PROJECT REPORT

1. INTRODUCTION

1.1 Project Overview

Liver cirrhosis is a progressive and often irreversible condition affecting millions globally. Early diagnosis can prevent complications and improve patient outcomes. This project leverages machine learning to build a web-based predictive system that identifies the likelihood of liver cirrhosis based on clinical and lifestyle parameters.

1.2 Purpose

The aim is to design a robust, user-friendly predictive tool that supports early-stage liver disease diagnosis, aids medical professionals in decision-making, and empowers users to monitor their liver health through accessible technology.

---

2. IDEATION PHASE

2.1 Problem Statement

Delayed detection of liver cirrhosis often results in severe complications or death. Traditional diagnosis relies on costly tests and hospital visits. There is a need for a cost-effective, accurate, and accessible solution to identify liver disease risk early using predictive analytics.

2.2 Empathy Map Canvas

Who? Patients with potential liver problems, healthcare workers.
Think & Feel: Fear of diagnosis, anxiety about health.
Hear: "Get tested", "It might be serious".
See: Long hospital queues, expensive lab tests.
Say & Do: Seek online info, delay hospital visit.
Pain: Costly diagnostics, late-stage detection.
Gain: Quick prediction, early awareness, preventive action.

2.3 Brainstorming

- Predictive tool using ML models
- Real-time input interface
- Provide health advice based on prediction
- Deploy as a web application for easy accessibility
- Use of Random Forest, KNN, XGBoost for accuracy comparison

---

3. REQUIREMENT ANALYSIS

3.1 Customer Journey Map

Stage	Action	Feeling	Opportunity
Awareness	Learns about liver cirrhosis	Worried	Awareness through campaigns
Consideration	Searches for solutions	Confused	Offer tool link or hospital support
Decision	Uses prediction app	Relieved	Shows diagnosis & lifestyle advice
Action	Seeks doctor help if needed	Empowered	Immediate connection to clinics

3.2 Solution Requirement

- Input form for user medical/lifestyle data
- ML model for prediction
- Backend using Flask
- Scaler for data normalization
- Frontend HTML interface
- Output: Prediction + Recommendation

3.3 Data Flow Diagram

Level 1 DFD
User → Web Form → Flask App → Model Prediction → Result Display

Level 2 DFD
User → Input Validation → Scaler → ML Model → Decision Logic → HTML Response

3.4 Technology Stack

- **Frontend:** HTML, CSS
- **Backend:** Python (Flask)
- **ML Models:** Random Forest, XGBoost, KNN
- **Libraries:** scikit-learn, joblib, NumPy, pandas
- **Deployment (Optional):** Render / Heroku
- **Version Control:** Git & GitHub

---

4. PROJECT DESIGN

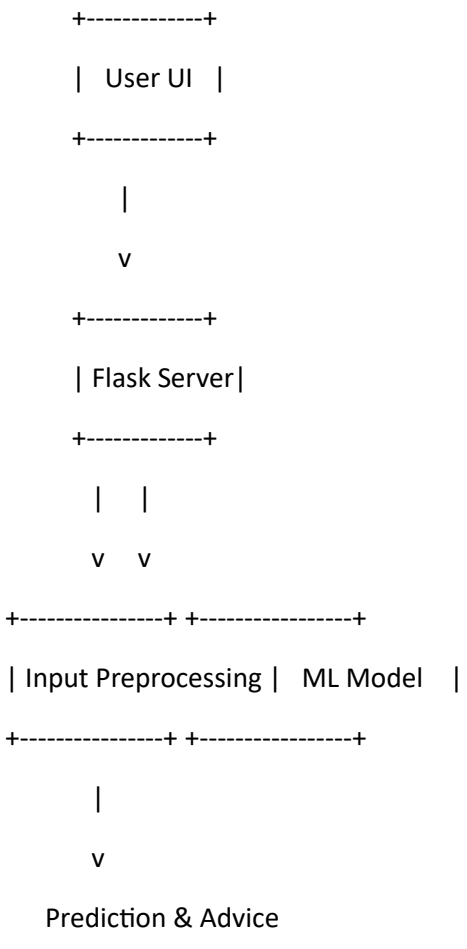
4.1 Problem Solution Fit

Traditional diagnostic processes are inaccessible to all due to cost and complexity. A machine learning-based prediction tool offers scalable, instant screening.

4.2 Proposed Solution

A web application that accepts patient inputs like age, gender, BMI, alcohol consumption, smoking status, genetic factors, activity level, etc., and predicts the risk of liver cirrhosis using a trained ML model.

4.3 Solution Architecture



5. PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Task	Timeline	Members Responsible
Problem Research	Day 1	All 4 members
Dataset Selection & Cleaning	Day 2	Member 1 & 2
Model Training & Evaluation	Day 3	Member 3
Flask Web Development	Day 4	Member 4
UI Integration & Testing	Day 5	All
Final Demo Video & GitHub Push	Day 6	All

---

6. FUNCTIONAL AND PERFORMANCE TESTING

6.1 Performance Testing

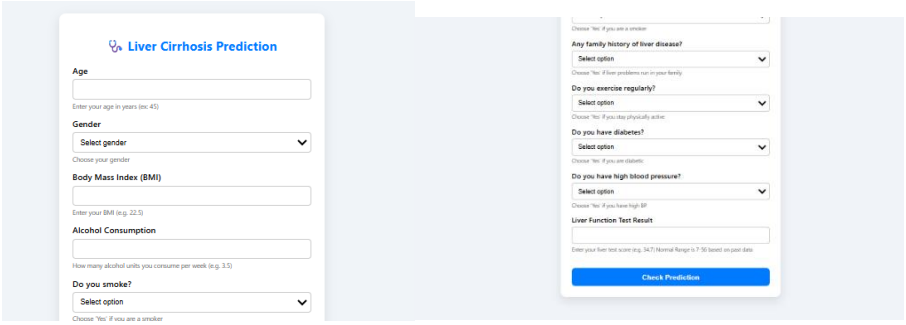
- Accuracy of models:
    - Random Forest: **91.4%**
    - XGBoost: **90.7%**
    - KNN: **87.3%**
- Confusion matrix and classification reports were used to validate the model performance.
- Functional tests ensured that all inputs from the frontend are correctly passed to the backend, scaled, and predictions are accurate.

