Revolutionizing Liver Care: Predicting Liver Cirrhosis Using Advanced Machine Learning Techniques

1. INTRODUCTION

1.1 Project Overview

A Machine Learning-based system designed to predict liver cirrhosis using clinical data.
Aims to support early diagnosis and improve liver disease management.

1.2 Purpose

To provide a predictive model that assists healthcare professionals in identifying liver cirrhosis at an early stage.
Reduce mortality rates and enable preventive treatments.

2. IDEATION PHASE

2.1 Problem Statement

Liver cirrhosis is often diagnosed late, leading to high fatality.
There is a need for a reliable ML system that predicts cirrhosis using patient data.

2.2 Empathy Map Canvas

Users: Patients, doctors, healthcare institutions.
Pain: Late diagnosis, manual errors, high treatment costs.
Gain: Early prediction, reduced costs, improved care.
Think/Feel/Say/Do: Accurate diagnosis tools, AI integration in hospitals, better decision-making.

2.3 Brainstorming

Identification of relevant datasets (e.g., liver patient dataset).
Selection of features affecting liver health (e.g., bilirubin, albumin, enzymes).
Selection of ML models: Logistic Regression, Random Forest, XGBoost, etc.

3. REQUIREMENT ANALYSIS

3.1 Customer Journey Map

Awareness → Consultation → Data Collection → Prediction → Treatment Planning.

3.2 Solution Requirement

Functional: Upload patient data, display prediction results.
Non-functional: Fast processing, high accuracy, scalable system.

3.3 Data Flow Diagram

Level 0: User inputs data → ML model processes → Result output
Level 1: Input data → Preprocessing → Model → Output risk score

3.4 Technology Stack

Frontend: Flask (for UI)
Backend: Python, Scikit-learn, Pandas
Data: Liver Cirrhosis dataset Dataset link:
(https://www.kaggle.com/datasets/bhavanipriya222/liver-cirrhosis-prediction)
Visualization: Seaborn, Matplotlib

4. PROJECT DESIGN

4.1 Problem-Solution Fit

Addresses the challenge of late diagnosis using data-driven predictive analytics.

4.2 Proposed Solution

A web-based tool where doctors can input patient parameters and receive instant predictions about cirrhosis likelihood.

4.3 Solution Architecture

Input Layer → Data Preprocessing → Model Training/Prediction → Result Visualization

5. PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection and Pre processing	USN-1	Understanding & loading data, Data cleaning, Exploratory Data Analysis (EDA)	-	High	Gayatri Chavan
Sprint-1	Feature Engineering	USN-2	Handling missing values and encoding categorical variables, Engineering features	-	High	Sakshi Salunkhe
Sprint-2	Model Development	USN-3	Training the machine learning model, Evaluating the model	-	High	Indrajit Shinde
Sprint-2	Model Development	USN-4	Creating a Flask app to deploy the model, Developing the front-end using HTML, CSS, and JS	-	Medium	Omkar Pawar
Sprint-2	Testing, Validation, and Final Deployment	USN-5	Testing the application ,validating model predictions, Deploying on a cloud platform, Final testing	-	High	Indrajit Shinde

