

CHAPTER – 5

SYSTEM DESIGN

CHAPTER 5

SYSTEM DESIGN

The system design for the project "Leveraging ANN for Targeted Drug Sensitivity Prediction on GDSC Data" is focused on creating an efficient, user-friendly platform for predicting drug sensitivity. The system integrates a comprehensive data preprocessing pipeline, a robust Artificial Neural Network (ANN) for prediction, and a web-based interface to ensure seamless interaction with users. The goal is to provide accurate IC50 predictions and visualization tools to support research and clinical applications. This chapter outlines the core design components, including use cases, architecture, data flow, sequence, and activity diagrams.

5.1 Architectural Diagram

The architecture of the system is designed in layers to ensure modularity, scalability, and maintainability. The main layers include:

- **Frontend Layer:** Handles user interactions through a web interface. Features include dropdowns for data input, real-time prediction display, and drug lookup functionality.
- **Backend Layer:** Processes user inputs, manages data preprocessing, and generates predictions using the ANN model. This layer is implemented using Flask.
- **Model Layer:** Incorporates a PyTorch-based ANN model to predict IC50 values based on preprocessed genomic data.
- **Data Preprocessing Layer:** Normalizes, scales, and encodes genomic data, ensuring compatibility with the model

The architecture diagram (Fig. 5.1) illustrates the interaction between these components.

