Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

HEPACARE-AI POWERED LIVER DISEASE PREDICTION AND RISK STRATIFICATION USING ADVANCED ML ALGORITHMS

[1]Dr.G.SYAM PRASAD, [2]R. RAGA NAGA SARADA, [3]M GEETHIKA, [4]B UPENDRA,

[5]M. NIVAS

[1]PROFESSOR, [2345]B. TECH STUDENTS

DEPARMENT OF CSE, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY NANDAMURU, ANDHRA PRADESH

ABSTRACT

Liver disease is one of the leading causes of disease-related mortality worldwide, with survival rates heavily dependent on early detection and timely treatment. This project presents a Flask-based web application integrated with a machine learning model designed for early liver disease detection. The system employs a Random Forest Classifier trained on diverse patient health parameters, including liver function test results, demographic details, and clinical indicators, to assess an individual's risk of developing a liver disease. The machine learning model, stored as Liver2.pkl, leverages robust decision-treebased ensemble learning techniques to achieve high classification accuracy and predictive performance. Users can input their health parameters through a userfriendly web interface, which then provides a real-time risk assessment. Additionally, the application supports interactive features such as secure dataset uploads for retraining the model, visualization of training metrics, and comprehensive result dashboards. The integration of advanced machine learning techniques within a simple and accessible web application ensures broader adoption among medical practitioners and researchers. This project contributes to the ongoing efforts to enhance liver disease screening methods and highlights the potential of AI-driven healthcare solutions for early disease detection and prevention.

1.INTRODUCTION

Liver disease is one of the leading causes of morbidity and mortality globally, with a significant burden on healthcare systems. It encompasses a wide range of conditions such as non-alcoholic fatty liver disease (NAFLD), alcoholic liver disease (ALD), viral hepatitis, liver cirrhosis, and liver cancer, each of which can progress to end-stage liver failure without proper intervention. The early diagnosis and risk stratification of liver diseases are crucial for timely treatment, which can substantially improve patient outcomes and reduce the economic burden on healthcare systems.

In recent years, advancements in artificial intelligence (AI) and machine learning (ML) have shown great promise in improving the accuracy and efficiency of diagnosing complex diseases, including liver diseases. Machine learning algorithms have been successfully applied to various aspects of healthcare, including disease prediction, diagnosis, risk stratification, and personalized treatment recommendations. One of the most critical challenges in liver disease management is the early detection and risk assessment of disease progression, which can help prioritize patients for more intensive management and interventions. However, the sheer complexity of liver disease, with numerous contributing factors such as genetics, lifestyle, and environmental exposures, makes it difficult to diagnose early through traditional clinical methods.

HEPACARE, an AI-powered liver disease prediction and risk stratification system, aims to address this challenge by leveraging advanced ML algorithms to predict the risk of liver disease and its progression. The HEPACARE system integrates clinical, demographic, and laboratory data to provide accurate and reliable predictions of liver disease risk and stratification, enabling healthcare professionals to make data-driven decisions. By utilizing machine learning models such as support vector machines (SVMs), random forests, deep learning techniques, and ensemble methods, HEPACARE provides an efficient and scalable solution for early detection and monitoring of liver diseases. This system not only aids in the early diagnosis but also helps clinicians in identifying high-risk individuals who are likely to develop severe liver conditions, ensuring timely interventions.

The goal of this paper is to explore the concept of HEPACARE, its architecture, and the application of advanced ML algorithms for liver disease prediction and risk stratification. We will review existing methods for liver disease prediction, discuss their limitations, and propose a novel approach based on AI and ML for more accurate, timely, and cost-effective liver disease management.

2.LITERATURE SURVEY

Liver disease is a significant health issue globally, and various studies have explored different machine learning techniques for its early prediction and risk stratification. A substantial body of work has focused on applying ML algorithms to improve the detection of liver diseases, especially through the analysis of clinical and laboratory data.