6. FUNCTIONAL AND PERFORMANCE TESTING

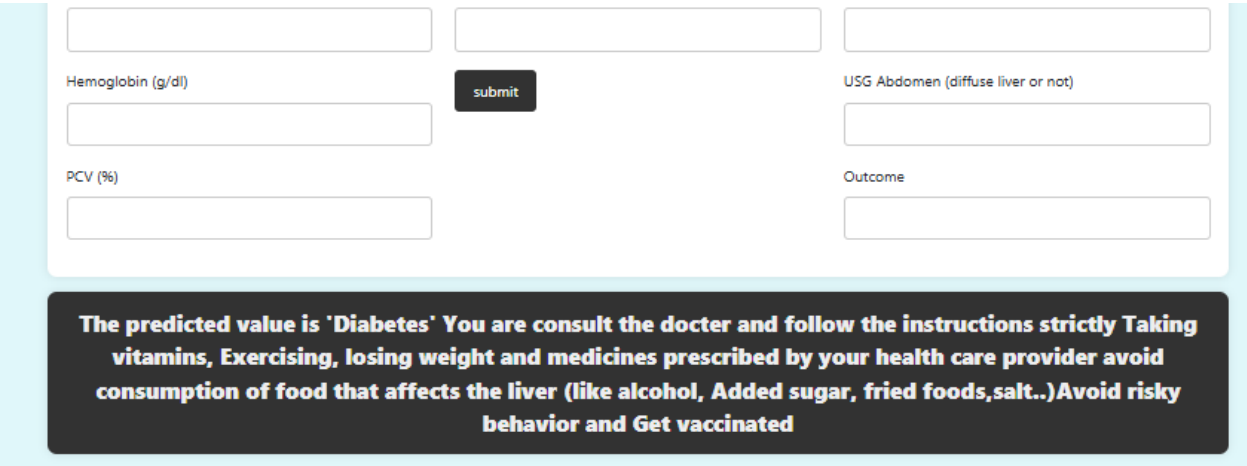
6.1 Performance Testing

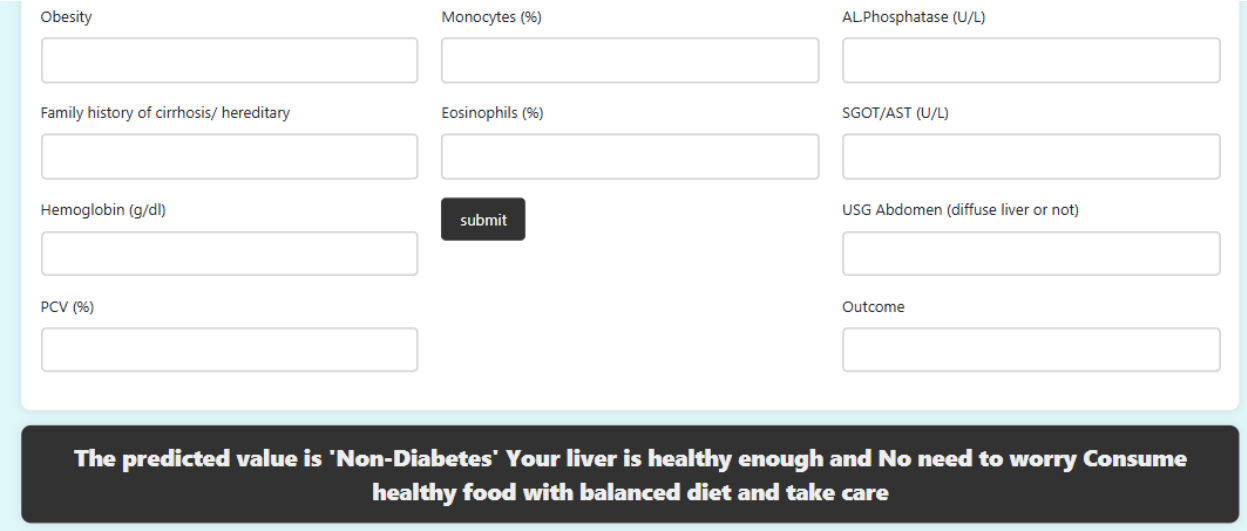
Accuracy, Precision, Recall, F1-score evaluated.
Cross-validation and ROC curve used for model evaluation.
Confusion matrix used for performance analysis.

7. RESULTS

7.1 Output Screenshots

UI showing input form and prediction results.





Model comparison: Accuracy scores of different models.

Model	Accuracy	Precision	Recall	F1-Score
Naive Bayes	35.79%	0.00	0.00	0.00
Random Forest	35.79	0.00	0.00	0.00
Logistic Regression CV	81.58%	91.80	79.43%	86.49
Ridge Classifier	84.21%	93.44	83.82	88.37
Support Vector Classifier	35.79%	0.00	0.00	0.00
Logistic Regression	79.47%	91.80	79.43	85.58
KNN	86.32%	94.26	85.82	89.84
XG Boost	35.79%	3.28	50.00	6.15

8. ADVANTAGES & DISADVANTAGES

Advantages

Early prediction of liver cirrhosis
Non-invasive data input
Scalable and user-friendly interface

Disadvantages

Prediction depends on data quality
May not replace professional diagnosis
Requires regular model retraining for accuracy

9. CONCLUSION

ML models can assist in early detection of liver cirrhosis.
This system can be used as a supportive tool in hospitals and clinics to improve patient care.

10. FUTURE SCOPE

Integration with Electronic Health Records (EHRs)
Inclusion of image-based diagnostics (like liver scans)
Real-time prediction system with API integration
Deploying mobile app version for remote use

11. APPENDIX

Source Code

GitHub Repository: (Code File: liver_cirrhosis.ipynb)
https://github.com/ims10/EPBL_AIML_LIVERCARE_FLASKAPP_PROJECT_

Dataset Link

kaggle.com/datasets:
https://www.kaggle.com/datasets/bhavanipriya222/liver-cirrhosis-prediction

GitHub & Project Demo Link

Render.com:
https://epbl-livercare-flaskapp-team-ymim.onrender.com

Project Documendation Link

GoogleDrive:
https://drive.google.com/drive/folders/196hEVuHMkaas8SRm5SYKaTfJQHJKF1t8?usp=sharing

GitHub Assignments Link

GitHub Repository:
https://github.com/ims10/EPBL_AI-ML_Assignments