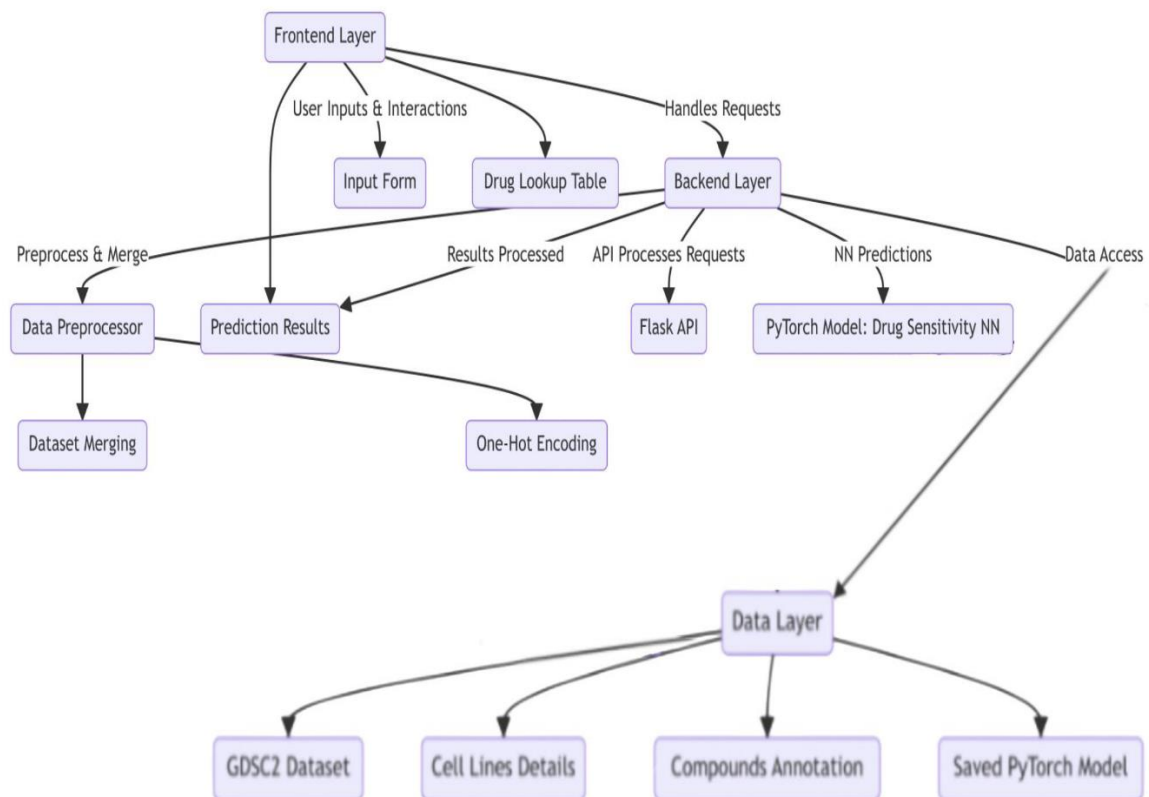


Figure 5.1: Architectural Diagram

5.2 Use Case Diagram and Description

The use case diagram illustrates the various interactions users have with the system. Key users include researchers, clinicians, and administrators. The diagram shows that users can perform actions such as submitting genomic data for IC50 prediction, accessing drug information through the lookup feature, and viewing prediction results and visualizations. Each interaction is designed to ensure usability and accessibility, with clear workflows that guide the user from data input to output interpretation. The use case emphasizes the core functionalities of the system, including its ability to simplify data input through dropdown menus and provide real-time predictions.

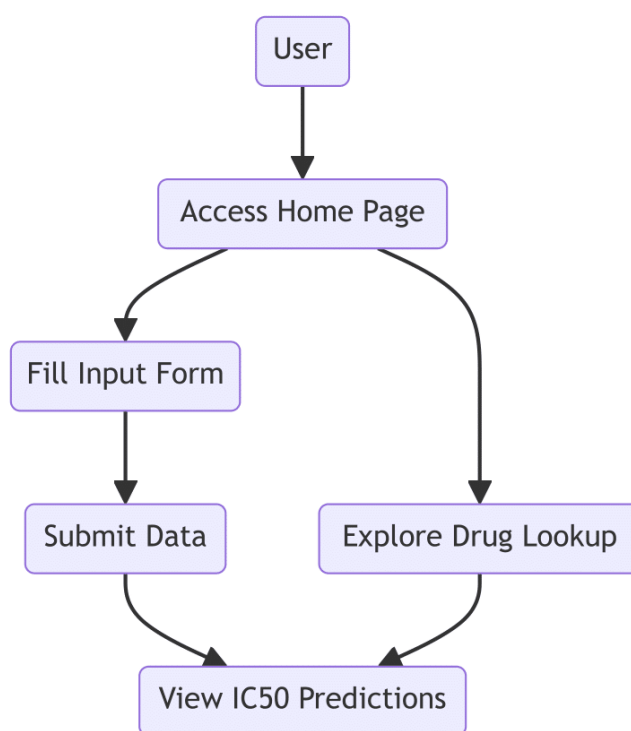


Figure 5.2: Use Case Diagram

5.3 Sequence Diagram

The sequence diagram provides a step-by-step representation of the interaction flow between the user, the web interface, the backend, and the ANN model. The user begins by entering genomic and drug-related data through the frontend interface. This input is then validated and preprocessed by the backend, ensuring compatibility with the model. Once processed, the data is passed to the ANN model, which generates the IC50 prediction. The prediction is returned to the backend and displayed to the user on the interface. This sequence ensures a seamless experience, highlighting the logical progression of data through the system while maintaining efficiency and accuracy.

