**Project Initialization and Planning Phase**

|  |  |
| --- | --- |
| Date | 12 July 2024 |
| Team ID | SWTID1720077079 |
| Project Title | Wild Blueberry Yield Prediction |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution)**

This project employs a machine learning system for accurate blueberry yield prediction, addressing farmer challenges with unpredictable weather and soil conditions. It includes data collection, advanced modeling, user-friendly interface integration, and rigorous testing. Required resources: high-performance computing, and a skilled team of data scientists, engineers, designers, and testers. Goal: Empower farmers with reliable yield forecasts to improve decision-making and operational efficiency.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | Develop a ML system for accurate blueberry yield prediction. |
| Scope | Includes data collection, advanced modeling and rigorous testing for operational reliability. |
| **Problem Statement** | |
| Description | This project aims to create a machine learning system to predict blueberry yields accurately. |
| Impact | Enhances farmers' ability to make informed decisions and improve operational efficiency by providing reliable yield forecasts. |
| **Proposed Solution** | |
| Approach | Utilize scikit- learn for developing predictive models integrating historical yield data, weather patterns, soil health, and pest dynamics. |
| Key Features | Includes comprehensive data preprocessing, advanced feature engineering, scikit- learn based model development, user-friendly interface integration, and rigorous testing for reliability. |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | CPU/GPU specifications, number of cores | Core i5, 11th gen  Nvidia GTX |
| Memory | RAM specifications | 8 GB |
| Storage | Disk space for data, models, and logs | 512 GB SSD |
| **Software** | | |
| Frameworks | Python frameworks | Flask |
| Libraries | Additional libraries | scikit-learn, pandas, numpy, matplotlib, seaborn, pickle |
| Development Environment | IDE, version control | Jupyter Notebook, Git, Spyder |
| **Data** | | |
| Data | Source, size, format | Kaggle dataset, 85 KB, CSV |