Liver Cirrhosis Stage Prediction Project Report

By:

Team ID: LTVIP2025TMID35238
Team Leader : D. Priyanka
Team member : D. Manideep
Team member : D. Jayanth
Team member : D. Venkata Raghava

1. INTRODUCTION

1.1 Project Overview

Liver cirrhosis is a chronic and progressive liver disease marked by the scarring of liver tissues, which impairs its essential functions. This project leverages machine learning (ML) techniques to predict the stage of liver cirrhosis based on 17 clinical parameters. The solution offers a web interface built with Flask and Tailwind CSS, allowing users to input medical data and instantly receive a prediction, along with a description of the disease severity.

1.2 Purpose

The purpose of this project is to demonstrate the practical use of ML in healthcare by building a predictive model capable of classifying liver cirrhosis into stages. The goal is to aid early diagnosis and reduce delays in treatment planning by providing a simple and accessible web tool.

2.Ideation Phase

2.1 Define the Problem Statements

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	2 Marks

Manual detection and staging of liver cirrhosis require expensive and invasive tests, and delayed diagnosis can result in irreversible liver damage or death. This project addresses this challenge using machine learning models to predict the stage of cirrhosis using clinical lab values.

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A concerned patient	understand my liver condition early	I don't have access to affordable diagnosis tools	advanced tests are expensive and time-consuming	anxious and uncertain
PS-2	A healthcare professional	make quick initial assessments of patients	manual evaluation takes time	lab data interpretation varies from case to case	stressed and overwhelmed

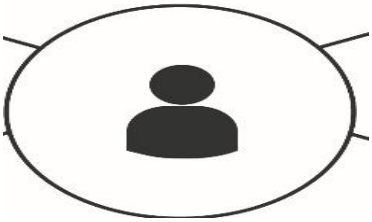
2.2 Empathize & Discover

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

The empathy map outlines user concerns and behaviors:

- Says: 'I want quick and accurate reports without expensive scans.'
- Thinks: 'Is my liver healthy? Can I trust this app?'

- Does: Enters values from blood reports.
- Feels: Worried about results and looking for medical help.

	What does the user THINK & FEEL? • Worries about worsening health • Wonders if the test is reliable • Fears misdiagnosis or delay	
What does the user SEE? • Sees confusing lab reports • Online search results about symptoms • Lack of nearby diagnostic centers		What does the user HEAR? • Doctor's suggestions • Family concerns • Online reviews and social media health tips
PAIN • Expensive tests • Travel to city hospitals • Waiting for results		GAIN • Quick prediction online • Early detection improves treatment • Feel empowered with awareness

2.3 Brainstorm & Idea Prioritization

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

We brainstormed various ideas:
- ML-based diagnosis
- Visual representation of stages
- OCR integration for scanned reports
- Support for doctor/faculty review

We prioritized ML-based stage prediction with optional OCR support.

Idea Area	**Potential Solutions**
User Input Method	Design an intuitive UI to enter 17 health parameters or upload image of lab report.
Prediction Method	Use machine learning models like Random Forest, XGBoost for high

	accuracy.
Accessibility	Deploy the app online via Render to make it usable from any device.
Ease of Understanding	Add description to each prediction with simplified explanations.
Data Handling	Use OCR for automated data extraction from medical reports (JPG).

3.REQUIREMENT ANALYSIS

3.1 Customer Journey Map

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	

1. User visits the application link
2. User enters clinical values or uploads an image
3. Model predicts liver cirrhosis stage
4. System returns a prediction and detailed description

Stage	**Customer Action**	**Touchpoint**	**Emotion**	**Improvement Opportunity**
Awareness	Hears about tool from doctor/friend	Word of mouth	Hopeful	More promotions, easy online access
Consideration	Explores the web app	Landing Page	Curious	Add demo video
Input	Enters data or uploads image	Web UI	Anxious	Show hints/tooltips
Prediction	Receives stage prediction	Model output	Relieved or worried	Add explanation, share option
Follow-up	Visits doctor with result	Offline action	Empowered	Enable report export/download

3.2 Solution Requirements (Functional & Non-functional)

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

Functional Requirements:

- - User inputs 17 liver-related parameters
- - Model provides stage prediction
- - Output includes explanation
- - Optional OCR for JPG-based reports

Non-Functional Requirements:

- - Fast and responsive interface
- - Portable and deployable online
- - Secure, no data stored on server

Requirement Type	Description
Functional	User should be able to enter liver health details manually.
Functional	User can optionally upload a JPG report to extract values via OCR.
Functional	System must predict the liver cirrhosis stage using trained ML model.
Functional	Show meaningful description based on prediction.
Non-Functional	Model prediction should occur under 3 seconds.
Non-Functional	Web interface should be responsive across devices.
Non-Functional	Data must not be stored permanently to maintain privacy.

