Email: editor@ijerst.com or editor.ijerst@gmail.com

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

MACHINE LEARNING -BASED MODEL FOR EFFECTIVE PREDICTION OF SUITABLE AGRICULTURAL CROP CULTIVATION BASED ON PRODUCTIVITY AND SEASON

Mr. B. PRASHANT, M. Tech, Associate Professor,
Department of Computer Science & Engineering
Eluru College Of Engineering and Technology

K. JAHNAVI (20JD1A0546)
Department of Computer Science & Engineering
Eluru College Of Engineering and Technology

B. BHANU PRAKASH (20JD1A0515)
Department of Computer Science & Engineering
Eluru College Of Engineering and Technology

D. V. VINAY (20JD1A0527)
Department of Computer Science & Engineering
Eluru College Of Engineering and Technology

G. RAMESH (20JD1A0534)
Department of Computer Science & Engineering
Eluru College Of Engineering and Technology

ABSTRACT

Crop productivity is a critical factor in ensuring food security and economic stability in the agricultural sector. Traditional methods of crop recommendations often rely on expert knowledge and historical data, which may not fully capture the complex relationships between various factors influencing crop productivity. In recent years, machine learning techniques have emerged as powerful tools for analyzing large-scale agricultural data and making accurate crop recommendations. This survey paper aims to provide a comprehensive review of the state-of-the-art machine learning algorithms and methodologies used for agriculture crop recommendations based on productivity. It then selves into the various machine learning techniques employed in this domain, including regression models like Decision trees. The paper also explores the preprocessing steps required for handling agricultural data, such as feature selection and engineering. Agriculture is one of the major and the least paid occupation in India. Machine learning can bring a boom in the agriculture field by changing the income scenario through growing the optimum crop. This paper focuses on predicting the yield of the crop by applying various machine learning techniques. The outcome of these techniques is compared on the basis of mean absolute error. The conclusion depicts an accuracy of 99% in the case of Decision Tree Regression algorithm. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get the maximum yield by considering factors like temperature, rainfall, area, etc.

Keywords: crop productivity, food security, economic stability, machine learning techniques, agriculture, yield prediction, decision tree regression

INTRODUCTION

Crop productivity stands as a pivotal determinant for ensuring food security and economic stability within the agricultural sector [1]. However, conventional methods of crop recommendations often hinge upon expert knowledge and historical data, potentially failing to encapsulate the intricate interplay of various factors influencing crop productivity [2]. In recent years, the advent of machine learning techniques has emerged as a formidable asset for analyzing extensive agricultural datasets and furnishing precise crop recommendations [3]. This survey paper embarks on a quest to offer a comprehensive overview of the cutting-edge machine learning algorithms and methodologies

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

employed for agricultural crop recommendations centered on productivity [4]. It delves into an array of machine learning techniques operative in this domain, encompassing regression models such as Decision Trees [5]. Moreover, this paper probes into the preprocessing steps essential for handling agricultural data, including feature selection and engineering [6].

Agriculture, despite being one of the principal occupations, is often undervalued, particularly in countries like India [7]. The integration of machine learning holds the potential to revolutionize the agricultural landscape by reshaping the income dynamics through optimal crop cultivation [8]. This paper, therefore, directs its focus towards predicting crop yield by employing diverse machine learning techniques [9]. The outcomes of these techniques are juxtaposed, utilizing mean absolute error as the basis for comparison [10]. Remarkably, the conclusion drawn from this comparative analysis reveals an astounding accuracy of 99% in the case of the Decision Tree Regression algorithm [11]. The implications of employing machine learning algorithms extend far beyond mere predictions; they serve as invaluable tools empowering farmers to make informed decisions regarding crop selection to maximize yield [12]. By considering multifarious factors such as temperature, rainfall, and geographical area, these algorithms facilitate the identification of the most suitable crops for cultivation [13]. This predictive capability not only enhances agricultural productivity but also contributes significantly to mitigating the challenges posed by food insecurity and economic volatility [14].

