International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 10, No.1 (Aug-2021)

http://dx.doi.org/10.12785/ijcds/100171

Two-Stage CBR Based Healthcare Model to Diagnose Liver Disease

Seema Sharma[1] and Deepti Mehrotra[1]

[1] *Amity University Uttar Pradesh, Noida, India*

Received 16 September. 2020, Revised 28 February. 2021, Accepted 27 February. 2021, Published 5Aug. 2021

Abstract: In India, liver disease is one of the major cause of deaths in all age groups and moreover its resistance to early detection is more disturbing fact. It is therefore important to have a decision support system for early prediction of liver disease so that proper medical assistance can be given to patients. A number of researches have been done to address this area. However, the current literature fails to offer a system, which can make such similar kind of early prediction. Hence, this research proposes a Two-stage CBR (TSCBR) model to enhance the accuracy of prediction of liver disease. Case Based Reasoning (CBR) is added to 5 different classification models including neural network (NN), decision tree based Random Forest (RF), logistic regression (LR), Support Vector Machine (SVM) and naive Bayes (NB) to enhance the improvement on the correct diagnosis of disease. The main focus of this research is to reduce the number of cases that are wrongly classified as not having liver disease. Every new case passes through two stages of proposed model to see whether the new case belongs to liver disease or not. To analyze the accuracy of the proposed research two stage CBR model is compared with already existing models to diagnose liver disease without CBR. 10-fold cross-validation approach is used to reduce biasness and to analyze the performance of proposed system. It was observed that naive Bayes with CBR model shows the highest accuracy of 0.83 among all the existing compared classification technique. The proposed technique decreases the number of false negative cases. Thus, resulting in enhanced number of early detections of liver cases. It has been observed that the accuracy of TSCBR is better than the existing system.

Keywords: Early Prediction Of Liver Diseases, Two Stage CBR Model, Addition Of CBR To Naïve Baye, 10-fold Cross-Validation Approach, Case Based Reasoning;

I. INTRODUCTION

Liver is among one of the critical and largest solid organs in the human body. It performs multiple essential functions in the body. Liver helps in metabolism, produces bile that is required for digestion and makes proteins, which is required for blood clotting. Also, it helps in detoxification and decomposition of red blood cells [1]. Diseases related to Liver results in Jaundice, Chronic fatigue and swelling in legs etc. Liver diseases normally occur due to infection caused by viruses, bacteria or fungi, due to inflammation, damaged liver cell, inherited liver disorders and due to toxin chemical exposures like alcohol, drug overdose and autoimmunity. The patients infected from liver disease are susceptible to other like hepatitis or serious lifelong illness like cirrhosis. According to World Health Organization, in India the 10th most common reason of death is liver related disease. It affects one in every 5th Indian. Worldwide leading cause of death is Liver Cirrhosis is at 14th and the liver cancerous tumor is 2nd [2].

Liver is able to function in partially damaged state, thus resulting early detection of liver diseases is difficult. Symptoms of liver disease rarely appear in early stages. Therefore, there is need for a system that can detect liver diseases in early stages. Proper diagnoses of liver disease at early stage require expertise of a doctor. Detection is also impacted by the other factors like physical and emotional state of the doctor or clinic workload pressure. Henceforth, the medical expert system is very much needed, to take decisions in the same way as humans do and consistently.

Case-Based Reasoning is a problem-solving methodology based upon cognitive model on how the human beings resolve the day-to-day problems. In day-to-day life, we are often face problems, in few occasions we have already experienced them and we most likely have solution to problem from earlier experience. Similarly, CBR suggests the probable solution to a problem with the utilization of previous experiences. It uses the existing solution of past similar cases in order to get the solution of

E-mail: seema.modgil@gmail.com

encountered issues. Furthermore, CBR can be viewed as process model that has 4 phases as shown in Figure1.

The CBR cycle [3], shown in Fig.1 involves Retrieve, Reuse, Revise and Retain phases. In Retrieval phase, similarity function is used to retrieve alike cases from the knowledge base to resolve the new problem in hand. Case base is the knowledge base and lies in the heart of CBR. It stores all the previous cases of CBR. The Reuse phase of CBR reuses the cases extracted in the retrieval phase to find out the solution of new confronted problem. Retrieved solutions from previous phase are used directly to solve the new problem if the retrieved case is exactly same to the new one. But if the retrieved cases are not same but are similar then their solutions are modified to fit the new problem. The solution suggested in previous phase is now tested on real scenario in Revise phase to check on the quality of solution. Feedback is further taken to revise the solution. The main objective of Retain phase is to keep cases in the "case base" and preserve knowledge. If the new solved problem is considerably different from the previous stored cases, then the problem and its suggested solution are kept in knowledge base for future reasoning. An expert system can be built based upon the CBR. It's a progressive system, with each time a problem is resolved a new experience in the form of case is retained in system for future references. Increase in number of cases in case base will result in performance bottleneck. Extensive research is ongoing for the maintenance of case bases [10].

