Project Design Phase-II

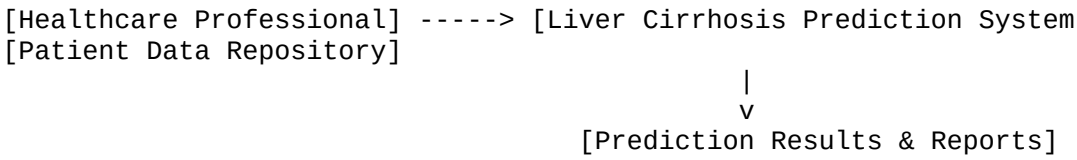
Data Flow Diagram & User Stories

Date: 28 June 2025
Team ID: LTVIP2025TMID45560
Project Name: Revolutionizing Liver Care: Predicting Liver Cirrhosis using Advanced Machine Learning Techniques
Maximum Marks: 4 Marks

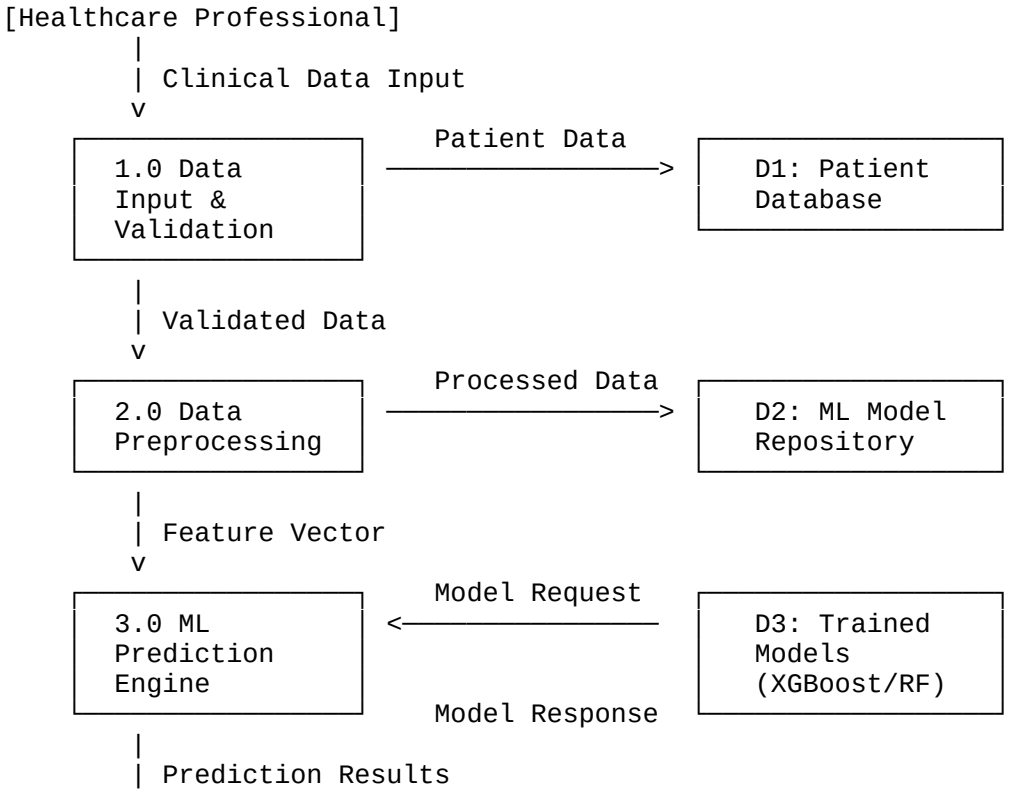
Data Flow Diagrams:

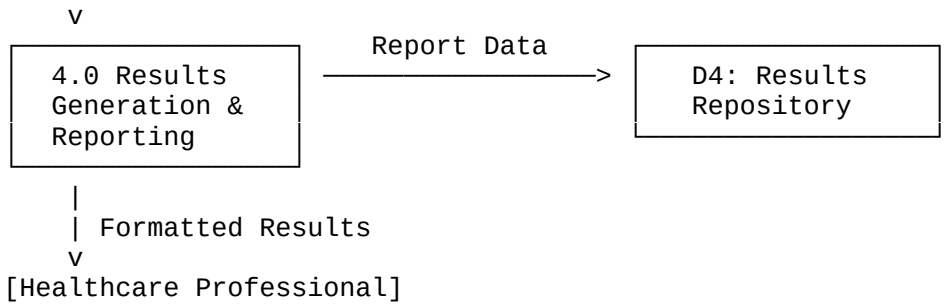
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Context Diagram (Level 0):

