Metaheuristic Optimization Review (MOR) *Vol. 01, No. 01, PP. 45-54, 2024*

Innovations in Machine Learning Models for Hepatitis Diagnosis and Disease Progression Prediction

Marwa M. Eid[1,*], Wei Hong Lim[2]

[1]Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura, Egypt

[2]Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia

Email: mmm@ieee.org; limwh@ucsiuniversity.edu.my

Abstract

Chronic liver disease (CLD) is a group of conditions for which up to half of the global population remains at risk and causes serious complications: liver cirrhosis and liver cancer. Therefore, early diagnosis and proper treatment of these diseases enhance the prognosis of patients suffering from CLD. This review paper explores how machine learning (ML) techniques are used in practice to diagnose, prognosis, and treat chronic liver diseases. Continuing with more specific examples of collected data from the results of several studies, their more comprehensive implementation is expected to improve the respective management processes and the detection of liver disease in patients more accurately. The review further discusses the various ML methods, including supervised and unsupervised learning, neural network, and ensemble learning, also applied to the estimation of risk felt by the patients, suggesting a course of treatment or how far the disease has progressed. While the inclusion of ML technology in the field of Hepatology is progressing well, some issues like model diversity, applicability of models, and concerns about ethics still pose challenges. This paper points out the importance of working in teams from various fields to develop appropriate mechanisms for dealing with these issues and adequately use ML for clinics. In conclusion, the results indicate that there is a possibility that ML will change the management of chronic liver diseases, which in turn will lead to the development of innovative treatment methods and better patient management.

Keywords: Persistent hepatic illnesses; Computational techniques; Identification; Prediction; Management and evidence-based methods

1. Introduction

Place the contents in this section before or after those that are still to be rewritten as deepfake. Do not stick to the order in which they were provided. Chronic liver disease is one of the conditions, which include hepatitis B and Hepatitis C, as well as the cancer of the liver cells, also known as hepatocellular carcinoma (HCC), that millions of people across the globe are suffering. These diseases create a significant burden in terms of the population afflicted, loss of life, and costs associated with healthcare; hence, improving the existing ways of diagnosing and treating these conditions is of utmost importance. However, contemporary medical science has made it possible to diagnose various liver diseases using modern imaging techniques and drugs, wherein these imaging techniques pose a substantial financial cost and some potential risk factors

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

associated with liver needle aspirations and serological tests. This has increased the need for early intervention, eficaz diagnosis, treatment, and managing confined resources effectively [1].

Artificial intelligence in medicine – machine learning (ML), where computers learn from data patterns – has recently started to create a paradigm shift in medical diagnostics and healthcare delivery. Using vast amounts of data, ML can also find relationships and indicators that are hard to find with regular analysis. In the case of liver diseases, ML has proven to be very efficient in relating to and interpreting different forms of data and information, such as serum biomarkers, clinical data, genetic materials, imaging and other forms of photography. These capabilities allow better forecasts of the disease's development processes and its risk factors, as well as the creation of therapies that meet the needs of every particular patient [2], [3].

In the last ten years, many researchers have explored the use of ML through its various forms in finding the causes, finding solutions and managing liver diseases. Enhancing their performance on the classification of liver diseases, which is often the case even when the image is not clear, deep learning, support vector machines (SVM), decision trees, and ensemble strategies, among other machine learning approaches, have been used. For example, imaging data is where trained ML models can be used to note liver changes that are not visible but are precursors of conditions such as HCC or fibrosis. MN-based models can also be used to predict risks by evaluating the effect caused by a combination of specific genetic markers and other clinical variables [4], [5].

In Mississippi, after these encouraging outcomes, the use of ML approaches for the management of patients with liver diseases remains a challenge to integrate into practices. Initiatives to actualize this include problems regarding the comprehensiveness of data, interpretability of developed models, and adequate cross-validation of the models in different settings. Additionally, the ever-changing nature of liver diseases, including their causes and rate of progression, will call for models that are even more complicated. Resolving these issues will call for concerted efforts of clinicians, data scientists, and researchers to produce ML mechanisms that will be accurate, interpretable, reliable and applicable to many healthcare systems [6], [7].

This review seeks to give an extensive insight into the recent developments in the use of machine learning in the diagnostics, treatment, and prognosis of liver diseases. It will summarize some critical literature, emphasizing the study methods conducted and how ML can be incorporated into health practices to improve early detection, prognosis, and tailored treatment. Furthermore, this review will address the present concerns and the prospects of this fast-growing area of study, stressing the importance of standard protocols, ethical issues, and the possibilities of ML in managing liver infections.

