Kalpa Publications in Computing
Volume 19, 2024, Pages 255–265
Proceedings of 6th International Conference on Smart Systems and Inventive Technology

Forecasting Agricultural Business Development based on Seasonal Demand using Machine Learning Techniques

Mahendhiran P D[1], Aakash V[2], Arees Muhmmed S[2], Rithanya S[2], Varun S[2]

[1]Assistant Professor, Department of Computer Science and Business Systems, Sri Krishna College of Engineering and Technology, Coimbatore-641008.
mahendhiranpd@skcet.ac.in

[2]UG Scholar, Department of Computer Science and Business Systems, Sri Krishna College of Engineering and Technology, Coimbatore-641008.
mahendhiranpd@skcet.ac.in, aakashv72003@gmail.com

Abstract

Agriculture, humanity's foundational activity, encompasses a rich tapestry of practices essential for sustaining life on our planet. From the cultivation of crops and the rearing of livestock to the management of forests and fisheries, agriculture stands as the bedrock of food production, economic development, and environmental stewardship. Across cultures and civilizations, agriculture has played a pivotal role in shaping societies, landscapes, and livelihoods, fostering connections between people and the land they cultivate. In the contemporary world, agriculture faces a myriad of challenges, from population growth and climate change to resource scarcity and ecology degradation. In light of the growing global population, surpassing 8.1 billion people, and climate change alters weather patterns and exacerbates extreme weather events, the pressure on agricultural systems to produce more food while conserving natural resources and mitigating environmental impact has never been more acute. Through the analysis of datasets sourced from the Government of India, encompassing critical factors such as pH, temperature, rainfall, humidity, and NPK substance, the study aims to provide actionable insights for agricultural stakeholders. At the heart of the project lies the recognition that informed decision-making is essential for driving efficiency, resilience, and profitability in agricultural operations. By empowering stakeholders with data-driven predictions and informed strategies, the project aims to enhance agricultural business development, enabling farmers, policymakers, and other stakeholders to make more informed choices. Whether it's optimizing crop selection, improving resource allocation, or mitigating risks associated with climate variability, the project endeavors to provide tools to the stakeholders and managing the complexities of modern life requires knowledge and experience agriculture. The predicted accuracy is 99.69%.

Keywords: Machine Learning Model, Agriculture Sector, Data Analysis, Data Training, Optimization, Random Forest Algorithm.

1 Introduction

Agriculture holds immense significance for India, serving as a cornerstone for food security, sustainable development, and poverty alleviation. However, the sector faces numerous challenges stemming from governmental policies, institutional frameworks, and economic factors. Government support primarily targets rice and wheat cultivation, often neglecting other crops and leading to an imbalanced agricultural ecosystem. Despite being the largest economic sector, agriculture struggles with inefficiencies and suboptimal productivity levels. Recognizing the need for improvement, researchers focus on enhancing agricultural planning and crop productivity to maximize yield and quality. Our project aspire to address mentioned challenges by leveraging machine learning methodologies to optimize crop yield. Among various methods explored, the Random Forest classification technique emerges as the most promising approach for enhancing crop production and yield rates. By employing Random Forest, we seek to improve crop selection methods and optimize agricultural output. The project's structure encompasses related work, methodology, results, and analysis, culminating in a comprehensive understanding of its impact on agricultural productivity and sustainability. In essence, our endeavor aligns with the broader objective of bolstering India's agricultural sector, driving economic growth, and ensuring food security for the nation's populace. Through innovative approaches and strategic interventions, we aspire to contribute towards a more resilient and prosperous agricultural landscape in India. According to the paper's structure, Section 2 summarizes related research and its connection to the present study. Section 3 describes the methodology applied in this research. Section 4 delves into results obtained and provides analysis. In the final section, the findings are summarized and the study is concluded.

2 Literature Survey

Sabena, S., Kalaiselvi, S., Anusha, B., & Ramesh, L. S. [1] explores the domain of classification with multiple labels. This study delves into the application of advanced classification techniques to scenarios where each instance may belong to multiple classes simultaneously, with these classes potentially exhibiting correlation. The context likely revolves around addressing the complexities inherent in categorizing data instances into multiple classes concurrently, particularly in fields such as social sciences and humanities where such multi-label classification problems frequently arise. By investigating the interrelation between labels and leveraging this correlation within the classification process, the paper aims to propose novel methodologies that enhance the accuracy and effectiveness of multi-label classification tasks. The insights derived from this research are anticipated to offer valuable contributions to the field, potentially improving the classification performance and aiding in the interpretation of complex datasets within the social sciences and humanities domains.

