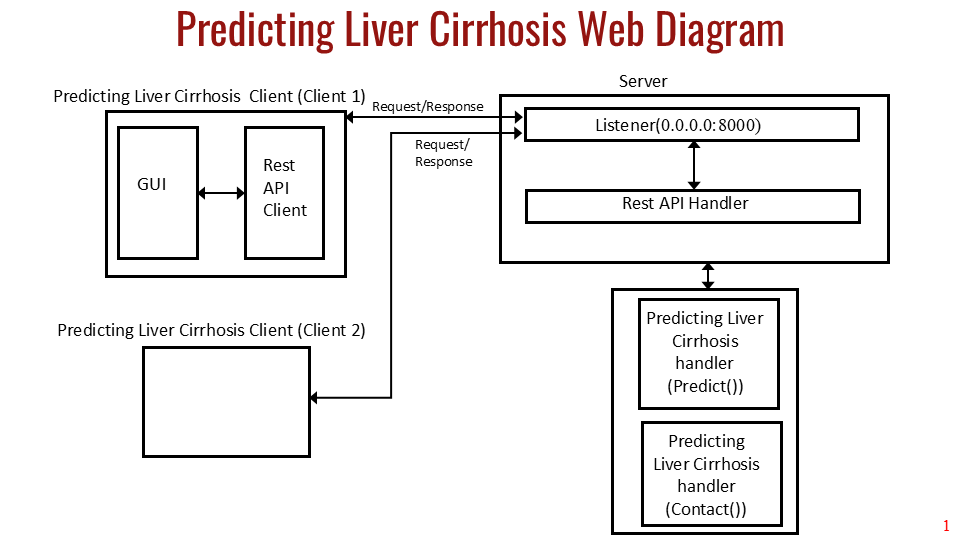
**Project Design Phase-II**

**Data Flow Diagram & User Stories**

|  |  |
| --- | --- |
| Date | 16 June 2025 |
| Team ID | LTVIP2025TMID35624 |
| 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.