2. Literature Review

Machine learning (ML) innovations have brought remarkable changes in the field of hepatology in recent times, more so in liver disease diagnosis and prognosis. Invasive methods such as liver biopsy are effective, but because they are associated with certain risks, there has been a shift to non-invasive ML-based approaches. Another advantage of longitudinal data and machine learning is that models have been designed to predict and manage liver diseases such as fibrosis, hepatitis, and hepatocellular carcinoma (HCC) based on multiple biomarkers, clinical data, or even genetic data. This literature review highlights recent publications focused on applying such different ML methods to improve the diagnostic accuracy, risk assessment, and treatment conditions that patients receive. Such developments indicate the transformative role ML can have in the management of liver diseases and the care of patients.

In recent years, as noted in [8], there has been an increase in the application of machine learning techniques as a non-invasive technique to stage chronic liver illnesses instead of invasive liver biopsy techniques. The study aimed to assess the performance of different classified machine-learning algorithms for advanced liver fibrosis prediction by utilizing and combining serum biomarkers and clinical data. Thirty-nine thousand five hundred sixty-seven patients with chronic hepatitis C were recruited into a prospective cohort and analyzed after stratifying them for METAVIR score into F0F2: mild to moderate fibrosis and F3F4: advanced fibrosis. In order to predict the dangerous factors of advanced fibrosis, the decision tree, genetic algorithm, and particle swarm optimization - multi-linear regression machine learning models were created. The quality of the models was evaluated by ROC curve analysis with best AUROC ranging from 0.73 to 0.76 and accuracy between 66.3%and 84.4%. Age, platelet count, AST and albumin levels were significant predictors of

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

advanced fibrosis. In conclusion, it appears that machine learning algorithms can be used to assess the risk of advanced liver fibrosis in patients with chronic hepatitis C.

The investigation in question aimed to assess a new hybrid artificial intelligence-based classifier for its ability to classify the prognosis of patients diagnosed with chronic hepatitis, as explained in the paper [9]. In the analysis, 19 biomarkers, taken from 155 hepatitis patients, which are publicly available at the University of California Irvine Machine Learning repository, were utilized. The number of features employed for the classification was limited due to the Lagrangian support vector machine (LSVM) geometric margin maximization criterion, which demonstrated that 16 out of the initial 19 factors were the most important for prognosis. Finally, several artificial intelligence classifier algorithms. They employed the critical prognostic factors as input variables. The results proved that the proposed hybrid method, which combined LSVM with a multilayer perceptron (MLP), produced results of 100% in both validation accuracy and area under the receiver operating characteristic (AUROC) curve. Such remarkable result emphasizes the usefulness of the hybrid artificial intelligence-based design in prognosis prediction among patients in clinical practice.

According to [10], in light of the increasing occurrence of hepatocellular carcinoma (HCC) cases in mainland China, there is a need for measures aimed at its early detection, especially in patients with prior hepatitis b virus (HBV) related liver cirrhosis and with low serum alpha-fetoprotein (AFP) level. The objective was to develop individual predictive models of hepatocellular carcinoma risk using machine learning ML approaches with personal characteristics, medical and treatment histories, and non-invasive biomarkers. 6980 patients from the archived cohort treated from 2012 to 2018 were studied, and available clinical and pre-treatment laboratory information was analyzed to determine the predictive factors for the development of HCC using ML and univariate regression methods. Twelve independent risk factors were identified in the predictive models built with Gaussian naive Bayes, extreme gradient boosting XGBoost, random forest, and LASSO regression. Risk factors distinguished were male sex, age above six decades, the rise of alkaline, phosphate, and male AFP levels and carcinoembryonic antigens, while risk factors such as high blood pressure, low calcium, and low hemoglobin were protective. The best outcome came from the XGBoost model, where the AUC was 0.832 in the training set and 0.829 in external validation, thus outperforming the conventional AUC-based method that relied on serum AFP measurements. It is evident that M€L models, especially XGBoost, offer new insights into the early detection and treatment of HCC.