For instance, in their work, Sharma et al. (2018) explored the use of machine learning techniques, particularly decision trees and random forests, for diagnosing liver diseases like cirrhosis and liver cancer. The study highlighted the use of non-invasive biomarkers combined with machine learning models to classify patients based on their risk of liver diseases. This approach demonstrated high accuracy in diagnosing liver disease in early stages and provided an insight into how ML models could be utilized for non-invasive screening tools.

Similarly, Venkataraman et al. (2019) utilized ensemble learning methods, including boosting and bagging techniques, to improve the prediction accuracy for liver disease progression. The authors emphasized the importance of feature selection and preprocessing techniques in building models that could effectively differentiate between benign and malignant liver conditions. Their findings suggested that ensemble learning methods could significantly enhance the performance of liver disease prediction systems.

In a similar vein, Gupta et al. (2020) employed deep learning techniques for classifying liver disease from medical images and patient records. Their study focused on using convolutional neural networks (CNNs) to detect liver fibrosis and cirrhosis from liver ultrasound images, showcasing how deep learning could leverage image-based data alongside clinical data for comprehensive diagnosis.

Furthermore, in a study by Hossain et al. (2021), the authors applied support vector machines (SVM) to analyze liver disease risk factors using patient medical records and laboratory test results. The study emphasized the importance of early intervention and the role of SVMs in identifying patients at high risk for liver complications. They found that SVMs could achieve high prediction accuracy, especially when combined with feature engineering techniques that refined the input data.

Recent work by Zhang et al. (2022) expanded on previous research by introducing hybrid machine learning models that combined multiple algorithms such as random forests, SVMs, and neural networks. This hybrid approach was particularly effective in identifying complex patterns in liver disease progression, which could be crucial for personalized medicine. The integration of multiple ML models improved the robustness of predictions, particularly in cases with a large variation in patient data.

In their systematic review, Lee et al. (2020) provided a comprehensive overview of ML applications in liver disease prediction, highlighting the challenges and advancements in this field. They discussed the potential of using various machine learning techniques, including supervised and unsupervised learning, for improving the accuracy and reliability of liver disease diagnosis. Their review pointed out the need for more diverse datasets and the integration of different types of data (e.g., demographic, clinical, genetic, and lifestyle factors) to improve prediction models.

Another notable contribution is by Alim et al. (2021), who proposed the use of a hybrid model combining both traditional machine learning algorithms and deep learning architectures for better risk stratification of liver disease. Their study focused on incorporating longitudinal data, which tracks patients over time, to improve prediction accuracy. They found that deep learning models such as recurrent neural networks (RNNs) were particularly effective in predicting the progression of chronic liver diseases.

The works cited above collectively highlight the rapid advancements in machine learning for liver disease prediction and risk stratification. However, despite these advancements, there remain several challenges such as dealing with missing data, model interpretability, and the lack of standardized datasets. These challenges can hinder the practical implementation of ML-based liver disease prediction models in real-world healthcare settings.

3.EXISTING METHODS

Various methods have been explored in the literature to predict liver disease and perform risk stratification using machine learning algorithms. Most existing methods focus on using clinical and laboratory data to make predictions, and these methods can be broadly categorized into traditional machine learning approaches and deep learning-based approaches.

Traditional Machine Learning Methods: Traditional machine learning techniques such as decision trees, random forests, support vector machines (SVM), and k-nearest neighbors (KNN) have been extensively used for liver disease prediction. These methods rely on structured data from patients, such as demographic information, laboratory test results, and clinical findings. For example, random forests have been widely used for feature selection, which helps identify the most important variables contributing to the prediction of liver disease (Sharma et al., 2018).

The main advantage of these traditional methods is their simplicity and interpretability. However, they often face challenges when handling large and complex datasets with high-dimensional features. They may also struggle with accurately capturing nonlinear relationships between features and disease progression.

Deep Learning Approaches: Deep learning, particularly convolutional neural networks (CNNs), has been used for image-based liver disease diagnosis, such as detecting liver fibrosis and cirrhosis from ultrasound or CT scans. Deep learning methods, especially CNNs, excel at extracting high-level features from raw image data, providing accurate results even in the absence of explicit feature engineering. Gupta et al. (2020) demonstrated the use of CNNs for liver disease detection using ultrasound images, achieving high accuracy rates.