Paul, M., Vishwakarma, S. K., & Verma, A. [2] focuses on utilizing data mining techniques to analyse soil behaviour and predict crop yield and likely explores the relationship between soil characteristics and crop productivity, recognizing the significance of soil behaviour in determining agricultural outcomes. By employing data mining approaches, the researchers aim to uncover patterns and correlations within soil data that could influence crop yield. The paper likely discusses the methodology employed, which may involve collecting soil data from various sources and utilizing data mining algorithms to extract meaningful insights. These algorithms could include techniques such as decision trees, neural networks, or clustering algorithms, tailored to the specific requirements of soil analysis and crop yield prediction.

Furthermore, the research may delve into the predictive modelling aspect, where the developed models are tested and validated using historical data on soil behaviour and corresponding crop yields. The accuracy and effectiveness of the predictive models are likely evaluated using performance metrics, allowing the researchers to assess the reliability of the proposed approach. In general, the paper is expected to offer insights on utilizing data mining techniques to better understand soil behavior and enhance crop yield prediction. The results of this study may have an impact on farming methods, assisting farmers and decision-makers in making well-informed choices to maximize crop yield.

Everingham, Yvette, Justin Sexton, Danielle Skocaj, and Geoff Inman-Bamber. [3] aims to achieve precise predicting the yield of sugarcane by employing the random forest algorithm. The research likely addresses the importance of accurate yield prediction in sugarcane production, a critical aspect for optimizing agricultural practices and enhancing sustainability. By utilizing a random forest algorithm, the authors aim to overcome the challenges associated with traditional methods of yield prediction, which may lack the ability to capture complex interactions between various factors influencing yield. The paper likely discusses the methodology employed, which involves training the random forest algorithm using historical data on sugarcane yield along with relevant environmental and management factors. The algorithm's ability to handle large datasets and nonlinear relationships makes it well-suited for capturing the complexities of sugarcane yield prediction. Furthermore, the research may present findings demonstrating the effectiveness of the random forest algorithm in accurately predicting sugarcane yield. These findings may include comparisons with other predictive models or traditional methods to highlight the advantages offered by the random forest approach. Overall, the paper likely contributes to advancing the field of sugarcane yield prediction by introducing a novel methodology based on the random forest algorithm. The results of this study may have an impact on agricultural professionals and sugarcane growers, empowering them to maximize yield outcomes and make better decisions while advancing sustainable sugarcane production methods.

Sharma, S. K., Sharma, D. P., & Gaur, K. [4] examines, with an emphasis on the Jaipur district, the effectiveness of Random Forest deep learning regression in predicting the yield of important crops within Rajasthan 3A agroclimatic zone. Utilizing historical data, such as weather conditions and previous yield records, is part of yield prediction. Within the semi-arid eastern plain, the study focuses on six main crops: barley, wheat, mustard, gram, groundnut, and moong. Rainfall, sunshine hours, temperature (minimum and maximum), and relative humidity were all obtained from the Agrometeorology Observatory of Sri Karan Narendra College of Agriculture, Jobner, Jaipur, between 1991 and 2020. The official bulletins of the Rajasthani government's Directorate of Economics and Statistics served as the source of crop yield data. The Random Forest regression, a supervised learning model, demonstrated robust performance with an accuracy of 92.3%. This methodology facilitates precise yield forecasting, aiding farmers and policymakers in strategic crop production and management planning for the region. Moreover, the paper offers scientific recommendations based on the study findings, aimed at benefiting farmers, policymakers, and other stakeholders involved in agricultural decision-making processes.

