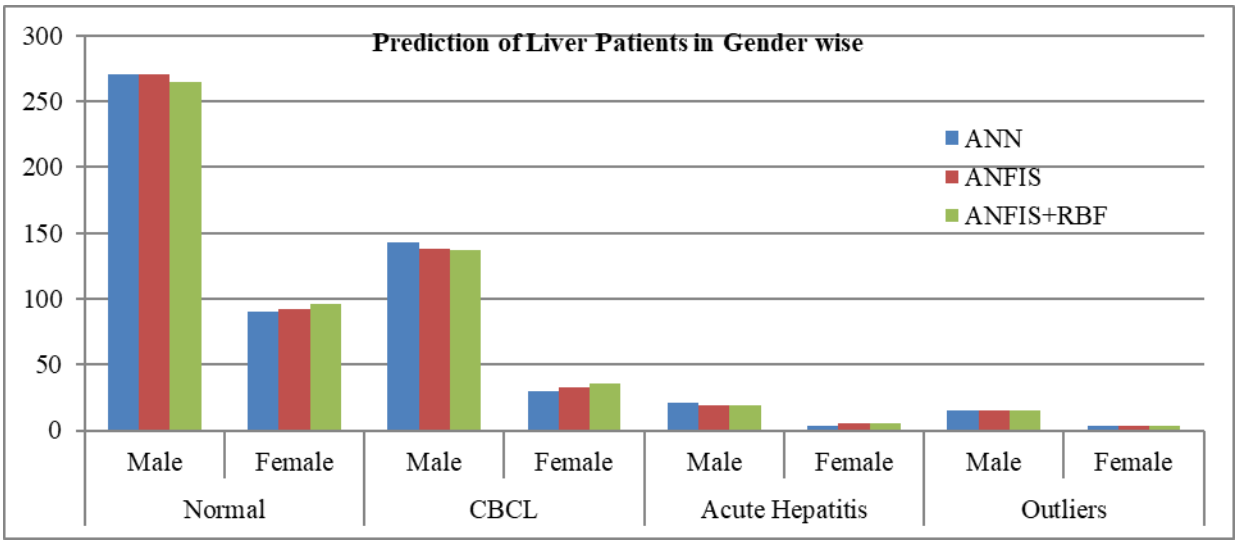

Figure 4. Prediction of Liver Patients by Gender Wise

Similarly, in kidney dataset, the diseases like Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure and Chronic Glomerulonephritis are also predicted. This is shown in table 3. From the results shown in figure 5, it was found that the proposed ANFIS +RBF algorithm works better in prediction than the existing ones.

Table 4. Prediction of Kidney Diseases

Kidney diseases	ANN	ANFIS	ANFIS +RBF
Normal	439	443	444
Acute Nephritic Syndrome	43	44	44
Chronic Kidney disease	42	40	40
Acute Renal Failure	16	15	15
Chronic Glomerulonephritis	43	41	40

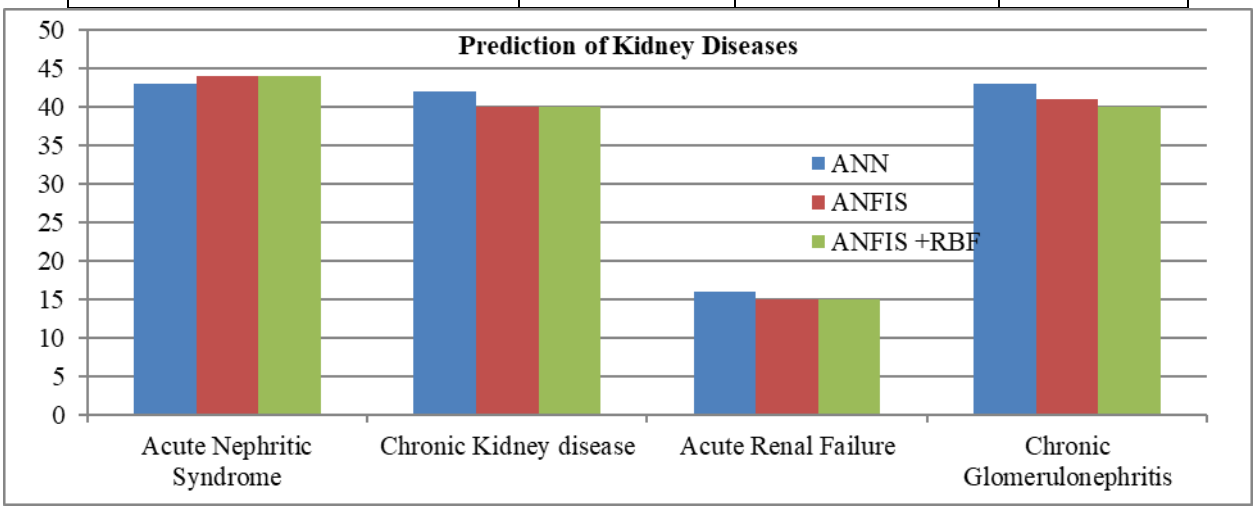

Figure 5. Prediction of Kidney Diseases

Conclusion:

Data mining methodologies have transformed the mounds of data into useful information for accurate predictions. Hence in this research work, a hybrid classification algorithm was proposed to predict kidney and liver diseases. In liver disease prediction, CBCL and acute hepatitis are predicted. In kidney disease prediction, Acute Nephritic Syndrome, Chronic Kidney disease, Acute Renal Failure and Chronic Glomerulonephritis diseases are predicted. The proposed algorithm is a combination of ANFIS and RBF. The performance of these algorithms are analyzed, by various performance measures like Accuracy Measure, Correctly Classified Instances, Incorrectly Classified Instances, TP Rate, Accuracy, Precision, F-measure and Execution time. From the experimental results, it was observed that the performance of the proposed hybrid algorithm is more efficient than existing algorithms. The proposed algorithm has higher accuracy with the minimum execution time. Hence it is concluded that the proposed algorithm is a best classifier compared with other neural network algorithms.