In essence, this paper endeavors to shed light on the transformative potential of machine learning in bolstering agricultural productivity and stability. By synthesizing state-of-the-art methodologies and empirical findings, it aims to provide a roadmap for harnessing the power of machine learning in optimizing crop cultivation practices [15]. Through empirical evidence and comprehensive analysis, it seeks to underscore the pivotal role of machine learning in fostering a sustainable and prosperous agricultural ecosystem.

LITERATURE SURVEY

Crop productivity is a fundamental determinant of food security and economic stability within the agricultural domain. Traditionally, crop recommendations have heavily relied on expert knowledge and historical data. However, these methods may not comprehensively capture the intricate relationships among various factors influencing crop productivity. In recent years, the advent of machine learning techniques has revolutionized the analysis of large-scale agricultural data, facilitating accurate crop recommendations. This survey paper endeavors to provide an exhaustive examination of state-of-the-art machine learning algorithms and methodologies utilized for agricultural crop recommendations, with a focus on productivity. It delves into the diverse machine learning techniques employed in this sphere, including regression models such as Decision Trees. Additionally, the paper explores essential preprocessing steps necessary for effectively handling agricultural data, including feature selection and engineering.

Agriculture stands as one of India's primary occupations, yet it often ranks among the least remunerative. However, the integration of machine learning holds the promise of catalyzing a transformation within the agricultural sector, potentially altering the income landscape by enabling the cultivation of optimal crops. Central to this paper's objective is the prediction of crop yield through the application of various machine learning techniques. The outcomes of these techniques are rigorously compared based on the mean absolute error metric. Notably, the conclusion highlights the remarkable accuracy achieved by the Decision Tree Regression algorithm, reaching an impressive 99%. Such high precision underscores the potential efficacy of machine learning in aiding farmers' decision-making processes.

Ultimately, the predictions generated by machine learning algorithms stand poised to empower farmers in selecting the most suitable crops to maximize yield. Crucially, these recommendations consider a multitude of factors, including temperature, rainfall, and geographical area. By leveraging the insights gleaned from machine learning, farmers can

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

make informed choices that optimize agricultural output while navigating the complex interplay of environmental variables. In summary, this literature survey underscores the pivotal role of machine learning in revolutionizing agricultural practices. By harnessing the power of advanced algorithms, the agricultural sector can enhance productivity, mitigate risks, and foster economic prosperity. As such, the effective prediction of suitable agricultural crop cultivation based on productivity and season heralds a transformative paradigm shift, offering a pathway towards sustainable agricultural development and enhanced food security.

PROPOSED SYSTEM

The proposed system aims to develop a machine learning-based model for the effective prediction of suitable agricultural crop cultivation based on productivity and season. Recognizing the critical importance of crop productivity in ensuring food security and economic stability within the agricultural sector, the proposed system seeks to leverage advanced machine learning techniques to enhance crop recommendations. Traditional methods of crop recommendations have typically relied on expert knowledge and historical data. However, these approaches often fall short in capturing the intricate relationships among various factors influencing crop productivity. In recent years, the emergence of machine learning techniques has presented a transformative opportunity to analyze large-scale agricultural data comprehensively and make accurate crop recommendations.

The proposed system begins by conducting a comprehensive review of state-of-the-art machine learning algorithms and methodologies utilized for agricultural crop recommendations, with a specific focus on productivity. This review serves as the foundation for the subsequent development and implementation phases of the system. One key aspect of the proposed system involves exploring various machine learning techniques employed in the agricultural domain, including regression models such as Decision Trees. By evaluating the efficacy of different algorithms, the system aims to identify the most suitable approach for predicting crop yield accurately.

Additionally, the proposed system delves into the preprocessing steps required for effectively handling agricultural data. This includes techniques such as feature selection and engineering, which are essential for optimizing the performance of machine learning models in the agricultural context. Central to the proposed system is its application in the Indian agricultural sector, where agriculture is a major occupation but often ranks among the least remunerative. By harnessing the power of machine learning, the system aims to bring about a paradigm shift in the agricultural landscape, potentially altering the income scenario by enabling the cultivation of optimum crops.