Patil, P., Athavale, P., Bothara, M., Tambolkar, S., & More, A. [5] focuses on using machine learning methods to choose appropriate crops and forecast crop productivity. It specifically targets districts in Maharashtra, aiming to enhance agricultural decision-making processes. The research utilizes weather and soil parameters to predict crop yield accurately. Regression techniques are applied for yield estimation, while multi-class classification is utilized for crop type prediction. The model that performs best for yield prediction is the Random Forest Regression, with an R-squared score of 0.96 and an amazing Mean Absolute Error (MAE) of 0.64. Meanwhile, the Naive Bayes classifier demonstrates exceptional accuracy in predicting crop types, boasting an accuracy rate of 99.39%.The suggested method provides farmers with insightful information that helps them choose crops and project yields in their particular farming setting. Improvement ideas are given to increase the predictions' accuracy and applicability. These include adding real-time data from Internet of Things sensors and growing the dataset to include variables like fertilizer and irrigation consumption. Additionally, it is suggested that a mobile application be created to give farmers easy access to the forecasts as well as extra services like price estimations based on market trends. All things considered, the study emphasizes how machine learning may transform crop management methods and stresses how crucial precise yield prediction is to improving farming practices and raising output in Maharashtra.

Hasan, R., Hossain, R., Ahmed, F., Islam, M. I., & Iftekhar, M. [6] addresses the challenge of aligning consumer demands with market requirements for food crops by leveraging weather forecasts with machine learning methods. The study focuses on developing a predictive model that utilizes machine learning algorithms and weather forecasts to optimize food crop production and meet market demands effectively. By integrating historical data on consumer preferences, market trends, and weather patterns, the model aims to provide valuable insights into crop selection, cultivation practices, and market strategies. Regression analysis is employed to predict consumer demands based on various factors such as population demographics, dietary trends, and socioeconomic indicators. Concurrently, weather forecasting data is incorporated to anticipate environmental conditions that may impact crop yields and market supply.The suggested strategy provides a comprehensive way to close the gap between what the market needs for food crops and what consumers want. Through the application of sophisticated analytical methods and up-to-date meteorological forecasts, agriculturalists and decision-makers can guarantee sufficient food production and distribution. The dissertation contributes to the field of agricultural economics and food security by providing a data-driven approach to address the complexities of matching supply with demand in the food market. It also emphasizes how predictive analytics and machine learning may improve agricultural supervisory and maximize resource allocation in the food production industry.

Saeed, U., Dempewolf, J., Becker-Reshef, I., Khan, A., Ahmad, A., & Wajid, S. A. [7] aims to forecast wheat yield in Pakistan's Punjab province by using meteorological data and MODIS NDVI (Normalized Difference Vegetation Index) through the Random Forests algorithm. The research aims to develop an accurate predictive model for wheat yield forecasting, utilizing readily available data sources such as weather parameters and satellite-derived vegetation indices. The study aims to offer significant perspectives for agricultural planning and decision-making in the area by incorporating various data sources. Because it is so good at managing large, complicated datasets and capturing nonlinear correlations between predictor variables and crop yield, the Random Forests method is used. The model makes use of MODIS NDVI data, which acts as a stand-in for vegetation growth and health, together with weather factors including temperature, precipitation, and humidity. The predictive model is trained and validated using historical wheat yield data obtained from agricultural surveys and satellite imagery. The model's prediction accuracy is assessed using performance metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The study's findings show how well the Random Forests algorithm predicts wheat yield using weather information and MODIS NDVI. The model achieves satisfactory accuracy levels, providing valuable information for crop monitoring, yield estimation, and agricultural management practices in Punjab province. The study showcases the possibility of combining satellite imagery with machine learning algorithms for crop production predictions and decision support in areas like Punjab, Pakistan, and advances remote sensing applications in agriculture overall.

Padma, T., & Sinha, D. [8] focuses on enhancing crop yield prediction through the application of an improved Random Forest algorithm. Creating a more precise and effective predictive model for crop production estimation is the main goal of the study. To achieve this, the study proposes enhancements to the traditional Random Forest algorithm to better capture the complexities of crop-yield relationships. The improved Random Forest algorithm incorporates novel features such as feature engineering, ensemble learning techniques, and hyperparameter optimization. These enhancements aim to improve the model's predictive performance by reducing bias and variance, thus yielding more reliable yield predictions. The research methodology involves collecting and pre-processing agricultural data, including historical yield records, weather parameters, soil characteristics, and agronomic practices. Subsequently, significant characteristics are extracted from the raw data using feature engineering techniques, and these features serve as input variables for the predictive model. Mean Absolute Error and Root Mean Square Error are two common assessment measures used to assess the effectiveness of the upgraded Random Forest model. To evaluate the effectiveness of the suggested improvements, comparisons between conventional Random Forest and other machine learning methods are carried out. The study's findings show that the enhanced Random Forest algorithm is capable of precisely forecasting crop output. The model outperforms traditional Random Forest and other baseline algorithms, showcasing its potential as a valuable tool for crop yield estimation. Overall, the research contributes to advancing the field of crop yield prediction by introducing enhancements to the Random Forest algorithm. The proposed improvements offer promising prospects for enhancing agricultural decision-making processes and optimizing crop management practices.