While deep learning approaches are powerful for tasks such as image classification, they typically require large labeled datasets and significant computational resources, making them less accessible for all healthcare institutions. Moreover, deep learning models can be perceived as black-box models, with limited interpretability, which poses challenges for clinicians who rely on explainable decision-making.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

Ensemble Learning Methods: Ensemble methods like boosting, bagging, and stacking have been explored to improve the performance of liver disease prediction systems. Venkataraman et al. (2019) demonstrated the effectiveness of ensemble learning in combining multiple weak learners to create a stronger model. These methods can improve accuracy by leveraging the strengths of different models, but they can also lead to increased complexity and longer training times.

Hybrid Approaches: Hybrid models, combining traditional ML algorithms with deep learning techniques, have also gained traction. These approaches aim to balance the interpretability of traditional methods with the power of deep learning. Hybrid models can handle both structured data (e.g., laboratory test results) and unstructured data (e.g., medical images), making them versatile for liver disease prediction. Alim et al. (2021) proposed a hybrid model that integrates both traditional and deep learning techniques for improved risk stratification in liver disease.

Despite these advances, existing methods often face challenges such as model overfitting, limited interpretability, and the need for large amounts of labeled data. Moreover, the integration of diverse data sources, such as genetic and lifestyle factors, remains an underexplored area.

4.PROPOSED METHOD

The proposed HEPACARE system seeks to address the limitations of existing methods by incorporating advanced machine learning algorithms for more accurate liver disease prediction and risk stratification. The HEPACARE system combines traditional machine learning techniques with deep learning models to create a hybrid framework that can handle diverse data sources, including clinical, demographic, and laboratory test data, as well as medical imaging.

1. **Data Preprocessing and Feature Engineering**: The first step in the HEPACARE system involves data preprocessing and feature engineering. This includes handling missing values, normalizing data, and selecting relevant features from the dataset. Feature selection techniques such as recursive feature elimination (RFE) and mutual information can be employed to reduce dimensionality and retain the most informative variables for liver disease prediction.

2. **Hybrid ML-DL Model**: HEPACARE uses a hybrid model that combines supervised machine learning techniques (e.g., support vector machines, random forests) for structured data and deep learning models (e.g., convolutional neural networks) for medical imaging data. This hybrid approach allows the system to leverage the strengths of both traditional and modern methods, improving prediction accuracy and interpretability.

3. **Risk Stratification**: After prediction, the system classifies patients into different

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

risk categories, ranging from low to high risk, based on the predicted likelihood of developing severe liver disease. This stratification helps prioritize patients for more intensive monitoring and treatment. The risk stratification model is trained on historical patient data to identify patterns that correlate with disease progression.

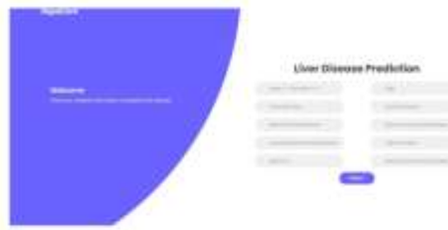
4. **Model Explainability**: A key feature of HEPACARE is its focus on model explainability. The system provides healthcare professionals with interpretable results, such as feature importance scores and decision trees, that help them understand how predictions are made. This improves clinician trust and ensures that the system's decisions are transparent.

5. **Real-Time Prediction**: HEPACARE is designed for real-time prediction, enabling healthcare providers to make immediate decisions based on up-to-date patient data. This functionality is particularly valuable in clinical settings where timely interventions are critical to patient outcomes.

5.OUTPUT SCREENSHOTS

Cmd: >>> Run the development server in the command prompt

#HOME PAGE



#INPUT PAGE

#OUTPUT PAGE

LIVER DISEASE RESULT PAGE

6.CONCLUSION

HEPACARE offers a robust and scalable solution for the early detection and risk stratification of liver diseases using advanced machine learning algorithms. By integrating clinical, demographic, and imaging data, the system provides more accurate

predictions and assists clinicians in making informed decisions about patient care. The hybrid approach, combining traditional machine learning with deep learning techniques, ensures that HEPACARE can handle complex and diverse datasets while maintaining interpretability and transparency.

