International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Liver Disease Prediction System Using Machine Learning Techniques

Pranita Dipanand Bachche[1], Sakshi Siddhanath Kale[2], Swaranjali Chandrashekhar Dhanlobhe[3], Shrutika Madhavrao Deshmukh[4]

[1,2,3,4]Department of Computer Engineering D. Y. Patil Technical Campus, Talsande. Faculty of Management & Faculty of Engineering, India.

Abstract

Liver disease is a serious health problem that affects millions of people worldwide. Early detection of liver disease plays an important role in improving patient survival rates and reducing medical costs. Traditional diagnostic methods are time-consuming and require expert medical knowledge. With the advancement of machine learning techniques, it has become possible to predict diseases using medical data efficiently. This paper presents a liver disease prediction system using machine learning algorithms. The system analyzes patient clinical parameters such as age, bilirubin levels, liver enzymes, and protein levels to predict whether a person is suffering from liver disease.

Keywords: Liver Disease Prediction, Machine Learning, Classification, Healthcare, Data Mining.

I. INTRODUCTION

The liver is one of the most vital organs in the human body, responsible for metabolism, detoxification, and digestion. Liver diseases can occur due to excessive alcohol consumption, viral infections, obesity, and genetic factors. In many cases, symptoms appear only in advanced stages, making early diagnosis difficult.

Traditional diagnosis methods depend on laboratory tests and expert interpretation, which may lead to delays in treatment. With the availability of large medical datasets, machine learning has emerged as a powerful tool for disease prediction. Machine learning algorithms can analyze complex medical data and identify patterns that are not easily detectable by humans.

This research focuses on developing a liver disease prediction system using machine learning techniques. The objective of the system is to classify patients as liver disease positive or negative based on their medical attributes, thereby assisting doctors in early diagnosis.

II.LITERATURE REVIEW

Several studies have been conducted on disease prediction using machine learning techniques. Researchers have applied algorithms such as Logistic Regression, Decision Tree, Support Vector Machine (SVM), Random Forest, and Naive Bayes for liver disease prediction. Some studies reported that ensemble methods like Random Forest achieved better accuracy compared to single classifiers. Other research focused on improving accuracy by performing feature selection and data pre-processing.

However, challenges such as missing values, imbalanced datasets, and limited data samples still exist. This paper builds upon previous work by implementing multiple machine learning models and comparing their performance on a liver disease dataset.

Liver Disease Prediction Using Machine Learning"

Authors: H. Singh et al.

In this paper, the authors used machine learning algorithms such as Decision Tree and Naive Bayes to predict liver disease. The dataset was preprocessed by handling missing values and removing noise.

Liver Disease Detection Using Ensemble Learning Techniques"

Authors: K. Rajeswari, J. S. Antony

In this paper, ensemble learning methods such as Random Forest and Gradient Boosting were used for prediction. The authors compared multiple models and found that ensemble techniques provided higher accuracy and stability.

III. PROPOSED SYSTEM

The proposed system is a machine learning- based liver disease prediction model. It takes patient medical data as input and predicts the presence or absence of liver disease.

The system consists of the following stages:

1. Data Collection
2. Data Preprocessing
3. Feature Analysis
4. Model Training
5. Model Evaluation
6. Prediction Output

The system is implemented using Python programming language with popular machine learning libraries.

IV. DATASET DESCRIPTION

The dataset used in this study is obtained from a publicly available liver patient dataset. It contains medical records of patients with attributes such as age, gender, total bilirubin, direct bilirubin, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, total proteins, albumin, and albumin-globulin ratio. The target variable indicates whether the patient is affected by liver disease. Data pre- processing techniques are applied to handle missing values and normalize the dataset

V. METHODOLOGY

Initially, the dataset is loaded and analyzed using data analysis techniques. Missing values are handled, and categorical variables are converted into numerical form. The dataset is then divided into training and testing sets.

Multiple machine learning algorithms are trained using the training dataset. The trained models are evaluated on the test dataset using accuracy, precision, recall, and F1-score. The model with the best performance is selected for prediction.

