SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
SCHOOL OF COMPUTING
DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

FIRST REVIEW

SOIL TESTING AND CROP YIELD PREDICTION USING MACHINE LEARNING

Project Category: Product

Guide Name:
Dr. R. Siva
Associate Professor

Students:
Risabh Rai (RA2111033010105)
Ashraya Agnihotri (RA2111033010149)

ABSTRACT

Soil is essential for successful agriculture, providing the nutrients needed to grow crops. Different soil types have unique properties that influence which crops can grow in them. Understanding these properties is crucial for determining which crops to plant in specific soils.

This project presents a model designed to assess soil fertility, recommend suitable crop seeds for fertile soils, and predict crop yield based on various soil characteristics. By analyzing predictions, we can recommend which crops are likely to perform best.

The study uses various ML algorithms like SVM, Random Forest, Naive Bayes, Linear Regression, MLP, and ANN for soil classification and crop yield prediction.

INTRODUCTION

Agriculture is essential for human survival, but as the global population grows, the demand for food increases, leading to the use of harmful pesticides that degrade soil quality. Machine Learning offers a transformative solution by enabling computers to predict crop yields and analyze soil fertility using data, helping farmers make informed decisions.

By collecting real-time soil and crop data, Machine Learning models, including SVM, ANN, Decision Trees, Naive Bayes, and Linear Regression, can classify soil fertility and predict crop yields, optimizing farming practices and reducing the need for harmful chemicals, ultimately leading to more sustainable agriculture

LITERATURE SURVEY

ID	Title	Methodology	Identification of Gaps and Limitations
1	**Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize** *Agriculture, Kuradusenge M, Hitimana E, Hanyurwimfura D, et al., 2023*	This study applied Random Forest, Polynomial Regression, and Support Vector Regressor to predict crop yields in Rwanda using weather data (rainfall, temperature) for Irish potatoes and maize. Random Forest provided the best results with high accuracy.	The research was limited to two crops in a specific district, without incorporating soil properties or other climatic factors beyond rainfall and temperature. Generalizability to other regions or crops is uncertain.
2	**Enhancing Agricultural Productivity: Predicting Crop Yields from Soil Properties with Machine Learning** *African Journal of Biological Sciences, Manju G, Sania Thomas, Binson V A, 2024*	The study utilized three machine learning models (k-NN, SVM, Neural Network) to predict crop yields for coconut, plantain, ginger, and tapioca based on soil properties from 500 samples. The k-NN model outperformed others, achieving an accuracy of 88.8%.	The study focused on a limited number of crops and specific soil properties, without considering external environmental factors like weather or broader geographical data.
3	**Impact of Regional Bans of Highly Hazardous Pesticides on Agricultural Yields: The Case of Kerala** *Agriculture & Food Security, Sethi A, Lin C Y, Madhavan I, et al., 2022*	This study analyzed the impact of regional pesticide bans on agricultural yields in Kerala using statistical analysis of pre- and post-ban yield data. Findings indicated significant changes in crop yield trends after the ban.	The study did not explore predictive models or long-term effects of alternative pesticides. The analysis was region-specific, limiting broader applicability.

ID	Title	Methodology	Identification of Gaps and Limitations
4	**Predicting Soil Nutrient Deficiency Using Remote Sensing and Machine Learning** *Computers and Electronics in Agriculture, Pantazi X E, Moshou D, Alexandridis T, et al., 2016*	The research employed remote sensing data coupled with machine learning models (Random Forest, SVM) to predict soil nutrient deficiencies in agricultural fields, achieving high prediction accuracy.	The study's reliance on remote sensing data may not capture all relevant soil properties, and its application is limited to the specific conditions under which the data was collected.
5	**Machine Learning-Based Prediction of Crop Yield: A Comparative Study** *Journal of Artificial Intelligence Research, Gandhi N, Armstrong L J, Petkar O, et al., 2016*	This comparative study evaluated the performance of several machine learning models (k-NN, Random Forest, Neural Network) using soil and weather data to predict crop yields.	The study did not account for economic factors or variations in farmer practices, which could significantly impact crop yields. Results may not generalize well to different regions or climates.
6	**Application of Convolutional Neural Networks for Plant Disease Detection** *IEEE Transactions on Neural Networks and Learning Systems, Zhang X, Qiao Y, Meng F, et al., 2020*	Convolutional Neural Networks (CNNs) were applied to detect plant diseases from image data, demonstrating high accuracy in classification and potential for early disease detection.	The study focused solely on disease detection without extending its findings to crop yield predictions.