CBR is widely accepted in diverse fields, especially in medical diagnosis. In this study, we propose a CBR based system to provide assistance to practitioners in expediting and verifying early diagnosis of liver disease by lowering uncertainties and potential risk factors.

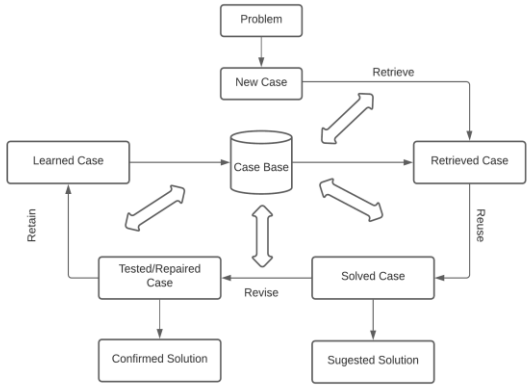

Figure 1. The CBR Cycle

The further objective is to avoid the biasness with the use of 10-fold cross-validation approach and assess the performance of proposed system. The paper has been structured as follows: Introduction is in section I, Related work is highlighted in section II, Section III explains the system design and Section IV includes an Assessment Platform to describe the data collection & preprocessing of collected data to make suitable input format. The Result and Discussion is further presented in Section V. Finally, in section VI conclusion is drawn for the proposed study.

II. Related Work

A number of researches have been conducted in this area [6,36]. Few most concerned and related works are presented in this section. The entire study in this section has been presented into two parts, Data Mining Models in liver disease diagnosis and CBR in medicine. Each part is further discussed in following sub-sections.

A. Data Mining Models in liver disease diagnosis

In past few decades, medical history of patients has been recorded electronically for quick reference and analysis. Therefore, many data mining and statistical models have been used to analyze large amount of data such as Bayesian network (BN) [11], classification and regression tree (CART), case-based reasoning (CBR) [4], support vector machines (SVM), artificial neural network (ANN), genetic algorithms, parametric and non-parametric statistics method. Ashley Spann et al. [12] have conducted a detailed literature survey on use of various machine learning techniques to diagnose liver disease and liver transplantation. Few among them have often been used for liver disease diagnosis & prognosis and are covered in the subsequent section.

An Artificial Neural Network (ANN) is a data processing approach, which is motivated by the working of human biological nervous systems for processing information. It is a collection of large number of highly interrelated processing nodes that are called as neurons. They work collectively to resolve a particular problem. Just the way human brain does, ANN learns by example. ANN has been widely used in medical diagnosis applications as it dynamically adapts itself according to the training dataset. For liver diagnosis, ANNs have been extensively used by Rong-yun Mai et al. [13] to diagnose liver cirrhosis with Hepatocellular Carcinoma, by Z. Xu to diagnose stage of cirrhosis[14] , by Alexandros Arjmand et al. [15] to find out alterations in liver in Histopathology images , by D Santhosh Reddy et al. [16] to diagnose and to categorize the ultrasound liver parenchyma texture into four different classes, by Cheng-Hsiung Wenga [19] to diagnose liver disease , Kawaguchi, T., Tokushige, K., Hyogo, H. et al. [8] have done data mining analysis on the diagnosis of nonalcoholic fatty liver disease-related hepatocellular carcinoma and concluded that patients treated with hepatectomy and a serum albumin level ≥3.7 g/dLG in Japan have more survival rate. ZhenjieYao et al. [17] have used densely deep learning model to detect liver disease. Authors have concluded using deep learning model has better performance over logistic regression and random forest. Abdar M. et al. proposed the model for

diagnosis of liver disease using multilayer perceptron neural network and boosted decision tree. Author has observed that performance of B-C5.0 is better that B-classification and regression tree (B-CART) and B-Chi square Automatic interaction detector (B-CHAID) [18].

Random Forest (RF) is a decision tree based learning algorithm, which builds multiple decision trees and then combines them all to have more accurate predictions. RF has been used widely in many medical diagnosis applications. To diagnose liver disease RF have been used by Chieh Chen Wu et.al [20] to predict fatty liver disease, by M. Aiswarya et al., M. R. Haque et al. and Nazmun Nahar [21-23] to diagnose liver disease and by S. Aman and B. Pandey [24] to build intelligent system which successfully classified Cholelithiasis, Liver Cirrhosis , Alcoholic Liver Damage and Primary Hepatoma.

machine to find out cirrhotic patients with or without minimal hepatic encephalopathy and achieved accuracy of 83%, by J.S. Sartakhti et al. [28] proposed a hybrid model, which is based upon non-linear SVM and simulated annealing (SVM-SA) for the hepatitis diagnosis.

B. CBR in medicine

CBR can be appropriate for any field where confronted problem can be solved using the previous analogous problem and its solution. It can be used for variety of problems like for classification, prediction, diagnosis, configuration and planning. From the literature it has been observed that CBR is successfully applied to medical applications too. CBR utilizes the similar method for diagnosis as the practitioner does for seeking an appropriate medical solution to the existing new problem. L. Gierl and R. Schmidt [29], I. Bichindaritz and C.