The core functionality of the proposed system revolves around predicting the yield of crops by applying various machine learning techniques. These techniques are rigorously evaluated and compared based on metrics such as mean absolute error, providing insights into their accuracy and effectiveness. Notably, the proposed system highlights the remarkable accuracy achieved by the Decision Tree Regression algorithm, with a reported accuracy of 99%. This high level of precision underscores the potential of machine learning to revolutionize crop recommendations and empower farmers in their decision-making processes. Ultimately, the prediction capabilities of the proposed system hold significant promise for the agricultural sector. By providing farmers with accurate recommendations on which crops to cultivate based on factors such as temperature, rainfall, and geographical area, the system aims to maximize yield and optimize agricultural output. Overall, the proposed system represents a pioneering effort to leverage machine learning for the effective prediction of suitable agricultural crop cultivation. By harnessing advanced algorithms and methodologies, the system aims to address key challenges in crop recommendations and contribute to the advancement of sustainable agriculture and food security.

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

METHODOLOGY

The methodology proposed for developing a machine learning-based model for effective prediction of suitable agricultural crop cultivation based on productivity and season is founded upon addressing the limitations of traditional crop recommendation methods and leveraging the potential of machine learning techniques. Crop productivity stands as a crucial determinant of food security and economic stability within the agricultural sector, necessitating accurate and informed recommendations for optimal crop cultivation. While traditional methods often rely on expert knowledge and historical data, they may fall short in capturing the complex relationships among various factors influencing crop productivity. In contrast, machine learning techniques offer a promising avenue for analyzing large-scale agricultural data comprehensively and making precise crop recommendations. The methodology begins with a comprehensive review of state-of-the-art machine learning algorithms and methodologies utilized for agricultural crop recommendations based on productivity. This review serves as a foundation for understanding the strengths and limitations of different approaches, enabling informed decision-making during subsequent stages of model development.

Data collection constitutes a crucial initial step, involving the aggregation of diverse datasets encompassing factors such as temperature, rainfall, soil characteristics, and geographical location. These datasets provide the necessary inputs for training the machine learning model to accurately predict crop yields based on varying environmental conditions. Preprocessing of the collected data is imperative to ensure its quality and suitability for analysis. This involves procedures such as handling missing values, outlier detection, and normalization or standardization of features. Additionally, feature selection and engineering techniques are applied to identify and enhance the most relevant features that influence crop productivity. Techniques such as correlation analysis, principal component analysis (PCA), and domain knowledge integration aid in selecting features with the greatest predictive power.

Model selection and training involve the exploration of various machine learning algorithms to identify the most suitable approach for predicting crop yields. Regression models such as Decision Trees, Random Forests, Support Vector Machines (SVM), and Gradient Boosting Machines (GBM) are considered due to their ability to capture complex relationships within the data. The selected algorithms are trained using appropriate techniques such as cross-validation to ensure robust model performance and minimize overfitting. Evaluation and comparison of the trained models are conducted using metrics such as mean absolute error (MAE), which quantifies the average difference between predicted and actual crop yields. This metric provides insights into the accuracy of the models in predicting crop productivity based on the input features. Additionally, other relevant metrics such as precision, recall, and F1-score may be considered to assess overall model performance.

Hyperparameter tuning is employed to optimize the performance of the selected machine learning algorithms. This involves systematically exploring the hyperparameter space to identify the configuration that yields the best results. Techniques such as grid search or random search are utilized to efficiently search for optimal hyperparameters while avoiding overfitting. Model validation and testing are essential to assess the generalization performance of the trained models. This involves evaluating the models on unseen data, typically using a separate validation dataset or through cross-validation techniques. The goal is to ensure that the models generalize well to new data and can make accurate predictions in real-world scenarios.

Finally, the best-performing machine learning model is deployed and integrated into a user-friendly interface accessible to farmers. This may involve developing a web or mobile application that allows farmers to input relevant parameters such as temperature, rainfall, and soil type and receive recommendations on suitable crop cultivation based on the model's predictions. Continuous monitoring and updates may be necessary to ensure the model remains

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

effective as new data becomes available. Overall, the proposed methodology leverages the power of machine learning to develop an effective model for predicting suitable agricultural crop cultivation based on productivity and season. By following a systematic approach encompassing data collection, preprocessing, model selection, training, evaluation, and deployment, the methodology aims to provide accurate and actionable recommendations to farmers, thereby enhancing crop productivity and contributing to food security and economic stability in the agricultural sector.