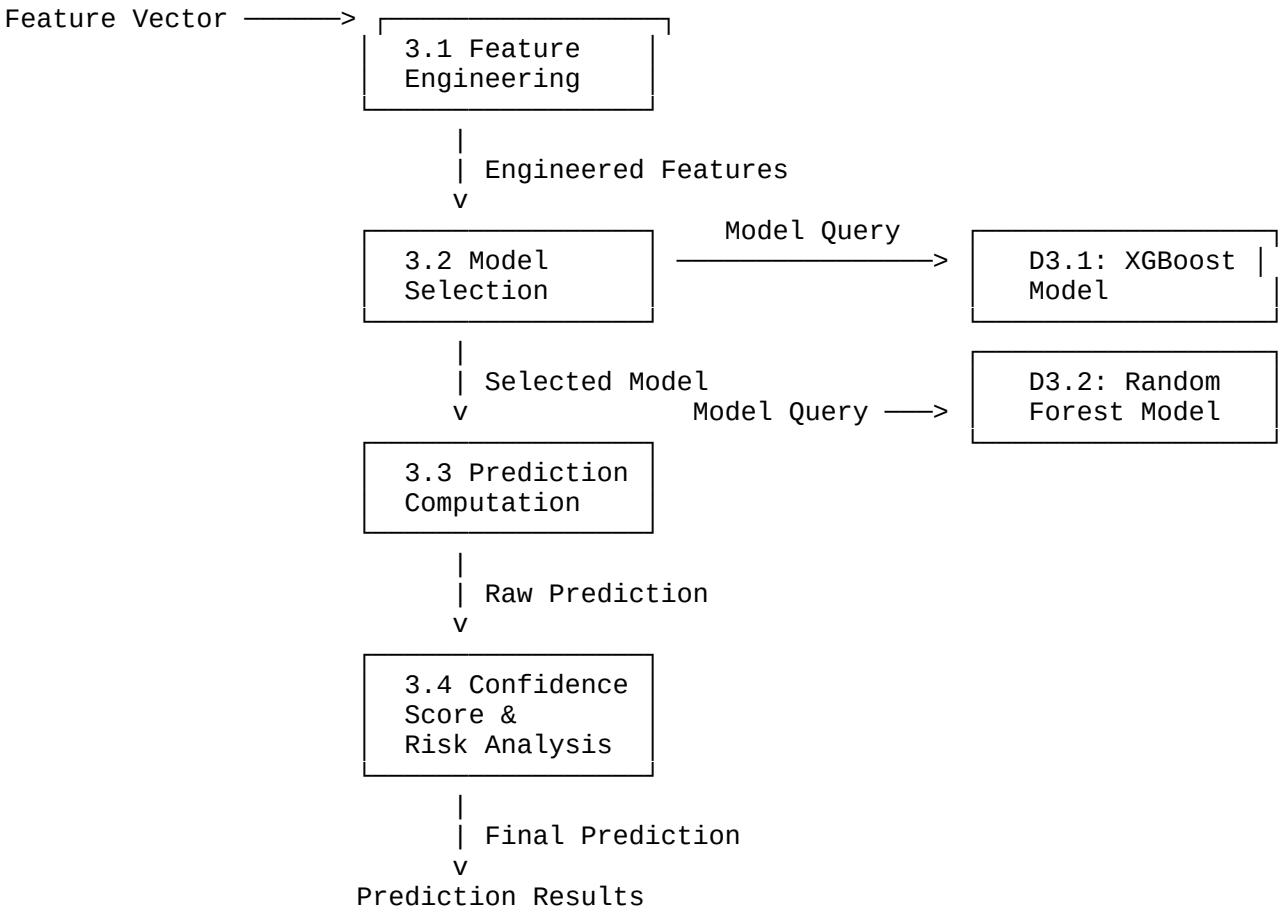


Level 1 Data Flow Diagram:





Level 2 Data Flow Diagram - ML Prediction Engine (Process 3.0):



User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
Healthcare Professional (Doctor)	Authentication & Access	USN-1	As a doctor, I can log into the system using my hospital credentials so that I can access	I can successfully authenticate and access the dashboard with appropriate doctor-level permissions	High	Sprint-1

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
			patient prediction tools.			
		USN-2	As a doctor, I can reset my password through a secure email link so that I can regain access if I forget my credentials.	I receive a secure reset link and can successfully change my password	Medium	Sprint-1
	Clinical Data Input	USN-3	As a doctor, I can input patient clinical data (age, gender, bilirubin, albumin, etc.) through a user-friendly form so that I can get cirrhosis predictions.	All required clinical parameters are captured with proper validation and range checking	High	Sprint-1
		USN-4	As a doctor, I can upload patient data via CSV/Excel file so that I can process multiple patients efficiently.	System accepts and validates bulk data uploads with error reporting	Medium	Sprint-2
		USN-5	As a doctor, I can save patient data as drafts so that I can complete entries later without losing information.	Data is automatically saved and can be retrieved for completion	Low	Sprint-2
	Prediction & Results	USN-6	As a doctor, I can get cirrhosis risk predictions within 3 seconds so that I can make timely clinical decisions.	Prediction results are displayed quickly with confidence scores and risk levels	High	Sprint-1
		USN-7	As a doctor, I can view detailed	Key contributing factors are highlighted with	Medium	Sprint-2

