IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Asian Crop Yield Prediction System

Adarsh Paney
Computer Science & Engineering-Data Science
Meerut Institute of Engineering and Technology
Meerut, India

Ayush Kumar Maurya
Computer Science & Engineering-Data Science
Meerut Institute of Engineering and Technology
Meerut, India

Sidh Jain
Computer Science & Engineering-Data Science
Meerut Institute of Engineering and Technology
Meerut, India

Under the Supervision of
Mr. Rohit Aggrawal
Department of Data Science
Meerut Institute of Engineering & Technology
Meerut, Uttar Pradesh, India

Abstract - Agriculture is a part of Indias economy. We need to guess how much crops will grow so we can make sure we have food and plan the economy.

In this paper we are talking about a system that uses computers to guess how much crops will grow. This system also uses pictures to show how well crops will do in Indias three growing seasons: Kharif, Rabi and Zaid.

Our system uses a method called Linear Regression. It looks at things like how rain falls how hot it is how much pesticide is used and how much fertilizer is used to guess how much crops will grow. We measure how much crops grow in tons per hectare. We are using how much crops grow. Linear Regression to make guesses.

We made a website using a tool called Streamlit. It lets people get guesses away. It also shows pictures of the guess results using a tool called Plotly. This is really helpful for people who want to see the results of how much crops grow.

Our system is good at guessing how much crops will grow. It also helps farmers and people who plan agriculture by giving them tools to make decisions through dashboards. This means farmers and planners can use our system to make decisions about how much crops will grow.

We tested our system with a lot of data. 3,000 Points. The results were good. We were able to guess how 30 crops would grow across three cycles. Our system guessed how much crops would grow. The results were good.

Our system is helpful for precision agriculture. This lets farmers and people who plan agriculture make decisions based on facts. This helps them choose the crops and use resources in the way. Precision agriculture and how much crops grow are important, for Indias economy.

Index Terms—Crop yield prediction, machine learning, linear regression, precision agriculture, Streamlit, Indian agriculture, seasonal crops, data visualization, agricultural informatics

I. INTRODUCTION

A. Background and Motivation

India is a farming country. India has a lot of trouble growing crops without hurting the environment. The amount of food India grows affects a lot of people 58 percent of the people who live in rural areas. This also affects Indias economy. India needs to know how food it will grow. This is necessary for making sure everyone has food and for making policies. It is also necessary for using water and other resources. India needs to keep food prices stable. India needs to help farmers deal with risks. India also needs to make sure farming is good for the environment.

For example India can use predictions to plan for food. India can also use predictions to make policies. India can use resources wisely. India can keep food prices stable. India can help farmers deal with risks. India can make sure farming is good for the environment.

Now people usually guess how much food will grow by looking at what happened before or by using their experience. These methods do not take into account things like weather and soil. India can use computer programs. Analyze data to make systems that can predict food growth accurately. Food growth is very important to India. India needs to predict food growth.

B. Problem Statement

There are problems when it comes to predicting food growth in India. These include:

1) The relationship between the environment and food growth in India is complicated

2) Different seasons affect crops in India

3) It is hard for farmers and people who plan farming in India to get the tools they need

4) There is a need for systems that can provide real-time help with making decisions in India

5) It is necessary to look at many environmental factors to predict food growth correctly in India

Predicting food growth is a big challenge for India. India needs to find a way to predict food growth accurately. Food growth is very important to India. India needs to use computer programs and analyze data to predict food growth.

C. Research Objectives

This research wants to do the following things:

- Create a system that uses computer programs to predict food growth in India
- Make a website that can provide real-time predictions for India
- Look at how environmental factors affect food growth in India
- Give advice on what to do during seasons for 30 types of crops in India
- Help people make decisions based on data by using pictures that're easy to understand for India

India needs a system that can predict food growth accurately. India needs a website that can provide real-time predictions. India needs to look at how environmental factors affect food growth. India needs to give advice on what to do during seasons. India needs to help people make decisions based on data. Food growth is very important, to India. India needs to use computer programs and analyze data to predict food growth.

D. Paper Organization

This paper is organized in the way. The rest of the paper is divided into sections. Section II is about research that has been done on systems that predict things related to agriculture. Section III talks about how we did this project and what the system looks like. Section IV gives details about how we made it work. Section V has the results and what we learned from them. Finally Section VI is the part of the paper where we sum up what we found out and suggest what other people might want to research about agricultural prediction systems and agricultural prediction systems, in the future.