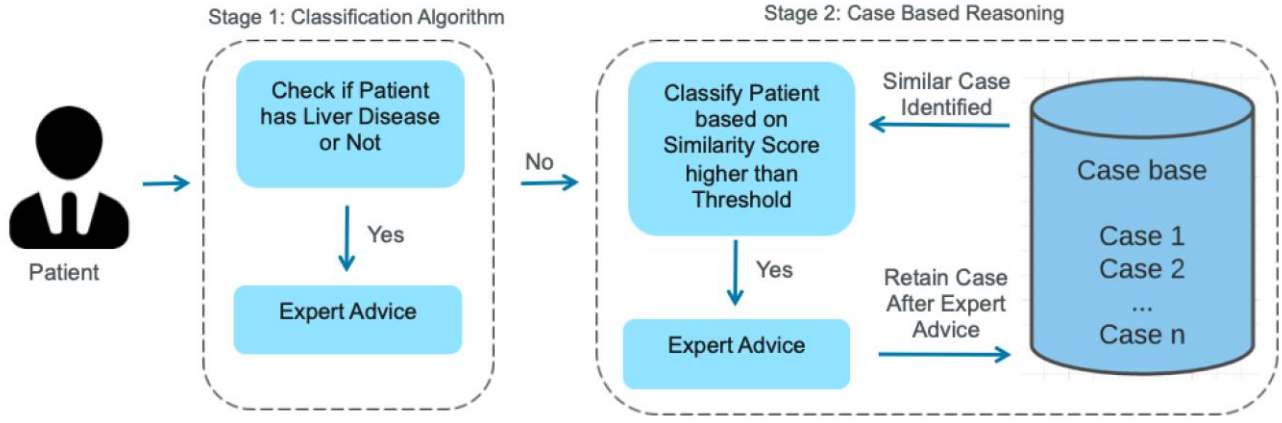

Figure 2 . Schematic diagram of two stage CBR Based health care model to Diagnose Liver Disease

Naïve Bayes (NB) performs the classification based on the Bayes theorem. NB has been used by A. Krishna et al. [25] for the classification of liver tumor and by Shapla Rani Ghosh and Sajjad Waheed [26] for diagnosis of liver disease., S. Aman [9] et al. has studied the usage of linear, nonlinear and decision tree-based classification algorithms to diagnose liver disease disorder. In their study authors have successfully concluded that use of CART algorithm to diagnosis liver disease has more accuracy rates as compare to LDA, DLDA, QDA, DQDA, NB, ANN.

Support Vector Machine (SVM) is a non-linear supervised learning algorithm. SVM makes hyper-planes in m dimensional space that divides instances of different class labels for classification where m is the total number of features in the dataset. Ali, L., Wajahat, I., Amiri Golilarz, N. et al [7] designed hybrid model which is based on linear discriminant analysis (LDA) for dimensionality reduction, support vector machine (SVM) for classification and genetic algorithm (GA) for SVM optimization to identify Hepatocellular carcinoma (HCC), by Chen, QF. Zou, TX., Yang, ZT. et al. [27] used support vector

Marling [30], Shahina Begum [31] did an extensive literature survey of use of CBR in medical diagnosis. Few among them are cited below. R. Vásquez-Morales et al. have proposed NN-CBR Twin model for the explanation of prediction made by neural network to identify chronic kidney disease [5]. Xiao-Ou Ping et al. [34] developed an expert system based on CBR to predict the reoccurrence of liver cancer. R. Ali et al. [33] suggested a hybrid case-based reasoning method for the recommendation of customized physical activities according to user's particular requirement and interests. On Contrary with traditional physical activity recommendation system, Author has merged case-based reasoning along with rule-based reasoning and preference-based reasoning. L. Campo et al. [32] developed a CBR based system specially designed for the usage of odonatologists for their treatment predictions. Abdelhak Mansoul and Baghdad Atmani [35] proposed a model that selects the best case out of all retrieved similar cases. Zhengxing Huang et al. [37] use CBR to predict the inpatient length of stay. Souad Guessoum et. al. [38] proposed an expert system based on CBR for the

identification of respiratory disease known as obstructive pulmonary disease.