The referenced article [11] examined the influence of rs12979860 IL28B gene polymorphism on the risk of developing hepatic fibrosis concerning hepatitis C virus infection in patients irrespective of underlying causes. The study compared the predictive performance of the data mining approach and traditional scores based on guidelines such as ascitic sarcopenia and aspartate aminotransferase to platelet ratio index (APRI) in diagnosing high-grade HCV-related fibrosis. The database encompassed 427 patients suffering from chronic hepatitis C, and a decision tree was constructed using data mining using the Reduced Error Pruning (REP) method. In addition, the Auto-WEKA utility was used to test the advanced fibrosis predictive algorithms and select the algorithm with the highest performance from the 39 available. The study results indicated that APRI and FIB-4 were sensitive-specific in the 0.523-0.831 and 0.415-0.917, respectively. The algorithm REP Tree based on principles of data mining had a sensitivity of 0.749, a specificity of 0.729, and an area under the curve of 0.796. Out of 16 attributes considered in the analysis, the IL28B genotype was the strongest predictor of advanced hepatitis C-related fibrosis. Prediction using Auto-WEKA highlighted the 'multilayer perceptron' (MLP) neural network as the best-performing algorithm with a sensitivity of 0.825, specificity of 0.811 and area under the ROC curve of 0.880. This confirms the better performance of MLP over APRI, FIB-4 and REP Tree in assessing advanced fibrosis in chronic hepatitis C patients.

According to the fight against hepatitis C described in [12], this disease has worldwide effects on more than 70 million people and causes about 400,000 deaths every year. Therefore, prompt treatment is necessary, especially in diabetic patients, in order to enhance the chances of survival. This research took advantage of Electronic Health Records Systems (EHRs) to better comprehend hepatitis C and its prognosis using statistical data analysis performed through digesting computers, which uncovered elements that physicians would have otherwise missed. The research also included an automatic diagnosis system termed Real-Time World Laboratory, which sought to evaluate EHR data from 615 clients to find blood donors' status, those suspected of being blood donors, patients diagnosed with hepatitis, those with fibrosis and those with cirrhosis. In achieving the end goal of disease prevention, a variety of ML techniques such as Support Vector

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

Machine (SVM), Multivariate Adaptive Regression Splines (MARS), Bayesian Generalized Linear Model (BGLM), Random Forest (RF) and Decision Trees (DT) were researched in seeking to find out which of the disease detection why the most successful. An accuracy performance matrix was used to determine the effectiveness of each of the models employed, all of which were implemented in R programming, pointing out how ML can be helpful in the early detection of hepatitis C and its treatment.

The study reported in [13] highlights the importance of liver disease diagnosis at its early stages because diagnosis in the later stages is often associated with death. In order to fulfill this goal, the study attempts to experiment with various machine learning algorithms on the Indian Liver Patients Dataset (ILPD) to find the best approach to improving the accuracy of diagnosis. Given that data quality is essential, rigorous preprocessing steps such as sampling, normalization and particle swarm optimization (PSO) were carried out on the dataset. The results showed that the J48 algorithm provided better results than the other algorithms in distinguishing between patients and non-patients. It is also worth noting that this research encompasses an etiological survey that analyses the impact of different factors on the dataset in order to facilitate the understanding of mechanisms of liver disease diagnosis and shows that machine learning can be used for the early diagnosis and treatment of patients at risk of liver disease.

In the referred article denoted as [14], a common diagnosis strategy for a health condition called hepatitis is presented using a new way of machine learning, which combines support vector machine (SVM) with simulated annealing (SA) techniques. Simulated annealing, a random optimization technique to solve many complex optimization problems, has been cherished recently for its advantage in predictive modeling support vector machines. The machine learning repository has provided the dataset used for this research. The classification accuracy was estimated using 10-fold cross-validation, obtaining a remarkable accuracy of 96.25%. This outcome indicates the feasibility of the proposed hybrid technique, making it a better option than other diagnosis classification techniques for hepatitis published in the studies.

As mentioned in [15], the early and definite diagnosis of hepatocellular carcinoma (HCC) is essential to assist in the customized treatment and follow-up for chronic hepatitis B (CHB) infected patients. This research aimed to create a model for the prediction of HCC based on the analysis of the patient's medical imaging and comparing it with the efficacy of liver stiffness measurement (LSM) and other clinical prognostic scores in CHB patients. Of the 1,215-screened patients, 434 CHB patients were recruited for a 5-year follow-up, and 96.3% had liver biopsies performed. The authors employed deep learning radiomics analysis on 2,170 images acquired using two-dimensional shear wave elastography (2D-SWE) and their corresponding B-mode ultrasound (US) images, with an intent to blend high-dimensional imaging characteristics and low-dimensional serological data to develop machine learning strategies for HCC prediction under the situation of skewed class distribution. The best performance-based model developed, HCC-R, utilized age and sex and image predictors of 2D-SWE and B-mode ultrasound images and made very accurate predictions. During the 5-year follow-up, 32 patients (7.4%) were diagnosed with HCC. The HCC-R model attained the areas under the receiver operating characteristic curve (AUC) of 0.981, 0.942 and 0.900 in the training, validation and testing cohorts, respectively and was way above the AUC obtained with the LSM, which ranged between 0.676 to 0.784, $p < 0.05$ also above the other clinical prognostic scores (AUC: 0.544 to 0.869). All of these features of the HCC-R radiomics model, including both biomechanical and morphological features of the patients, show its ability to predict the occurrence of HCC with high accuracy, leading to improved treatment and management of follow-up of CHB patients.