Sapkal, K. G., & Kadam, A. B. [9] focuses on using soil data to forecast crop types using a Random Forest classifier. The main goal of the project is to use soil-related factors to create a predictive model for crop prediction that is both accurate and efficient. The study emphasizes the importance of soil characteristics in determining crop suitability and productivity. The research methodology involves collecting soil data from various agricultural fields, including several attributes such as soil pH, moisture content, nutrient levels, and texture. These soil attributes are then used as input variables for training the Random Forest classifier. Because of its capacity to handle high-dimensional datasets and capture nonlinear interactions between crop kinds and input factors, the Random Forest method was selected. The dataset used to train the model includes labeled samples of several crop varieties along with the corresponding properties of the soil. Recall, accuracy, precision, and F1-score are among the common classification measures used to assess the Random Forest classifier's performance. To guarantee the model's resilience and generalizability, cross-validation procedures are utilized. The study's findings show how well the Random Forest classifier predicts different crop kinds using information about the soil. The model's excellent classification accuracy shows that it has the potential to be a useful tool for agricultural settings' crop prediction. Overall, by utilizing machine learning methods and soil data, the research advances the field of crop prediction. With the help of the suggested Random Forest classifier, farmers and other stakeholders will be able to make well-informed decisions about the management and choice of crops depending on the properties of the soil.

Rao, M. S., Singh, A., Reddy, N. S., & Acharya, D. U. [10] aims into the use of machine learning techniques to forecast crop yields with the goal of creating a trustworthy predictive model necessary for agricultural planning and decision-making. The research methodology involves gathering and pre-processing data related to crop yields, weather conditions, soil attributes, and agricultural practices. The relevant characteristics are then extracted from the data using feature engineering techniques, which act as input variables for the predictive model. A range of machine learning techniques, such as random forests, decision trees, and support vector machines, are assessed using common evaluation metrics like F1-score, accuracy, precision, and recall. The results show how successful machine learning is at predicting crop yield, with the random forest method outperforming other algorithms on a variety of datasets. Overall, by demonstrating how machine learning approaches can improve crop prediction accuracy, this study advances agricultural science and provides farmers and policymakers with useful information for maximizing crop production strategies.

3 Methodology

Interpreting the data and results from the proposed solution involves a systematic analysis of the performance of ML methodology, including Random Forest, Using the dataset provided, crop yield is forecasted. The initial step involves pre-processing the data, ensuring that missing values are appropriately handled and features are scaled to avoid bias during model training. This evaluation facilitates comparison of the models, enabling the identification of the best-performing model in terms of predictive accuracy and error minimization. Additionally, feature importance examination is done to comprehend the significance of input features namely pH, temperature, rainfall, humidity, and NPK content in predicting crop yield. To determine the purpose and key objectives this process includes collection of data, Cleaning the data, Evaluation and Visualizing the data. User input attributes for crop prediction model is shown in Fig 1.

S.no	Attribute	Description
1.	Nitrogen	Soil should ideally contain 10-50 mg/kg potassium for various crops.
2.	Potassium	Soil potassium levels vary from 0.3% to 2.5%, affecting plant growth differently.
3.	Phosphorus	Soil phosphorus mainly as orthophosphate, usually ranging 500-800 mg/kg.
4.	Temperature	Optimal soil temperature for germination: 68-86°F (20-30°C); spring wheat: 37°F (3°C).
5.	Humidity	Ideal RH for plant growth: 60-70% and low RH increases evapotranspiration, affecting growth.
6.	pH	Ideal pH for most plants: 5.5-7.5 and some thrives outside this range.
7.	Rainfall	Rainfall varies widely in India: abundant in tropics, lower in arid regions.
8.	Label	Crop prediction model generates crop names as output. eg: Wheat, Rice, Coffee, Mango etc.