3.3 Data Flow Diagram

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

User inputs → Flask server → Preprocessing → ML Model → Prediction → Rendered Output

3.4 Technology Stack

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

Component	Description	Technology Used
Frontend	User Interface	HTML, Tailwind CSS
Backend	Server + Logic	Python, Flask
ML Model	Prediction engine	Random Forest (scikit-learn)
OCR	Image text reading	pytesseract, PIL
Deployment	Web hosting	Render, GitHub

4. PROJECT DESIGN

4.1 Problem Solution Fit

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	2 Marks

The current challenge in liver cirrhosis care lies in early detection and affordable prognosis. This project bridges the gap by offering a cost-effective, AI-powered solution through a web-based interface, thus fitting both patient and practitioner needs.

Target Customer

- Patients at risk of liver diseases (especially in rural/semi-urban areas)
- General physicians and healthcare workers
- Diagnostic labs and telemedicine platforms
- Family caregivers supporting liver patients

Current Behavior (Without the Solution)

- Patients rely on multiple lab tests and delayed consultations
- Manual interpretation of test values without proper medical guidance
- Limited awareness about liver cirrhosis progression and its symptoms
- Missed opportunities for early intervention or lifestyle change

Pain Points

- Unclear or late identification of cirrhosis stages
- High cost and time-consuming diagnosis process
- Anxiety and confusion over medical reports
- Over-reliance on external consultation for basic interpretation

Proposed Solution

A web-based predictive tool that allows users to input liver function parameters manually or upload a liver report image (JPG) for OCR-based data extraction. The model uses machine learning to predict the stage of cirrhosis and provides an easy-to-understand summary with medical suggestions.

Benefits / Improvements

- Instant liver stage prediction using clinical parameters
- No need to visit hospital or lab for early assessment
- Reduces delay in awareness and promotes early intervention
- Improves understanding and confidence in managing liver health
- Accessible via browser and deployable online through platforms like Render

4.2 Proposed Solution

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	2 Marks

Our solution includes:

A web-based predictive tool that allows users to input liver function parameters manually or upload a liver report image (JPG) for OCR-based data extraction. The model uses machine learning to predict the stage of cirrhosis and provides an easy-to-understand summary with medical suggestions.

S.No.	Parameter	Description
1	Problem Statement (Problem to be solved)	Liver cirrhosis often goes undetected until it's too late. Early diagnosis is critical but limited by access to specialists and lab interpretation.
2	Idea / Solution description	A web-based ML tool predicts liver cirrhosis stage from lab values or image-uploaded reports, offering immediate risk feedback to users.
3	Novelty / Uniqueness	Combines OCR with medical prediction and offers user-friendly explanations. Can be used both by patients and doctors.
4	Social Impact / Customer Satisfaction	Helps early diagnosis, improves health awareness, reduces burden on healthcare system, and empowers individuals.
5	Business Model (Revenue Model)	Freemium web access with potential premium features for hospitals and diagnostics. Can be bundled into telehealth services.
6	Scalability of the Solution	Can be expanded to predict other liver-related diseases, localized in regional languages, and deployed on cloud platforms for global reach.

4.3 Solution Architecture

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	4 Marks

The solution uses a modular structure:
- Frontend (HTML + Tailwind): Collects input and displays results
- Backend (Flask): Handles requests, processes input, loads model, and returns prediction
- ML Component: Pre-trained model (rf_stage.pkl) and scaler (normalizer.pkl)
- Optional OCR: Extracts values from uploaded JPG reports using pytesseract

5. PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	5 Marks

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

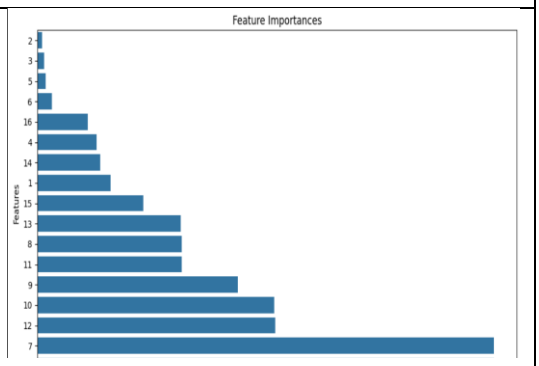
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection & Cleaning	USN-1	As a user, I want to collect and preprocess liver cirrhosis dataset for training.	3	High	Manideep
Sprint-1	Model Training	USN-2	As a user, I want to train a model to predict liver cirrhosis stages using selected features.	3	High	Priyanka