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
			prediction explanations so that I understand which factors contributed to the risk assessment.	their relative importance		
		USN-8	As a doctor, I can compare predictions from different ML models (XGBoost vs Random Forest) so that I can validate results.	Side-by-side comparison of model predictions with accuracy metrics	Low	Sprint-3
	Report Generation	USN-9	As a doctor, I can generate comprehensive medical reports including prediction results so that I can document patient assessments.	Professional PDF reports with patient data, predictions, and recommendations	High	Sprint-2
		USN-10	As a doctor, I can export prediction results to integrate with Electronic Health Records (EHR) systems.	Results exported in HL7 FHIR format for EHR integration	Medium	Sprint-3
Medical Technician	Data Entry	USN-11	As a medical technician, I can input laboratory test results so that doctors can use them for predictions.	Lab values are properly validated and stored with appropriate units	High	Sprint-1
		USN-12	As a medical technician, I can verify and edit patient demographic information so that prediction	Data validation ensures consistency and completeness	Medium	Sprint-1

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
			accuracy is maintained.			
	Quality Control	USN-13	As a medical technician, I can flag unusual lab values for doctor review so that data quality is maintained.	Outlier detection with notification system for abnormal values	Medium	Sprint-2
System Administrator	User Management	USN-14	As an administrator, I can manage user accounts (create, modify, deactivate) so that system access is properly controlled.	Complete user lifecycle management with role-based permissions	High	Sprint-1
		USN-15	As an administrator, I can audit user activities so that system security and compliance are maintained.	Comprehensive audit logs with search and reporting capabilities	High	Sprint-1
	System Monitoring	USN-16	As an administrator, I can monitor system performance and model accuracy so that service quality is maintained.	Real-time dashboards showing system health and prediction metrics	Medium	Sprint-2
		USN-17	As an administrator, I can update ML models so that prediction accuracy remains current with latest research.	Version-controlled model deployment with rollback capabilities	Low	Sprint-3
	Data Management	USN-18	As an administrator, I can backup and restore patient	Automated backup systems with point-in-time recovery options	High	Sprint-1

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
			data so that information is protected against loss.			
Hospital IT Manager	Integration	USN-19	As an IT manager, I can integrate the system with existing hospital information systems so that workflows are streamlined.	API endpoints for seamless integration with HIS/EHR systems	Medium	Sprint-3
		USN-20	As an IT manager, I can configure system security settings so that HIPAA compliance is maintained.	Comprehensive security configuration with encryption and access controls	High	Sprint-1
	Analytics	USN-21	As an IT manager, I can generate usage analytics and prediction statistics so that system value can be measured.	Dashboard showing utilization metrics and clinical outcome improvements	Low	Sprint-3
Research Coordinator	Data Analysis	USN-22	As a research coordinator, I can export anonymized prediction data so that clinical research can be conducted.	De-identified data export with statistical summaries for research	Low	Sprint-3
		USN-23	As a research coordinator, I can analyze prediction trends and model performance so that clinical insights can be derived.	Analytical tools for trend analysis and model performance evaluation	Low	Sprint-3

Data Flow Narrative:

Process 1.0 - Data Input & Validation:

Healthcare professionals input patient clinical and laboratory data through web forms. The system validates data ranges, handles missing values, and stores validated information in the Patient Database. Invalid inputs trigger error messages with guidance for correction.

Process 2.0 - Data Preprocessing:

Raw patient data undergoes preprocessing including normalization, feature scaling, and encoding of categorical variables. The system handles missing values using appropriate imputation techniques and prepares feature vectors for ML model consumption.

Process 3.0 - ML Prediction Engine:

The core prediction engine loads trained models (XGBoost, Random Forest) and processes feature vectors to generate cirrhosis risk predictions. The system provides confidence scores and risk level categorization.

Process 4.0 - Results Generation & Reporting:

Prediction results are formatted into comprehensive medical reports, stored for future reference, and presented to healthcare professionals through interactive dashboards and exportable formats.

Data Stores:

Data Store	Description	Content
D1: Patient Database	Secure storage of patient clinical data	Demographics, lab results, medical history, clinical parameters
D2: ML Model Repository	Storage of trained machine learning models	Serialized XGBoost and Random Forest models, model metadata
D3: Trained Models	Active models used for predictions	Current production models with version information
D4: Results Repository	Storage of prediction results and reports	Historical predictions, generated reports, audit trails

---

External Entities:

1. **Healthcare Professional (Doctor):** Primary user who inputs patient data and interprets prediction results
2. **Medical Technician:** Supports data entry and quality control processes
3. **System Administrator:** Manages system operations, security, and maintenance
4. **Hospital IT Manager:** Oversees integration and compliance requirements
5. **Research Coordinator:** Utilizes system data for clinical research purposes
6. **Patient Data Repository:** External systems providing patient information
7. **Electronic Health Records (EHR):** Integration target for prediction results

Data Flow Summary:

The system follows a structured data flow from clinical data input through machine learning prediction to results delivery, ensuring data integrity, security, and clinical utility throughout the process. Each process is designed to support healthcare workflows while maintaining HIPAA compliance and medical data standards.