References:

1. Abeer Y. Al-Hyari," Chronic Kidney Disease Prediction System Using Classifying Data Mining Techniques", library of university of Jordan, 2012.
2. Amit Kumar Sharma, Sudesh Kumar Prajapat, Mohammed Aslam, A Comparative Study between Naïve Bayes and Neural Network (MLP) Classifier for Spam Email Detection, International Journal of Computer Applications, National Seminar on Recent Advances in Wireless Networks and Communications, NWNC-2014, page no 12-16
3. Antonio SalgadoCastillo, Tahimy González Rubio, Evaluation of ANN and SVM for the classification and prediction of patients with diabetic neuropathy, International Journal of Computer Science and Information Technologies, Vol. 3 , 2012, page no 54 – 57
4. Chaitrali S. Dangare, Sulabha S. Apte, International Journal of Computer Applications (0975 – 888), Volume 47– No.10, June 2012, page no 44-48
5. Kincade, K. (1998). Data mining: digging for healthcare gold. Insurance & Technology, 23(2), IM2-IM7.
6. Krishnaiah. V, Narsimha. G, Subhash Chandra. N, Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques, International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, page no 39 – 45
7. Michael E. Brier, Prasun C. Ray, Jon B. Klein, Prediction of delayed renal allograft function using an artificial neural network, Nephrol Dial Transplant (2003) 18: page no 2655–2659
8. Milley, A. (2000). Healthcare and data mining. Health Management Technology, 21(8), 44-47.
9. Neha Sharma, Hari Om, Extracting Significant Patterns for Oral Cancer Detection Using Apriori Algorithm,

Intelligent Information Management, 2014, 6, page no 30-37

10. Suman Bala, Krishan Kumar, A Literature Review on Kidney Disease Prediction Using Data Mining Classification Technique, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.7, July- 2014, pg. 960-967

11. Sumathy, Mythili, Praveen Kumar, Jishnujit T M, Ranjith Kumar. K, Diagnosis of Diabetes Mellitus based on Risk Factors, International Journal of Computer Applications (0975 – 8887), Volume 10– No.4, November 2010, page no 1-4

12. Negar Ziasabounchi, Iman Askerzade, ANFIS Based Classification Model for Heart Disease Prediction, International Journal of Electrical & Computer Sciences, Vol:14 No:02, April 2014 page no 7-12

13. Taksala Devapriya A, Deepesh B, Development of an ANFIS Neural Network System for Early Diagnosis of Critical Health Condition of a Patient, IOSR Journal of Electronics and Communication Engineering, Volume 4, Issue 6 (Jan. - Feb. 2013), Page no 57-63

14. Veenu Mangat, Dimple Rani, Design of an efficient data classification method for disease diagnosis, International Journal of Engineering, Business and Enterprise Applications, Dec., 2012-February, 2013, page no. 100-104

15. Milan Kumari, Sunila Godara, Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, International Journal of Computer Science and Technology, Vol. 2, Issue 2, June 2011, page no 304-308

16. Ananda Kumar. K, Punithavalli. M, Efficient Cancer Classification using Fast Adaptive Neuro-Fuzzy Inference System (FANFIS) based on Statistical Techniques, International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence, page no 132-137

17. Hyontai Sug, Improving the Prediction Accuracy of Liver Disorder Disease with Oversampling, Applied Mathematics in Electrical and Computer Engineering, ISBN 978-1-61804-064-0, page no 331-335

18. Karthik. S, Priyadarishini. A, Anuradha. J and Tripathi. B. K, Classification and Rule Extraction using Rough Set for Diagnosis of Liver Disease and its Types, Advances in Applied Science Research, 2011, 2 (3): page no 334-345

19. Liver Anatomy and Functions, innerbody.com copyright (c) 1999 - 2015 howtomedia, inc. all rights reserved http://www.innerbody.com/image_digeov/card10-new2.html

20. Rong-Ho Lin, Chun-Ling Chuang. A hybrid diagnosis model for determining the types of the liver disease, Elsevier publications, Computers in Biology and Medicine 40 (2010) , page no 665–670

21. Sadie Kara, Aysegul Given, Ayse OztUrk Onerc, "Utilization of artificial neural networks in the diagnosis of optic nerve diseases", Elsevier Publication, Computers in Biology and Medicine 36, 2006 page no 428–437

22. Salmalian. K, Nariman-Zadeh. N, Haftchenari. H, Jamali. A, Design of

hybrid genetic- SVD ANFIS networks for fatigue life modelling of GRP composites, Springer Link, Fuzzy Information and Engineering, September 2009, Volume 1, Issue 3, pp 271-287

23. Yashika Birdi, Tanjyot Aurora , Parul Arora, Study of Artificial Neural Networks and Neural Implants, International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 1 Issue: 4, page no 258-262