Figure 1: user input attributes for crop prediction model

Visualizations, including scatter plots and bar charts, are employed to illustrate the relationships between input features and predicted crop yield, aiding in the extraction of actionable insights for agricultural business development. Based on the evaluation results and feature importance analysis, the most suitable model is selected, and its robustness is validated using out-of-sample data or cross-validation techniques. Ultimately, the interpretation of the data culminates in a comprehensive summary of findings, including key insights and recommendations for leveraging the results to forecast agricultural business development based on seasonal demand. Through this approach, stakeholders can make informed decisions to optimize agricultural practices and enhance crop productivity. A sample of expected labels based on several inputs is shown in Fig 2.

A	B	C	D	E	F	G	H
N	P	K	Temperature	Humidity	pH	Rainfall	Label
90	42	43	20.879744	82.002744	6.502985	203.9832	Rice
85	58	41	22.00094	80.319644	7.038096	226.655537	Rice
76	51	18	26.169459	71.920763	6.247207	79.8492542	Maize
118	33	30	24.131096	67.158363	6.380401	173.322839	Coffee
78	42	42	20.130175	81.604873	7.628473	262.71734	Rice
12	27	26	29.093827	45.566159	5.323071	96.2352043	Mango
40	143	201	24.972561	82.728286	6.476757	66.7001628	Grapes
26	121	201	22.19109412	90.025751	6.162034	112.3126628	Apple
0	18	14	29.77149434	92.007199	7.207991	114.4161786	Orange
61	64	52	43.30204933	92.834054	6.641098	110.5622291	Papaya
101	58	18	25.66891439	81.381033	6.652143	78.5959581	Cotton
62	49	37	24.21744605	82.85284	7.479248	166.1365886	Jute

Figure 2: sample of expected label based on several inputs

4 Proposed System

The supervised machine learning method known as the Random Forest algorithm is used various problems including regression and classification alike. It works by creating many decision trees using various data subsets, then aggregates predictions through voting to yield an enhanced solution. Employing the bagging method, Random Forest trains models on varied datasets, amalgamating their outcomes to improve overall performance. By constructing decision trees from data samples and amalgamating predictions through voting, Random Forest enhances accuracy. A conceptual framework for crop forecasting is shown in Fig 3.

This method selects the most optimal solution based on the aggregated predictions, ensuring robust results for the system. Its efficacy lies in its ability to mitigate overfitting and handle complex datasets, making it a preferred choice for various applications. Through this approach, Random Forest leverages the collective wisdom of multiple models to deliver superior predictive accuracy compared to individual decision trees or other conventional methods. To achieve the cross-validation process

- Split the data
- Iterate through folds
- Train the model
- Validate the model
- Repeat
- Aggregate results

The dataset used in the paper is sourced from the Government of India and forecasting comprises critical agricultural parameters, including pH levels, temperature, rainfall, humidity, and NPK content. However, specific details regarding the dataset's size, time period covered, data collection methods, and preprocessing steps are not provided. Transparency regarding the dataset's quality, representativeness, and potential biases is essential for strengthening the credibility and reproducibility of the research outcomes. Including such information would provide readers with a comprehensive understanding of the dataset's relevance to the research objectives.

5 Result and Analysis

Our project aims to develop a user-friendly web application to assist farmers and users in predicting crop outcomes based on seasonal weather patterns. The application will provide valuable insights and guidance, helping users make knowledgeable choices about crop planning and administration. Validated results are shown in Fig 5 to Fig 8.

