PREDICTION OF LIVER DISEASE USING MACHINE LEARNING ALGORITHMS

Guide : Prof . R A Manikandan

Professor

School of Engineering

Computer Science-AI&ML

Malla Reddy University ,India.

K.Siddartha
B-Tech
School of Engineering
Malla Reddy University
2111cs020523@mallareddyuniversity.ac.in

K,Sindhu
B-Tech
School of Engineering
Malla Reddy University
2111cs020524@mallareddyuniversity.ac.in

B.Sai Varun Reddy
B-Tech
School of Engineering
Malla Reddy University
2111cs020525@mallareddyuniversity.ac.in

N.Sindhu
B-Tech
School of Engineering
Malla Reddy University
2111cs020526@mallareddyuniversity.ac.in

A.Harish
B-Tech
School of Engineering
Malla Reddy University
2111cs020527@mallareddyuniversity.ac.in

B.Siva Prasad
B-Tech
School of Engineering
Malla Reddy University
2111cs020528@mallareddyuniversity.ac.in

ABSTRACT

The escalating prevalence of liver diseases globally underscores the need for effective diagnostic tools and predictive models. This research addresses the increasing incidence of liver diseases by leveraging machine learning algorithms to predict the risk of liver disease based on various factors such as blood test results, alcohol consumption, and dietary habits. Using Python programming language and the Scikit-learn library, a predictive model is developed employing classification algorithms including Logistic Regression and K-Nearest Neighbors. The system prompts users to input their blood test details, and the model, trained on a diverse dataset, provides a risk assessment for liver disease. The study compares the performance of these algorithms to identify the most accurate model for risk prediction. The developed system holds promise for early detection and intervention in individuals at risk of liver diseases, contributing to improved public health outcomes.

INTRODUCTION

1 . Background Information:

Liver diseases contribute significantly to global mortality, with the alarming prevalence of liver failures, especially among the high-risk population in India. Given the deceptive nature of liver functionality, early detection becomes imperative to prevent serious damage, considering the anticipated projection that India might become the World Capital for Liver Diseases by 2025. This research addresses the urgency of early diagnosis by harnessing the power of machine learning (ML). Utilizing diverse ML algorithms such as Logistic Regression and Support Vector Machines, we develop a predictive model to identify individuals at risk of liver diseases well in advance. The study highlights the pivotal role of ML in providing timely interventions, emphasizing its potential to mitigate the

escalating impact of liver diseases and improve public health outcomes.

2 . Objective

Early Detection Enhancement: The primary objective of this research is to enhance the early detection of liver diseases through the application of innovative technologies. By leveraging data mining tools and machine learning algorithms, the study aims to develop a robust predictive model capable of identifying potential liver disease cases at an early stage, thereby reducing complications and improving patient outcomes.

Technology-Aided Liver Cirrhosis Management: The research seeks to contribute to the alleviation of delays in the detection and treatment of liver cirrhosis by proposing an advanced technological solution. The envisioned system is designed to be user-friendly and accessible, offering an efficient means for both patients and healthcare professionals to facilitate timely intervention and management of liver diseases.

Building on Previous Research Contributions: This study aims to build upon and contribute to the existing body of knowledge in the field of liver disease prediction. By advancing the works of previous researchers, the objective is to refine and extend current methodologies, incorporating novel insights and methodologies to further enhance the accuracy and reliability of liver disease prediction models.

Public Health Impact: The overarching goal is to make a meaningful contribution to public health by developing a scalable and effective tool for early detection and management of liver diseases. The research aims to provide valuable insights that can inform healthcare policies and practices, ultimately reducing the global burden of liver diseases and associated mortality rates.

3. Dataset

The dataset employed in this proposed system is integral to the application of machine learning concepts for liver disease prediction. The dataset includes crucial parameters obtained from blood test reports, encompassing age, gender, and various biochemical markers such as total Bilirubin, direct Bilirubin, Alkaline Phosphatase, Alamine Amino transferase, Aspartate Aamino transferase, total proteins, albumin, and A/G ratio. These attributes collectively form the input features required for training and testing machine learning models. The system prompts users to input their details, mimicking a blood test report, and subsequently utilizes the training dataset to train various models. The model exhibiting the highest accuracy during testing becomes the most accurate and is employed for predicting the final result, categorizing individuals as either at risk or not at risk of liver disease. The evaluation of results is conducted using metrics such as accuracy and a confusion matrix, ensuring a comprehensive assessment of the model's performance. This dataset structure aligns with the requirements of the machine learning algorithms, enabling effective training and validation of the predictive models.

LITERATURE SURVEY

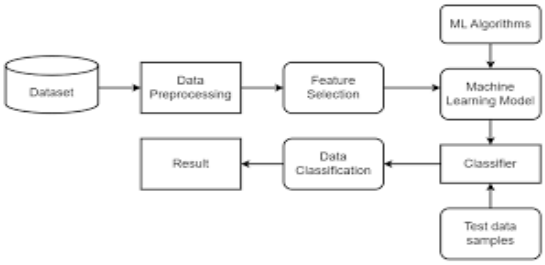
Logistic Regression.

