Project Design Phase-II
Data Flow Diagram & User Stories

Date	31 January 2026
Team ID	LTVIP2026TMIDS82279
Project Name	Rising Waters: A Machine Learning Approach to Flood Prediction
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.

Example:
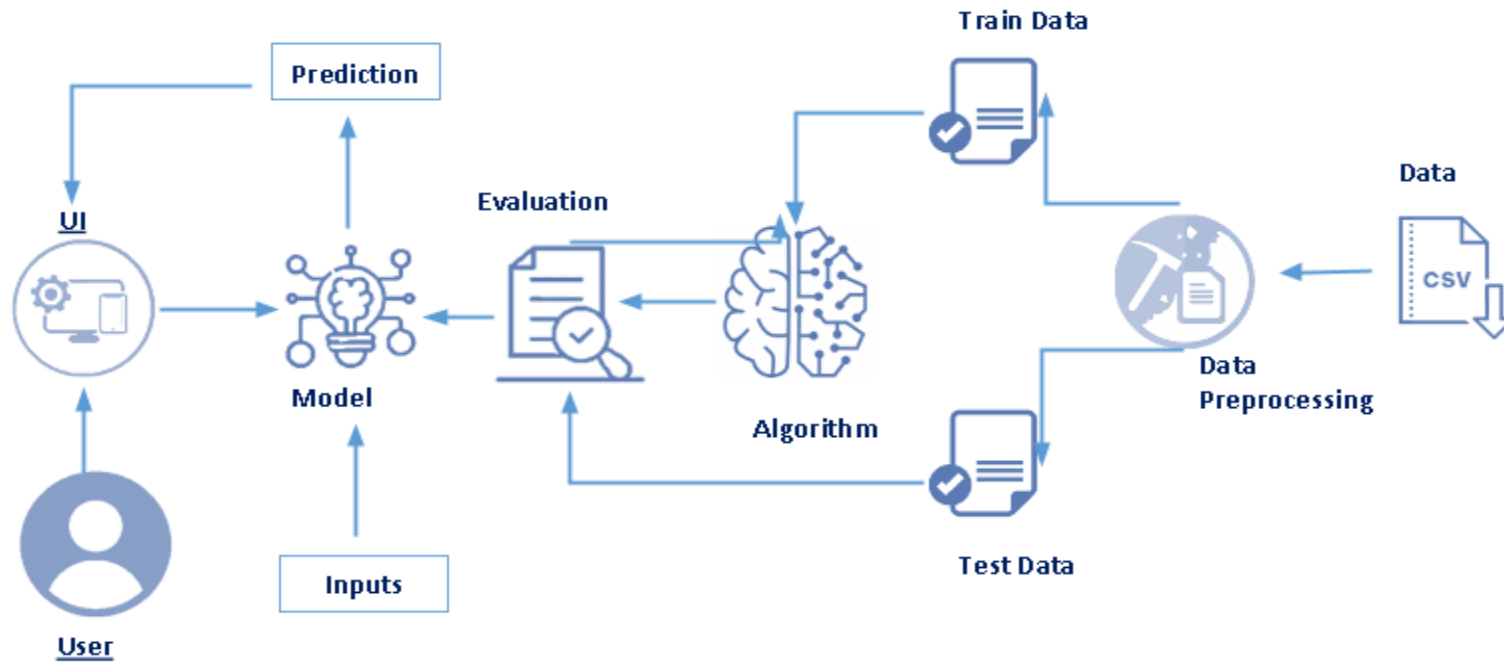
Flood Prediction using Machine Learning is a vital application that aims to forecast and predict flood occurrences with high accuracy. By analyzing historical weather data, river levels, terrain information, and other relevant factors using machine learning algorithms, this project helps in early warning and mitigation of potential flood events. The goal is to provide timely alerts and actionable insights to authorities, communities, and individuals to minimize the impact of floods on lives and infrastructure.

User Stories – Flood Prediction System

a.Data Collection & Upload

US-01:
As a *Data Analyst*, I want to upload flood-related CSV datasets so that the system can use historical environmental data for training.

US-02:
As a *System Administrator*, I want the system to validate uploaded datasets so that incorrect or corrupted files are rejected.

---

b.Data Preprocessing

US-03:
As a *Data Analyst*, I want the system to clean missing and inconsistent data so that the model can be trained accurately.

US-04:
As a *Data Analyst*, I want the dataset to be automatically split into training and testing data so that model performance can be evaluated properly.

---

c.Model Training & Algorithm

US-05:
As a *Machine Learning Engineer*, I want to train the XGBoost algorithm using the prepared training dataset so that the system learns flood prediction patterns.

US-06:
As a *Machine Learning Engineer*, I want to tune model parameters to improve prediction accuracy and reduce overfitting.

---

d.Model Evaluation

US-07:
As a *System Analyst*, I want to evaluate the trained model using accuracy, precision, recall, and confusion matrix so that I can measure its performance.

US-08:
As a *Project Stakeholder*, I want to compare multiple algorithms so that the best-performing model can be selected.

---

e.User Interface & Prediction

US-09:
As a *User (Farmer/Government Officer)*, I want to enter rainfall, temperature, humidity, and other environmental values into the form so that I can check flood risk.

US-10:
As a *User*, I want to receive a clear prediction result (High Risk / Low Risk) so that I can take preventive action.

US-11:
As a *User*, I want the system to provide quick real-time predictions so that I can make timely decisions.