**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 03 February 2025 |
| Team ID | xxxxxx |
| Project Title | Implementing AgriPrediction using Machine Learning |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification Template**

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | This project aims to build a machine learning-based AgriPrediction system for forecasting crop yield, soil health, and weather conditions to assist farmers in making informed decisions. |
| Data Collection Plan | The data will be collected from multiple sources, including government agricultural databases, weather APIs, satellite imagery, soil reports, and farmer surveys. |
| Raw Data Sources Identified | · **Government Agricultural Data**: Historical crop yield, soil quality, and farming practices.  · **Weather APIs**: Real-time and historical weather data (temperature, humidity, rainfall).  · **Satellite Imagery**: Data on soil moisture, vegetation health, and climate patterns.  · **Soil Reports**: Information on soil pH, nutrients, and fertility.  · **Farmer Surveys**: Real-time field data on pests, irrigation, and farming practices. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Government Agricultural Data | Historical data on crop yield, soil quality, and farming practices. | Govt. Agri Data | CSV | 500 MB | Public |
| Weather Data (OpenWeather API) | Real-time and historical weather data, including temperature, humidity, and rainfall. | OpenWeather API | JSON | 5 GB | Free API (with rate limits) |
| Soil Test Data | Data on soil composition, pH levels, and nutrient concentrations. | Soil Data Repository | Excel | 50 MB | Private (with access) |
| Satellite Imagery | Remote sensing data for soil moisture, vegetation health, and climate . | NASA Earth Data | GeoTIFF | 20 GB | Public |