III. System Model

reconsidered and again predicted by using CBR- added mode. The aim of study is to propose a CBR based method to reduce the number of cases those are misclassified as healthy cases. The cost of handling cases those are

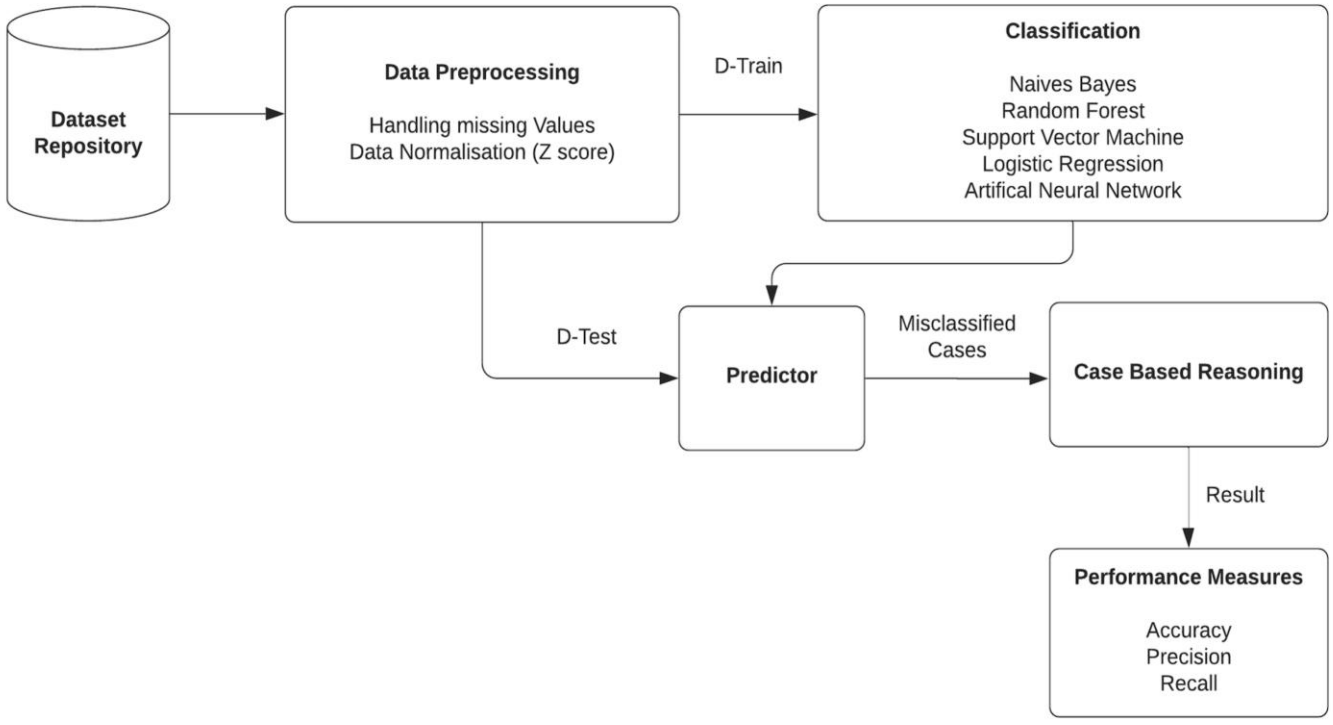

Figure 3. Flowchart of two stage Liver Diagnosis Process

The main motivation behind this study is to improve the predictions of diagnosis of patients with liver disease. The schematic diagram of the proposed model has been shown in Figure 2.The timely diagnosis of liver disease not only lowers the burden on hospitals but also controls the mortality rate that would be caused due to this disease. Correct diagnosis of any disease depends upon practitioner's experience as well as on correct analysis of clinical results. In order to assist physician to expedite diagnosis process many researchers have developed different models for different diseases. The efficiency of these models varied depending upon the data used. It has been proved by many researchers in past combining multiple classifications models help to improve accuracy of diagnosis model.

Two-stage CBR (TSCBR) based model is proposed to improve the accuracy of a liver medical diagnosis system. Flowchart of the process is shown in Figure 3.In first stage, 5 different classification algorithms are used to divide the patients into healthy (without liver disease) and unhealthy (with liver disease) category. In second stage, patients considered healthy but those actually not healthy are misclassified as healthy cases are more as compare to handling cases those are misclassified as unhealthy. Specifically, for the diseases whose early diagnosis is difficult.

IV. Assessment Platform

This section includes few important aspects of this research such as the collection procedure of the input data and the preprocessing techniques of collected data to prepare it as input data. Each of the parts is further discussed in the following subsections.

A. Collection of Dataset

For experimental work, Secondary data was accessed from free machine learning repository known as UCI (University of California at Irvine). Samples were collected from Andhra Pradesh state in India. The dataset includes total of 11 features and data from 583 patients. Among 583 patients' samples, 416 samples are from patients suffering from liver disease and rest of 167 samples are for patients without liver disease. The features

and their value/range in the dataset are listed below in Table-1.

