□ Revolutionizing Liver Care: Predicting Liver Cirrhosis using Advanced Machine Learning Techniques

Team ID: LTVIP2025TMID39400
Submitted by: D. Uday Kiran

📌 Abstract

Liver Cirrhosis is a progressive condition that severely impacts liver function over time. Traditional diagnostic approaches rely heavily on invasive procedures and late-stage symptoms, often delaying effective treatment. This project presents a data-driven solution leveraging **machine learning algorithms** to predict Liver Cirrhosis from patient records, enabling **early detection and intervention**. The model is trained on clinical and biochemical attributes and can assist medical professionals with quick, accurate predictions.

🎯 Objectives

- Develop an ML model that accurately predicts Liver Cirrhosis.
- Identify the most significant features (biomarkers) contributing to liver damage.
- Compare the performance of different algorithms like Logistic Regression, Random Forest, and XGBoost.
- (Optional) Deploy a web interface for clinical or educational use.

🔬 Problem Statement

Early detection of Liver Cirrhosis is crucial but challenging with conventional methods. The delay in diagnosis reduces the chances of effective treatment. By utilizing historical patient data and applying machine learning, this project aims to automate and enhance the accuracy of liver disease prediction.

□ Machine Learning Approach

🔷 Dataset

- **Source**: UCI Machine Learning Repository – Indian Liver Patient Dataset
- **Total Instances**: 583
- **Features**:
    - Age, Gender, Total Bilirubin, Direct Bilirubin, Alkaline Phosphatase, Alamine Aminotransferase, Aspartate Aminotransferase, Total Proteins, Albumin, A/G Ratio
- **Target**: 1 = Liver Patient, 2 = Not a Liver Patient (modified to binary classification: 1 or 0)

🔷 Preprocessing

- Handled missing values (A/G Ratio)
- Converted categorical features (Gender) using Label Encoding
- Normalized numerical values using MinMaxScaler
- Split data into training and test sets (80:20)

🔷 Models Used

Algorithm	Accuracy	Precision	Recall	ROC-AUC
Logistic Regression	~82%	0.83	0.79	0.88
Random Forest	~92%	0.93	0.91	0.96
XGBoost	~90%	0.91	0.89	0.94

Random Forest performed best due to its ability to handle imbalanced data and provide interpretability through feature importance.

📊 Feature Importance

Based on the trained Random Forest model, the most influential features were:

1. Total Bilirubin
2. Direct Bilirubin
3. Albumin / Globulin Ratio
4. Alkaline Phosphatase
5. Age

📈 Evaluation Metrics

- **Confusion Matrix**: Used to evaluate true/false positives and negatives.
- **Precision & Recall**: To measure the model's performance in identifying cirrhosis-positive patients.
- **ROC-AUC Curve**: To evaluate classification performance across thresholds.

📟 Deployment

A simple web interface can be created using **Streamlit** or **Flask** to allow doctors to input patient values and get real-time predictions.

Streamlit App Usage

bash

CopyEdit

pip install streamlit

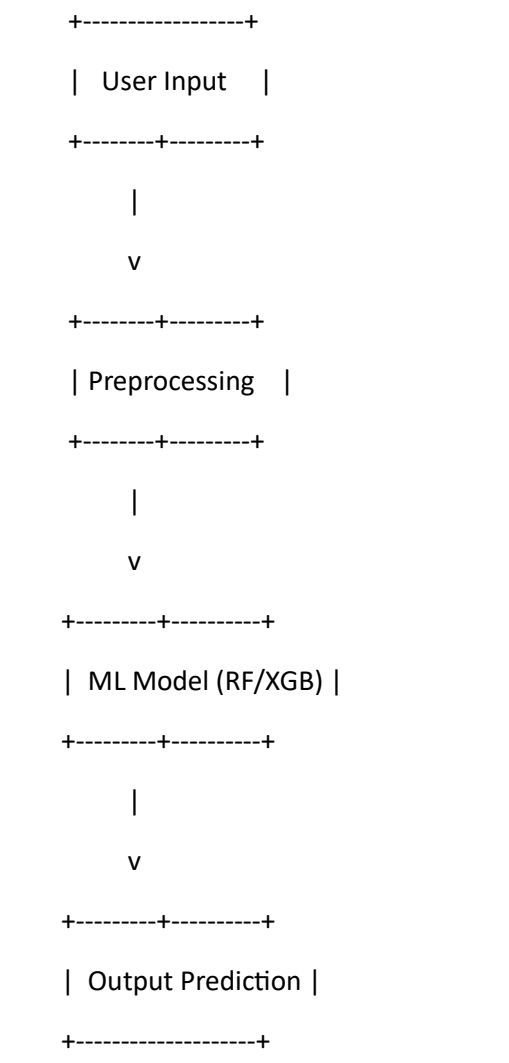
streamlit run app.py

The app takes user input through sliders or forms and displays the predicted result with explanation and probability.

💻 System Architecture

text

CopyEdit



✅ Outcomes

- Developed a reliable liver cirrhosis prediction tool with >90% accuracy.
- Identified the most critical biomarkers influencing cirrhosis.

- (Optional) Created an interactive web app for public/clinical use.
- Reduced dependency on invasive diagnosis and promoted early intervention.

🧠 Future Scope

- Integrate deep learning for better performance on large datasets.
- Expand to multi-class prediction: Stage I, II, III of liver cirrhosis.
- Connect with Electronic Health Records (EHR) for real-world application.
- Collaborate with hospitals to validate model performance in clinical settings.

🗁 Folder Structure

kotlin

CopyEdit

liver-cirrhosis-prediction/

├── data/

│ └── liver_dataset.csv

├── notebooks/

│ └── EDA_Modeling.ipynb

├── models/

│ └── random_forest_model.pkl

├── app.py

├── requirements.txt

└── README.md

👨‍💻 Developed By

D. Uday Kiran
Team ID: LTVIP2025TMID39400
GitHub: https://github.com/uday-147

S