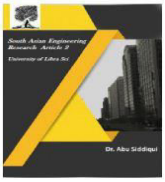
International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal



LIVER DISEASE PREDICTION AND CLASSIFICATION USING MACHINE LEARNING

[1]E AMARNATH REDDY, [2]M TEJASWI REDDY , [3]K SRAVANI, [4]M RAMOJI

[123]ASSISTANCT PROFESSOR, [4]UG SCHOLAR, DEPARTMENT OF CSE

BRILLIANT INSTITUTE OF ENGINEERING & TECHNOLOGY,

ABDULLAPURMET(V&M) RANGA REDDY DIST-501505

ABSTRACT:

As liver disease rates are surging due to factors like excessive alcohol consumption, inhalation of contaminated gas, drug use, and ingestion of tainted or pickled foods, doctors are now able to leverage medical expert systems for automated forecasting. Thanks to the rapid progress in machine learning technologies, we can now achieve early predictions of liver disease, significantly enhancing the early detection of this serious condition. This advancement not only improves healthcare accessibility but also means that medical expert systems can function effectively in remote areas. Given that the liver plays a crucial role in detoxifying the body, timely detection is essential for effective recovery.

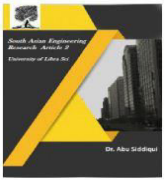
Various machine learning techniques are being applied to liver disease diagnosis, including supervised, unsupervised, and semi-supervised methods. Techniques such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), K-Means clustering, neural networks, and decision trees each present unique capabilities in terms of accuracy, precision, and sensitivity. This paper aims to provide a comprehensive overview and comparative analysis of these machine learning approaches currently utilized in the medical field for diagnosing and predicting liver disease, focusing specifically on their accuracy, sensitivity, precision, and specificity.

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

I. INTRODUCTION

The incidence of liver disease has surged dramatically in recent years, putting significant pressure on healthcare systems globally. Factors such as excessive alcohol intake, environmental degradation, substance abuse, and contaminated food sources have intensified this troubling health crisis. To combat the growing number of liver disease cases, it's essential to develop an effective and automated medical expert system capable of early diagnosis.

Machine learning, a rapidly evolving facet of artificial intelligence, has made remarkable advancements with great potential to transform the healthcare sector. Medical professionals can leverage predictive models powered by machine learning algorithms to identify liver disease in its initial stages. Early detection is vital for timely treatment, intervention, and improved patient outcomes. Moreover, implementing these medical expert systems in remote and underserved areas could help bridge the healthcare gap, providing access to accurate diagnoses and prognoses.



A broad array of machine learning techniques—including supervised, unsupervised, semi-supervised, and reinforcement learning—has been applied to diagnose liver disease. Specific methodologies such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), K-Mean clustering, neural networks, and decision trees have shown varying levels of effectiveness regarding accuracy, precision, sensitivity, and specificity. Understanding the relative performance of these various approaches is crucial for selecting the most appropriate strategy based on specific diagnostic needs.

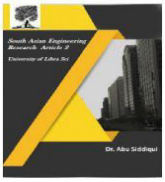
Several researchers have explored diverse machine learning algorithms for predicting liver disease. This work aims to deliver a comprehensive review and comparative analysis of these methodologies. By synthesizing the existing literature, this study seeks to highlight the potential benefits of utilizing machine learning for enhanced detection and prediction of liver disease. The evaluation criteria will focus on accuracy, sensitivity, precision, and specificity to illustrate the strengths and weaknesses of different approaches.

By harnessing machine learning technologies, healthcare providers can enhance the accuracy and efficiency of liver disease diagnoses, paving the way for earlier treatments and better patient outcomes. The findings of this study may serve as a foundation for future advancements in the field, fostering collaborations between clinicians and data scientists to develop robust and reliable medical expert systems for liver disease prognosis. Ultimately, these innovations could significantly impact the healthcare system and alleviate the burden that liver disease imposes on individuals and society as a whole.

II. LITERATURE SURVEY

In Ramana's article, "A Critical Study Of Selected Classification Algorithms For Liver Disease Diagnosis," the research assesses various classification techniques used to analyze multiple liver patient datasets. The study focuses on several algorithms, including the Naive Bayes classifier, C48, Back Propagation Neural Network, and Support Vector Machines. To evaluate these algorithms, the author considers four key metrics: Accuracy, Precision, Sensitivity, and Specificity. Classification methods are quite prevalent in various automatic medical diagnosis technologies. Identifying issues in liver patients can be challenging, as the liver can continue to function adequately even when partially impaired.