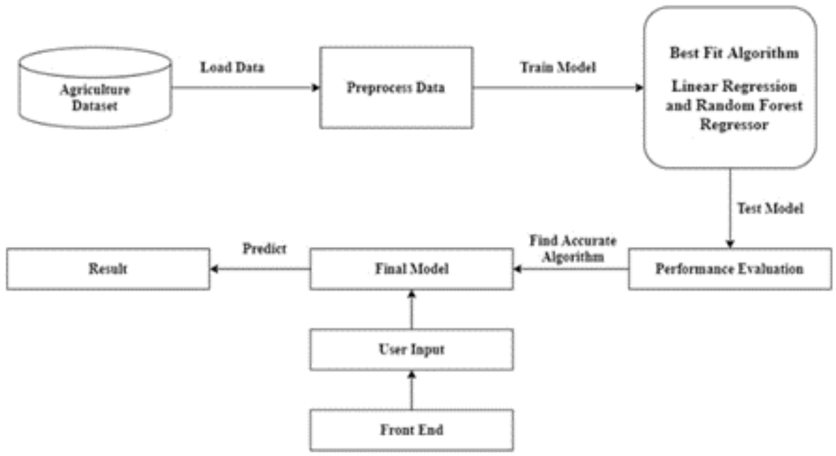

Figure 3: Conceptual Framework for crop

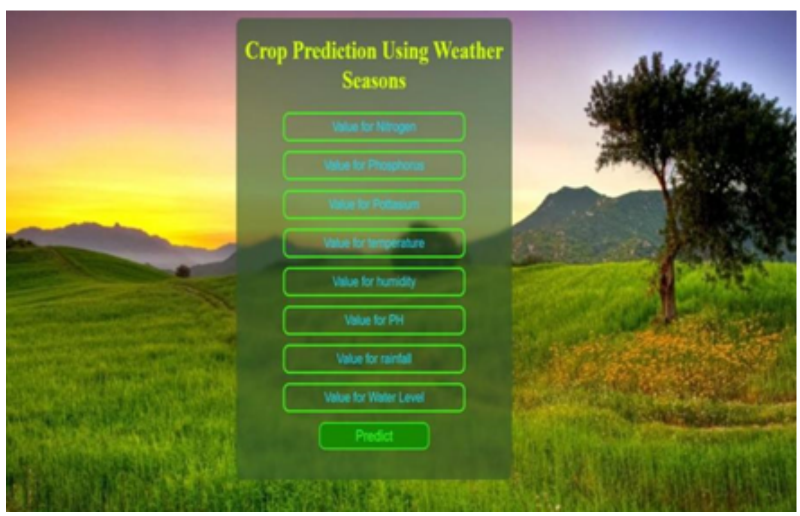

Figure 4: user interface homepage

Figure 5 : passing input parameters

Figure 6: Predicted result after entering parameters

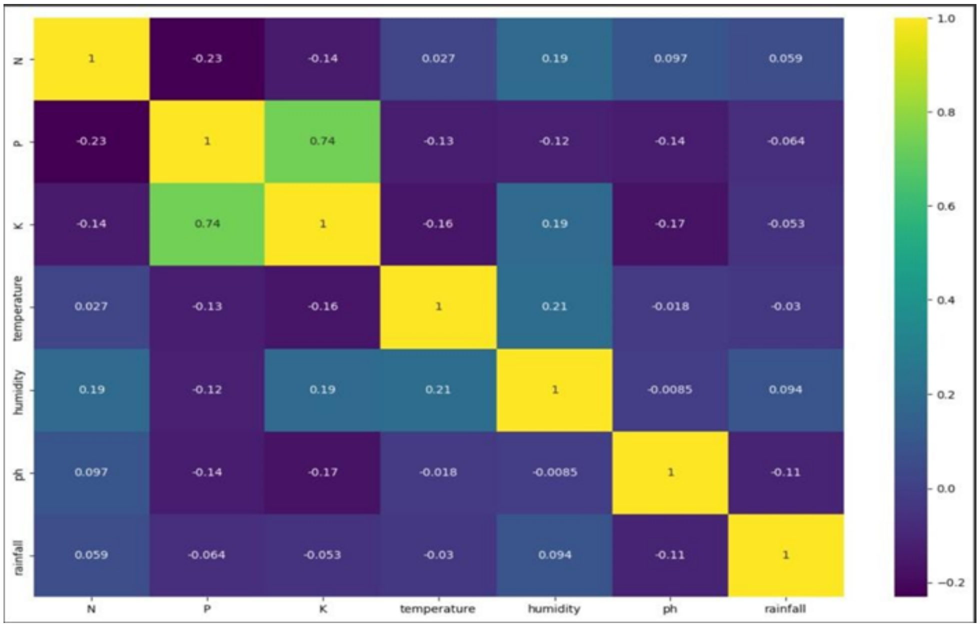

Figure 7: graphical representation of type of crops

We chose Python language for our system, with the Flask framework to display results on a web page. Users, including both farmers and general users, can access the home page to input details like NPK levels, humidity, temperature, rainfall, pH and water level. Upon completing the form, users can press the predict button, which sends a request to the server. The system then utilizes the Random Forest algorithm to provide a prediction based on the input data. To improve the performance bagging method is used which is producing more decision trees for the data subset which in turn provide single output with high accuracy. Tuning hyperparameters, feature engineering is used here to this increase performance by scaling or encoding categorial variable and ensembling methods. This user-friendly web application streamlines the prediction process, aiding users in deciding crop planning and its supervision with knowledge. The result page displays the predicted crop yield according to the input data given by the user. And shows the predicted value for the selected factors such as N, P, K, pH, temperature, rainfall, humidity. With this information, users may decide on their farming methods with knowledge.

```
accuracy_score(y_test,predict_r) * 100

99.69696969696969
```

