Liver Disease Prediction Project Documentation

## 👥 Team Members:

- Goriparthi Poojitha

- Gandham Sowbhagya

- Gajjala Mahesh Reddy

- Ganga Harsha Vardhan

# 1. INTRODUCTION

## 1.1 Project Overview

# 2. TECHNOLOGIES USED

- Python

- Scikit-learn (sklearn)

- Random Forest Classifier

- Pandas

- Google Colab

- GitHub

# 3. PROJECT ARCHITECTURE

1. Data Collection and Preprocessing (Google Colab)

2. Model Training using Random Forest (100% accuracy)

3. Saving Model and Normalizer as .pkl files

5. Real-time Prediction Display

# 4. IMPORTANT LINKS

GitHub Repository: https://github.com/MaheshReddy-7/REVOLUTIONIZING-LIVER-CASE-PREDICTING-LIVER-CIRRHOSIS-USING-ADVANCED-MACHINE

Demo Video: https://drive.google.com/file/d/11Ns\_XK2pfkDLAqfhMbiMjf6H\_ripH7J5/view?usp=drivesdk

# 6. NOTES & NEXT STEPS

Thank you for reviewing this project. We hope it contributes meaningfully to early detection and awareness of liver cirrhosis using technology.

# 7. IDEATION PHASE

## 7.1 Problem Statement

Liver cirrhosis is a life-threatening condition that can remain undetected until advanced stages. Manual diagnosis relies heavily on multiple biochemical tests and expert interpretation. This project aims to automate liver disease risk prediction using machine learning to assist early diagnosis and awareness.

## 7.2 Empathy Map Canvas

User: Healthcare professionals and patients.

Goals: Early diagnosis, reduced diagnostic delays, and better treatment outcomes.

Pain Points: Manual analysis is time-consuming and may miss subtle indicators.

Motivations: Use technology to simplify and accelerate health assessments.

## 7.3 Brainstorming

# 8. REQUIREMENT ANALYSIS

## 8.1 Customer Journey Map

User opens web app → Enters medical values → ML model predicts risk → Displays result.

## 8.2 Solution Requirement

- Clean and reliable liver disease dataset

- Feature selection and preprocessing

- Random Forest model training and tuning

- Saving the model and normalizer

## 8.3 Technology Stack

- Python (Data preprocessing, model training)

- Scikit-learn (Random Forest Classifier)

- Google Colab (Model training environment)

- GitHub (Version control and sharing)

# 9. PROJECT PLANNING & SCHEDULING

Week 1: Data Collection and Preprocessing

Week 2: Model Selection and Training

Week 4: Testing, Optimization, and Deployment

# 10. FUNCTIONAL AND PERFORMANCE TESTING

Model Accuracy: 100% on training data (may require validation for real-world use).

Tested prediction logic using known patient data.

# 11. RESULTS

The app predicts whether a person is at high or low risk of liver cirrhosis based on various liver-related biochemical attributes. The system simplifies early risk assessment and enables patients to consult doctors promptly.

# 12. ADVANTAGES & DISADVANTAGES

Advantages:

- Accurate prediction using Random Forest

- Reduces manual workload in hospitals

Disadvantages:

- Dataset may not represent all demographics

- Model needs validation with larger datasets

- Only works with numerical inputs

# 13. FUTURE SCOPE

- Incorporate additional diagnostic parameters like ultrasound or imaging results

- Train on larger, diverse datasets

- Develop a mobile app version for remote access

- Integrate with hospital systems for real-time usage

# 14. APPENDIX

GitHub Repository: https://github.com/MaheshReddy-7/REVOLUTIONIZING-LIVER-CASE-PREDICTING-LIVER-CIRRHOSIS-USING-ADVANCED-MACHINE

Demo Video: https://drive.google.com/file/d/11Ns\_XK2pfkDLAqfhMbiMjf6H\_ripH7J5/view?usp=drivesdk