In the paper referenced in section [16], the authors focus on improving hepatitis B and C diagnostics by analysis of the data obtained from questionnaires of patients treated in different regions of infectious diseases in Slovakia. The research was experimental and was designed to address two problems: the costs of treatment recommended according to the already established diagnostics and the strategies aimed at improvement of hepatitis diagnostics by exploration and prediction on additional patient data. Exploratory data analysis was carried out to confirm or disprove the existence of relations between input attributes, such as age and gender, and the resulting diagnosis. Besides, some useful sets of rules for decision-making in the course of the impossible diagnosis were obtained through predictive data mining that can be used in clinical practice at the initial stages of clinical problem solving. Treatment decisions were also analyzed economically using the estimated treatment costs. They emphasized the need for prompt and effective diagnosis in order to prevent hepatitis from being diagnosed late in patients.

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

The publication identified with the mark [17] explains how hepatitis C virus (HCV) is among the most serious global health issues that confront the population today, infecting nearly 2 - 3% of the population, causing chronic hepatitis C. The fact that no vaccine is effective against the virus, making preventive and therapeutic measures, including the use of interferon-alpha (IFNA), ribavirin (RBV) and their new direct-acting antiviral (DAA) agents necessary aggravate the situation. The research focused on the complete nucleotide sequence of HCV and, like all other methods of analysis, the Gini Index and Chi-squared analysis helped machine learning algorithms that identified genetic determinants of therapy outcomes. Because of the extensive bioinformatics study, several nucleotide features were found to correlate with the treatment outcome in patients infected with HCV subtypes 1a and 1b. In particular, with these algorithms, 75.00% accuracy was obtained in predicting non-responders out of the responders in subtype 1a versus 85.00% in subtype 1b. They also accurately separated treatment responders and relapses at 82.50 % and 84.17 %, respectively. Based on the detected nucleotide features, decision trees were devised, which provided a valuable tool to separate different treatment response categories and contribute to optimizing individualized therapy.

According to the documentary evidence gathered in [18], cancer is still considered one of the most ill-famed causes of death throughout the globe, with hepatic cancer being on the rise, recording 841,080 cases in 2018, which formed about 5 % of the total cancer cases reported that particular year. HCC is the most common type of liver cancer, and it is one of the cancers causing death at an alarming rate. Hence, it ranks second in the overall statistics of cancers. Another epidemiological factor involved in liver cancer incidence, notably 50 percent of HCC cases in most countries, is the chronic infection with the hepatitis B virus. This staggering figure of one out of every 12 deaths in the world being attributed to liver cancer is a result of these aggressive viruses. The present study seeks to evaluate the survival rate among hepatitis patients by offering a dataset that provides an inclusive set of information on serum hepatitis. The first step is to refine the content of the information so that it will be suitable for follow-up analysis and fulfill the predictive model requirements. After that, various algorithms will be used to forecast results and assess the performance of the models developed. It will be possible to compare the efficiency of performance of each algorithm concerning the relation of actual figures to the postulated outcomes where patients were either survivors or non-survivors tagged "Live" or "Die." Finally, it is hoped that the model will be able to assess an individual patient's risk concerning the survival duration with respect to the values of the variables investigated.

In the investigation referred to in [19], scientists sought to perform a risk analysis using computational intelligence (CI) based methods, specifically machine learning (ML), to build a 28-day mortality predictive model in patients with acute-on-chronic liver failure (ACLF) to measure mortality risks as well as treatment outcomes. The study included ACLF patients from six tertiary care centers. Features for developing the Cleveland Clinic smoking cessation ML model were selected using LASSO regression. The evaluation of models included the area under the curve (AUC) and accuracy measures, and Shapley additive explanations were used to explain the predictions of the ML model. Of 736 patients, 587 were assigned to the training set, and 149 were assigned to an outer validation set. The most critical variables were age, hepatic encephalopathy, total bilirubin, prothrombin time activity (PTA) and creatinine. The XGBoost (XGB) model offered the best prognostic predictions in patients with ACLF, with the highest AUC and AUC accuracy. Delong's test determined that the XGB model was superior to conventional scoring systems, including the Child-Pugh score, MELD and CLIF SOFA, CLIF C OF, and CLIF C ACLF. The final model XGB-ML employed trait assessments measured at baseline and on days 3, 7, and 14; interventions also contributed significantly to the model's predictive capability, thus assisting clinicians in the treatment evaluation aspect.