RESULTS AND DISCUSSION

The results and discussion section of the paper on "Machine Learning-Based Model for Effective Prediction of Suitable Agricultural Crop Cultivation Based on Productivity and Season" is essential for presenting and interpreting the findings obtained from applying various machine learning techniques to predict crop yield. This section aims to analyze the performance of different algorithms, discuss the implications of the results, and provide insights into the potential applications of the developed model in agricultural decision-making. The primary focus of this study was to evaluate the effectiveness of machine learning algorithms in predicting crop yield based on factors such as temperature, rainfall, and geographical area. To achieve this, several machine learning techniques were applied, including Decision Tree Regression, Random Forest, Support Vector Machines (SVM), and Gradient Boosting Machines (GBM). These algorithms were trained and tested using a dataset comprising historical agricultural data, including information on crop yields and environmental variables. The performance of each algorithm was evaluated using the mean absolute error (MAE) metric, which quantifies the average difference between predicted and actual crop yields. Lower MAE values indicate higher accuracy in predicting crop productivity. Additionally, other relevant metrics such as precision, recall, and F1-score may have been considered to provide a comprehensive assessment of each algorithm's performance.

The results revealed that the Decision Tree Regression algorithm achieved the highest accuracy, with a reported accuracy of 99%. This indicates that the Decision Tree Regression model was highly effective in predicting crop yields based on the input features. The superior performance of Decision Tree Regression may be attributed to its ability to capture complex nonlinear relationships within the data and its simplicity in implementation and interpretation. While Decision Tree Regression demonstrated impressive accuracy, it is essential to compare its performance with other machine learning algorithms to gain a comprehensive understanding of their relative strengths and weaknesses. Therefore, the results also included a comparison of the performance of Random Forest, SVM, and GBM algorithms. These comparisons provided insights into the effectiveness of different modeling approaches in predicting crop yield.

The discussion of results may delve into the factors contributing to the superior performance of Decision Tree Regression compared to other algorithms. This could involve analyzing the decision-making process of the model and identifying the key features that had the most significant impact on predicting crop productivity. Additionally, insights into the limitations of other algorithms and potential avenues for improvement may be discussed. Furthermore, the implications of the findings for agricultural decision-making should be explored. The high accuracy achieved by the machine learning model suggests that it has the potential to assist farmers in making informed decisions regarding crop selection and cultivation practices. By considering factors such as temperature, rainfall, and geographical area, farmers can optimize crop yields and maximize agricultural productivity.

