Smart Agriculture Data Analytics System

Team Name: The Cacheuals

Team Members:

Chandana Aluri

Gayathri Netapalli

Vydehi Ramineni

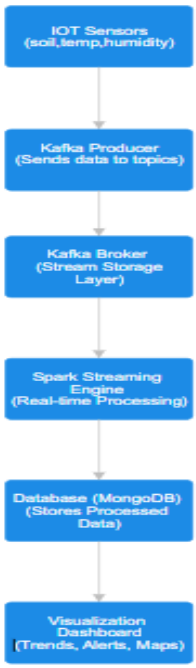
1. Project Idea:

This project will focus on creating a smart agricultural monitoring and analytics system that utilizes IoT sensors, Big Data processing, and the cloud to store data. The system will collect real-time environmental data from IoT sensors to measure soil moisture, temperature, humidity, and rainfall. The data will then be sent to Apache Kafka for real-time streaming and will be processed with Apache Spark to indicate if irrigation is needed or to identify crop stress conditions. The results will be stored in a database and visualized in a way that can be understood by farmers, adding to their ability to use data for making decisions to improve crop yield, decrease water usage, and optimize fertilizer utilization.

2.Tools and Technologies:

- IoT Sensors / data simulator: To gather or generate real-time soil and climate data.
- Apache Kafka: for a reliable streaming layer from sensors to processing.
- Apache Spark Streaming: for real-time processing and decision-making.
- MongoDB or AWS DynamoDB: for processed sensor data storage.
- Python: for processing and simulation.
- Power BI / Tableau / Web Dashboard (Flask or React): for visualization of analytics results.
- Cloud platform (AWS or GCP): for deployment, storage and scalability.

3. Architecture/Methodology:



4. Explanation of the Data Flow Diagram:

- Step-1: IoT Sensors: Gather information about moisture levels in the ground, air temperatures, air humidity, and precipitation from a farm.
- Step-2: Kafka Producer: Sends the data produced by the sensor to the Kafka topic for the streaming data pipeline in real-time.
- Step-3: Kafka Broker: Serves as the intermediary system that can handle continuous stream of data at high velocities on an ongoing basis.
- Step-4:Spark Streaming Engine: Will provide real-time processing of the data, the cleaning of it and the detection of irrigation or fertilizer needs.
- Step-5:Database (MongoDB / DynamoDB): Will store the data processed and historical data for wrap-up analysis in due time.
- Step-6: Visualization Dashboard: Where a farmer can observe environmental parameters, crop health indicators and predictive insights.

5. Project Goals:

- Create an IoT-based data simulator that will generate real-time environmental readings such as soil moisture level, temperature, and humidity.
- Implement a Kafka-based streaming pipeline that will collect, queue, and transmit data from multiple simulated sensors simultaneously.
- Create real-time data processing in Apache Spark whereby data can be cleaned, aggregated and analyzed to create irrigation and fertilizer recommendations.
- Store the processed recommendations and sensor reading in a scalable cloud Database (MongoDB or AWS DynamoDB) for further analysis.
- Create a web-based or BI dashboard that will show environmental trends, irrigation alerts, and crop condition insights in real-time.
- Create a predictive analytics module which predicts crop water needs and growing conditions using historical data and weather trends.