JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AGRICULTURAL DECISION SUPPORT SYSTEM USING MACHINE LEARNING AND DATA ANALYTICS

[1] **Mr.M. JAYAKANDAN**
[1]Assistant Professor,

[2] **Ms. A.ARUNADEVI**
[2]PG Student II-M.Sc Computer Science,

[3]**Mr.M.S.MOHAMED NAWFAL**
[3] PG Student II- M.Sc Computer Science

[4]**Ms.K.NISHALINI**
[4]PG Student II-M.Sc Computer Science

PG & RESEARCH DEPARTMENT OF COMPUTER SCIENCE,
E. G.S. PILLAY ARTS & SCIENCE COLLEGE-NAGAPATTINAM-611 002, TAMIL NADU, INDIA.

ABSRACT:

The Agricultural Decision Support System is a machine learning-based approach designed to revolutionize agricultural practices through advanced data analysis and predictive modeling. By focusing on predicting crop yields and efficiently managing soil nutrients, it aims to address key challenges faced by farmers. Using historical and real-time data, the system provides actionable recommendations for optimal nutrient management strategies and accurate yield predictions. This empowers farmers with data-driven insights, thereby enhancing productivity, sustainability, and overall farm efficiency.

KEYWORDS: Agricultural Decision Support System, Machine Learning, Predictive Modeling, Crop Yield Prediction, Soil Nutrient Management.

I. Introduction

The system's architecture incorporates a user-friendly application interface that facilitates easy access to its features. A prominent aspect of this interface is the integrated chatbot, which is available on the homepage. This chatbot serves as a virtual assistant, ready to address any questions and doubts users may have regarding agriculture and farming. Whether it's about specific crop treatments, soil conditions, or seasonal farming practices, the chatbot provides instant, reliable information, making the farmer's job easier and more informed.

One of the significant strengths of the Agricultural Decision Support System is its flexibility and scalability. As agricultural datasets expand over time, the application is designed to adapt and grow alongside them. This ensures that the yield predictions remain accurate and relevant, no matter how large or diversified the dataset becomes. This adaptability is crucial for keeping up with the dynamic nature of farming and environmental conditions. Overall, the project aspires to integrate cutting-edge technology with practical agricultural knowledge, ultimately fostering a future where farming is both highly productive and ecologically sustainable. Through continuous improvements and updates based on user feedback and new data, the Agricultural Decision Support System aims to remain at the forefront of agricultural innovation.

The current state of agricultural yield prediction is plagued by limitations that hinder optimal farm management. Traditional methods often rely on generic models and historical averages, neglecting crucial factors like specific soil composition (beyond NPK), planting seasonality (rabi/kharif), and field size. This limited data scope results in inaccurate predictions, leading to inefficient resource allocation and potentially missed opportunities for maximizing yield potential. Furthermore, generic models are prone to overfitting, which reduces their effectiveness on individual farms with unique characteristics. Existing approaches also present user-friendliness challenges, requiring technical expertise to interpret complex data analysis tools. Finally, they fail to offer initial guidance on potential factors influencing yield, making it difficult for farmers to make informed decisions about resource allocation and crop management strategies.

In essence, current methods leave farmers flying blind, unable to fully leverage the power of data to optimize their operations. This project proposes a web application that bridges this gap, empowering farmers with a more comprehensive and data-driven approach to agricultural enhancement. However, the fragmented nature of agricultural data, often siloed across various sources, further complicates the ability to gain holistic insights. This project not only addresses the limitations of existing prediction models but also aims to integrate seamlessly with existing data sources, creating a unified platform for informed decision-making.

II. Related Work

Some farm management software platforms like Granular or John Deere's Field Connect offer yield prediction tools and integrate with chatbots for basic information retrieval or automated tasks. These chatbots likely won't have the same level of sophistication as a dedicated yield prediction chatbot, but they can provide basic support. Companies like Ward Labs or Logan Labs offer soil testing services and recommendations for fertilizer and amendments based on the analysis. While they don't use chatbots, they provide results and recommendations. Finding a system that combines yield prediction, soil supplement recommendations, and advanced chatbot interaction is likely less common. This is an area of ongoing development, and dedicated agricultural chatbots are still emerging.