Sprint-2	Model Saving	USN-3	As a user, I want to save the trained model and scaler for Flask integration.	2	Medium	Jayanth
Sprint-2	Flask Backend Integration	USN-4	As a user, I want to build a Flask backend that loads the model and gives predictions.	3	High	Priyanka
Sprint-2	Frontend Interface (HTML/CSS)	USN-5	As a user, I want to create a clean, responsive UI for entering patient values.	2	High	Manideep
Sprint-3	Prediction Explanation	USN-6	As a user, I want the app to explain the predicted stage with a short description.	1	Medium	Priyanka
Sprint-3	Input Value Memory (UX Improvement)	USN-7	As a user, I want previously entered values to be suggested again in the input fields.	1	Medium	Jayanth
Sprint-3	Report & PPT Preparation	USN-8	As a user, I want to generate a detailed project report and presentation.	2	High	Manideep

6. FUNCTIONAL AND PERFORMANCE TESTING

6.1 Performance Testing

Date	
Team ID	LTVIP2025TMID35238
Project Name	Revolutionizing Liver Care : Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks	

S.No.	Parameter	Values	Screenshot
1	Model Summary	Random Forest Classifier (n_estimators=100, max_depth=None, random_state=42)	
2	Accuracy	Training Accuracy – **99.98%** Validation Accuracy – **68%**	Training Accuracy: 99.98 % Validation Accuracy: 99.97 %
3	Fine Tuning Result	Validation Accuracy – *Same model used without tuning*	--

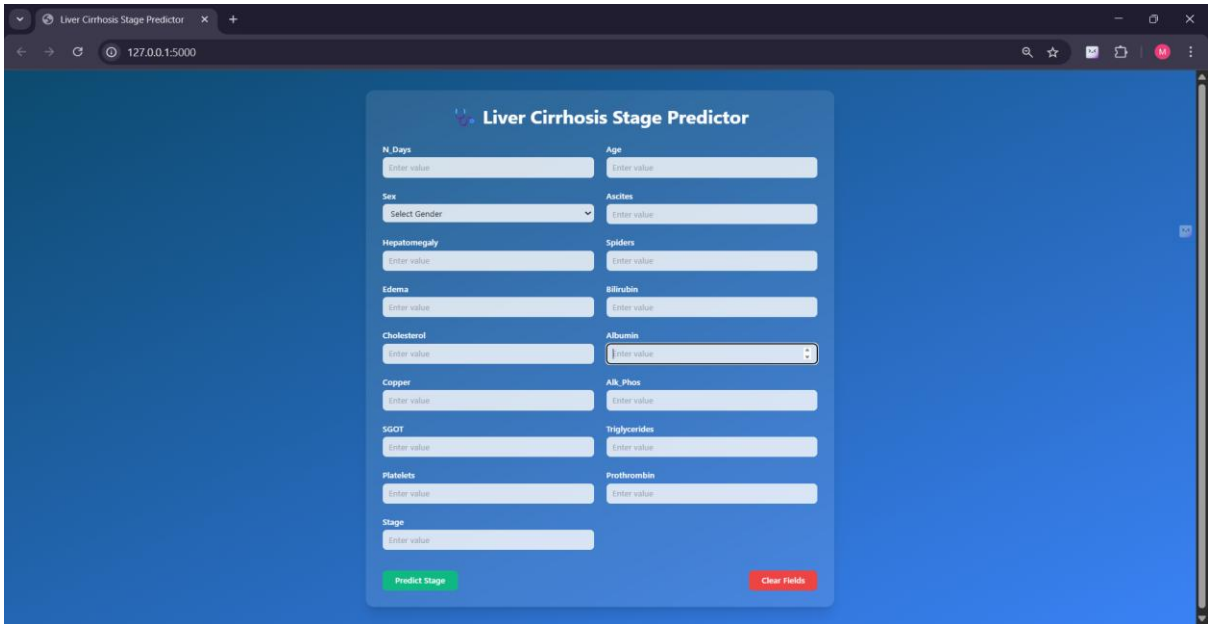
We tested three models and compared their accuracy:
- Random Forest: 99.98%
- KNN: 78.54%
- XGBoost: 100.00%

