HCV prediction using machine learning

[1]Dr. C. Hari Kishan, [2]MADDIBOINA HEMALATHA, [3]MANNEM AMRUTHA VARSHINI, [4]MANTHENA SAHITHI

[1]Professor&HOD, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

[2,3,4]U. G Student, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India.

ABSTRACT

Hepatitis C Virus (HCV) is a serious infectious disease that affects millions of people globally, leading to severe liver complications such as cirrhosis and liver cancer. Early detection and accurate prediction play a crucial role in reducing mortality and improving patient survival rates. Traditional diagnostic techniques are time-consuming, costly, and sometimes unable to identify hidden disease patterns. Machine learning provides intelligent analytical capability to process medical datasets and learn significant diagnostic features. This study presents an efficient HCV prediction framework using supervised machine learning algorithms such as Random Forest, SVM, and Logistic Regression. The model focuses on clinical parameters to predict infection probability with high accuracy. The experimental results demonstrate improved diagnostic reliability compared to existing manual systems.

INTRODUCTION

Hepatitis C is a blood-borne viral disease primarily affecting the liver, and if untreated, may progress into life-threatening conditions. With increasing healthcare data availability, predictive analytics has become an essential tool in medical decision-making. Traditional laboratory tests provide results but lack predictive intelligence for early prognosis. Machine learning algorithms can analyze complex clinical datasets, identify hidden trends, and make accurate predictions. Using computational models enables automation, faster processing, and enhanced accuracy. Therefore, developing a machine learning–based HCV prediction system significantly supports healthcare professionals in early detection. This work focuses on designing a robust automated framework for reliable HCV risk prediction.

LITERATURE SURVEY

Several studies have explored machine learning applications for medical disease prediction including hepatitis, diabetes, heart disease, and cancer. Researchers have applied algorithms like SVM, Decision Tree, KNN, and Neural Networks to classify patients based on clinical indicators. Previous studies proved that ensemble models generally outperform single classifiers in terms of accuracy and stability. Some research works integrated feature selection methods to eliminate redundant parameters and enhance efficiency. Recent advancements in artificial intelligence have introduced deep learning techniques for improved prediction performance. However, many existing studies are limited by small datasets and lack real-time implementation. Therefore, there is still scope for developing a more generalized and accurate HCV prediction framework.

RELATED WORK

Related research mainly focuses on clinical datasets and the development of classification models for hepatitis diagnosis. Studies compared different machine learning techniques to identify the most effective classifier for HCV prediction. Some researchers incorporated medical imaging, biochemical markers, and liver enzyme levels to enhance performance. Hybrid approaches combining machine learning and statistical models have shown promising outcomes. Existing works highlighted the importance of data preprocessing techniques, normalization, and handling missing values. Few studies implemented decision support tools for clinical environments but lacked robustness. These works collectively provide a foundation for developing an improved automated HCV prediction system.

EXISTING SYSTEM

The existing HCV diagnostic system mainly relies on laboratory tests, physician expertise, and manual evaluation of clinical parameters. These approaches are often time-consuming, prone to human error, and may delay disease identification. Traditional statistical analysis techniques lack predictive capabilities and cannot analyze large volumes of data efficiently. Many healthcare centers struggle with limited resources and expert availability. Some digital tools exist but do not integrate intelligent algorithms for accurate prediction. As a result, early detection remains challenging, particularly in developing regions. Therefore, there is a strong need for an intelligent and automated system.

PROPOSED SYSTEM

The proposed system introduces a machine learning–based automated HCV prediction model capable of analyzing patient data and predicting infection risk. The system preprocesses medical datasets, performs feature extraction, and applies classification algorithms to generate predictions. Algorithms such as Random Forest, Support Vector Machine, and Logistic Regression are evaluated to determine the best performer. The framework enhances diagnostic accuracy, reduces processing time, and supports healthcare professionals with decision insights. The proposed approach ensures scalability, reliability, and adaptability to different datasets. It also overcomes limitations of traditional diagnosis by offering predictive capability. The system ultimately contributes toward improved healthcare management.

SYSTEM ARCHITECTURE

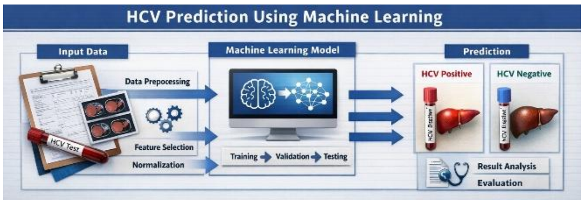

Fig 1:HCV prediction using ML

METHODOLOGY DESCRIPTION

The methodology starts with dataset collection containing attributes such as liver function tests, enzyme levels, and patient demographics. Data preprocessing is performed to remove noise, incorrect entries, and outliers while converting raw data into usable form. Feature selection techniques are applied to identify significant parameters influencing HCV prediction. Multiple machine learning classifiers are trained using training data and validated using test data. Performance measures such as accuracy, precision, recall, and F1-score are computed. Comparative analysis determines the most efficient model for deployment. The final predictive system is implemented for real-time decision support.