The study presented in [20] focused on understanding the molecular mechanisms of hepatocellular carcinoma (HCC) caused by the hepatitis B virus (HBV), employing both bioinformatics and deep learning techniques. It is essential to comprehend the molecular basis of prevention, diagnosis, and treatment of HBV-associated HCC since 50-80% of the cases of HCC in the world are due to HBV infection. The research commenced with a thorough analysis of the GEO database, purely focused on the gathering of relevant gene sets associated with HBV-HCC and then differential and weighted gene coexpression network analysis (WGCNA) were employed to find the genes that had altered expression and were influential in tumors. A single study identified three key genes, RACGAP1, ECT2, and NDC80, using three deep learning methods: Lasso, random forest, and support vector machine (SVM). The diagnostic models built using these genes were accurate with AUC values of 0.976 (RACGAP1), 0.969 (ECT2), 0.976 (NDC80) in the training set, and 0.878, 0.731, 0.915 in the validation set, respectively. All three genes exhibited high expression in patients

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

with poor outcomes, indicating their close association with the progress of HBV-HCC disease. In addition, molecular docking studies showed that the identified vital genes could interact with anti-HCC treatments Lenvatinib, Regorafenib, and Sorafenib, highlighting their potential effectiveness. Therefore, RACGAP1, ECT2, and NDC80 serve as valuable markers in the diagnosis of HBV-HCC and hold potential in targeted drug development.

Hepatocellular carcinoma (HCC) is a primary malignant tumor of the liver, and about 50–80% of cases worldwide are associated with hepatitis B virus (HBV) infection. Moreover, other causative factors can worsen HCC development when they coexist with HBV. As discussed in the paper [21], it is vital to understand the molecular mechanisms of HBV-induced hepatocellular carcinoma (HBV-HCC) if we are to improve the approaches to prevention, diagnosis, and treatment. The study was oriented on bioinformatics and deep learning to investigate those mechanisms. As such, gene sets relevant to HBV-HCC were first acquired from the GEO database followed by differential analysis and then weighted gene coexpression network analysis (WGCNA) to pinpoint genes showing aberrant expressions concerning tumors. Subsequently, deep learning techniques such as Lasso, random forest, and support vector machine (SVM) were used for crucial gene identification: RACGAP1, ECT2, and NDC80. In HBV-HCC diagnostic models trained using the genes above, the ROC curves achieved an AUC of 0.976 (RACGAP1), 0.969 (ECT2), and 0.976 (NDC80) in the respective training set while achieving 0.878, 0.731, and 0.915 in the respective validation sets. Higher expression of these genes was associated with adverse outcomes in patients with liver cancer, on the other hand, indicating the involvement of these genes in the progression of HBV-HCC. In addition, molecular docking also showed that these genes effectively interact with cancer medicine, such as lenvatinib, Regorafenib, and Sorafenib, to target cancer, suggesting these genes could be used for diagnostic purposes and therapeutic targeting.

Noncommunicable illnesses of the liver have been noted to tend toward increased mortality in many nations, particularly after the year nineteen, which marked the COVID-19 global health crisis, due to numerous cases of liver infections incurred from excessive alcohol, gas pollution and contaminated food, drugs and medications. According to the paper [22], to control the spread of liver disease, particularly Hepatitis C, which is becoming an epidemic, a collection of multimodal data sources has been created and used to develop classification and prediction models with machine learning techniques. The research employed algorithms such as decision trees, support vectors and others to stratify patients on the severity of liver-related illnesses and the degree of Hepatitis C. Particular attention was given to the Unconventional Consensus Classifier Algorithm, being a tool that was used to stratify the stages of the disease. The project also incorporated sophisticated data preprocessing techniques and feature engineering, enhancing the predictive power of the models. This way, medical practitioners can find practical solutions to improve the care and management of patients through early diagnosis and treatment scheduling to optimize the treatment effects.

Looking at Table 1, it is easy to conclude that machine learning can tremendously change the practice of hepatology. With complex mathematical models and extensive data, researchers can create efficient and dependable designs for predicting, categorizing, and treating patients with liver conditions in the early stages of the disease. Such designs assist physicians in making rational decisions, enhancing patient care, and alleviating the effects of liver illness.

