RES MILITARIS Social Science Journal

PREDICTION AND CLASSIFICATION OF LIVER DISEASE USING MACHINE LEARNING

GUGLOTH SURESH, NAVEENBABU KATEPALLI, ANGOTHU RAM BABU

Dept of CSE,
Priyadarshini Institute of Science and Technology for Women Khammam.

ABSTRACT:

Due to the quick rise in liver illness caused by excessive alcohol use, contaminated gas inhalation, drug use, tainted food, and pickled food packaging, a doctor can make an automatic forecast with the aid of a medical expert system. Early liver disease prediction is now attainable because to the consistent advancements in machine learning technology, allowing for simple early identification of the fatal condition. This will make healthcare more beneficial, and a medical expert system can be employed in a remote location. The liver is vital to life and promotes the body's ability to rid itself of poisons. Early detection of the condition is therefore crucial for recovery. many machine learning techniques, including supervised, unsupervised, and semi-supervising, bolstering SVM, KNN, K-Mean clustering, neural networks, decision trees, and other learning techniques for diagnosing liver disease provide varying accuracy, precision, and sensitivity. The goal of this paper is to provide an overview and comparative analysis of all machine learning techniques currently being used in the medical field for the diagnosis and prediction of liver disease. The analysis is based on accuracy, sensitivity, precision, and specificity.

Keywords- SVM, KNN, K-Mean clustering, Neural networks, Decision trees,accuracy, sensitivity, precision, and specificity.

I. INTRODUCTION

The prevalence of liver illness has rapidly increased in recent years, placing a heavy load on healthcare systems around the world. The negative consequences of excessive alcohol consumption, environmental damage, drug abuse, and tainted food have all exacerbated this serious health problem. Therefore, to address the alarming rise in liver disease cases, the creation of an efficient and automated medical expert system capable of early prediction becomes crucial.

Machine learning, a rapidly developing area of artificial intelligence, has made impressive strides and has enormous promise for revolutionising the healthcare industry. Predictive models can now be used by medical experts to spot liver illness in its early stages by utilising the power of machine learning algorithms. Early detection is essential for prompt treatment, intervention, and better patient outcomes.Additionally, by incorporating such medical expert systems in remote and disadvantaged locations, it is possible to close the healthcare gap and give access to precise diagnosis and prognosis.

Algorithms for supervised, unsupervised, semi-supervised, and reinforcement learning have all been widely used for the diagnosis of liver disease among the wide range of machine learning techniques available. Different techniques, including Support Vector Machines (SVM), k-Nearest Neighbours (KNN), K-Mean clustering, neural networks, and decision trees, have demonstrated varied degrees of effectiveness in terms of accuracy, precision, sensitivity, and specificity. To choose the best strategy based on certain diagnostic criteria, it is essential to comprehend the relative performance of different strategies.

In order to predict liver illness, several researchers have used a variety of machine learning algorithms. This work intends to give a thorough survey and comparative analysis of these methods. This study aims to shed light on the potential advantages of utilising machine learning for improved liver disease detection and prediction by reviewing and synthesising the available literature. Accuracy, sensitivity, precision, and specificity will be the main focus of the evaluation criteria in order to shed light on the advantages and disadvantages of various strategies.

Medical personnel can increase the precision and effectiveness of diagnosing liver disease by utilising machine learning technology, opening the door to early therapies and better patient outcomes. The results of this study may lay the groundwork for future developments in the field, encouraging partnerships between medical professionals and data scientists to create robust and trustworthy medical expert systems for the prognosis of liver disease. In the end, such advancements can make a considerable impact on the healthcare industry and help reduce the burden that liver disease places on both people and society at large.

II. LITERATURE SURVEY

According to Ramana's article, "A Critical Study Of Selected Classification Algorithms For Liver Disease Diagnosis," The classification of several liver patient datasets is evaluated in this research along with the chosen classification techniques. The Naive Bayes classifier, C48, Back propagation Neural Network algorithm, and Support Vector Machines are the classification algorithms that are taken into consideration here. Accuracy, Precision, Sensitivity, and Specificity are the four factors used to evaluate these algorithms. In many different automatic medical diagnosis technologies, classification approaches are fairly common. As the liver will continue to operate correctly even when it is partially damaged, problems with liver patients are difficult to identify at an early stage.

