

Title: Intelligent Crop Yield Forecasting Using Big Data Analytics

Team Members: Abhinav Tote- G24AI1059, Harsh Goyal- G24AI1074, Harshal Ingle-G24AI1087, Omar Khan-G24AI1094

Explanation: This project is all about smart crop yield forecasting, using big data analytics to pull together various environmental data sources like weather patterns, soil traits, and satellite data. By employing data fusion and predictive modeling techniques, the application aims to deliver precise, real-time crop yield predictions. This helps farmers make the most of their resources, boost productivity, and make well-informed choices for sustainable farming.

Key Technology Challenges:

- Managing and merging large, diverse datasets from different sources like weather, soil, satellite imagery
- Ensuring data quality through cleaning, preprocessing, and addressing any missing or inconsistent information.
- Creating scalable predictive models that can update in real-time.
- Protecting proprietary agricultural data while adhering to data privacy regulations.

Technology Stacks Involved:

- Hadoop for distributed storage and big data processing.
- Apache Spark for fast in-memory data processing and analytics.
- Python for data manipulation, machine learning, and scripting.
- TensorFlow for building and training deep learning predictive models.
- Google Earth Engine for accessing and processing satellite and geospatial data.

List of Deliverables:

- A functional demo of the crop yield prediction application that integrates real-time environmental data.
- A codebase that includes data fusion modules, predictive modeling scripts, and data preprocessing pipelines.
- A report that compares model performance and accuracy across various algorithms and datasets.