Logistic Regression is a statistical model used for binary classification tasks. It employs a logistic function to estimate parameters and model the probability of a binary outcome. The algorithm computes the log odds for one outcome based on independent variables, converting them into probabilities. Logistic Regression is versatile and widely used, particularly for predicting liver disease risk, providing insights into relationships between input features and outcomes**.**

K-Nearest Neighbor

K-Nearest Neighbors (KNN) is an intuitive machine learning algorithm for classification and regression. Its non-parametric, lazy learning approach defers model construction, making it versatile. Using distance metrics like Euclidean distance, KNN classifies data points based on the majority class of their k nearest neighbors. Emphasizing simplicity and adaptability, KNN holds promise for predicting liver disease risk from clinical and biochemical parameters.

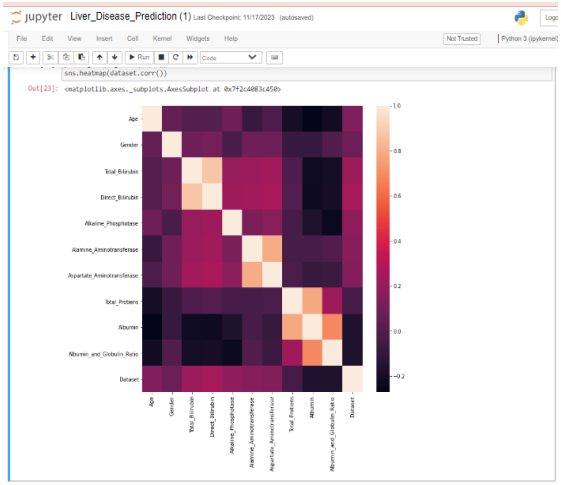
ARCHITECTURE



Accuracy & Precision for ML Methods

Model	Accuracy	Precision
Logistic Regression	68%	72%
K-Nearest Neighbors	68%	73%

Correlation graph between each attribute



DESCRIPTION OF MODULES:

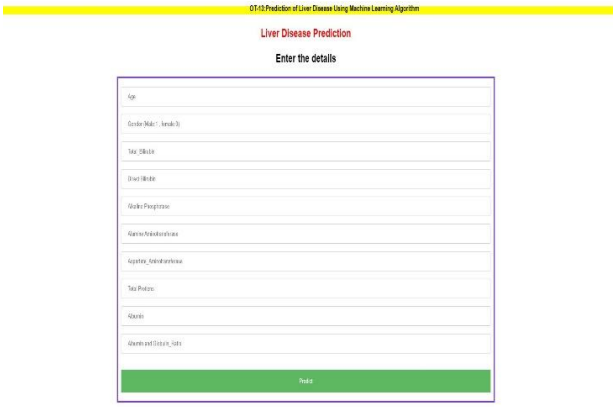
Modules present in this Liver disease prediction system project are:

1.User Module

- This module serves as the entry point for users, allowing them to interact with the system.
- Users, upon visiting the web application, input relevant details from their blood test reports as required by the application.
- The user interface prompts the submission of these details, facilitating seamless interaction.
- Following the submission, users can access and view the prediction results generated by the system.

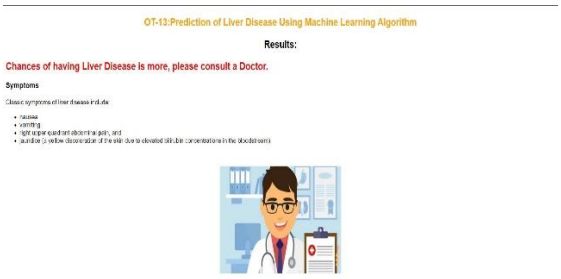
2. Prediction Module

- This module is responsible for constructing the classification model essential for disease prediction.
- Users input relevant features and attributes through the User Module, which are then processed by the Prediction Module.
- The Prediction Module employs machine learning algorithms to build and train the classification model based on the provided input.
- Upon completion of the training process, the model is utilized for predicting the likelihood of disease, providing valuable insights into potential health risks.
-



UI of Liver Disease Predictor

Test case screenshot



Output

CONCLUSION.

In conclusion, our project, introduces an efficient and user-friendly system with a focus on diagnosing liver diseases through machine learning techniques. By employing Logistic Regression and K-Nearest Neighbor(KNN), we aimed to enhance predictive accuracy. Our results indicate that KNN outperformed Logistic Regression, offering a more accurate prediction of liver disease risk. The incorporation of a Graphical User Interface (GUI) further enhances the project's applicability as a medical tool for hospitals and medical staff. This research contributes to the field of healthcare by providing a reliable and accessible method for early diagnosis of liver diseases, ultimately improving patient outcomes and facilitating informed medical decisions.

REFERENCES.

1. https://thesai.org/Downloads/Volume14No2/Paper_99-Liver_Disease_Prediction_and_Classification_using_Machine_Learning.pdf
2. https://www.ijert.org/crop-prediction-using-machine-learning-approaches
3. https://www.mdpi.com/2073-431X/12/1/19
4. Prediction and analysis of liver diseases using data mining written Shambel Kefelgen, Pooja Kamat
5. Software based prediction of liver disease with feature selection and classification techniques witten Jagdeep Singh, Sandeep Bagga and Ranjodh Kaur