"Prediction Of Different Types Of Liver Diseases Using Rule Based Classification Model," Kumar, Yugal, and G. Sahoo, Different liver illnesses are categorised in this document based on laboratory test findings and analysis. We therefore created a Rule Base Classification 4Model (RBCM) to forecast various forms of liver illnesses in order to simplify this difficult process. The suggested approach combines various data mining techniques with rules. This paper's goal is to suggest a rule-based categorization model with machine learning methods for the forecasting of various liver disorders. A data collection with 12 attributes was created, containing the records of 583 patients, 441 of them were men and the remaining patients were women. SVM, or Support Vector Machine The suggested model for the prediction of liver illnesses uses Rule Induction (RI), Decision Tree (DT), Naive Bayes (NB), and Artificial Neural Network (ANN) data mining approaches with K-cross fold technique. In order to analyse the liver disease data set and the independence of attributes, statistical methods (such as ANOVA and the Chi square test) are employed to evaluate the performance of these data mining techniques. Out of 583 patients, 416 have liver disorders, while the remaining 167 are healthy. Among all techniques (RI, SVM, ANN, and NB) with all parameters (Accuracy 98.46%, Sensitivity 95.7%, Specificity 95.28%, and Kappa 0.983), the suggested model with decision tree (DT) technique delivers the better results.accuracy (82.33%), sensitivity (68.03%), specificity (91.28%), and kappa (0.801) performance.Your query lacked specific information on the datasets that were utilised.The researchers used Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) as two evaluation metrics to compare the performance of the regression models. By calculating the difference between projected values and actual values, these metrics are frequently used to assess the precision of regression models. It is also discovered that the worst performance of the rule-based classification model (SVM, Accuracy 82.33%, Sensitivity 68.03%, Specificity 91.28%, and Kappa 0.801 and the accuracy of chi square test is 76.67%) is almost similar to the best performance of the model without rules (RI, Accuracy 82.68%, Sensitivity 86.34%, Specificity 90.51%, and Kappa 0.619). This The research shows that the suggested rules-based classification model and the model without rules significantly differ in their ability to predict liver illnesses, with the rule-based classification model using the decision tree (DT) technique producing the most accurate results. A useful tool for making medical decisions is this model.

"Prediction Of Liver Fibrosis Stages By Machine Learning Model: A Decision Tree Approach" was proposed by Ayeldeen, Heba, Olfat Shaker, Ghada Ayeldeen, and Khaled M. Anwar. It claims that the use of information systems and tactical tools in the medical fields is always expanding. Medical decision-making benefits from the use of automated models, which enable doctors to make quick and accurate diagnoses or even predictions. Different statistical mining and machine learning tools can be utilised to make use of the knowledge or even at the early phases of knowledge collection. One of the worries is, for example, determining if the patient with the Hepatitis C virus also has liver fibrosis or not.What stage of fibrosis is present if the forecast comes true? A completely integrated system is required to easily get this knowledge without expensive diagnostic regular laboratory procedures. To predict the level of liver fibrosis in individuals, we therefore utilised a machine learning approach model based on decision tree classifier in this work. The accuracy of the decision tree classifier, according to the results, is 93.7%, which is greater than the range reported in recent studies conducted under comparable circumstances.

III. PROPOSED SYSTEM

The suggested Early system diagnosis of liver illness is critical and very important because it will aid in the disease's treatment and recovery. Additionally, it is quite challenging to accurately predict a disease's recovery in the early stages of the illness. The attributes for the machine learning methods will be Total Bilirubin, Direct Bilirubin, Alkaline Phosphotase, Alamine Aminotransferase, Aspartate Aminotransferase, Total Protiens, Albumin, Albumin and Globulin Ratio. To improve the accuracy of our models, we will train these attributes. A kind of artificial intelligence called machine learning enables computers to think like people and make decisions on their own without the need for human interaction. Machine learning has made significant progress in the diagnosis of

differences as a result of the rapid growth of artificial intelligence.many diseases. Additionally, machine learning algorithms improve performance and prediction.