The document titled "Prediction of Different Types of Liver Diseases Using Rule-Based Classification Model" by Kumar, Yugal, and G. Sahoo offers a detailed categorization of liver illnesses based on laboratory test results and comprehensive analyses. To streamline this complex process, the authors have developed a Rule-Based Classification Model (RBCM) aimed at predicting various liver disorders. This innovative approach integrates multiple data mining techniques alongside established rules. The primary objective of the paper is to propose a machine learning-driven rule-based classification model for the forecasting of different liver diseases. For this research, a dataset comprising 12 attributes was created, featuring records from 583 patients—441 male and 142 female. The RBCM utilizes several data mining techniques, including Support Vector Machine (SVM), Rule Induction (RI), Decision Tree (DT), Naive Bayes (NB), and



Artificial Neural Network (ANN), implemented with a K-cross fold validation method. To evaluate the performance of the data mining techniques, statistical methods like ANOVA and the Chi-square test were employed, helping analyze the liver disease dataset and the independence of its attributes. Out of the 583 patients, 416 were diagnosed with liver disorders, while 167 were deemed healthy. Among the various techniques (RI, SVM, ANN, and NB), the model based on the Decision Tree method showed the best results with an accuracy of 98.46%, a sensitivity of 95.7%, specificity of 95.28%, and a Kappa value of 0.983. The paper highlights that the evaluation metrics used, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), were crucial for comparing the performances of the regression models. These metrics assess how closely projected values align with actual values. Interestingly, the rule-based classification model had a minimum accuracy of 82.33%, with sensitivity at 68.03%, specificity at 91.28%, and a Kappa value of 0.801. This performance was comparable to the model without rules, which achieved slightly better accuracy (82.68%) but lower sensitivity (86.34%) and a Kappa of 0.619. This research clearly demonstrates that there is a significant difference in predicting liver diseases between the proposed rule-based classification model and the model without rules, with the decision tree technique yielding the most accurate outcomes. Ultimately, this model serves as a valuable tool for informed medical decision-making regarding liver disorders..

The research paper titled "Prediction Of Liver Fibrosis Stages By Machine Learning Model: A Decision Tree Approach" was authored by Ayeldeen, Heba, Olfat Shaker, Ghada Ayeldeen, and Khaled M. Anwar. It highlights the growing role of information systems and strategic tools in healthcare. Automated models play a crucial role in medical decision-making, assisting healthcare professionals in making swift and precise diagnoses or predictions. Various statistical mining techniques and machine learning tools can be employed to harness knowledge even during the preliminary data collection phases. A significant concern in this field is identifying whether a patient with Hepatitis C also suffers from liver fibrosis and, if so, determining the stage of fibrosis. To facilitate access to this essential information without requiring costly laboratory tests, an entirely integrated system is necessary. Therefore, in this study, we utilized a machine learning model based on a decision tree classifier to predict liver fibrosis levels in patients. The decision tree classifier achieved an impressive accuracy of 93.7%, surpassing the performance reported in recent similar studies.

III. PROPOSED SYSTEM

The early diagnosis of liver disease is crucial as it significantly contributes to the treatment and potential recovery of patients. However, accurately predicting a patient's recovery in the initial stages of the illness can be quite challenging. The key indicators we will focus on, using machine learning techniques, include Total Bilirubin, Direct Bilirubin, Alkaline Phosphatase, Alamine Aminotransferase, Aspartate Aminotransferase, Total Proteins, Albumin, and the Albumin/Globulin Ratio. By training our models on these attributes, we aim to enhance their accuracy. Machine learning, a form of artificial intelligence, enables computers to process information and make decisions independently, mimicking human thought processes. The rapid

advancement of artificial intelligence has led to significant improvements in diagnosing various diseases through machine learning algorithms, which continually enhance their performance and predictive capabilities.

1. The key difference between Machine Learning Algorithms (MLAs) and traditional predictive models is that MLAs analyze previously collected data to uncover unique patterns among variables, allowing them to generate predictions more effectively.

2. Research shows that MLAs significantly enhance the accuracy of identifying individuals at risk for certain diseases.

3. Supervised learning techniques often require guidance from a supervisor, teacher, or instructor. This support simplifies the process for the algorithm, enabling it to learn and predict based on a training set of labeled data patterns.

4. Common algorithms in this space include SVM, Naive Bayes, ANN, and K Means Clustering..

IV. RESULTS

Fig. 1.3 -Back propagation

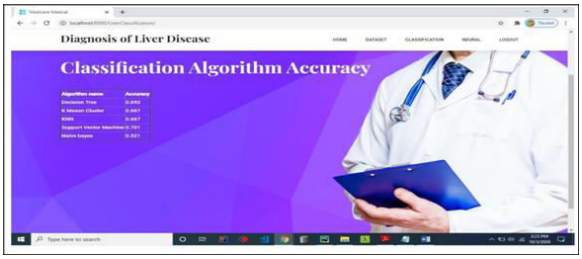


Fig. 1.2 -Classification reports

V. CONCLUSION

This work provides a foundational insight into previously published research that has utilized various machine learning algorithms for the detection and diagnosis of liver disease. Through this survey and analysis, it has become clear that certain algorithms, such as decision trees, J48, and artificial neural networks (ANN), deliver greater accuracy in identifying and predicting liver conditions. Furthermore, the effectiveness of different algorithms can vary depending on the specific context; however, the dataset and feature selection are essential factors for achieving precise predictions. The study also summarizes the different machine learning techniques

employed by various researchers, highlighting that each approach comes with its own advantages and disadvantages influenced by factors like datasets and feature selection.

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