Table1. Features used in liver dataset

Feature	Range/Value
Gender of patient (Gender)	[0 – Male & 1 – Female]
Age of patient (Age)	[4 to 90]
Direct Bilirubin (DB)	[0.1 to 19.7]
Total Bilirubin (TB)	[0.4 to 75]
Aspartate aminotransferase (Sgpt alamine)	[10 to 2000]
Almine aminotransferase (Sgpot aspartate)	[10 to 4929]
Alkaline phosphate (Alkphos)	[63 to 2110]
Albumin (ALB)	[0.9 to 5.5]
Total Proteins (TP)	[2.7 to 9.6]
Liver Disease classification field	[1-liver Disease & 2- without Liver Disease]

B. *Preprocessing of dataset*

Dataset contains few missing values. Missing values can be handled in many ways. In this study, missing numeric values in dataset are handled by replacing with the mean value. Mean value is computed by taking average of the corresponding feature values of all the samples that belong to same class. After handling missing values, dataset is normalized using *z-score* method. Z_{Score} standardized the entire attribute within the range of -1 to 1. Let f is individual value of each feature, m is mean and s is standard deviation of the feature. Then z-score value of each attribute is calculated as in equation (1) where f is observed value, m is sample mean value and s is standard deviation.

$$Zscore = \frac{f-m}{s} \qquad (1)$$

C. *Working Model*

Working Model of two-stage CBR (TSCBR) based healthcare model to diagnose liver disease has been shown in the above Figure 4 further.

After normalization, Dataset is split into two sets $Data\_Train$ and $Data\_Test$ in the 70-30 ratio respectively which is further used for training and testing of the classification algorithms respectively. In the first step of proposed model, data is analyzed using 5 different algorithms namely Random Forest (RF), non-linear algorithm support vector machine (SVM), Naive Bayes (NB), Logistic Regression (LR) and Neural network (NN). In second step, misclassified patients as healthy are reconfirmed by Retrieval only CBR. CBR contains a knowledge base of past cases that is known as "case base". In this study, a case is represented by feature vector. To

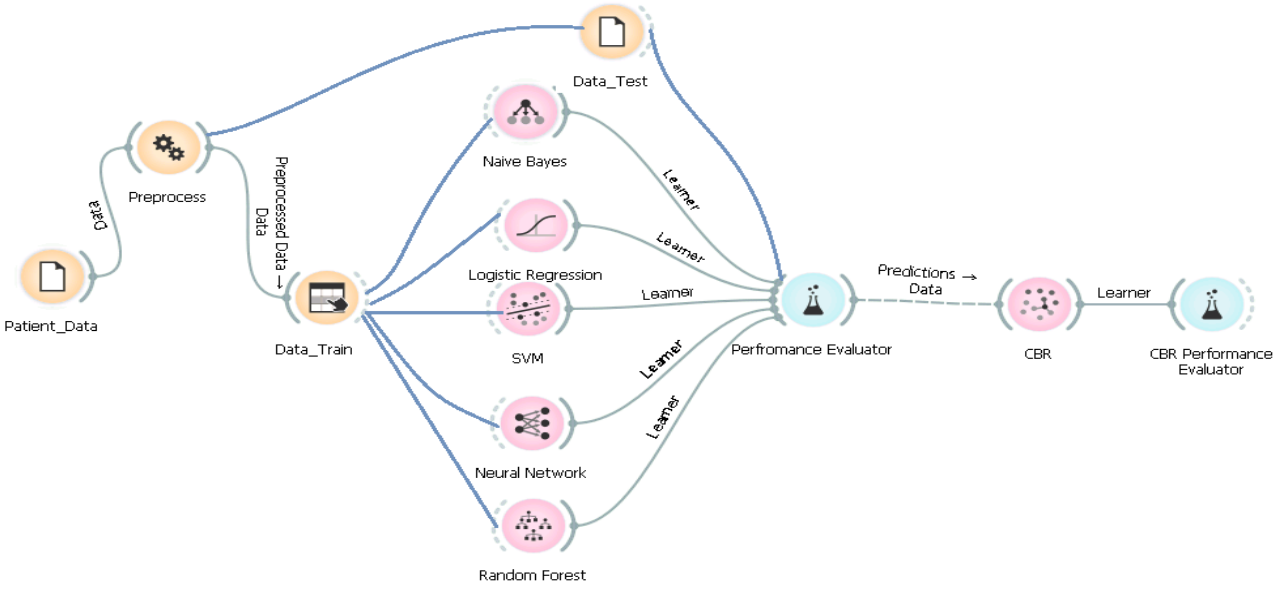

Figure 4. Working Model of two-stage CBR (TSCBR) based healthcare model to diagnose liver disease

solve a new confronted case, similar cases are extracted from case base by using a function to calculate a similarity between cases. In this model, we have used Euclidian distance function as shown in equation 2 for retrieving similar cases from CBR. Similarity Score in equation (2) represents the degree of similarity between two-cases. Euclidean distance Ed is computed by taking under root of adding all the squared differences of corresponding elements between two vectors u and v of size n.

$$Ed(u,v) = \sqrt{\sum_{i=0}^{i=n}(u_i - v_i)} \quad (2)$$

V. Results and Discussions

The study implements classification algorithms and CBR-added hybrid model to diagnose a liver disease. 10-fold cross-validation approach is applied to eliminate the biasness of models. Performances of models are evaluated on the basis of 3 measures Accuracy, Precision and Recall.