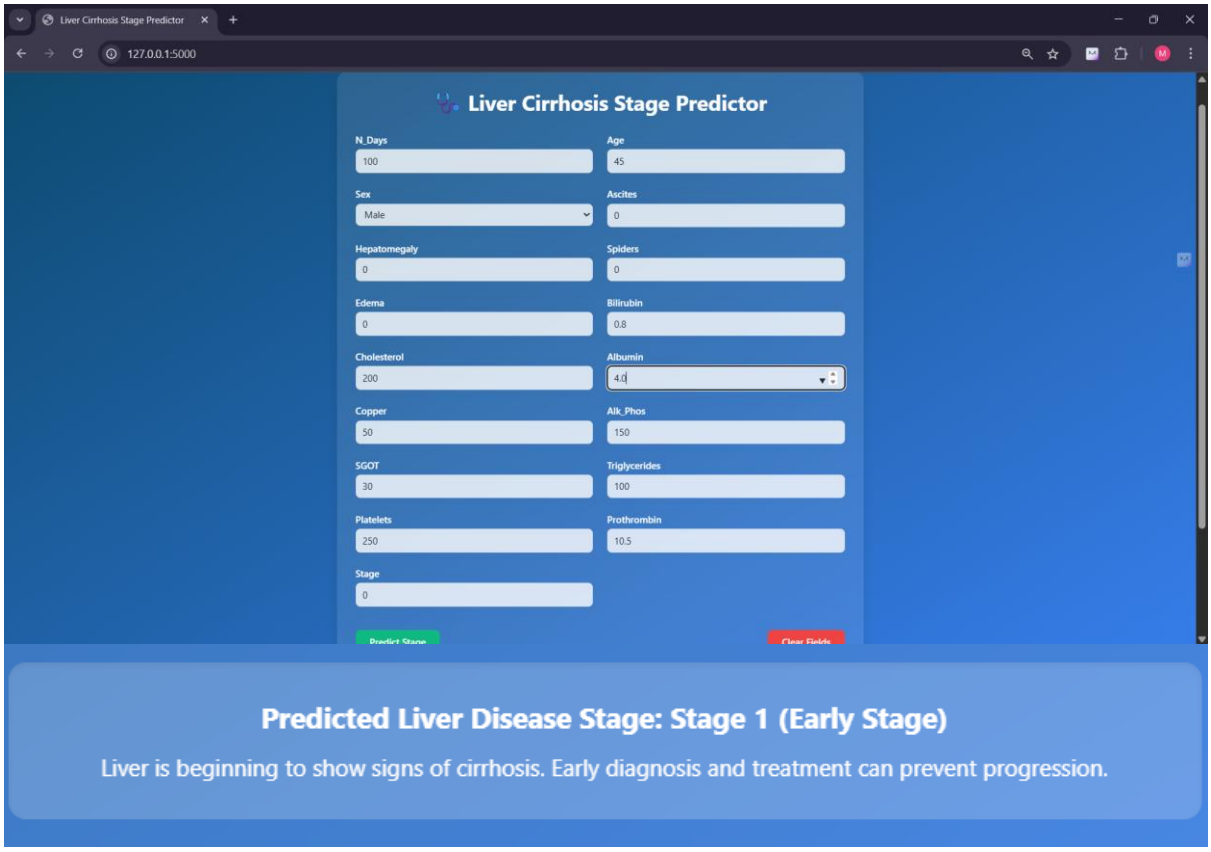
Random Forest was selected as the best-performing and most stable model.