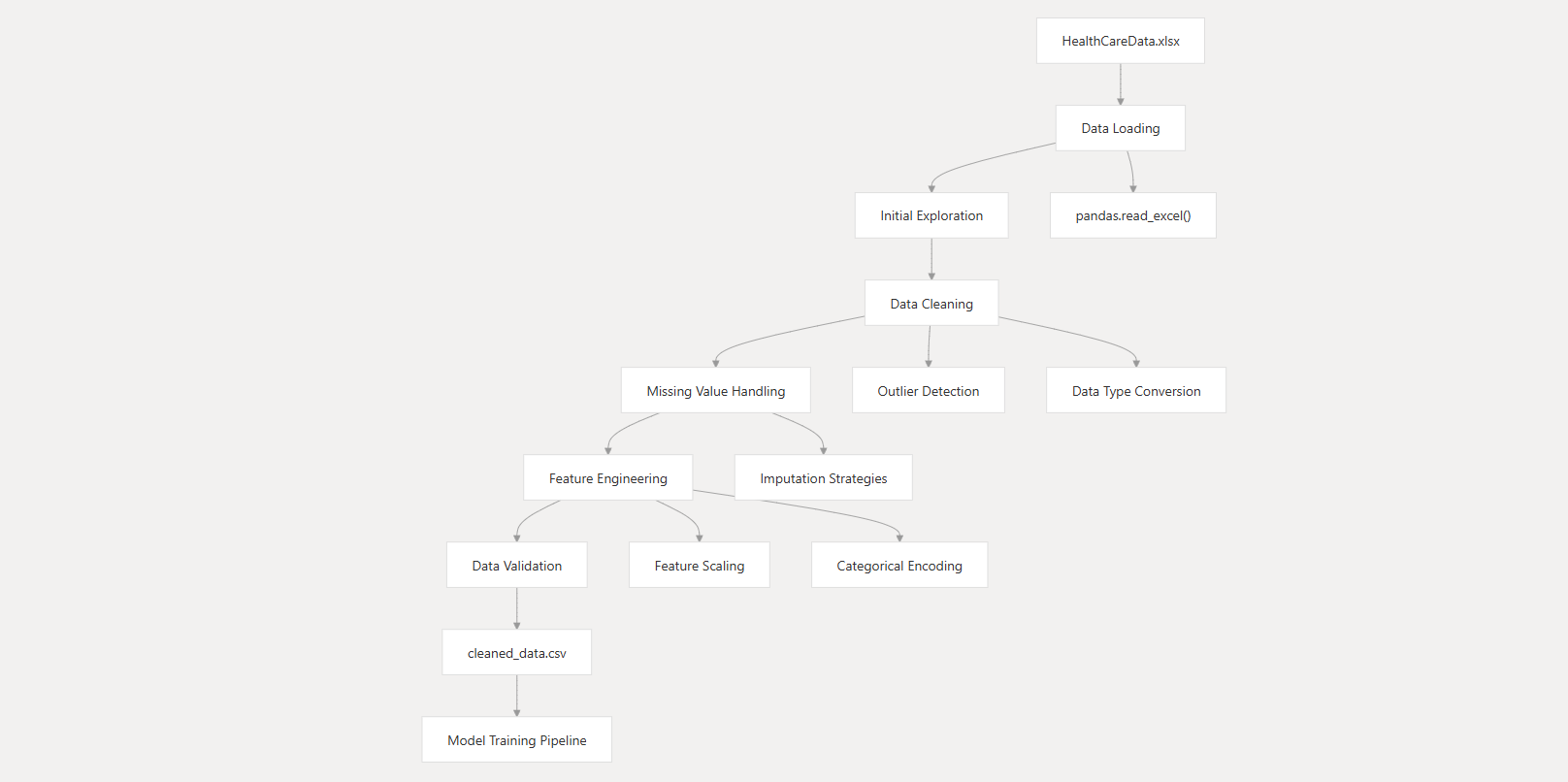


**Data Pipeline and Processing**

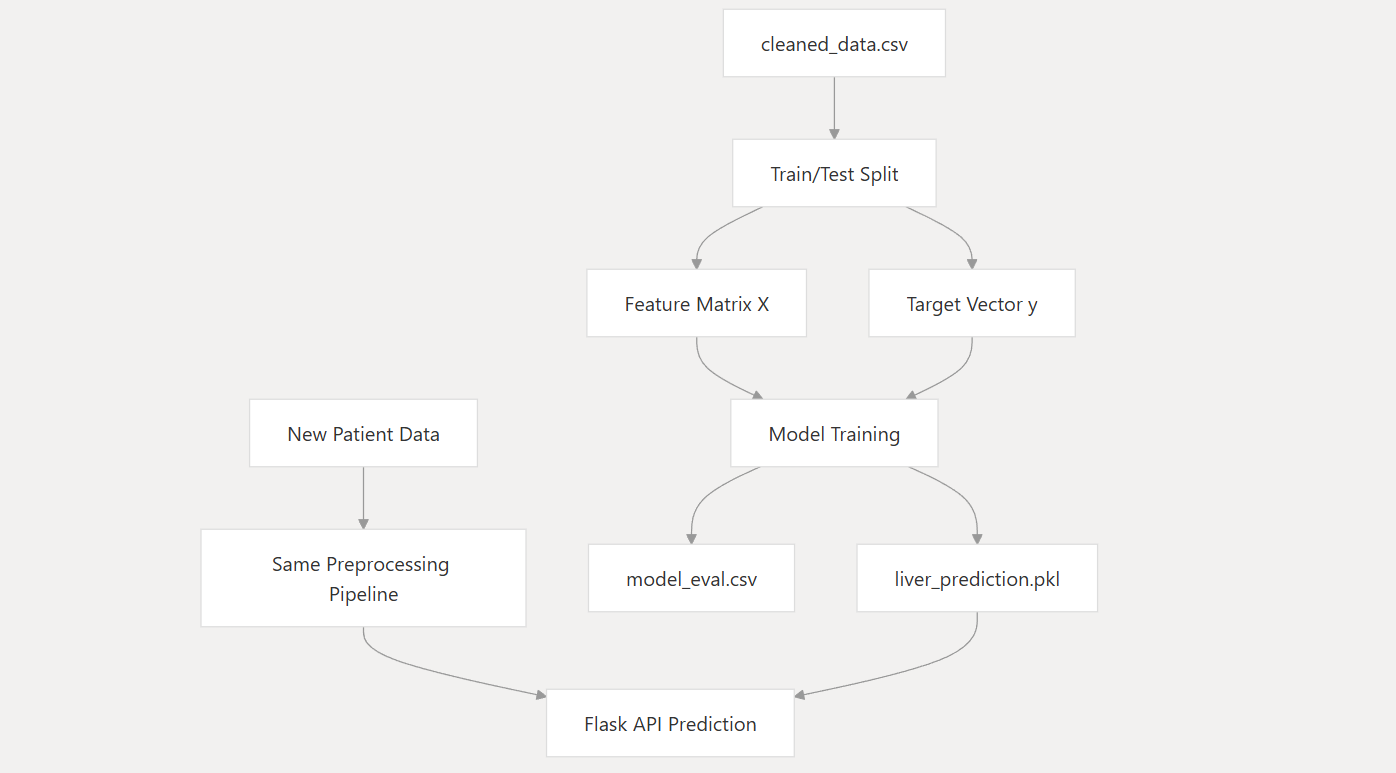
The data pipeline and processing components responsible for transforming raw healthcare data into clean, model-ready datasets for liver cirrhosis prediction. The pipeline handles data ingestion, cleaning, preprocessing, feature engineering, and validation to prepare data for machine learning model training and inference.