Accuracy is the measure that gives us the proportion of samples that is correctly predicted over the total number of samples. Precision gives the ratio of true positives over total number of true positive and false positive. The main aim of research is to build the model to decrease the false positive. It would cost more if the model fails to diagnose the sick person for a disease when early detection is difficult. It would be costlier than conducting more test on a healthy person. Recall gives the ratio of true positive over total number of true positive and false negatives.

Figure 5 depicts the accuracy comparison between models without using CBR and the proposed two stage CBR models. It can be observed for all cases the proposed two-stage model performs better. Results of models are further evaluated using Precision and Recall.

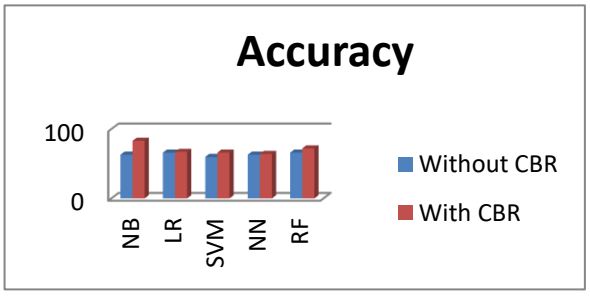

Figure 5. Accuracy comparison of different models

Precision represents the proportion of results that are significant and recall is the portion of all relevant results that are accurately classified. These two performance measures play an important role in medical CBR based system as it is very crucial to recognize ill patients so that the treatment is not delayed. Figure 6 compares the model on the precision measures.

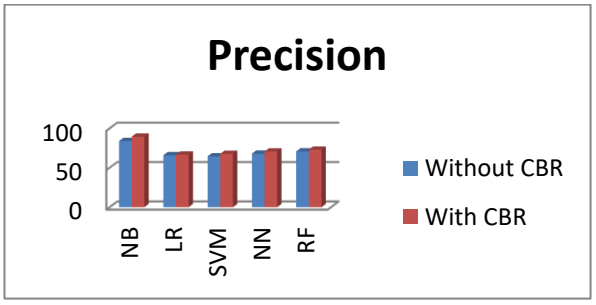

Figure 6. Precision comparison of different models

Figure. 7 display comparison of various models on the basis of recall. It can be observed that the proposed model gives better performance using precision and recall.

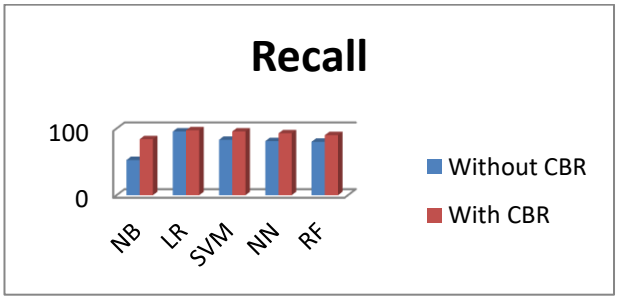

Figure 7. Recall performance measure for different classification algorithm

Table 2 compares the performance of two stage liver diagnosis model with already existing models.

Table 2. Comparison of using CBR with already existing techniques

First Phase	Accuracy	Precision	Recall	Second Phase	Accuracy	Precision	Recall
NV	63	84	52.7	NV - CBR	83	89.5	83.9
LR	66	66	95.5	LR – CBR	67	66.8	97.3
SVM	59.8	64.5	83	SVM-CBR	66	67.7	95.5
NN	63	67.9	81	NN-CBR	64	70.7	92.8
RF	66	70.8	80	RF-CBR	72	73	90

VI. Conclusion

Liver disease can be go undiagnosed in early stage due to the fact liver can work even if it is damaged partially. From the literature it has been observed that more research is needed to focus on false negative cases. Medical diagnosis system is important for early diagnosis of such diseases. Classification algorithms performance varies from one dataset to other. The prediction accuracy of classification algorithms can be improved by using hybrid models. In this paper, two stage CBR added model

(TSCBR) has been proposed to enhance the prediction of diagnosis of liver disease. Observed results have been compared by accuracy, precision and recall. Thus, the proposed technique fulfills the objective of this research work.

References

[1] https:"//www.webmd.com/digestive-disorders/liver-and-hepatic-diseases#3"

[2] https:"//timesofindia.indiatimes.com/life-style/health-fitness/health-news/is-liver-disease-the-next-major-lifestyle-disease-of-india-after-diabetes-and p/articleshow/ 58122706.cms " , 2019

[3] A. Aamodt, and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *AI Communications* 7, pp. 39-59, 1994.

[4] S. Sharma and D. Mehrotra, "Building CBR based diagnosis system using jCOLIBRI," *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Noida, 2017, pp. 634-638,doi:10.1109/CONFLUENCE.2017.7943229.

[5] G. R. Vásquez-Morales, S. M. Martínez-Monterrubio, P. Moreno-Ger and J. A. Recio-García, "Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning," in *IEEE Access*, vol. 7, pp. 152900-152910, 2019, doi: 10.1109/ACCESS.2019.2948430

[6] A. Singh, and B. Pandey, "Intelligent techniques and applications in liver disorders: a survey", *Int. J. Biomedical Engineering and Technology*, Vol. 16(1), pp.27–70, 2014.

