Question 1

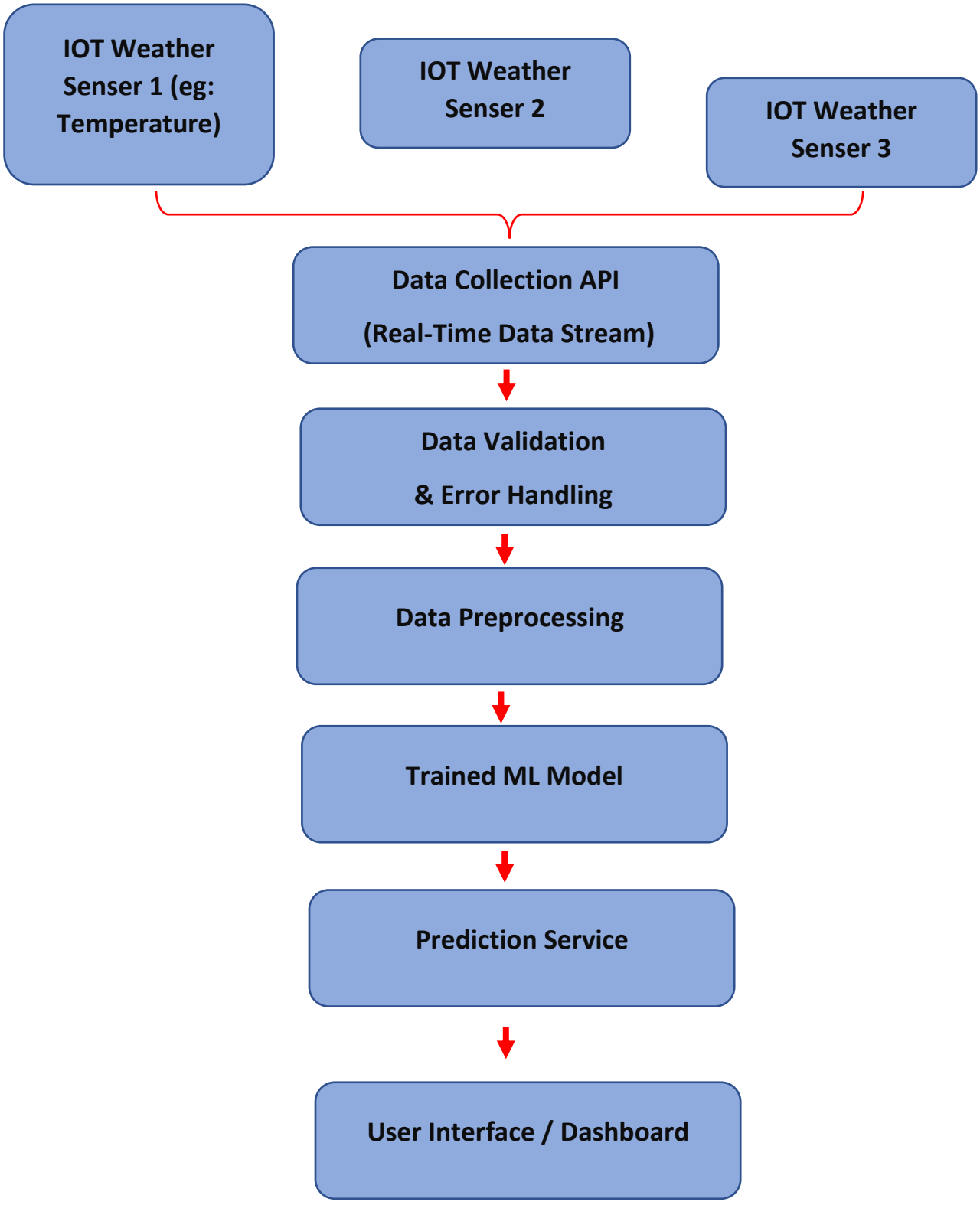
Part - 2

Weather Forecasting

Team Name: **DataMavericks**

To design a system that predicts the daily probability of rain for the next 21 days using real-time data, we need a robust architecture that handles data collection, preprocessing, model inference, and result delivery. Additionally, the system must address potential issues such as sensor malfunctions or missing data.

System Diagram



Component Descriptions

1. IoT Weather Sensors:
   - These devices collect real-time weather data such as temperature, humidity, wind speed, cloud cover, and pressure.
   - The data is transmitted to the system at 1-minute intervals.
   - Challenge: Sensors may malfunction or provide incomplete data.

2. Data Collection API:
   - An API ingests real-time data from IoT sensors and streams it into the system.
   - Challenge: Ensure reliability and scalability to handle continuous data streams.

3. Data Validation & Error Handling:
   - This component checks for missing or erroneous values in the incoming data. Missing values are imputed using techniques like mean or median imputation.
   - Malfunctioning sensor data is flagged and replaced with fallback values based on historical trends or default ranges.

4. Data Preprocessing:
   - The raw data is preprocessed to make it suitable for the machine learning model. Steps include:
     - Scaling numerical features.
     - Encoding categorical variables (if any).
     - Handling outliers to ensure robust predictions.
   - This step ensures consistency with the preprocessing done during model training.

5. Trained ML Model:
   - The pre-trained Logistic Regression model is deployed here to predict rain probabilities.
   - The model uses the processed real-time weather data as input and outputs the probability of rain for each of the next 21 days.

6. Prediction Service:
   - This service aggregates predictions and prepares them for user consumption.
   - It generates a daily report showing rain probabilities for the next 21 days.

7. User Interface / Dashboard:
   - A user-friendly interface displays predictions in an easily interpretable format (e.g., graphs, tables). Farmers can access daily rain probabilities to plan their activities effectively.