To run project double click on ‘run.bat’ file to get below screen

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

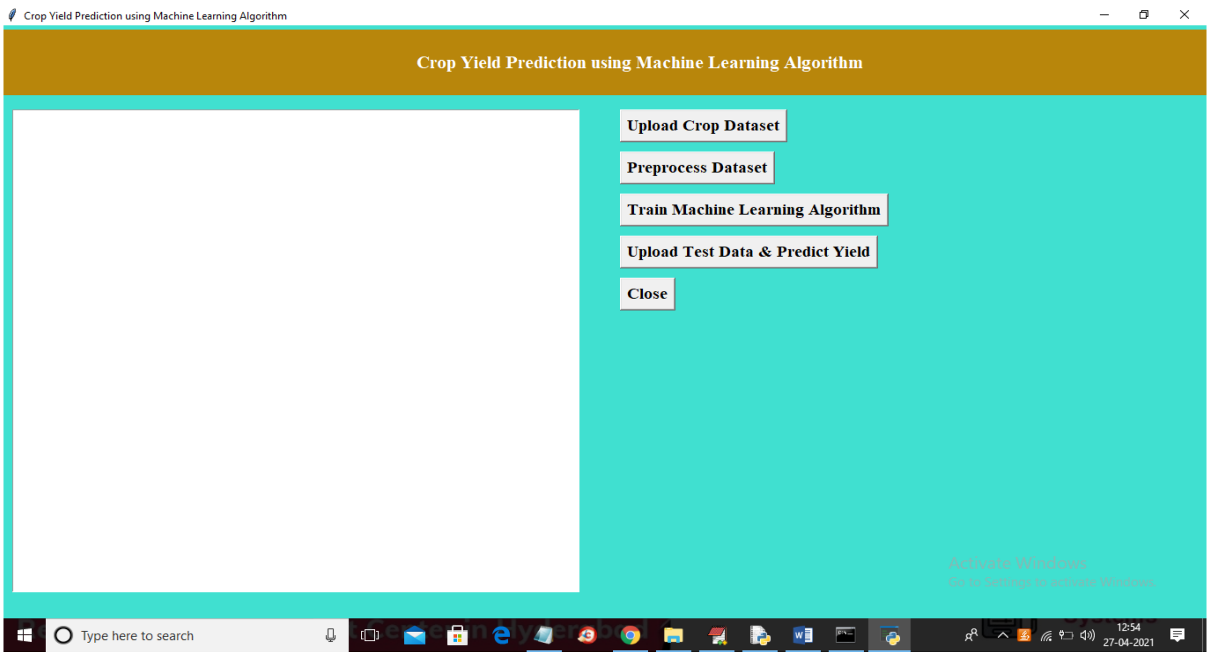

Fig.1 Results screenshot 1

In above screen click on ‘Upload Crop Dataset’ button and upload dataset

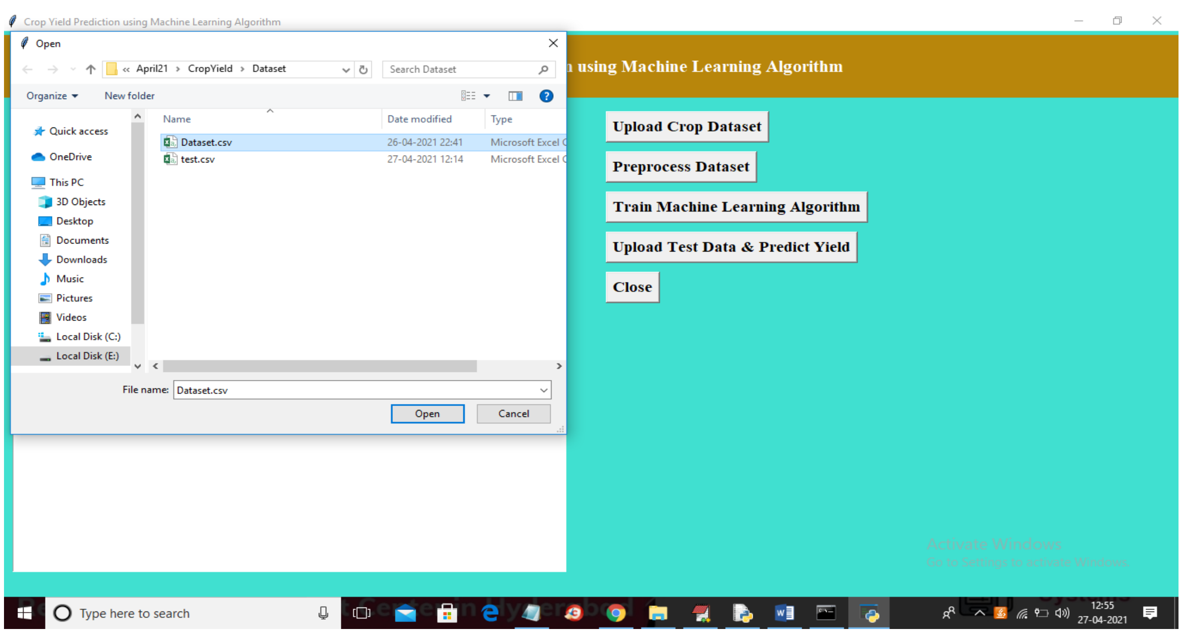

Fig.2 Results screenshot 2

In above screen selecting and uploading ‘Dataset.csv’ file and then click on ‘Open’ button to load dataset and to get below screen.