Table 1: Summary of Literature Review

Study	**Objective**	**Methodology**	**Key Findings**
[8]	Assess the performance of ML algorithms for advanced liver fibrosis prediction.	Utilized serum biomarkers and clinical data, employed decision tree, genetic algorithm, and particle swarm optimization - multi-linear regression models	Age, platelet count, AST, and albumin levels were significant predictors of advanced fibrosis.
[9]	Assess a new hybrid AI-based classifier for chronic hepatitis prognosis	Utilized 19 biomarkers from 155 patients, employed LSVM and MLP	The hybrid method achieved 100% accuracy and AUROC

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

[10]	Develop individual predictive models of HCC risk using ML	Utilized clinical and pre-treatment laboratory information, employed Gaussian naive Bayes, XGBoost, random forest, and LASSO regression	XGBoost model outperformed conventional methods in predicting HCC risk
[11]	Assess the influence of IL28B gene polymorphism on the risk of hepatic fibrosis and compare ML and traditional scores.	Utilized decision tree and Auto-WEKA, compared APRI, FIB-4, and MLP	MLP outperformed APRI, FIB-4, and REP Tree in assessing advanced fibrosis
[12]	Use ML to improve hepatitis C diagnosis and prognosis	Utilized EHR data, employed SVM, MARS, BGLM, RF, and DT	ML can be helpful in the early detection of hepatitis C
[13]	Improve the accuracy of liver disease diagnosis using ML	Utilized ILPD, employed J48 and other algorithms	J48 algorithm provided better results in distinguishing between patients and non-patients
[14]	Diagnose hepatitis using a hybrid SVM-SA technique	Utilized machine learning repository, employed SVM-SA	Achieved 96.25% accuracy in hepatitis diagnosis
[15]	Create a model for HCC prediction using ML	Utilized 2D-SWE and B-mode ultrasound images, employed deep learning radiomics analysis	The HCC-R model achieved high accuracy in predicting HCC
[16]	Improve hepatitis B and C diagnostics using data mining	Utilized patient questionnaires, employed exploratory data analysis and predictive data mining	Identified relations between input attributes and diagnosis, obtained applicable decision rules
[17]	Identify genetic determinants of therapy outcomes in HCV infection	Utilized complete nucleotide sequence of HCV, employed Gini Index, Chi-squared analysis, and machine learning algorithms	Identified nucleotide features correlating with treatment outcomes
[18]	Evaluate survival rate among hepatitis patients using ML	Utilized serum hepatitis dataset, employed various ML algorithms	Compared the efficiency of different algorithms in predicting survival
[19]	Build a 28-day mortality predictive model for ACLF patients	Utilized clinical data from ACLF patients, employed LASSO regression and XGBoost	XGBoost model outperformed conventional scoring systems in predicting mortality
[20]	Understand the molecular mechanisms of HBV-induced HCC using	Utilized GEO database, employed differential and WGCNA, and deep learning	Identified essential genes (RACGAP1, ECT2, NDC80) associated with HBV-HCC

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024

	bioinformatics and deep learning.	techniques (Lasso, random forest, SVM)	and their potential as diagnostic markers and therapeutic targets
[21]	Understand the molecular mechanisms of HBV-induced HCC using bioinformatics and deep learning.	Utilized GEO database, employed differential and WGCNA, and deep learning techniques (Lasso, random forest, SVM)	Identified essential genes (RACGAP1, ECT2, NDC80) associated with HBV-HCC and their potential as diagnostic markers and therapeutic targets
[22]	Develop classification and prediction models for liver diseases using ML	Utilized multimodal data sources, employed decision trees, support vectors, and Unconventional Consensus Classifier Algorithm	Stratified patients based on disease severity and Hepatitis C stage, improved predictive power of models

Machine learning in hematology is an excellent addition to the early diagnosis, risk stratification, and treatment management of liver diseases. The studies reviewed in this chapter show how ML can find patterns or risk factors that other methods lack, from non-invasive prediction models of liver fibrosis to hybrid algorithms for predicting better prognoses of hepatitis and HCC. Nevertheless, all these encouraging advancements are countered by challenges such as the quality of the data, interpretability of the models, clinical validations of the models, etc. Towards the end of this article, we discuss the future research direction focusing on algorithm improvements, increasing the size of databases, and creating more powerful and flexible ML solutions for clinical practice. Overall, ML can be very useful in preventing and treating chronic liver diseases at their early stages.