[7] Ali, L., Wajahat, I., Amiri Golilarz, N. et al. " LDA–GA–SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine. Neural Comput & Applic (2020)", *Springer*, https://doi.org/10.1007/s00521-020-05157-2

[8] Kawaguchi, T., Tokushige, K., Hyogo, H. et al. ,"A Data Mining-based Prognostic Algorithm for NAFLD-related Hepatoma Patients: A Nationwide Study by the Japan Study Group of NAFLD," Sci Rep 8, 10434 (2018). https://doi.org/10.1038/s41598-018-28650-0

[9] A. Singh, and B.Pandey, "Liver disorder diagnosis using linear, nonlinear and decision tree classification algorithms", *International Journal of Engineering and Technology (IJET)*, Vol 8 No 5, pp. 2059-2069, 2016.

[10] Khan, M.J., Hayat, H. & Awan, I. , "Hybrid case-base maintenance approach for modeling large scale case-based reasoning systems", *Hum. Cent. Comput. Inf. Sci.* 9, 9 (2019). https://doi.org/10.1186/s13673-019-0171-z

[11] M. Imani and S. F. Ghoreishi, "Bayesian Optimization Objective-Based Experimental Design," *2020 American Control Conference (ACC)*, Denver, CO, USA, 2020, pp. 3405-3411, doi: 10.23919/ACC45564.2020.9147824.

[12] Ashley Spann, Angeline Yasodhara, Justin Kang, Kymberly Watt, Bo Wang, Anna Goldenberg, Mamatha Bhat, "Applying Machine Learning in Liver Disease and Transplantation: A Comprehensive Review", 06 January 2020 https://doi.org/10.1002/hep.31103

[13] Rong-yun Mai, Jie Zeng ,Yi-shuai Mo1, Rong Liang ,Yan Lin, Su-su Wu ,Xue-min Piao ,Xing Gao ,Guo-bin Wu , Le-qun Li , Jia-zhou Ye, "Artificial Neural Network Model for Liver Cirrhosis Diagnosis in Patients with Hepatitis B Virus-Related Hepatocellular Carcinoma", *Therapeutics and Clinical Risk Management* ,17 July 2020 Volume 2020:16 Pages 639—649

[14] Z. Xu, X. Liu, X. E. Cheng, J. L. Song and J. Q. Zhang, "Diagnosis of cirrhosis stage via deep neural network," *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Kansas City, MO, 2017, pp. 745-749, doi: 10.1109/BIBM.2017.8217748.

[15] Alexandros Arjmand , Constantinos T. Angelis , Vasileios Christou,Alexandros T. Tzallas , Markos G. Tsipouras, Evripidis Glavas , Roberta Forlano ,Pinelopi Manousou and Nikolaos Giannakeas ,"Training of Deep Convolutional Neural Networks to Identify Critical Liver Alterations in Histopathology Image Samples", *Appl. Sci. 2020*, 10, 42; doi:10.3390/app10010042

[16] D. S. Reddy, R. Bharath and P. Rajalakshmi, "Classification of Nonalcoholic Fatty Liver Texture Using Convolution Neural Networks," *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom), Ostrava*, 2018, pp. 1-5, doi: 10.1109/HealthCom.2018.8531193.

[17] Zhenjie Yao , Jiangong Li , Zhaoyu Guan , Yancheng Ye , Yixin Chen, " Liver disease screening based on densely connected deep neural networks" , *Elsevier*, Volume 123,2020,Pages 299-304,ISSN 0893-6080,https://doi.org/10.1016/j.neunet.2019.11.005.

[18] Abdar, M., Yen, N.Y. & Hung, J.CS.," Improving the Diagnosis of Liver Disease Using Multilayer Perceptron Neural Network and Boosted Decision Trees". *J. Med. Biol. Eng.* 38, 953–965 (2018). https://doi.org/10.1007/s40846-017-0360-z.

[19] W. Cheng-Hsiung, T. C.-K. Huangb, R.-P. Han, "Disease prediction with different types of neural network classifiers", *Elsevier*, 2015.

[20] C. C. Wu, W. C. Yeh, W. D. Hsu, M.M. Islam, P. A. A. Nguyen, T. N. Poly et al., "Prediction of fatty liver disease using machine learning algorithms", *Computer Methods and Programs in Biomedicine*, vol. 170, pp. 23-29, 2019.

[21] M. Aiswarya, S. Srinivas and A.G. H. Narayanan, "Illustration of Random Forest and Naïve Bayes Algorithms on Indian Liver Patient Data Set", *International Journal of Pure and Applied Mathematics*, vol. 119 (10), pp. 585-595, 2018.

[22] N. Nahar and F. Ara, "Liver disease Prediction by using Different Decision Tree Techniques", *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, vol.8 (2), 2018.