Proposed methodology

The proposed web application offers a multifaceted agricultural enhancement solution by integrating yield prediction, soil supplement recommendation, and chatbot interaction. The system leverages user-provided field data to deliver targeted recommendations and predictions specific to the context of Indian agriculture. Users can input various soil nutrient levels, including but not limited to macro-nutrients like Nitrogen (N), Phosphorus (P), and Potassium (K) (collectively known as NPK).

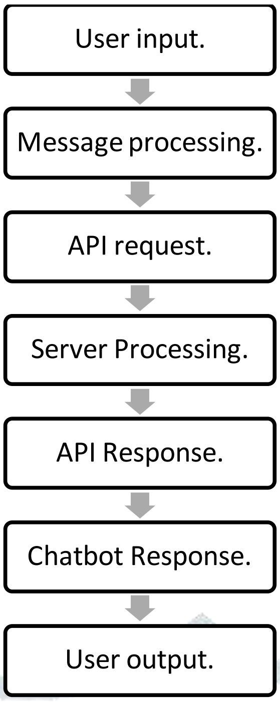
Additionally, users can specify the presence of other relevant elements that may influence crop growth in their region. To ensure accurate predictions, users can define the field area under consideration. Furthermore, the system allows users to specify the desired prediction timeframe by entering the target year and **season (rabi or kharif)**. This selection is crucial as India's climate necessitates distinct planting cycles for these two seasons. Finally, to account for regional climatic variations, geolocation data is collected, including the user's city within Tamil Nadu. This comprehensive data collection approach empowers the application's core functionalities to deliver valuable insights tailored to the user's specific agricultural needs and chosen planting season.

- **Enhanced Prediction Accuracy:** Going beyond basic field data, the application allows users to specify details like additional soil elements, field area, and planting season (rabi or kharif). This richer data creates a more comprehensive picture for the prediction model, potentially leading to **more accurate yield forecasts** compared to generic models.
- **Reduced Overfitting Risk:** Traditional methods relying on Random Forest Regressors (RFRs) are prone to overfitting, especially when dealing with limited data. This application focuses on the specific region of Tamil Nadu and incorporates user-provided data. This approach keeps the model focused on **relevant data points**, reducing the risk of overfitting and improving generalizability to various farm conditions within the region.
- **Improved User Experience:** The user-friendly chatbot interface simplifies interacting with the application. Farmers can ask questions, receive guidance on potential nutrient issues, and gain insights without needing the technical knowledge required for fine-tuning RFRs.
- **Informed Decision-Making:** By applying a formula based on user-provided nutrient levels, the application offers **initial guidance** for farmers by highlighting potential deficiencies or excesses. This information can be used to prioritize soil testing, inform fertilizer application decisions, and ultimately contribute to better resource allocation and crop management strategies.
- **Potential for Future Improvement:** While real-time feedback loops aren't used, the user-specified data points become valuable over time. As more users interact with the application, the model can be retrained with a richer dataset, leading to **potentially even more accurate predictions** in future iterations.

III SYSTEM ARCHITECTURE

Components:

- Farmer: Interacts with the UI to provide input and receive results.
- User Interface (UI): Facilitates data entry by the farmer and displays analysis results.
- Data Preprocessing Unit: Prepares user input data for further processing.
- Database: Stores user data, historical agricultural data, and soil property information.
- Chatbot (Optional): Offers conversational guidance to farmers.
- Machine Learning Model: Analyzes data to predict yield and recommend soil amendments using RandomForestRegressor with MinMaxScaler, ColumnTransformer, and OneHotEncoder.
- Report Generation: Creates a report containing analysis results and visualizations.