**Overview and Data Flow**

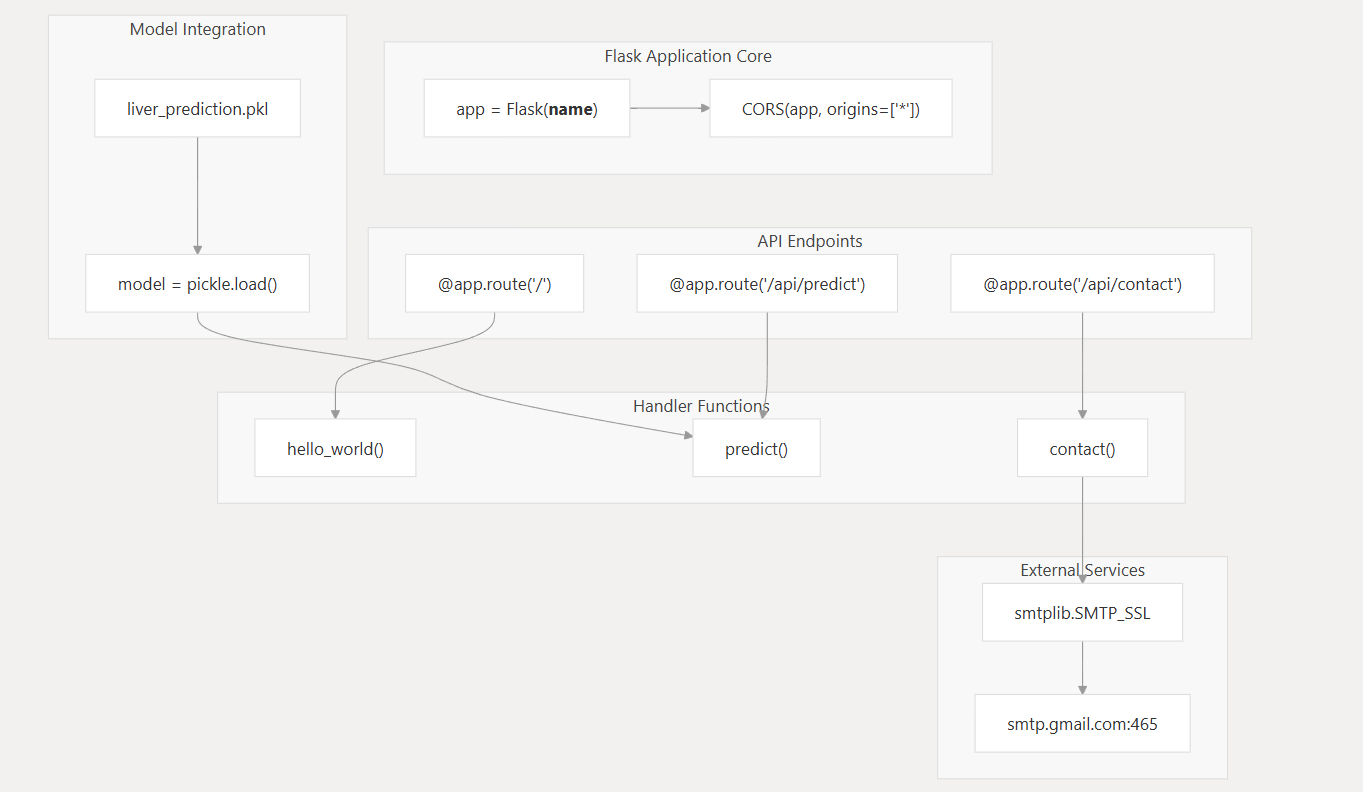
The data pipeline transforms raw healthcare data through multiple processing stages to create clean, standardized datasets suitable for machine learning algorithms. The pipeline processes patient demographic information, medical history, laboratory test results, and clinical measurements to predict liver cirrhosis occurrence.