PRODUCT VISION

Purpose: Make farming better with tools for soil testing and crop predictions.

Target Audience: Farmers who need help with their soil and crop decisions.

Unique Value Proposition: Offers simple, effective tools that use the latest technology to provide useful insights.

Long-Term Impact: Helps farmers grow more food efficiently and sustainably, supporting global food needs.

EXISTING SYSTEMS

1) IBM Watson Decision Platform for Agriculture:
This platform integrates AI, IoT, and weather data to offer insights into soil conditions and crop health, helping farmers make data-driven decisions.
Reference: https://newsroom.ibm.com/IBM-watson?item=30660

2) Microsoft Azure FarmBeats:
Azure FarmBeats is an end-to-end solution that aggregates data from sensors, drones, and satellites to provide actionable insights into farming operations.
Reference: https://www.microsoft.com/en-us/research/project/farmbeats-iot-agriculture/

3) John Deere's See & Spray Technology:
This innovative system uses computer vision and Machine Learning to identify and apply herbicides precisely on weeds, minimizing chemical usage.
Reference: https://www.deere.com/en/sprayers/see-spray-ultimate/

PROBLEM STATEMENT

In modern agriculture, accurately assessing soil fertility and predicting crop yield are crucial for optimizing farming practices and ensuring sustainable food production. Traditional methods for evaluating soil quality and predicting crop performance are often:

- time-consuming
- labor-intensive
- and may lack precision

Consider a scenario where a farmer wants to plant a new crop on their field. Traditionally, they would manually test soil samples in a lab to assess nutrient levels and suitability for the chosen crop. This process can take several weeks and may not always reflect current soil conditions accurately.

In real-time, using a Machine Learning-based system, the soil data can be analyzed instantly by the system, providing immediate feedback on soil fertility and recommending the best crop to plant based on current conditions.

OBJECTIVES

Purpose: Utilize machine learning to assess soil and forecast crop fertility.

Tasks:
- Analyze soil samples for nutrient levels and quality.
- Predict soil fertility to guide crop selection and fertilization.
- Optimize agricultural practices and enhance crop yields.

Approach: Combine machine learning with soil analysis.

Goal: Boost decision-making and crop production with data-driven insights.

PROPOSED SYSTEM

Model: A machine learning model that uses historical data, soil characteristics, and yield projections to recommend optimal crop rotation strategies for individual fields.

Goal: To enhance resource efficiency, improve soil health, and minimize pest threats.

Benefits:
- Efficient resource usage, including water and fertilizers
- Better soil health
- Decreased pest risks

Approach: Strategic crop rotation practices that are achieved through machine learning.

Target: Achieve sustainable and efficient farming practices.

PRODUCT BACKLOG

ID	Title	Epic	User Story	Priority	Acceptance Criteria	Functional Requirement	Non-Functional Requirement
1	Soil Data Collection	Data Acquisition	As a data scientist, I need accurate soil data to train the ML model for predicting crop yield.	Must	Soil samples are collected with details like pH, moisture, and nutrient levels.	Collect and document soil data from multiple locations.	Ensure minimal contamination and accurate reporting.
2	Data Preprocessing	Data Preparation	As a data engineer, I need to clean and preprocess soil data to ensure it's suitable for ML analysis.	Must	Raw data is cleaned, missing values are handled, and features are normalized.	Implement data cleaning and normalization procedures.	Complete preprocessing within a reasonable timeframe.
3	Model Selection	Model Development	As an ML specialist, I need to choose the best algorithm for predicting crop yield based on soil data.	Must	The best-performing model is selected based on accuracy and efficiency.	Evaluate and compare various ML algorithms.	Ensure reproducibility and transparency in model selection.

ID	Title	Epic	User Story	Priority	Acceptance Criteria	Functional Requirement	Non-Functional Requirement
4	Model Training	Model Training	As an ML engineer, I need to train the selected model using historical soil and crop yield data.	Must	The model is trained and validated, with performance metrics documented.	Implement training routines and validation procedures.	Optimize training time for timely project delivery.
5	Prediction Interface	User Interface	As a user, I want an interface where I can input soil data and receive crop yield predictions.	Should	The interface accepts soil data inputs and displays accurate predictions.	Develop a user-friendly input interface and prediction display.	Ensure response time for predictions is less than 2 seconds.
6	Model Evaluation and Tuning	Model Optimization	As an ML engineer, I need to evaluate and fine-tune the model to improve prediction accuracy.	Must	Model is tested on different datasets, and performance is optimized.	Implement evaluation and tuning procedures.	Tuning should not exceed 2 iterations to stay on schedule.

ARCHITECTURE/BLOCK DIAGRAM

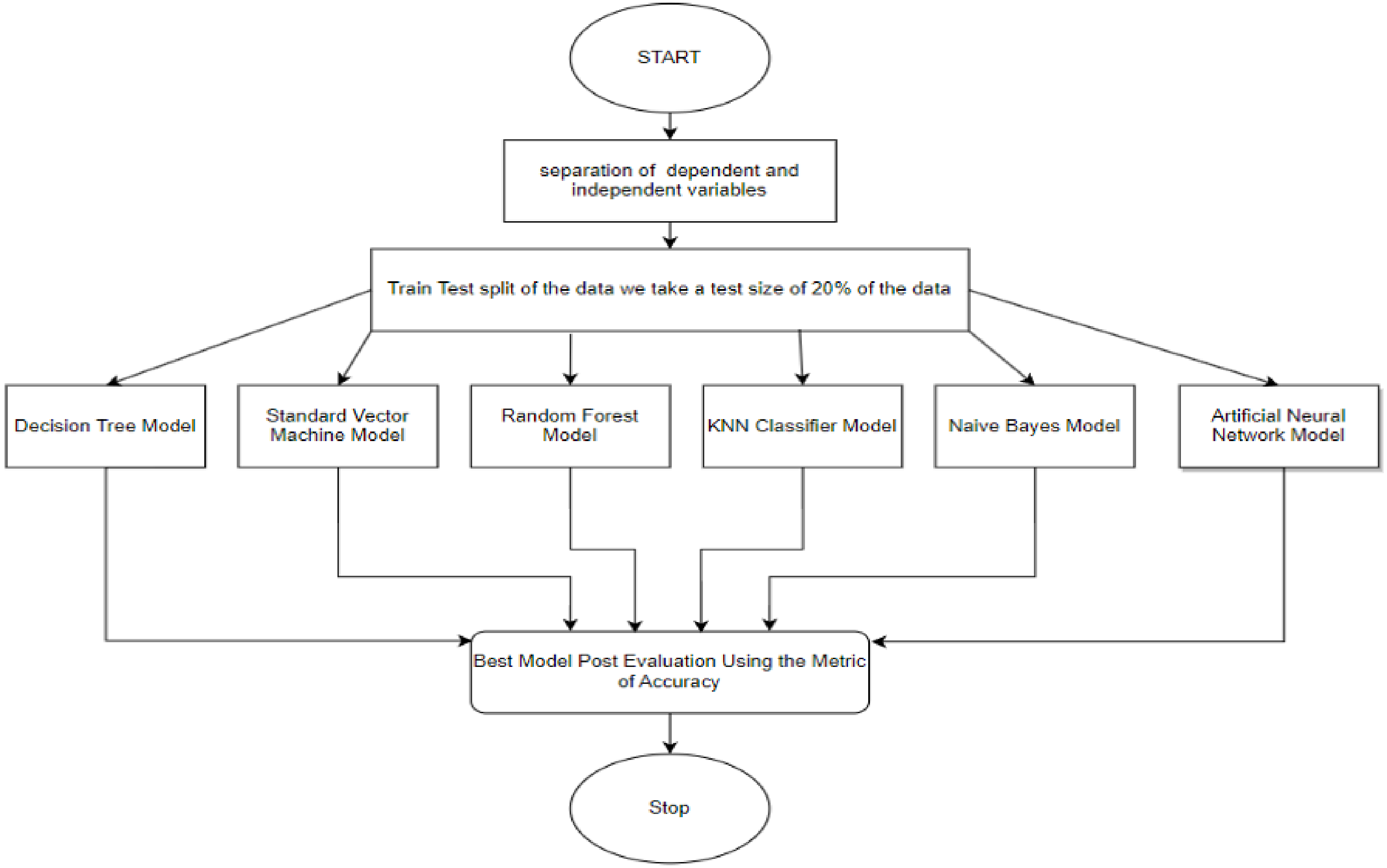


Figure 4.1.1: Architecture Diagram of Crop Prediction System

ARCHITECTURE/BLOCK DIAGRAM

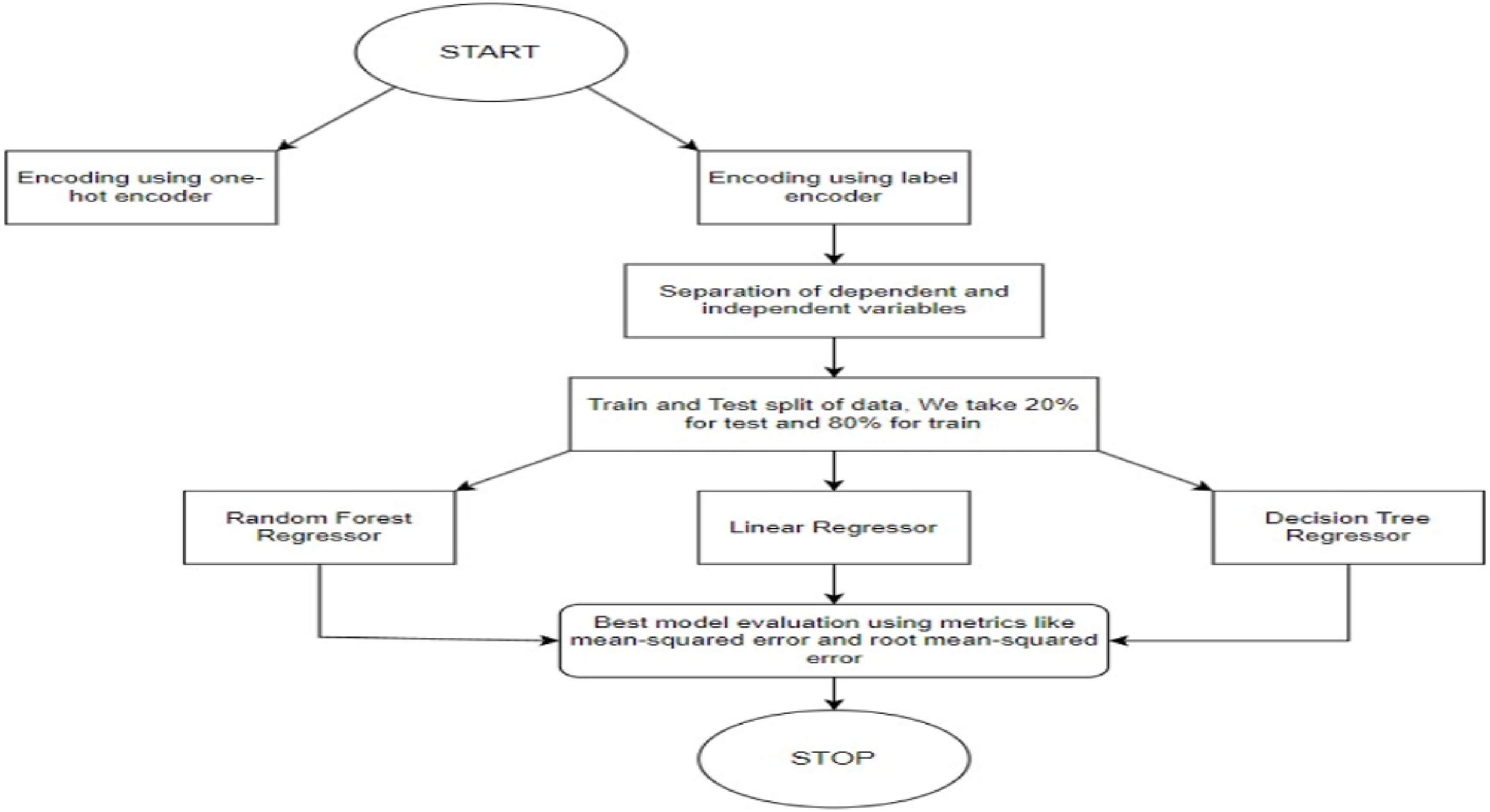
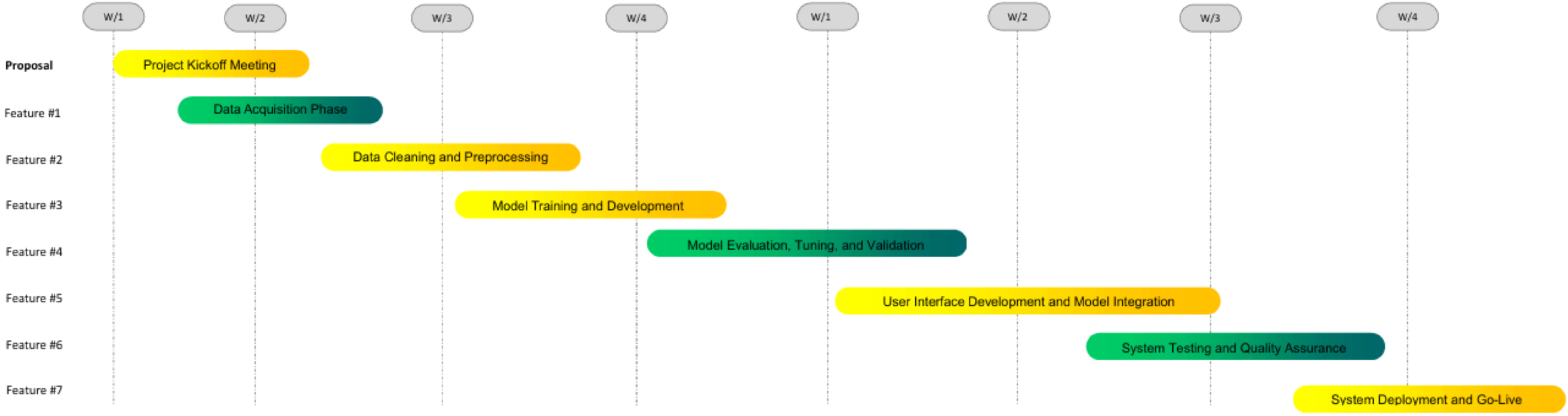


Figure 4.1.2: Architecture Diagram of Yield Prediction System

PROJECT TIMELINE

Release Plan/ Roadmap



REFERENCES

1) Jagdeep, Shalu, Vijayalakshmi(2021). Soil Analysis and Crop Fertility Prediction using Machine Learning. IJIRAE::International Journal of Innovative Research in Advanced Engineering, Vol: VIII,41-49.

2) SK AL Zaminur Rahman, Kaushik Chandra Mitra and S.M Mohidul Islam, “Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series”, 2018 IEEE, 21st International Conference of Computer and Information Technology (ICCIT), pp.978-1-5386-9242-4/ 18 .

3) Ashwini Rao, Janhavi U, Abhishek Gowda NS, Manjunatha and Mrs.Rafega Beham A, “Machine Learning in Soil Classification and Crop Detection”, IJSRD- International Journal for Scientific Research and Development, Vol-4, Issue 01, 2016, ISSN: 2321-0613.

THANKYOU

QUESTIONS AND FEEDBACK