Figure 8: Performance evaluation of the system using metric accuracy

In this model we trained 500 decision trees to build a random forest. The accuracy of the model achieved is 99.69 %.For optimization of the agriculture and business growth the practices includes Utilize machine learning techniques here we used a Random forest algorithm which helps in Data driven decision making to identification of patterns. Inclusion of real-time data integration i.e. data sources to provide real time information and advanced techniques such as deep learning techniques with continuous improvement will result in refining the agricultural processes.

6 Conclusion

This research paper represents a significant step forward in enhancing agricultural decision-making processes. By leveraging advanced machine learning methodologies, particularly the Random Forest algorithm, the project has shown that it is capable of producing highly accurate crop output forecasts. Through complete data analysis sourced from the Government of India, encompassing critical parameters such as pH, temperature, rainfall, humidity, and NPK content, valuable insights have been gleaned to inform agricultural business development strategies. The findings of the project underscore the importance of data-driven approaches in optimizing resource allocation, crop selection, and risk management in the agricultural sector. By empowering stakeholders with actionable insights derived from machine learning models, the project aims to drive efficiency, resilience, and profitability in agricultural operations. Moving forward, the insights generated by this project have the potential to inform policy decisions, drive innovation, and foster sustainable growth in the agricultural industry. Soil and water management, crop selection and rotation, integrated pest management, proper fertilization, climate-smart agriculture, new technology adoption practices help in enhancing productivity. By continuing to take full advantage of machine learning algorithms and agricultural information, stakeholders able to navigate the complexities of the agricultural landscape with confidence, ultimately contributing to a more prosperous and sustainable future for the sector.

References

Sabena, S., Kalaiselvi, S., Anusha, B., & Ramesh, L. S. (2016). An Multi-Label Classification with Label Correlation. Asian Journal of Research in Social Sciences and Humanities, 6(9), 373-386.

Paul, M., Vishwakarma, S. K., & Verma, A. (2015, December). Analysis of soil behaviour and prediction of crop yield using data mining approach. In 2015 International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 766-771). IEEE.

Everingham, Yvette, Justin Sexton, Danielle Skocaj, and Geoff Inman-Bamber. "Accurate prediction of sugarcane yield using a random forest algorithm." Agronomy for sustainable development 36 (2016): 1-9.

Sharma, S. K., Sharma, D. P., & Gaur, K. (2023). Crop yield predictions and recommendations using random forest regression in 3a agroclimatic zone, Rajasthan. Journal of Data Acquisition and Processing, 38(2), 1635.

Patil, P., Athavale, P., Bothara, M., Tambolkar, S., & More, A. (2023). Crop Selection and Yield Prediction using Machine Learning Approach. Current Agriculture Research Journal, 11(3).

Hasan, R., Hossain, R., Ahmed, F., Islam, M. I., & Iftekhar, M. (2021). Use of machine learning and weather predictions to bridge the gap between consumer demands and the market requirements of food crops (Doctoral dissertation, Brac University).

Saeed, U., Dempewolf, J., Becker-Reshef, I., Khan, A., Ahmad, A., & Wajid, S. A. (2017). Forecasting wheat yield from weather data and MODIS NDVI using Random Forests for Punjab province, Pakistan. International journal of remote sensing, 38(17), 4831-4854.

Padma, T., & Sinha, D. (2023). Crop Yield Prediction Using Improved Random Forest. In ITM Web of Conferences (Vol. 56, p. 02007). EDP Sciences.

Sapkal, K. G., & Kadam, A. B. (2024). Random Forest Classifier For Crop Prediction Based On Soil Data. Journal of Advanced Zoology, 45.

Rao, M. S., Singh, A., Reddy, N. S., & Acharya, D. U. (2022). Crop prediction using machine learning. In Journal of Physics: Conference Series (Vol. 2161, No. 1, p. 012033). IOP Publishing.