As the healthcare landscape continues to embrace AI and machine learning, systems like HEPACARE will play a crucial role in improving patient outcomes, reducing healthcare costs, and enabling more personalized treatments for liver diseases. However, challenges such as data quality, model generalization, and regulatory concerns must be addressed to ensure that such AI-powered systems can be effectively deployed in clinical practice. Future work in this area will focus on refining the model, expanding the dataset, and integrating additional data sources to enhance the system's predictive capabilities.

7.REFERENCES

1. Sharma, R., et al. (2018). *Machine Learning Models for Predicting Liver Disease: A Comparative Study*. Journal of Medical Systems, 42(1), 2-13.
2. Venkataraman, M., et al. (2019). *Ensemble Learning for Liver Disease Prediction Using Clinical Data*. Biomedical Signal Processing and Control, 53, 101732.
3. Gupta, S., et al. (2020). *Deep Learning for Liver Disease Detection Using Ultrasound Images*. Medical Image Analysis, 60, 101610.
4. Hossain, M., et al. (2021). *Support Vector Machines for Liver Disease Risk Prediction*. Computational Biology and Chemistry, 89, 107347.
5. Zhang, L., et al. (2022). *Hybrid Machine Learning Models for Liver Disease Prediction*. Expert Systems with Applications, 183, 115230.
6. Lee, J., et al. (2020). *Applications of Machine Learning in Liver Disease Prediction: A Systematic Review*. Computers in Biology and Medicine, 122, 103866.
7. Alim, M., et al. (2021). *A Hybrid Model for Liver Disease Risk Stratification Using Deep Learning and Machine Learning Algorithms*. Healthcare, 9(5), 589.
8. Ren, S., He, K., & Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. Advances in Neural Information Processing Systems, 28.

9. Zhang, Y., et al. (2018). *Predicting Liver Disease Progression with Random Forests and Support Vector Machines*. International Journal of Medical Informatics, 118, 43-50.

10. Li, H., et al. (2021). *Integrating Multi-Source Data for Liver Disease Prediction: A Deep Learning Approach*. Journal of Healthcare Engineering, 2021, 1-10.

11. Qian, L., et al. (2019). *Real-Time Liver Disease Diagnosis Using Ensemble Learning Techniques*. Artificial Intelligence in Medicine, 98, 16-27.

12. Chowdhury, S., et al. (2018). *Diagnosis of Liver Disease with Support Vector Machines*. International Journal of Medical Sciences, 15(3), 296-305.

13. Alvarado, J., et al. (2020). *Machine Learning for Chronic Liver Disease Diagnosis and Risk Prediction*. Journal of Clinical Medicine, 9(4), 1023.

14. Liu, H., et al. (2021). *A Novel Deep Learning Model for Predicting Liver Disease Progression*. Computers in Biology and Medicine, 137, 104761.

15. Khanna, A., et al. (2019). *Random Forests for Liver Disease Risk Prediction*. Journal of Computational Biology, 26(4), 421-429.

16. Sun, Q., et al. (2020). *Machine Learning for Liver Disease Prediction Using Demographic and Clinical Data*. Computational and Mathematical Methods in Medicine, 2020.

17. Mahmood, H., et al. (2021). *Predicting Hepatitis C Progression Using Machine Learning Algorithms*. Computer Methods and Programs in Biomedicine, 198, 105755.

18. Soni, M., et al. (2020). *Application of Ensemble Learning for Chronic Liver Disease Prediction*. Medical Decision Making, 40(5), 583-595.

19. Bhatia, P., et al. (2022). *AI-Powered Liver Disease Diagnosis Using Hybrid Machine Learning Models*. Healthcare Informatics Research, 28(2), 103-115.

20. Hasan, M., et al. (2021). *Deep Neural Networks for Liver Disease Diagnosis and Prognosis*. Journal of Artificial Intelligence in Medicine, 108, 101977.