1. The main distinction between the Machine Learning Algorithm (MLA) approach and the conventional predictive model is that MLAs learn from the data already collected in order to identify unique patterns between variables and produce predictions.

2. It has been demonstrated that MLAs increase the accuracy of spotting disease-risk individuals.

3. Learning techniques that are supervised sometimes involve the assistance of a supervisor, teacher, or instructor. It is made easy for the algorithm from input to output as well as to learn and predict since it includes a training set of patterns linked to label data.

4. SVM, Naive Bayes, ANN, and K Means Clustering are the algorithms.

IV. RESULTS

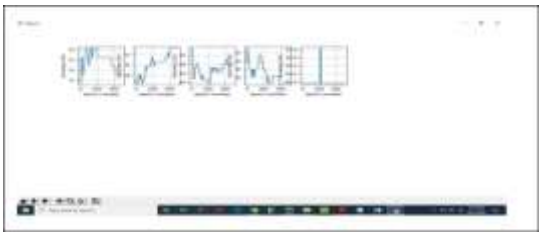

Fig. 1.3 -Back propagation

Fig. 1.2 -Classification reports

V. CONCLUSION

This effort gives us a fundamental understanding of previously published papers that used various machine learning algorithms to detect and diagnose liver disease. With the help of this survey and study, it has become obvious that certain machine learning algorithms, like decision trees, J48, and ANN, offer more accuracy in the identification and prognosis of liver illness. Additionally, different algorithms perform differently depending on the situation, but most significantly, the dataset and feature selection are crucial for producing accurate predictions. Additionally, the study provides a summary of the many machine learning approaches employed by various authors. Each machine learning technique has both positive and negative effects depending on the datasets and feature selection, among other factors.

VI. REFERENCES

[1] Ramana, Bendi Venkata, N. B. Venkateswarlu, and M. Surendra Prasad Babu. "A critical analysis of specific classification algorithms for the diagnosis of liver disease." 3.2 (2011): 101– 114 in the International Journal of Database Management Systems.

[2] Ramana, N. B. Venkateswarlu, Bendi Venkata, and MS Prasad Babu. "Liver classification using a modified rotation forest." International Journal of Engineering Research and Development, 6, no. 1, 2012, pp. 17–24.

[3] G. Sahoo, Yugal, and Kumar. "Prediction of different types of liver diseases using rule-based classification model." Health Care Technology 21, no. 5 (2013): 417-

[4] Khaled M. Anwar, Olfat Shaker, Ghada Ayeldeen, and Ayeldeen. "Decision tree approach for the prediction of liver fibrosis stages by machine learning model." 1-6 in 2015 Third World Conference on Complex Systems. IEEE, 2015.

[5] Sindhuja, D. and R. JeminaPriyadarsini, "A survey on classification techniques in data mining for analysing liver disease disorder." (2016) 483–488 in International Journal of Computer Science and Mobile Computing.

[6] "Comparison of machine learning approaches for prediction of advanced liver fibrosis in chronic hepatitis C patients." Hashem, Somaya, et al. The 15.3 (2017) issue of IEEE/ACM Transactions on Computational Biology and Bioinformatics contains pages 861–868.

[7] Sontakke, S., Lohokare, J., and Dani (February 2017). machine learning for liver disease diagnosis. 129–133) in 2017 International Conference on Emerging Trends & Innovation in ICT. IEEE. Ma, Han, Cheng-fu Xu, Zhe Shen, Chao-hui Yu, and You-ming Li are all mentioned in

[8]. A cross-sectional study on non-alcoholic fatty liver disease in China using machine learning approaches for clinical prediction modelling. The 2018 BioMed Research International

[9] Joseph Chakkalakal Mathew, Jacob, Joel, J. Mathew, and E. Isaac. "Using machine learning techniques to diagnose liver disease." 2018's Int Res J Eng Technol 5, no. 4.

[10] Ambika L Gusha S, Manjunath Varchagall, and Sivakumar D The study is titled "Chronic Liver Disease Prediction Analysis Based on the Impact of Life Quality Attributes." (2019). Volume 7, Issue 6S5, April 2019, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878

[11] Mehta Banu H. "Liver Disease Prediction Using Machine-Learning Algorithms" Volume 8 Issue 6, August 2019, International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249-8958