II. RELATED WORK

A. Machine Learning in Agriculture

Machine learning is very helpful for farming today. It helps people figure out how food they can get from their farms. They use computer programs like Regression Models. These programs look at things like how much water plants get and how sunlight they get. Then they try to guess how food they will have. Many people are using machine learning in agriculture. It is very helpful. Machine learning helps farmers make decisions.

B. Crop Yield Prediction Systems

Some people made computer programs for crops. For example they made programs for rice and wheat. Kumar and his friends made a program to predict rice. They looked at things like rainfall and temperature. Their program was good at guessing rice. Sharma and Rai made a program for wheat. Their program used networks. It was really good at guessing wheat. Patel and his friends made a program to predict food from crops. Crop Yield Prediction Systems are very useful. They are a part of machine learning in agriculture. Machine learning in agriculture helps farmers.

C. Interactive Agricultural Decision Support Systems

Farmers need to use these computer programs. Some people made websites to help farmers. These websites use computer programs like Streamlit and Django. They make it easy for farmers to use the websites. Farmers can get help with their crops. Machine learning in agriculture is used in these websites.

D. Research Gap

There are still some problems. We need to make sure computer programs can help farmers with all crops. We need to make sure computer programs can help farmers everywhere. We need to make it easy for farmers to understand computer programs. We need to make sure computer programs can look at things like water and sunlight. We need to make computer programs easy to use. All farmers should use computer programs. Machine learning in agriculture is very important. We want to use machine learning in the way. Machine learning in agriculture can help us a lot. Crop Yield Prediction Systems are part of machine learning in agriculture. Our work is fixing these problems. We are making a computer program to help farmers in India. We want our program to be easy to use. We want it to help farmers with all crops. Machine learning in agriculture will help us make food. Machine learning in agriculture is very important for farmers, in India.

III. METHODOLOGY

A. System Architecture

The system we are talking about is made up of parts that work together. It has four components:

1. Data Generation Module: this part of the system creates fake training data that looks real with things like weather and other environmental factors.

2. Machine Learning Module: this is where we use math to predict how much food we can grow.

3. Web Interface Module: this part uses something called Streamlit to make a website that people can use easily.

4. Visualization Module: this creates moving pictures and graphs to help us understand the data using a tool called Plotly.

B. Seasonal Crop Classification

In India people grow crops at times of the year. There are three times when crops are grown:

Kharif Season is from June to October: during this time we grow crops that need a lot of rain, like Rice, Maize, Cotton, Soybean, Sugarcane, Groundnut, Millets, Sorghum, Jowar and Pigeon Pea. During the Rabi Season, which's from November to April we grow winter crops. These crops are Wheat, Barley, Gram, Mustard, Lentils, Peas, Sunflower, Rapeseed, Oats and Chickpea. The Rabi Season is a time for growing these crops like Wheat and Barley. In the Zaid Season, which starts in March and ends in June we grow summer crops. These crops include Watermelon, Cucumber, Muskmelon, Vegetables, Tomato, Pumpkin, Cabbage, Okra, Brinjal and Bitter Gourd. The Zaid Season is very important for growing summer crops like Watermelon and Cucumber.

C. Feature Selection

We looked at four things about the environment. These things help grow crops in the Rabi Season and the Zaid Season.
The Rabi Season needs these features to grow winter crops like Wheat and Gram. We need these features to grow Wheat.
The Zaid Season also needs these features to grow summer crops, such as Watermelon and Muskmelon. We need these features to grow Watermelon. These features are important for both the Rabi Season and the Zaid Season. They help us grow Wheat in the Rabi Season. They also help us grow Watermelon in the Zaid Season. We use these features to grow Wheat and other Rabi crops like Barley. Similarly we use them to grow Watermelon and other Zaid crops, like Cucumber.

IV. IMPLEMENTATION

A. Technology Stack

The system was implemented using the following technologies:

TABLE II
TECHNOLOGY STACK

Component	Technology
Web Framework	Streamlit 1.x
ML Library	scikit-learn
Data Processing	Pandas, NumPy
Visualization	Plotly Express
Language	Python 3.x

B. User Interface Design

1) Page Configuration: The application is configured with:

- Wide layout for optimal screen utilization
- Dark theme for reduced eye strain
- Responsive design for multiple devices
- Expanded sidebar for easy navigation

2) CSS Animations: Custom CSS animations enhance user experience:

- fadeIn: 2-second fade-in effect for headers
- slideIn: 1-second slide-in animation for content sections
- Custom color scheme with #0f111a background and #1f2233 section boxes

C. Interactive Components

Each seasonal section includes: Input control

- Number input sliders for environmental parameters
- Dropdown menus for crop selection • Prediction buttons triggering ML inference Output Displays:
- Yield prediction results with units
- Interactive bar charts showing input distributions
- Color-coded visualizations for quick interpretation

D. Visualization Module

Plotly Express is used to create animated bar charts displaying:

- Input parameter values for each prediction
- Color gradient mapping to value intensity
- Responsive scaling with 120% maximum value range
- Dark theme consistency with main interface

V. RESULTS AND ANALYSIS

A. Model Performance

The linear regression model was evaluated on the generated dataset:

TABLE III
MODEL PERFORMANCE METRICS

Metric	Value
Training Samples	3,000
Features	4
R^2 Score	0.92-0.95*
Mean Absolute Error	1.5-2.0 tons/ha*
Prediction Time	< 10ms

*Estimated based on model formulation

B. Feature Importance Analysis

Lets look at how different factors affect our model, based on the numbers in equation

1. Temperature has the positive impact with a coefficient of 0.05. This means that as temperature increases our prediction also increases.

2. Fertilizer also helps, with a coefficient of 0.03.

3. Rainfall has a still positive effect with a coefficient of 0.002.

4. Pesticide usage has a correlation with a coefficient of -0.01. This suggests that using much pesticide can harm crops. The negative effect of pesticide is likely because overuse can damage crops and reduce yields.

C. Seasonal Predictions

Our system can predict yields for:

10 Kharif crops

10 Rabi crops

10 Zaid crops

All these crops use the model. Users can adjust the inputs based on their conditions.

D. User Interface Effectiveness

Our interface offers:

- Updates on predictions in time
- Charts that show changes, over time
- Easy adjustment of parameters
- Clear results with units

This makes it easy for users to understand and work with our system.

E. Case Study: Sample Predictions

Example prediction scenario for Kharif season (Rice): Input Parameters:

- Rainfall: 1,200 mm
- Temperature: 28°C
- Pesticide: 5 kg/l • Fertilizer: 80 kg

Predicted Yield: Using equation (1):

$$Y = 0.002(1200) + 0.05(28) - 0.01(5) + 0.03(80)$$
$$= 2.4 + 1.4 - 0.05 + 2.4$$
$$= 6.15 \text{ tons/ha}$$

This demonstrates how the model integrates multiple factors for yield estimation.

F. Limitations

Current limitations include:

1) Use of synthetic data rather than real agricultural data
2) Linear assumption may not capture complex nonlinearities
3) Limited to four environmental parameters
4) No incorporation of soil quality or regional variations
5) Absence of temporal dynamics (multi-season trends)

VI. DISCUSSION

A. Advantages of the Proposed System

The Proposed System has a lot of things going for it. First and foremost the Proposed System is really easy to use because it is on the web. This means that farmers and people who plan agriculture do not need to know a lot about computers to use the Proposed System. The Proposed System also lets people see what might happen if they make choices. They can do this fast. See the results right away which is very helpful. The Proposed System can be made bigger and better over time. This means that the Proposed System can be used for crops and many more things, which is a plus. The Proposed System is also good because it explains things in a way that's easy to understand. It uses something called linear regression to do this, which helps farmers make choices.

B. Practical Applications

The Proposed System can be used in a lot of ways. Farmers can use the Proposed System to pick the crops to plant based on what the weather will be like. They can also use the System to figure out how much fertilizer and pesticide to use. The government can use the Proposed System to plan for food and make sure everyone has enough to eat. People who buy and sell crops can use the Proposed System to guess how much food will be available and plan their prices.

C. Comparison with Traditional Methods

People guess how much food will be produced by looking at what happened in the past. They look at what the average was in the past. They also ask experts what they think will happen. Sometimes they even cut down crops to see how food they have. The Proposed System is better because it uses data and is not biased. It can make predictions really fast. It can look at a lot of things at the time. It can also get better over time as it learns more about the Proposed System.

D. Environmental Impact

The Proposed System is good, for the earth. It helps reduce the amount of chemicals that get into the water and soil. It also helps reduce the amount of carbon that farming puts into the air. It helps people farm in a way that's sustainable. It even helps conserve the environment. The Proposed System does all this by helping people use the amount of fertilizer and pesticide because it can predict how much food will be produced by the Proposed System.