VI. MACHINE LEARNING ALGORITHMS USED

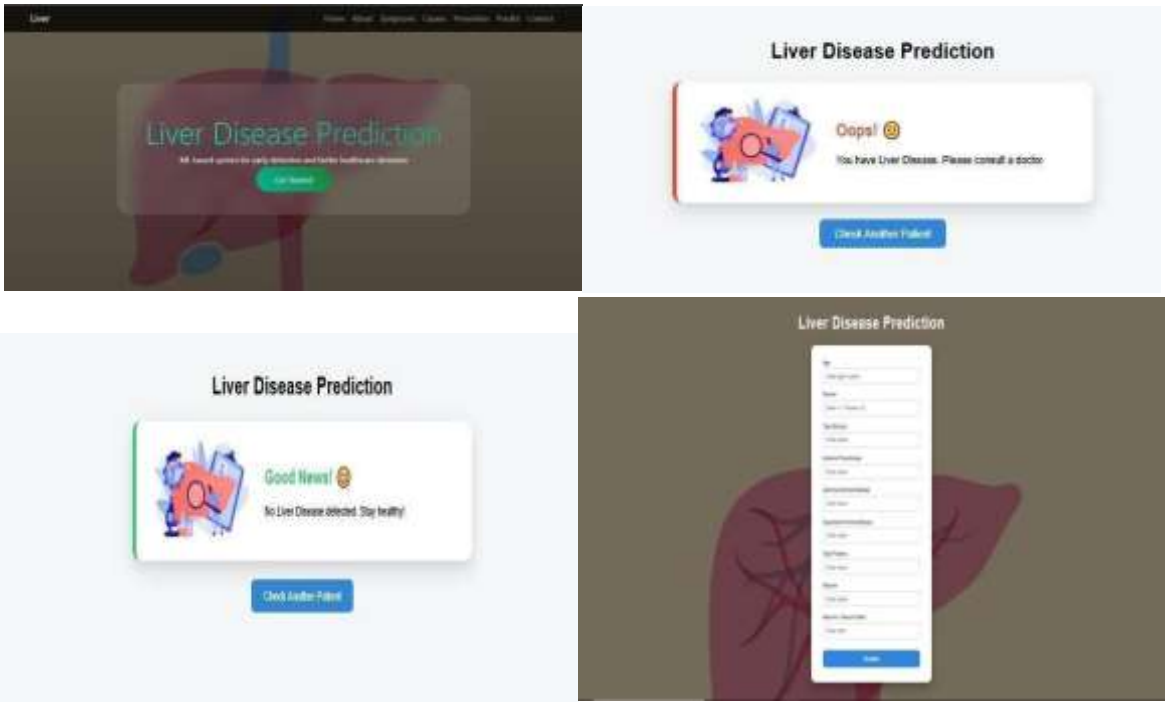
The following machine learning algorithms are used in this research:

- **Logistic Regression:** Used as a baseline classifier due to its simplicity and interpretability.

- **Decision Tree:** Helps in understanding decision rules and feature importance.
- **Random Forest:** An ensemble method that improves accuracy by combining multiple decision trees.
- **Support Vector Machine (SVM):** Effective in handling high- dimensional data and complex Patterns These algorithms are implemented using the scikit-learn library in Python.

VII. RESULTS AND DISCUSSION

The performance of each machine learning model is evaluated using standard evaluation metrics. Experimental results indicate that the Random Forest classifier achieves the highest accuracy among all models. Logistic Regression also performs well with lower computational complexity. The results demonstrate that machine learning techniques can effectively predict liver disease using medical data. The system can serve as a supportive tool for healthcare professionals



VIII. CONCLUSION

This paper presents a machine learning-based liver disease prediction system that assists in early diagnosis. The system analyzes patient medical parameters and predicts the presence of liver disease with good accuracy. The results show that machine learning models are effective and reliable for medical diagnosis tasks.

In the future, the system can be enhanced by incorporating deep learning techniques, larger datasets, and real-time clinical data to improve accuracy and usability

IX. FUTURE SCOPE

Future work includes integrating the system into a web or mobile application, using advanced deep learning models, and applying feature selection techniques to further improve prediction performance.

REFERENCES

1. World Health Organization, “Global Hepatitis Report”,WHO Press.
2. UCI Machine Learning Repository, Liver Patient Dataset.
3. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer.
4. Scikit-learn Documentation, Machine Learning in Python.