3. Conclusion

To summarize, we present it as more than just a refinement; the use of machine learning (ML) in the diagnosis, prognosis, and treatment of chronic liver diseases is an improvement in hepatology. With the use of enormous amounts of data, ML has been able to put such algorithms into action, which are capable of recognizing intricate relationships that enhance liver disease detection and management. This review discusses published data about ML, emphasizing its various applications, such as improving diagnosis rates, prognosis, and, most importantly, treatment, which results in better healthcare delivery. Nevertheless, the redeeming feature of ML in the room for the ML tools while in the clinical stage is the fact that there are challenges. Issues concerning data integrity, how the model is used, and different liver diseases need to be resolved for practical efficacy and application of ML techniques within the populations studied. Moreover, there is an urgent call to implement transparent processes and rules on moral grounds in developing and using ML systems for healthcare purposes. It is imperative to stress that all these problems must be resolved through an intensive cooperative effort of clinicians, data scientists, and researchers to ensure that the ML-based tools are precise, safe to use, and obtained for the desired audience. As far as managing liver diseases with ML is concerned, the expectations for the future are very high. Additional research and innovativeness will enhance the existing approaches and include even more advanced models appropriate for the complex nature of liver diseases, making it possible to implement targeted therapy. With the growing strides made in ML technologies, there is no doubt that they will transform the care of chronic liver diseases, their diagnosis and treatment, thus improving the quality of care offered to patients and health outcomes. Through these and other efforts, the health sector can optimally exploit machine learning to deal with the global menace of chronic liver diseases.

References

[1] S. Feng et al., “Current Status and Analysis of Machine Learning in Hepatocellular Carcinoma,” J Clin Transl Hepatol, vol. 11, no. 5, p. 1184, Oct. 2023, doi: 10.14218/JCTH.2022.00077S.

[2] H. W. Lee, J. J. Y. Sung, and S. H. Ahn, “Artificial intelligence in liver disease,” J Gastroenterol Hepatol, vol. 36, no. 3, pp. 539–542, Mar. 2021, doi: 10.1111/JGH.15409.

[3] A. Abdulrahman, .. Z. M., A. M. Zaki, F. H. H.Rizk, M. M. Eid, and E.-S. M. EL EL-Kenawy, "Exploring Optimization Algorithms: A Review of Methods and Applications," Journal of Artificial Intelligence and Metaheuristics, vol. 7, no. 2, pp. 08–17, 2024, doi: 10.54216/JAIM.070201.

[4] M. Ai-Hasani, L. R. Sultan, H. Sagreiya, T. W. Cary, M. B. Karmacharya, and C. M. Sehgal, "Machine learning improves early detection of liver fibrosis by quantitative ultrasound radiomics," IEEE Int Ultrason Symp, vol. 2022, p. 10.1109/ius54386.2022.9957180, 2022, doi: 10.1109/IUS54386.2022.9957180.

[5] M. Li, X. Song, C. Xu, and C. Yu, "Response to the clinical significance and prognostic value of 3-mercaptopyruvate sulfurtransferase in hepatocellular carcinoma," Liver International, vol. 42, no. 7, pp. 1696–1696, Jul. 2022, doi: 10.1111/LIV.15280.

[6] G. Chongo and J. Soldera, "Use of machine learning models for the prognostication of liver transplantation: A systematic review," World J Transplant, vol. 14, no. 1, p. 88891, Mar. 2024, doi: 10.5500/WJT.V14.I1.88891.

[7] A. C. Filiberto, I. L. Leeds, and T. J. Loftus, "Editorial: Machine Learning in Clinical Decision-Making," Front Digit Health, vol. 3, p. 784495, Nov. 2021, doi: 10.3389/FDGTH.2021.784495/BIBTEX.

[8] S. Hashem et al., "Comparison of Machine Learning Approaches for Prediction of Advanced Liver Fibrosis in Chronic Hepatitis C Patients," IEEE/ACM Trans Comput Biol Bioinform, vol. 15, no. 3, pp. 861–868, May 2018, doi: 10.1109/TCBB.2017.2690848.

[9] L. Parisi, N. Ravichandran, and M. L. Manaog, "A novel hybrid algorithm for aiding prediction of prognosis in patients with hepatitis," Neural Comput Appl, vol. 32, no. 8, pp. 3839–3852, Apr. 2020, doi: 10.1007/S00521-019-04050-X/TABLES/5.