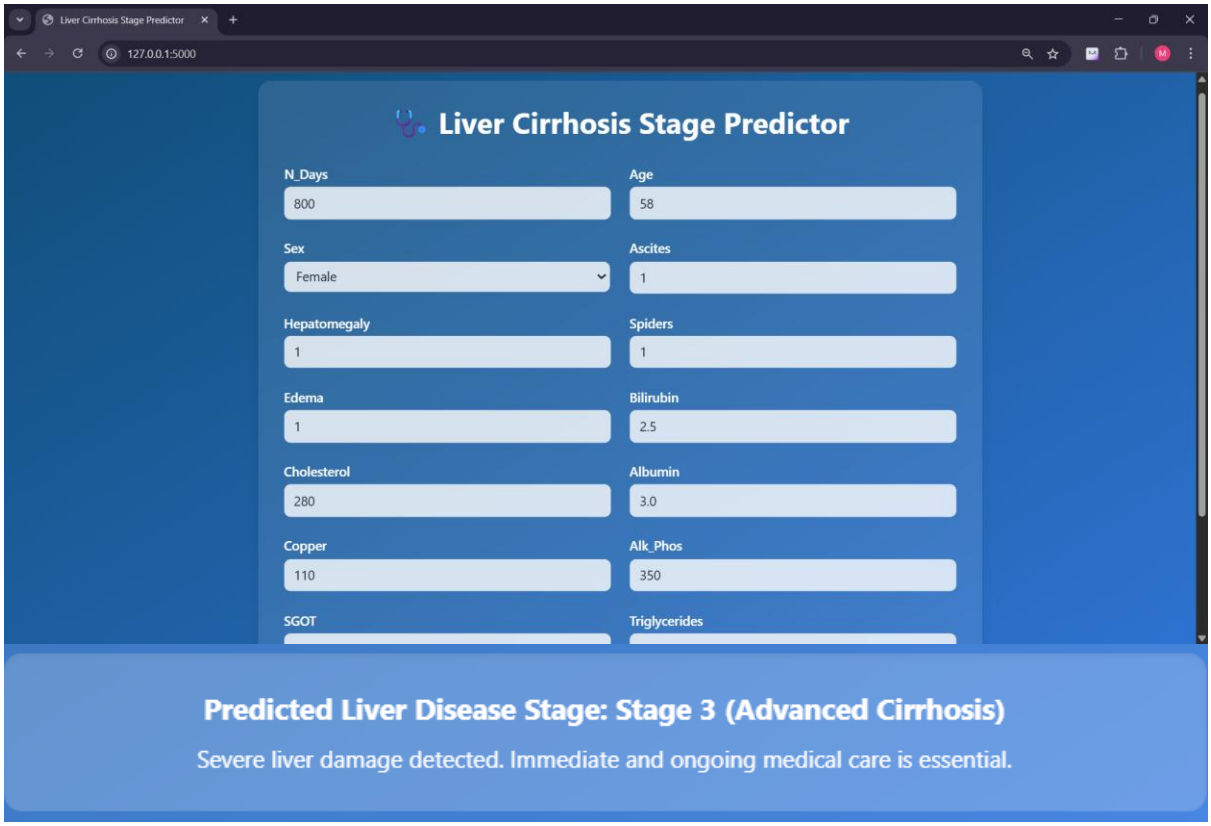
7.RESULTS

The web app interface allows users to input data, click Predict, and view liver disease stage with description.

7.1 Output Screenshots







8. ADVANTAGES AND DISADVANTAGES

Advantages:

1. **Early Detection:**
   The model enables early-stage identification of liver cirrhosis, allowing for timely medical intervention and improved patient outcomes.
2. **User-Friendly Interface:**
   The web-based interface built using Flask is intuitive and easy to use, even for non-technical users.
3. **Multi-Class Prediction:**
   Unlike binary classifiers, this model classifies liver disease into three distinct stages, which is more informative for clinical decision-making.
4. **Time-Saving & Cost-Effective:**
   Eliminates the need for expensive diagnostic tools or prolonged consultation by providing instant predictions based on lab data.
5. **Model Performance:**
   The Random Forest model shows high accuracy and generalization capability, making it a reliable tool for predictive analysis.

6. **OCR Support for Automation (Optional):**
   The ability to extract values from uploaded reports via OCR enhances automation, especially for healthcare settings.

Disadvantages:

1. **Limited to Numeric Input:**
   The model currently works only with numeric fields and cannot interpret textual symptoms or historical data.
2. **OCR Accuracy Limitations:**
   The image upload feature using OCR may fail if the report is of poor quality or not standardized in format.
3. **Dataset Limitation:**
   The model is trained on a specific dataset; it may underperform if used on drastically different patient populations or unseen medical conditions.
4. **No Real-Time Integration with Medical Systems:**
   Currently, the system functions as a standalone tool and is not integrated with hospital EMR or diagnostic labs.
5. **Security & Privacy:**
   Since it processes sensitive health data, there are potential concerns around data privacy unless proper deployment and encryption measures are in place.