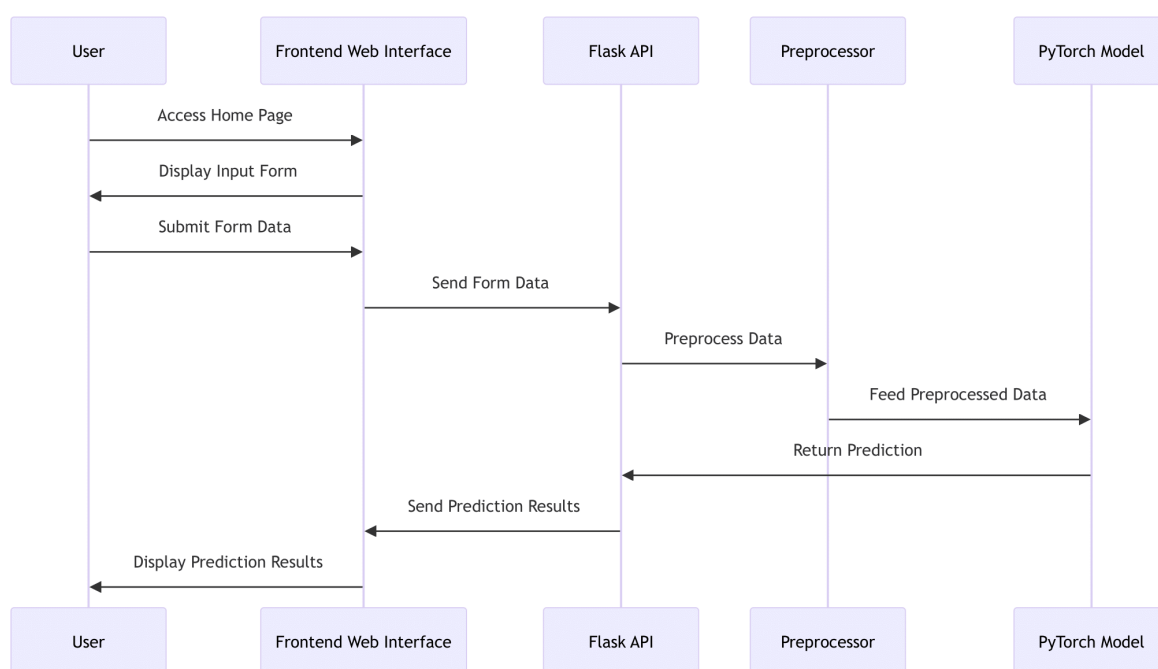


Figure 5.3: Sequence Diagram

5.4 Activity Diagram

The activity diagram provides a high-level overview of the system's workflow, starting with user interaction. Users input genomic and drug-related data, which is then validated by the system. The data preprocessing step ensures that all inputs are normalized, scaled, and encoded to meet the requirements of the ANN model. Once the data is ready, the ANN model processes it to generate IC50 predictions. These predictions are displayed on the user interface, allowing users to analyze the results. The activity diagram captures the iterative nature of model training and prediction refinement, emphasizing the system's focus on delivering accurate and reliable outputs.

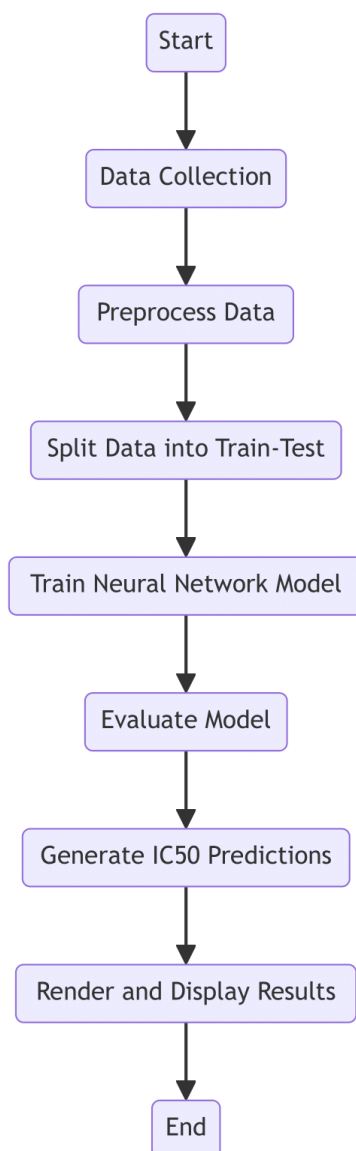


Figure 5.4: Activity Diagram

5.5 Data Flow Diagram

The data flow diagram demonstrates the movement of data across different components of the system. The process starts with user input, where genomic data and drug details are entered via the web interface. The data is then validated and preprocessed in the backend. Preprocessing involves normalization, scaling, and encoding, which prepare the data for input into the ANN model. The preprocessed data is fed into the model, which analyzes it and outputs the IC50 predictions. Finally, the predictions are sent back to the web interface for user display. The diagram emphasizes the system's ability to handle data efficiently and maintain a smooth flow from input to output.