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

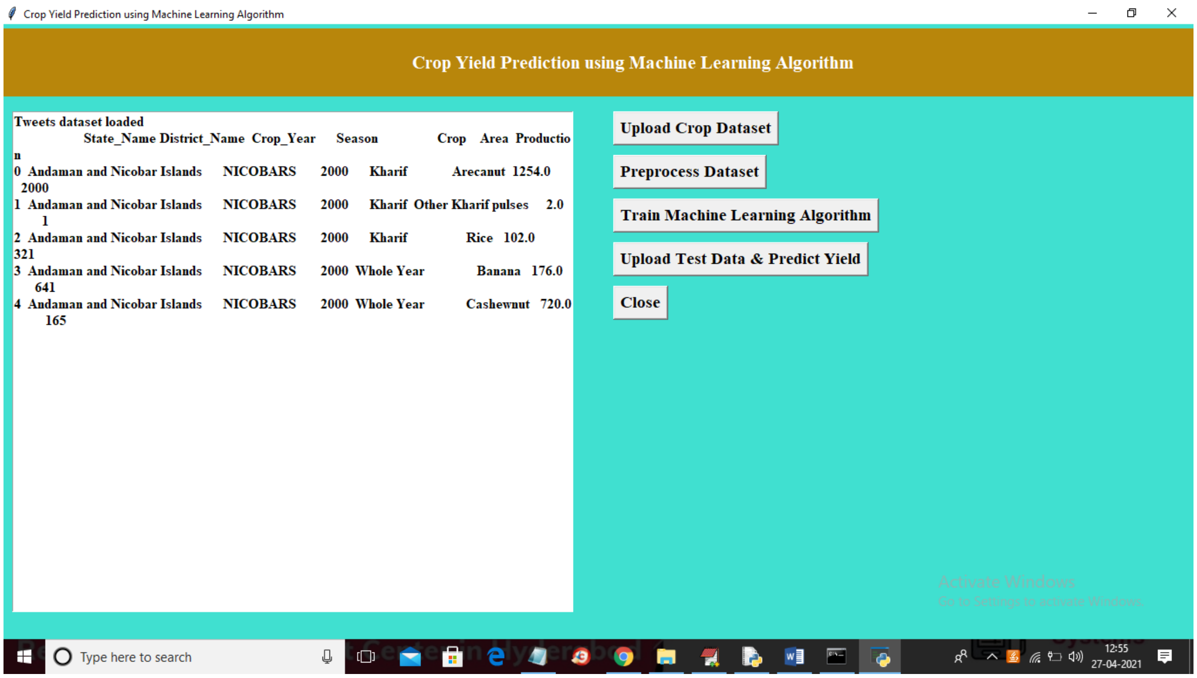

Fig.3 Results screenshot 3

In above screen dataset loaded and we can see dataset contains some non-numeric values and ML will not take non-numeric values so we need to preprocess dataset to convert non-numeric values to numeric values by assigning ID to each non-numeric value. So click on 'Preprocess Dataset' button to process dataset.