IV.EXPERIMENTAL RESULTS DISCUSSION AND OBSERVATION

Understanding the Components:

Percentage Reduction: This represents the estimated decrease in crop yield potential expressed as a percentage.
Optimal Range Start/End: These values define the "sweet spot" range for a specific nutrient in your soil. They are typically obtained from agricultural research or established guidelines for your specific crop and region.
Actual Value: This is the value obtained from your soil test for the particular nutrient being analysed.
Formula for Deficiency (Actual Value < Optimal Range Start):
Percentage Reduction = (Optimal Range Start - Actual Value) / (Optimal Range Start) * 100%
Let's break it down step-by-step:

- **Optimal Range Start - Actual Value):** This subtracts the actual nutrient level from the minimum acceptable level within the optimal range. This difference represents the degree of deficiency.
- **Divide by (Optimal Range Start):** This normalizes the difference by the minimum acceptable level, putting it in perspective of the optimal range.
- **Multiply by 100%:** This converts the result into a percentage, making it easier to interpret the potential yield reduction.

Formula for Excess (Actual Value > Optimal Range End):
Percentage Reduction = (Actual Value - Optimal Range End) / (Optimal Range End) * 100%
This formula follows the same logic but focuses on the situation where the actual nutrient level exceeds the desirable maximum within the optimal range.
Example with Units:
Let's consider the previous example of potassium with an optimal range of 120 mg/kg to 150 mg/kg and a measured value of 100 mg/kg:

1. **(Optimal Range Start - Actual Value):** (120 mg/kg - 100 mg/kg) = 20 mg/kg (degree of deficiency)
2. **Divide by (Optimal Range Start):** 20 mg/kg / 120 mg/kg = 0.1667
3. **Multiply by 100%:** 0.1667 * 100% = 16.67%

Therefore, based on the formula and assuming a linear relationship between nutrient availability and yield, the potential yield reduction due to potassium deficiency is estimated to be 16.67%.

Soil nutrient range:

For values below the optimal range:

$$\text{Percentage Reduction} = \left(\frac{\text{Optimal Range Start} - \text{Actual Value}}{\text{Optimal Range Start}}\right) \times 100\%$$

For values above the optimal range:

$$\text{Percentage Reduction} = \left(\frac{\text{Actual Value} - \text{Optimal Range End}}{\text{Optimal Range End}}\right) \times 100\%$$

In these formulas:

- Optimal Range Start is the lower limit of the optimal range for the element.
- Optimal Range End is the upper limit of the optimal range for the element.
- Actual Value is the actual value of the element.

- **Element:** The chemical element representing the essential plant nutrient (e.g., Nitrogen (N), Phosphorus (P), etc.).
- **Symptoms:** A description of the visual signs exhibited by the crop when a particular nutrient is deficient. (e.g., Yellowing of lower leaves, stunted growth).
- **Possible Fertilizers:** A list of fertilizers commonly used to provide the corresponding nutrient to the soil. (e.g., Urea, Ammonium sulphate, etc.).
- **Application:** Recommendations for application rates and timing of the fertilizers for optimal nutrient delivery. (e.g., Apply 40-60 kg N/ha before planting, etc.).
- **Crop:** The specific crop species for which the nutrient deficiency information applies.

Results and Recommendations (Rice)