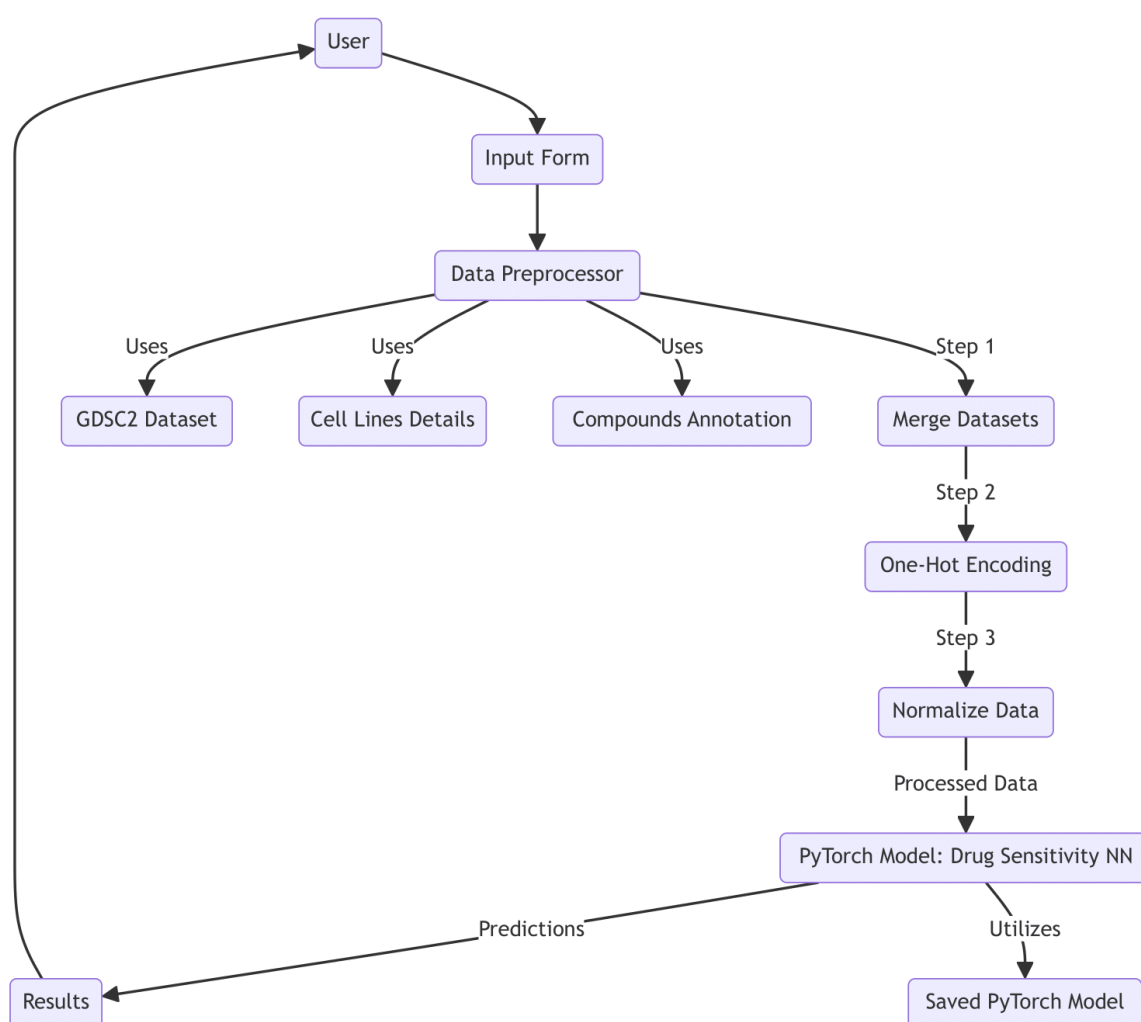


Figure 5.5: Data Flow Diagram