[10] Y. Xu et al., "Development of machine learning-based personalized predictive models for risk evaluation of hepatocellular carcinoma in hepatitis B virus-related cirrhosis patients with low levels of serum alpha-fetoprotein," Ann Hepatol, vol. 29, no. 6, p. 101540, Nov. 2024, doi 10.1016/J.AOHEP.2024.101540.

[11] H. I. Shousha, A. H. Awad, D. A. Omran, M. M. Elnegouly, and M. Mabrouk, "Data Mining and Machine Learning Algorithms Using IL28B Genotype and Biochemical Markers Best Predicted Advanced Liver Fibrosis in Chronic Hepatitis C," Jpn J Infect Dis, vol. 71, no. 1, pp. 51–57, 2018, doi: 10.7883/YOKEN.JJID.2017.089.

[12] A. Kukkar and J. K. Sandhu, "An Automatic Diagnosis System Based on Machine Learning Models for Predicting Hepatitis C from Blood Samples," Signals and Communication Technology, vol. Part F3355, pp. 91–108, 2024, doi: 10.1007/978-3-031-56818-3_6/FIGURES/10.

[13] S. Tanvir Fayaz, G. S. Tejanmayi, Y. Kanaka Ruthvi, S. Vijaya Shetty, S. U. Shenoy, and G. Bhat, "Prediction of Liver Patients Using Machine Learning Algorithms," Lecture Notes in Electrical Engineering, vol. 789, pp. 135–145, 2022, doi: 10.1007/978-981-16-1338-8_12/TABLES/4.

[14] J. S. Sartakhti, M. H. Zangooei, and K. Mozafari, "Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA)," Comput Methods Programs Biomed, vol. 108, no. 2, pp. 570–579, Nov. 2012, doi: 10.1016/J.CMPB.2011.08.003.

[15] J. Jin et al., "Deep learning radiomics model accurately predicts hepatocellular carcinoma occurrence in chronic hepatitis B patients: a five-year follow-up," Am J Cancer Res, vol. 11, no. 2, p. 576, 2021, Accessed: Oct. 27, 2024. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC7868753/

[16] A. Lukáčová, F. Babič, Z. Paraličová, and J. Paralič, "How to increase the effectiveness of the hepatitis diagnostics by means of appropriate machine learning methods," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9267, pp. 81–94, 2015, doi: 10.1007/978-3-319-22741-2_8/FIGURES/6.

[17] A. H. Kayvanjoo, M. Ebrahimi, and G. Haqshenas, "Prediction of hepatitis C virus interferon/ribavirin therapy outcome based on viral nucleotide attributes using machine learning algorithms," BMC Res Notes, vol. 7, no. 1, pp. 1–11, Aug. 2014, doi: 10.1186/1756-0500-7-565/TABLES/2.

[18] C. Vijayalakshmi and S. Pakkir Mohideen, "Survival prediction model with lion optimization for Hepatitis-B patients using supervised and unsupervised learning algorithm," Measurement: Sensors, vol. 32, p. 100958, Apr. 2024, doi: 10.1016/J.MEASEN.2023.100958.

[19] S. Qiu et al., "Predicting the 28-day prognosis of acute-on-chronic liver failure patients based on machine learning," Digestive and Liver Disease, Jul. 2024, doi: 10.1016/J.DLD.2024.06.029.

[20] Y. Anyin, L. Jianping, L. Mengru, Z. Hong, Z. Xulei, and W. Lianping, "Integrating bioinformatics and machine learning methods to analyze diagnostic biomarkers for HBV-induced hepatocellular carcinoma," Diagn Pathol, vol. 19, no. 1, pp. 1–10, Dec. 2024, doi: 10.1186/S13000-024-01528-8/TABLES/3.

[21] Y. Anyin, L. Jianping, L. Mengru, Z. Hong, Z. Xulei, and W. Lianping, "Integrating bioinformatics and machine learning methods to analyze diagnostic biomarkers for HBV-induced hepatocellular carcinoma," Diagn Pathol, vol. 19, no. 1, pp. 1–10, Dec. 2024, doi: 10.1186/S13000-024-01528-8/TABLES/3.

[22] A. Babu, S. Sri Krishna, K. S. Saran, and B. Aarthi, "Categorization and Forecasting of Hepatitis C Diagnosis via an Unconventional Consensus Classifier," International Journal of Advance Research, pp. 2454–132, 2024, Accessed: Oct. 27, 2024. [Online]. Available: https://www.ijariit.com

DOI: https://doi.org/10.54216/MOR.010104
Received: August 22, 2023 Revised: December 09, 2023 Accepted: January 09, 2024