[23] M. R. Haque, M. M. Islam, H. Iqbal, M. S. Reza and M. K. Hasan, "Performance Evaluation of Random Forests and Artificial Neural Networks for the Classification of Liver Disorder", *International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Rajshahi, pp. 1-5, 2018.

[24] A. Singh, A. and B. Pandey, "A New Intelligent Medical Decision Support System Based on Enhanced Hierarchical Clustering and Random Decision Forest for the Classification of Alcoholic Liver Damage", Primary Hepatoma, Liver Cirrhosis, and Cholelithiasis, *J Healthc Eng.*; vol. 1469043, 2018.

[25] Krishna, D. Edwin and S. Hariharan, "Classification of liver tumor using SFTA based Naïve Bayes classifier and support vector machine", *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur*, pp. 1066-1070, 2017.

[26] S. R. Ghosh and S. Waheed, "Analysis of classification algorithms for liver disease diagnosis", *journalbinet.com*, vol. 05 (01), pp. 361-370, 2017

[27] Chen, QF., Zou, TX., Yang, ZT. et al. Identification of patients with and without minimal hepatic encephalopathy based on gray matter volumetry using a support vector machine learning algorithm. Sci Rep 10, 2490 (2020). https://doi.org/10.1038/s41598-020-59433-1

[28] J.S. Sartakhti, M.H. Zangooei, K. Mozafari, "Hepatitis disease diagnosis using a novel hybrid method based on support vector machineand simulated annealing (SVM-SA)", *Comput. Methods Programs Biomed*, vol. 108, pp. 570–579, 2015.

[29] S. Begum, "Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments", *IEEE Transactions of Systems, Man and Cybernetics - PART C: Applications and Reviews*", vol.41(4), 2011.

[30] Bichindaritz and C. Marling, "Case-based reasoning in the health sciences: What's next?" *Artif. Intell. Med.* 36(2), pp.127–135, 2006.

[31] L. Gierl and R. Schmidt, "CBR in medicine, in Case-Based Reasoning Technology", *From Foundations to Applications' Springer-Verlag*: New-York, pp.273–298, 1998.

[32] L. Campo, J. Ignacio, Aliaga, F. JuaN, D. Paz, A. E. Garcia, J. Bajo, G. Villarubia, and J. M. Corchado, "Retreatment Predictions in Odontology by means of CBR Systems" , *Research Article, Computational Intelligence and NeuroScience, Hindawi Publishing*, 2016..

[33] R. Ali, M. Afzal, M. Hussain, M. Ali, M. H. Siddiqi, S. Lee, B. H. Kang, "Multimodal hybrid reasoning methodology for personalized wellbeing services", *Computers in Biology and Medicine, Elsevier*, vol. 69, pp.10-28, 2016.

[34] X.-O. Ping, Y.-J. Tseng, Y.-P. Lin, H.-J. Chiu, F. Lai, J.-D. Liang, G.-T. Huang, P.-M. Yang, "A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients", *Computers in Industry, Elsevier*, vol. 69, pp.12-21, 2015.

[35] A. Mansoul and B. Atmani, "Clustering to Enhance Case-Based Reasoning", *Modeling and Implementation of Complex Systems'*, DOI 10.1007/978-3-319-33410-3_10, 2016.

[36] L. sequira, R. Prabhu, S. S. Mayya, D. S. P. Nagaraju, E. S. Devi, B. S. Nayak, A. George, "Status of Chronic Kidney Disease in India"– Narrative Review, vol. 2(1), 2016.

[37] Z. Huanga, J. M. Juarezb, H. Duana, H. Lia, "Length of stay prediction for clinical treatment process using temporal similarity", *Elsevier*, vol. 41(2), pp. 274–283, , 2014.

[38] S. Guessouma, M. T. Laskrib, J. Lieberc, "RespiDiag: A Case-Based Reasoning System for the Diagnosis of Chronic Obstructive Pulmonary Disease", vol. 41(2), pp.267–273, 2014.

Seema Sharma received the B.Tech. degree in information technology from Punjab Technical University, Jalandhar, in 2003, and the ME degree in information technology from Punjab Engineering College (PEC) in 2006. She is currently pursuing the Ph.D. degree with Amity University Uttar Pradesh, Noida, India. She is an Assistant Professor with the Amity School of Engineering and Technology, Amity University Uttar Pradesh. Her research interests are in the field of data mining and machine learning.

Dr. Deepti Mehrotra did Ph.D. from Lucknow University and currently she is working as Professor in Amity school of Engineering and Technology, Amity University, Noida, earlier she worked as Director of Amity School of Computer Science, Noida, India. She has more than 20 year of experience in research, teaching and content writing. She had published more than 150 papers in international refereed Journals and conference Proceedings. She is editor and reviewer for many books, referred journal and conferences. She is regularly invited as resource person for FDPs and invited talk in national and international conference. She enjoys guiding research scholar and has many Ph.D. and M.Tech students.