Element	Report Value	Status	Normal Rate, ppm	Symptoms	Possible Fertilizers	Application
Nitrogen (N)	100.0	↓	150 - 250	Yellowing of lower leaves, stunted growth	Urea, Ammonium sulfate, Calcium ammonium nitrate (CAN)	Apply 40-60 kg N/ha before planting or as a top dressing during active growth.
Phosphorus (P)	45.0	✔	40 - 80	-	-	-
Potassium (K)	35.0	↓	100 - 200	Weak stems, lodging	Potassium chloride, Potassium sulfate	Apply 40-80 kg K2O/ha as a basal dose before planting or as a top dressing during active growth.
Sulfur (S)	9.0	↓	10 - 20	Yellowing of leaves, poor growth	Ammonium sulfate, Gypsum	Apply 20-30 kg S/ha before planting or as a top dressing during active growth.
Zinc (Zn)	1.0	✔	1 - 2	-	-	-
Iron (Fe)	73.0	✔	20 - 80	-	-	-
Manganese (Mn)	5.0	✔	1 - 5	-	-	-
Copper (Cu)	1.0	✔	0.1 - 1	-	-	-
Calcium (Ca)	1100.0	↑	400 - 1000	Blossom end rot, poor fruit quality	Apply gypsum as a soil amendment before planting or as a top dressing during active growth."	Apply 20-40 kg Mg/ha as a foliar spray or soil application during active growth.
Magnesium (Mg)	190.0	✔	50 - 200	-	-	-
Sodium (Na)	4.0	✔	4 - 20	-	-	-

(Sample recommendations page of Rice)

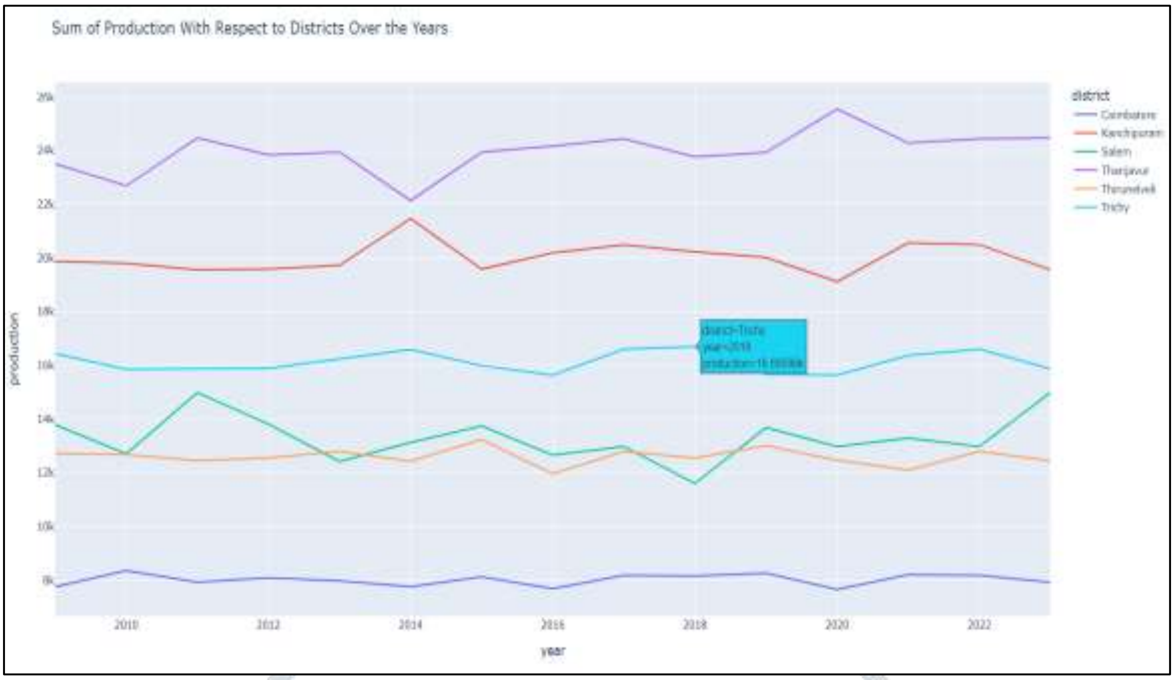
Soil nutrient range:

The below table is the range taken for the evaluation of the soil nutrient range for the optimal growth and production of the corresponding crops (Rice, Groundnut, Bengal Gram, Black Gram) taken for the yield predictions.