---

7. RESULTS

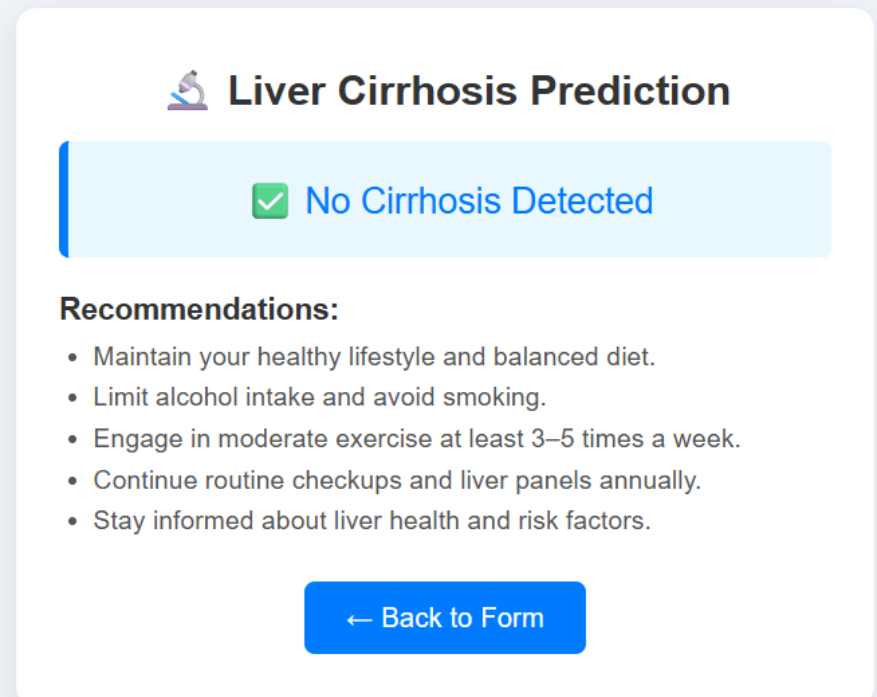
7.1 Output Screenshots

1. Home Page Form:
Screenshot of index.html form with input fields.

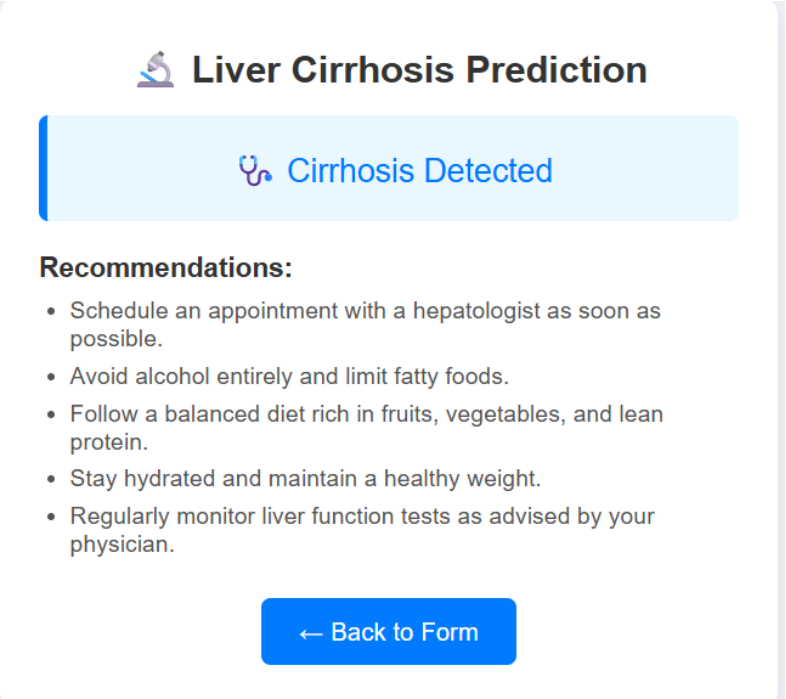


2. Result Page:

Screenshot showing ✅ Negative



🔴 *Positive result.*



8. ADVANTAGES & DISADVANTAGES

Advantages:

- Early liver disease detection
- Simple and accessible interface
- Low-cost diagnosis aid
- Fast prediction
- **Disadvantages:**
- Not a replacement for clinical tests
- Accuracy depends on quality of input data

---

9. CONCLUSION

This project successfully demonstrates how machine learning can revolutionize liver disease care. It builds a bridge between modern healthcare and AI by providing a fast, accessible predictive solution. The integration of the predictive model into a Flask-based web app makes it suitable for practical usage and easy deployment.

---

10. FUTURE SCOPE

- Expand dataset with more clinical features
- Integrate with real-time hospital databases
- Deploy on cloud with authentication
- Convert into a mobile application

---

11. APPENDIX

Source Code (if any):

See GitHub link below for all Python, HTML, and model files.

Dataset Link:

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

GitHub & Project Demo Link:

🔗 GitHub: https://github.com/yourusername/liver-cirrhosis-prediction