*Liver Cirrhosis Prediction using Machine Learning*

**Phase 1: Brainstorming and Ideation**

**Objective**:

To explore the problem domain, identify potential solutions, and finalize the core project idea.

**Key Points**:

* Discussed health issues related to liver cirrhosis.
* Identified a real-world need for early-stage prediction systems.
* Explored existing solutions and their limitations.
* Decided on leveraging ML for prediction using medical datasets.
* Selected key features (bilirubin, albumin, etc.) for further analysis.

**Phase 2: Requirement Analysis**

**Objective**:  
To gather functional and non-functional requirements and understand technical feasibility.

**Key Points**:

* Identified stakeholders: doctors, patients, researchers.
* Functional: User uploads data → System predicts cirrhosis risk.
* Non-functional: Accuracy, reliability, scalability.
* Dataset: Indian Liver Patient Dataset (ILPD).
* Tools: Python, scikit-learn, pandas, Jupyter Notebook.

**Phase 3: Project Design**

**Objective**:  
To create a visual and technical blueprint of the system components and their interaction.

**Key Points**:

* Designed system architecture: data preprocessing → model training → prediction UI.
* Data flow diagrams (DFD) and system component blocks.
* Considered modular code design for scalability.
* Chose logistic regression and random forest for baseline models.

**Phase 4: Project Planning (Agile Methodologies)**

**Objective**:  
To break down the project into iterative tasks and sprints following Agile practices.

**Key Points**:

* Defined 4 sprints:
  + Sprint 1: Data collection and cleaning.
  + Sprint 2: Model experimentation.
  + Sprint 3: Evaluation and optimization.
  + Sprint 4: Final integration and documentation.
* Weekly stand-ups and retrospectives were simulated.
* Used Kanban board for tracking progress (Trello).

**Phase 5: Project Development**

**Objective**:  
To implement the project components in code, integrating data pipeline and ML models.

**Key Points**:

* Cleaned dataset (handled missing values, encoded categorical variables).
* Applied normalization techniques (MinMax, Standard Scaler).
* Trained models (Logistic Regression, Random Forest, SVM).
* Performed hyperparameter tuning using GridSearchCV.
* Built UI using Streamlit for easy user interaction.

**Phase 6: Functional and Performance Testing**

**Objective**:  
To validate the system’s performance, accuracy, and usability before deployment.

**Key Points**:

* Used accuracy, precision, recall, F1-score for performance evaluation.
* Confusion matrix used to analyze prediction errors.
* Cross-validation ensured model stability.
* Manual testing of UI components and input/output consistency.
* Performance benchmarks met project objectives.