Sum of Production with Respect to Districts:

Nutrient (in ppm)	Rice	Groundnut	Bengal Gram	Black Gram
Nitrogen (N)	150 - 250	40 - 50	25 - 35	25 - 35
Phosphorus (P)	40 - 80	20 - 30	40 - 50	40 - 50
Potassium (K)	100 - 200	30 - 40	20 - 30	20 - 30
Sulphur (S)	10 - 20	10 - 15	10 - 15	10 - 15
Zinc (Zn)	1 - 2	1 - 2	1 - 2	1 - 2
Iron (Fe)	20 - 80	40 - 80	40 - 80	40 - 80
Manganese (Mn)	1 - 5	2 - 5	2 – 5	2 - 5
Copper (Cu)	0.1 - 1.0	0.2 - 1.0	0.2 - 1.0	0.2 - 1.0
Calcium (Ca)	400 - 1000	1000 - 2000	1000 - 2000	1000 - 2000
Magnesium (Mg)	50 - 200	200 - 400	200 - 400	200 - 400
Sodium (Na)	0 - 20	0 - 20	0 - 20	0 - 20

- **Function:** px.line (from plotly.express)
- **Purpose:** This line chart delves deeper, illustrating the total production for each district over the years. It allows for analyzing production patterns and identifying high-yielding or low-yielding districts.
- **Customization:** Similar to the previous line plot, this uses px.line with data likely grouped by year and district. The x-axis represents years, the y-axis represents production, and color is used to differentiate between districts (if there are many). A title provides context, such as "Sum of Production with Respect to Districts Over the Years".



CONCLUSION:

The Agricultural Decision Support System integrates multiple technologies to provide a robust and user-friendly platform for predicting crop yields and managing soil nutrients. By leveraging Flask for the web interface, Pandas and NumPy for data manipulation, Scikit-Learn for machine learning, and Plotly for visualization, the system offers comprehensive tools to enhance agricultural productivity and sustainability. Together, these technologies enable the development of a powerful system for real-time crop yield prediction and visualization.

REFERENCES

[1].Mr.M.Jayakandan and Dr.A.Chandrabose,"Land Weber Iterative Supervised Classification and Quantized Spiking Network for Crime Detection Emotion Analysis", International Journal of Intelligent Systems and Applications in Engineering, Volume 12, No 21S, (**2024**), Page No: 2219-2224, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[2]. Mrs.P.Ananthi and Dr.A.Chandrabose, "The Socio-technical Opportunities and Threats of Crowdsensing", The Scientific Temper, Volume: 15 (spl-1), 2024, Page No.:291-297, ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[3]. Mr.M.Jayakandan and Dr.A.Chandrabose, "Emotion Analysis Using Iterative Supervised Classification Algorithm for Crime Detection", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (**2024**), Page No: 2225-2231, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[4]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Exploring Learning-Assisted Optimization for Mobile Crowd Sensing", The Scientific Temper, Volume 15 (spl-1), 2024, Page No: 283-290. ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[5]. Mr.M.Jayakandan and Dr.A.Chandrabose, "An ensemble-based approach for sentiment analysis of covid-19 Twitter data using machine learning and deep learning techniques", The Scientific Temper Volume: 15 (spl-1), **2024**, Page No:114-120, ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[6]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Volterra Integral Equation and Logistic Drop-Offloading for Collaborative Mobile Fog Crowd Sensing", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (2024), Page No: 2186-2192, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[7].Mr.M.Jayakandan and Dr.A.Chandrabose, "Machine Learning Classifications for Automatic Sentiment Analysis on Twitter", Indian Journal of Natural Sciences Volume:15, Issue: 86, **October-2024**, Page No: 81858-81867, ISSN: 0976 – 0997 (**G-II_WEB OF SCIENCE**).

[8]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Utilizing Mathematical Modelling and Offloading to Conduct Crowdsensing in A Collaborative Setting", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (2024), Page no: 2193-2197, ISSN NO: 2147-6799 (**G-II-SCOPUS**).