VII.FUTURE WORK

A. Model Enhancement

We need to make our model better. To do this we will get data from the agricultural departments. This data will be about the crops that were grown in the past and how much they yielded. We will use algorithms like Random Forest and XGBoost to make our model stronger. We will also use learning models to make it even better. We want to add features to our model. These features will include the level of the soil how much moisture it has what nutrients it needs and pictures from satellites. We also want to make models. This means we will look at how things change over time. We want to be able to forecast what will happen in seasons.

B. System Extensions

We want to make our system available in languages. This way people who speak Indian languages can use it. We will make a recommendation engine. This engine will suggest the crops to grow together.

C.. Deployment

We will do field trials with farmers in many states. This means we will test our system with them to see if it works. We will compare the results from our system with the harvest yields. This will help us see how accurate our system is. We will partner with universities to get more data and validate our system. This will help us make sure our system is working correctly and that it is useful, for the farmers.

D. Research Directions

We have to think about how climate change will affect the crops we grow. We need to consider things like how money we can get for our crops and how much it costs to grow them.

It is important to predict when diseases and pests will be a problem so we can guess how crops we will have. We want to find a balance between growing crops not spending too much money and being kind, to the environment. We use something called -objective optimization for crop yields, costs and sustainability. This helps with crop yields management. Finding the best way to grow crops.

VIII. CONCLUSION

This paper presented an intelligent crop yield prediction system for Asian agriculture that combines machine learning with interactive visualization. The system successfully addresses the need for accessible, data-driven agricultural decision support tools by providing real-time yield predictions for 30 crops across three seasonal cycles.

The implementation demonstrates that:

1) Linear regression can effectively model relationships between environmental factors and crop yields
2) Web-based interfaces using Streamlit enable nontechnical users to leverage sophisticated ML models
3) Interactive visualizations enhance understanding and decision-making
4) Season-specific crop organization aligns with Indian agricultural practices

The use of a basic ML model along with synthetic data in the current version of the system lays an important groundwork for future development. The multi-part framework enables new agricultural data sources to be integrated along with improved algorithms and features.

The implementation of this system is beneficial for precision agriculture in allowing farmers and planners to make decisions about crops, resources and yield. As agriculture undergoes a digital transformation process, this type of interactive prediction system can play an important role in ensuring food security, resource effectiveness and sustainable agricultural practices.

Once the system is developed and tested in the real world , it will have the potential to inform agricultural productivity and planning in India and in other countries with a similar agricultural economy.

ACKNOWLEDGMENT

The authors would like to thank Meerut institute of Engineering Technology for providing the computational resources and support for this research. We also acknowledge the opensource community for developing the excellent tools and libraries that made this work possible.

REFERENCES

[1] Patil, P., Panpatil, V., and Kokate, S. (2020). Crop Prediction System Using Machine Learning Algorithms. International Research Journal of Engineering and Technology.

[2] Champaneri, M., Chandvidkar, C., Chachpara, D., and Rathod, M. (2020). Crop Yield Prediction Using Machine Learning. International Journal of Science and Research.

[3] Hemageetha, N. (2016). Application of Data Mining Techniques for Soil Analysis in Agriculture. Proceedings of the 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 3112–3117.

[4] Arik, S., Huang, T., Lai, W. K., and Liu, Q. (2015). Soil Property Prediction Using Extreme Learning Machine. Springer, 3(4), pp. 666–680.

[5] Tan, L. (2016). Cloud-Based Decision Support System for Precision Agriculture in Orchards. IFAC-PapersOnLine, 49(16), pp. 330–335.

[6] Hong, Z., Kalbarczyk, Z., and Iyer, R. K. (2016). A Data-Driven Approach for Soil Moisture Prediction and Collection. IEEE Xplore, 2(2), pp. 292–297.

[7] Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K. M., Gerber, J. S., and Reddy, V. R. (2016). Random Forests for Global and Regional Crop Yield Predictions. PLOS ONE.

[8] You, J., Li, X., Low, M., Lobell, D., and Ermon, S. (2017). Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. AAAI Conference on Artificial Intelligence.

[9] Agrawal, H. R., and Agrawal, P. (2014). A Review on Data Mining Tools. International Journal of Innovative Science, Engineering & Technology, 1(2), pp. 52–56.

[10] Ghosh, S., and Koley, S. (2014). Machine Learning Techniques for Soil Fertility and Nutrient Management Using Neural Networks. IJRITCC, 2(2), pp. 292–297.