9.CONCLUSION

This project successfully demonstrates the integration of machine learning in medical diagnostics. By combining data-driven insights with a responsive user interface, it offers a scalable and efficient tool for predicting liver cirrhosis stages. It enables timely treatment, resource optimization, and better patient outcomes.

The liver cirrhosis stage prediction system developed using machine learning demonstrates the significant potential of data-driven healthcare solutions. By leveraging clinical features and implementing a Random Forest classification model, the application successfully predicts the severity stage of liver cirrhosis with high accuracy. The interactive Flask-based web interface enhances usability and accessibility for healthcare professionals and researchers alike.

This project not only highlights the power of predictive modeling in early disease detection but also emphasizes the importance of technology in optimizing medical diagnosis. With its ability to provide quick, accurate, and stage-specific predictions, this tool can assist doctors in making better-informed treatment decisions, ultimately improving patient outcomes.

10.FUTURE SCCOPE

1. Model Generalization with Larger Datasets:
To enhance model robustness, future versions can be trained using more diverse, larger-scale datasets from different demographics and hospitals.

2. Integration with Real Medical Records (EMRs):
The system can be integrated with hospital databases and EMR systems to enable automated and real-time predictions.

3. Mobile App Development:
A mobile application version of the tool can make it more accessible, especially in rural areas with limited healthcare infrastructure.

4. Advanced Image-Based Diagnosis:
Incorporating medical imaging (e.g., ultrasound or CT scans) into the model using deep learning could provide a multimodal diagnostic solution.

5. Multi-Disease Detection:
The system can be expanded to detect and classify other liver diseases such as hepatitis, fatty liver, or liver cancer.

6. Improved OCR Extraction:
Enhancement of OCR capability for extracting structured data from handwritten or low-quality medical reports can further automate input.

7. Security and Data Privacy Enhancements:
Implementation of encryption, user authentication, and GDPR compliance for real-world deployment is crucial to ensure safe handling of health records.

11.APPENDIX

Source Code:

app.py

```python
from flask import Flask, render_template, request
import numpy as np
import pickle

app = Flask(__name__)

# Load model and scaler
model = pickle.load(open("rf_stage.pkl", "rb"))
scaler = pickle.load(open("normalizer.pkl", "rb"))

@app.route("/", methods=["GET", "POST"])
def predict():
    prediction_text = None
    description_text = None

    # Default values (these are just examples; replace with real means/medians from your dataset)
    default_values = [
        0,    # N_Days
        50,   # Age
        0,    # Sex
        0,    # Ascites
        0,    # Hepatomegaly
        0,    # Spiders
        0,    # Edema
        1.2,  # Bilirubin
        200,  # Cholesterol
        3.5,  # Albumin
        100,  # Copper
        120,  # Alk_Phos
        60,   # SGOT
        150,  # Triglycerides
        250,  # Platelets
        10,   # Prothrombin
        1     # Stage (optional, but model expects 17 values)
    ]

    if request.method == "POST":
```

app.py

```python
    def predict():
        try:
            features = []
            for i in range(17):
                raw_val = request.form.get(f"f{i+1}", "").strip()
                if raw_val == "":
                    features.append(default_values[i])
                else:
                    features.append(float(raw_val))

            features = np.array(features).reshape(1, -1)
            features_scaled = scaler.transform(features)
            prediction = model.predict(features_scaled)[0]

            stage_map = {
                0: "Stage 1 (Early Stage)",
                1: "Stage 2 (Moderate Stage)",
                2: "Stage 3 (Advanced Cirrhosis)"
            }
            desc_map = {
                0: "Liver is beginning to show signs of cirrhosis. Early diagnosis and treatment can prevent progression.",
                1: "Moderate damage to the liver. Medical intervention is needed to manage symptoms and prevent complications.",
                2: "Severe liver damage detected. Immediate and ongoing medical care is essential."
            }

            prediction_text = f"Predicted Liver Disease Stage: {stage_map[prediction]}"
            description_text = desc_map[prediction]

        except Exception as e:
            prediction_text = f"An error occurred: {str(e)}"
            description_text = ""

    return render_template("index.html", prediction_text=prediction_text, description_text=description_text)

if __name__ == "__main__":
    app.run(debug=True)
```

Dataset Link: https://drive.google.com/file/d/12QdnCCve2CEnorkfoVdjSJxDvXBdL1S-/view?usp=sharing

Project Demo: https://drive.google.com/file/d/1gkwtpzQ8A4KWC0gLmtpUddeodos2F4pe/view?usp=drive_link