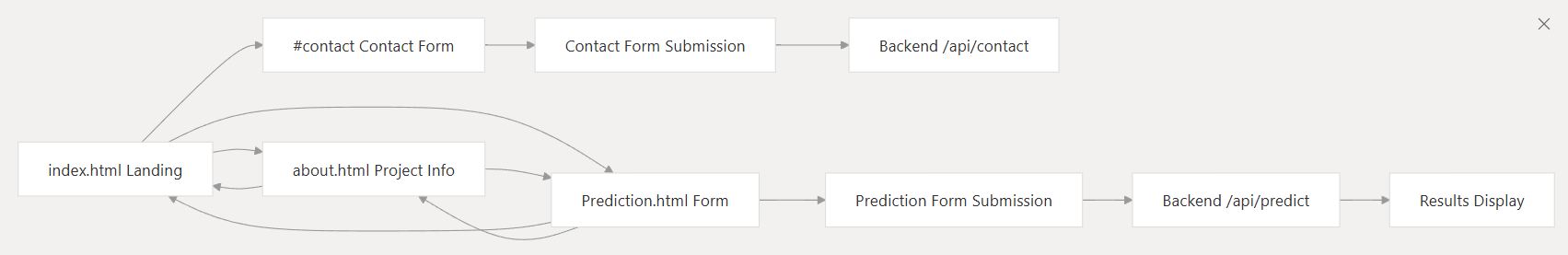
**Integration with Machine Learning Pipeline**



**Backend Application Architecture**

****

**User Flow**

****

**User Stories**

| **User Type** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Acceptance criteria** | **Priority** | **Release** |
| --- | --- | --- | --- | --- | --- | --- |
| Developer | Data Collection & Preprocessing | USN-1 | As a developer, I want to understand and clean the dataset, and handle missing values | Dataset is clean, missing values handled, and ready for analysis | High | Sprint-1 |
| Developer | Feature Engineering | USN-2 | As a developer, I want to perform EDA, encoding of categorical features, and feature creation | Features are encoded and created for model input | High | Sprint-2 |
| Developer | Model Development | USN-3 | As a developer, I want to train the model, tune hyperparameters, and evaluate its performance | Best model is selected and evaluated with metrics like accuracy, F1-score | High | Sprint-3 |
| Developer | Model Deployment | USN-4 | As a developer, I want to create a Flask API and build a frontend UI using HTML/CSS/JS | UI is responsive and backend API returns correct predictions | High | Sprint-4 |
| Developer | Testing & Final Deployment | USN-5 | As a developer, I want to test the entire system, deploy to cloud, and document the process | Model runs end-to-end on cloud with working UI and complete documentation | High | Sprint-5 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |