National Institute of Technology Nagaland

Department : Computer Science and Engineering

Agricare

Phase- II

Project review-1

Project Guide : **Dr. Shouvik Dey**

Rahul Shah 2020105251

Satyam Raj 2020105253

Date : March 11, 2024

CONTENTS

1. Introduction
2. Objective
3. Application
4. Dataset
5. Previous work
6. Proposed Design
7. Implementation
8. Result
9. Conclusion
10. Future Work
11. References

INTRODUCTION

- A major contribution to India's economy is made by Agriculture.
- A wrong or misguided decision by a farmer can have a lot of impact on the agricultural economy of that region and India in a broader sense.
- They usually take historical parameters and ancestral farming patterns into consideration without knowing that crop depends on weather, present-day, and soil conditions.
- A combination of machine learning algorithms, historical and scientifically collected soil parameters, weather data, state-wise crop production can predict the most appropriate crop.

OBJECTIVE

Predicting the crop yield according to weather conditions and previous year production yield based on user location.

APPLICATION

- Data Driven Decision Making
- Financial Analytics for a Crop
- Predecting the crop yield

DATA SET

Dataset gathering and analysis

- In this step, all the required dataset is gathered.

1. Agro-climatic Database

- This mainly constitutes the record for weather conditions for various districts of India for year 1997-2014. This data consists of 2300 observations for **Rice** crops.

- T_i = Temperature for i^{th} month
- SH_i = Soil Humidity i^{th} month
- RH_i = Relative Humidity for i^{th} month

T1	T2	T3	T4	SH1	SH2	SH3	SH4	RH1	RH2	RH3	RH4
34.87	34.29	30.52	29.78	14.22	17.4	18.31	15.69	41.5	52.06	67.25	61.56
36.8	30.95	29.65	29	12.02	20.51	20.87	19.04	32.62	72.94	79.5	76.12
37.56	34.44	29.92	28.73	11.29	16.85	19.23	18.49	28.25	51.44	72.81	75.44
35.9	32.23	31.94	30.95	14.22	18.07	18.07	13.98	39.94	60.44	61.56	52
35.43	31.35	30.59	31.55	15.26	19.41	18.92	13	43.31	67.56	69.19	46.62

DATA SET

Continue...

- WS_i = Wind Speed for i^{th} month
- SW_i = Soil Wetness(at 100cm depth) for i^{th} month
- RF_i = Cummulative Sum Rainfall for i^{th} month

WS1	WS2	WS3	WS4	SW1	SW2	SW3	SW4	RF1	RF2	RF3	RF4
2.65	2.34	2.93	2.01	0.38	0.38	0.46	0.47	47.46	105.47	131.84	63.28
3.15	2.45	1.7	1.88	0.34	0.54	0.56	0.55	63.28	221.48	163.48	126.56
3.03	2.71	1.88	1.45	0.33	0.39	0.59	0.54	0	179.3	189.84	89.65
2.99	2.55	2.3	1.76	0.36	0.42	0.45	0.42	47.46	126.56	100.2	15.82
2.16	2.41	1.97	1.78	0.34	0.47	0.52	0.41	84.38	184.57	89.65	15.82

DATA SET

2. Crop Yield database

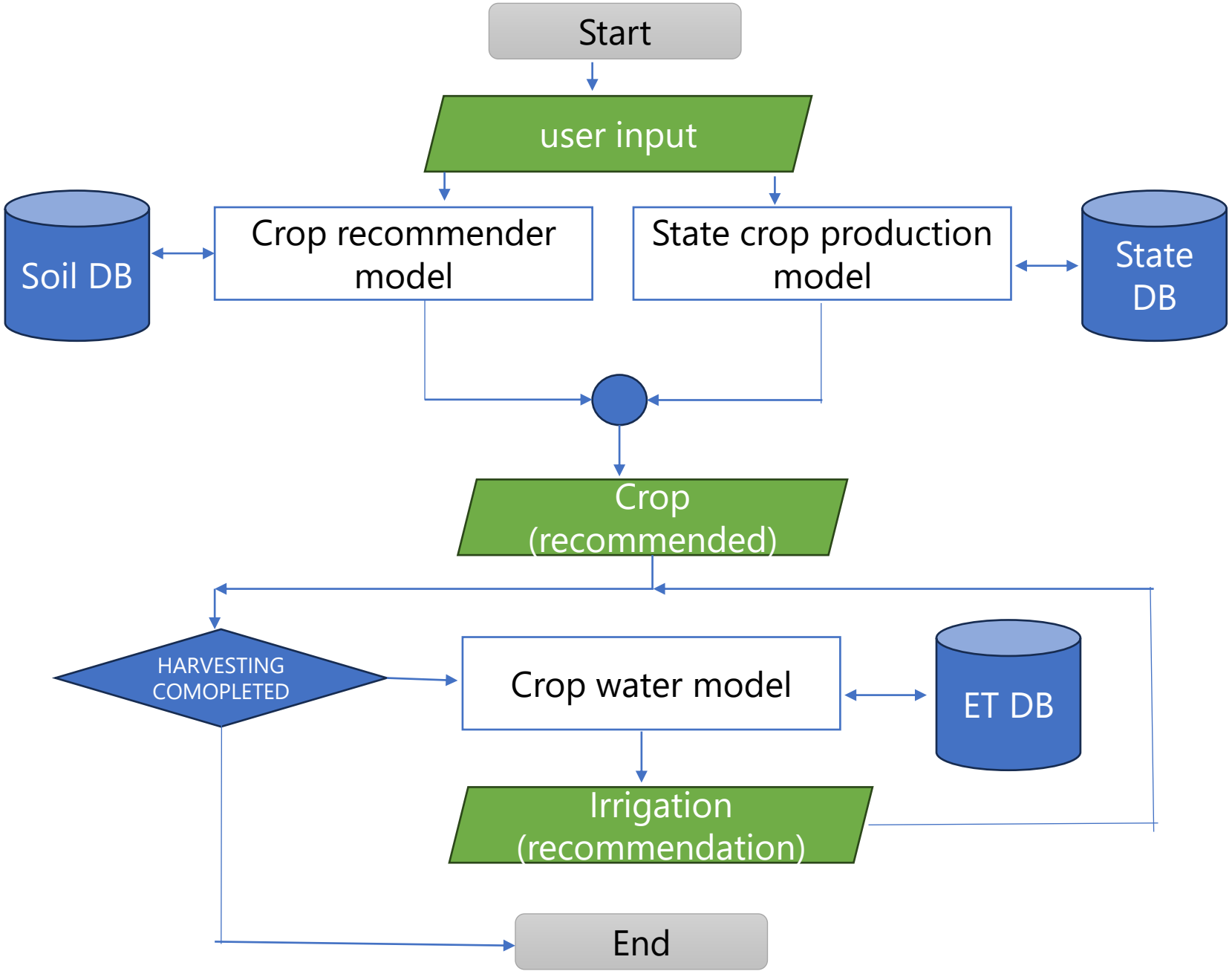
- This database consists of yield of crop from the various district of state of Uttar Pradesh for years 1997-2014

- Season : Agriculture season- Summer, Kharif
- Area in Hectares
- Production in Tonnes
- Yield in Tonnes/Hectares

Season	Crop	Area	Production	Yield
Kharif	Rice	935	2073	2.217112299
Kharif	Rice	1536	3478	2.264322917
Kharif	Rice	2142	3781	1.765172736
Kharif	Rice	1322	2502	1.892586989

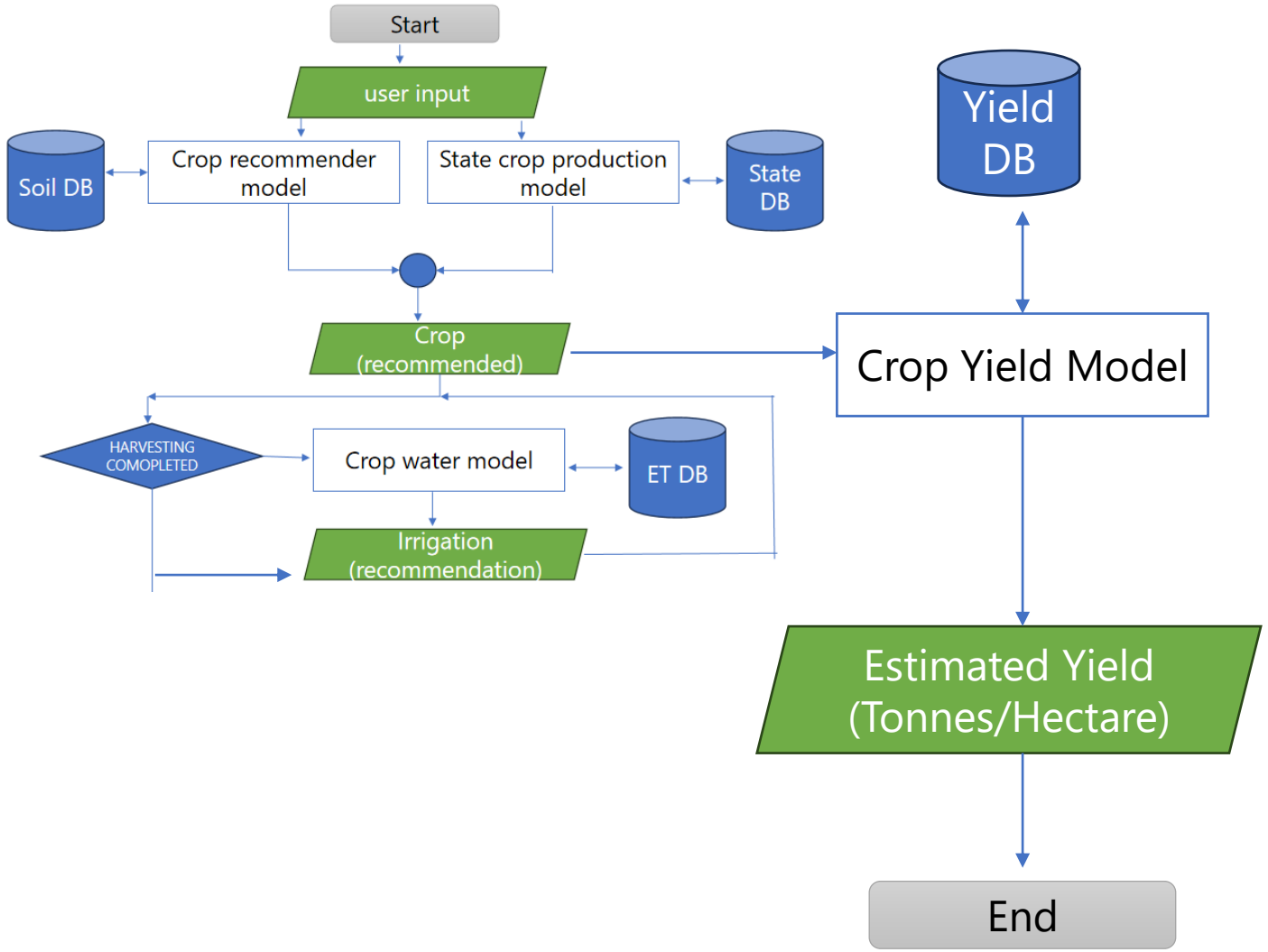
PREVIOUS WORK

- We had designed 3 machine learning models to support our purpose.
- They work on three different database.
- The user gives input(such as land location for Weather API and State production, and Soil composition such as NPK and pH value)
- We apply our ML models.
- Output the best suitable crop
- Apply crop water model to monitor the water required for irrigation at different stages of farming



PROPOSED DESIGN

- We create a new ML model to predict the crop yield for a particular crop(rice).
- This model takes input such as : temperature, rainfall, humidity, and soil moisture.
- Predicts the estimated yield based on historical data.



STEPS

A. Pre-processing

- Elimination of disturbances and outliers present in the CSV dataset
- We apply median, and mean of the whole column. Using SKLearn, we can conveniently clean the dataset.

B. Feature Selection

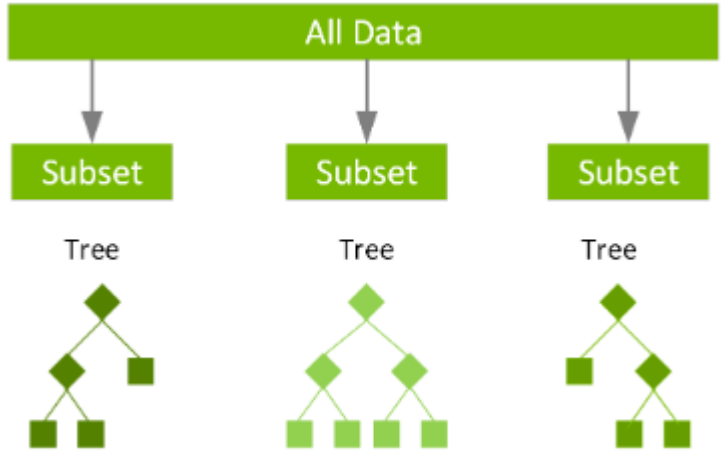
- Select Temperature, Humidity(relative & surface), soil moisture, windspeed as dependent features.
- Select yield as our target variable.

C. Choosing Machine Learning Model:

- Using pandas, a data frame was created that helped read the CSV file.
- Using Random Forest Regression model on Crop Yield databases

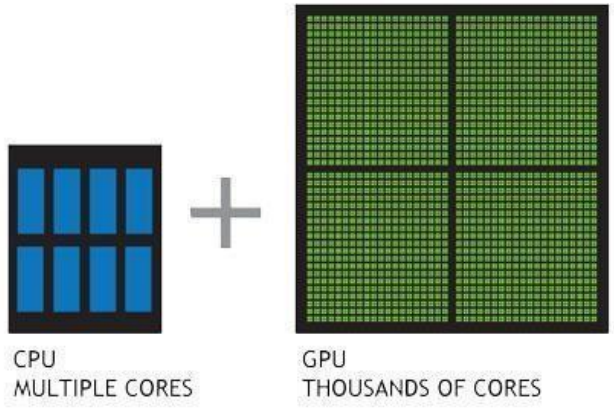
IMPLEMENTAION

Crop Yield Model



- This model takes gets trained on the Crop yield database.
- **XG Boost** is used along with **Random forest regressor**.
- **XG Boost** : Extreme Gradient Boosting, is a scalable, distributed gradient-boost decision tree (GBDT) machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems.

- **Random forest** uses a technique called "**bagging"** to build full decision trees in parallel from random bootstrap samples of the data set. The final prediction is an **average** of all of the decision tree predictions



- **Gradient Boosting** is based on "**boosting**" or improving a single weak model by combining it with a number of other **weak models** in order to generate a collectively strong model

RESULT

We test our model with Radom Forest Regressor Algorithm.

The calculated Root mean square error and R2 score are given below.

Root Mean Square Error (RMSE)	R2 Score
0.1296967329500163 2	0.477095143599374 3

CONCLUSION

- Our project is designed to support farmers.
- By inputting soil data and weather information, the system recommends the optimal crop for production, considering the specific soil features.
- We calculate the esitmated crop yield for rice according to weather conditions , soil moisture at crop root level, and historical data.
- Project is implemented only on certain aspects of agriculture. Precision Farming is a vast subject.

FUTURE WORK

- Creation of a user-friendly UI Application for seamless interaction.
- Continuous refinement of Crop Yield Model
- Utilization of a Weather API for real-time information on rainfall precipitation, and soil humidity.

REFERENCES

1. Thomas van Klompenburg, Ayalew Kassahun, Cagatay Catalb (2020) Crop yield prediction using machine learning: A systematic literature review, Computers and Electronics in Agriculture 177 (2020) 10570
2. Ayush Shah, A., Dubey, A., Hemnani, V., Gala, D., Kalbande, D.R., 2018. Smart Farming System:Crop Yield Prediction Using Regression Techniques In: Vasudevan, H., Deshmukh, A., Ray, K. (eds) Proceedings of International Conference on Wireless Communication . Lecture Notes on Data Engineering and Communications Technologies, vol 19. Springer, Singapore
3. Paul, M., Vishwakarma, S.K., Verma, A., 2015. Analysis of soil behaviour and prediction of crop yield using data mining approach. In: 2015 International Conference on Computational Intelligence and Communication Networks (CICN). IEEE
4. Ramesh Medar, Vijay S. Rajpurohit, Shweta Shweta (2019) Crop Yield Prediction using Machine Learning Techniques IEEE 5th International Conference for Convergence in Technology (I2CT)

Thank You