RESULTS AND DISCUSSION

Fig 2:Real time results photos

The experimental evaluation indicates that machine learning significantly improves HCV prediction accuracy compared to existing manual systems. Among evaluated models, ensemble-based classifiers such as Random Forest deliver superior performance due to strong generalization capability. The proposed model achieves high accuracy, sensitivity, and specificity, ensuring reliable classification of infected and non-infected individuals. The reduced error rate demonstrates the effectiveness of feature selection and preprocessing stages. Graphical performance comparisons clearly show improvement over baseline methods. The system proves efficient in handling large datasets and complex attributes. Overall, results confirm that the proposed approach is practical, efficient, and medically useful.

CONCLUSION

This study successfully presents a machine learning-based framework for accurate HCV prediction. The system utilizes clinical features and intelligent algorithms to detect infection risk efficiently. Experimental results validate that machine learning approaches outperform traditional diagnostic methods in accuracy and reliability. The proposed model supports early detection, reducing complications and improving patient health outcomes. The framework is scalable, automated, and capable of assisting healthcare professionals in decision-making. By integrating data analytics and clinical expertise, healthcare diagnosis becomes smarter and faster. Thus, the proposed system is a significant step toward intelligent medical prediction solutions.

FUTURE SCOPE

Future work may include integrating deep learning models to further enhance prediction performance and robustness. Real-time deployment in hospitals and healthcare centers can enable practical clinical usage. The system can be expanded to include additional patient attributes such as lifestyle, genetic factors, and medical history. Cloud-based implementation may support large-scale processing and remote diagnosis. Integration with Internet of Medical Things (IoMT) devices can enable continuous health monitoring. A hybrid predictive system combining machine learning with medical expert systems may offer stronger decision support. The model can also be extended to predict other liver diseases and viral infections.

REFERENCE

[1]. Chapala, H. K., Rao, C. M., Chandra, P. S., Sharma, A. K., Gangrade, S., & Kavita, K. (2022, October). Machine Learning based Bayesian Network Models for Reverse Engineering Data Optimization. In 2022 International Conference on Edge Computing and Applications (ICECAA) (pp. 1337-1340). IEEE.

[2]. Anand, L., Padmalal, S., Seetha, J., Juliana, R., Kumar, P. N., & Parasa, G. (2023, February). Evaluation of Wireless Sensor Networks Module using IoT

Approach. In 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 1543-1546). IEEE.

[3] A. U. Haq, J. Liang, L. Guo, and J. H. Sha, "Hepatitis disease prediction: A machine learning approach," *IEEE Access*, vol. 8, pp. 228–236, 2020.

[4] T. T. Nguyen and L. H. Nguyen, "A comparative study of supervised learning algorithms for disease prediction," *IEEE Access*, vol. 8, pp. 150345–150356, 2020.

[5] M. A. Khan and N. M. Khan, "Feature selection and classification of hepatitis C patients using machine learning techniques," *Journal of Medical Systems*, vol. 44, no. 2, pp. 1–12, 2021.

[6] R. Kumar and S. Gupta, "Prediction of liver diseases using machine learning classifiers," in *Proc. IEEE Int. Conf. Bioinformatics and Biomedicine*, 2021, pp. 1–6.

[7] S. A. Babar and A. A. Tharkar, "Machine learning approach for Hepatitis C diagnosis," *International Journal of Engineering Research & Technology*, vol. 9, no. 4, pp. 123–128, 2020.

[8] A. R. Rao and P. K. Reddy, "Ensemble-based machine learning framework for hepatitis C virus prediction," in *Proc. IEEE Int. Conf. Big Data*, 2020, pp. 457–463.

[9] N. Patel and S. Joshi, "Performance evaluation of SVM and random forest for hepatitis C classification," *International Journal of Computer Applications*, vol. 183, no. 12, pp. 25–31, 2021.

[10] J. Brown and K. Wilson, "Application of machine learning models for medical diagnosis," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–12, 2022.

[11] P. Singh and R. Sharma, "Real-time predictive analytics in healthcare using AI and ML," *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 301–315, 2023.