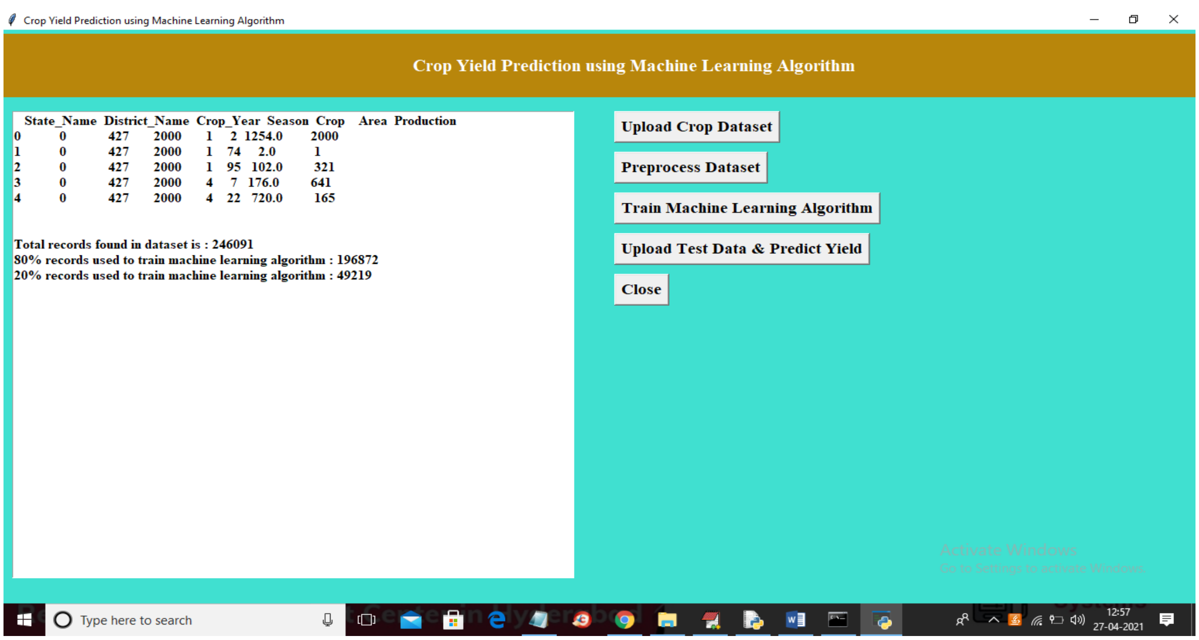

Fig.4 Results screenshot 4

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

In above screen all non-numeric values converted to numeric format and in below lines we can see dataset contains total 246091 records and application using (80%) 196872 records to train ML and using (20%) 49219 records to test ML prediction error rate (RMSE (root mean square error)). Now click on 'Train Machine Learning Algorithm' button to train Decision Tree Machine learning algorithm on above dataset and then calculate prediction error rate.

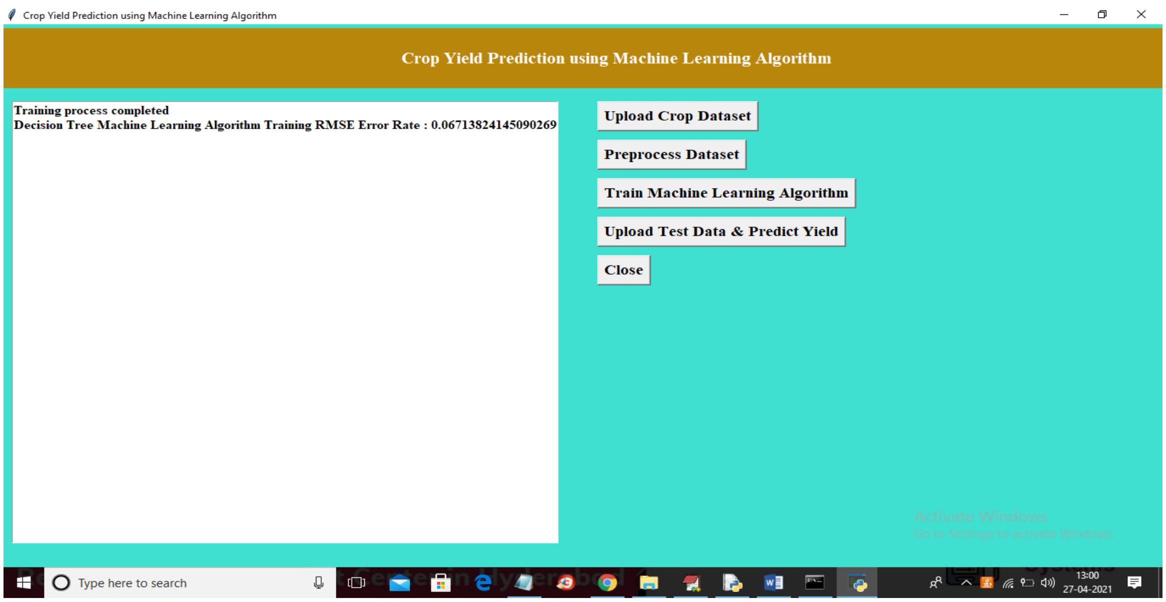

Fig.5 Results screenshot 5

In above screen ML is trained and we got prediction error rate as 0.067% and now Decision Tree model is ready and now click on 'Upload Test Data & Predict Yield' button to upload test data and then application will predict production.

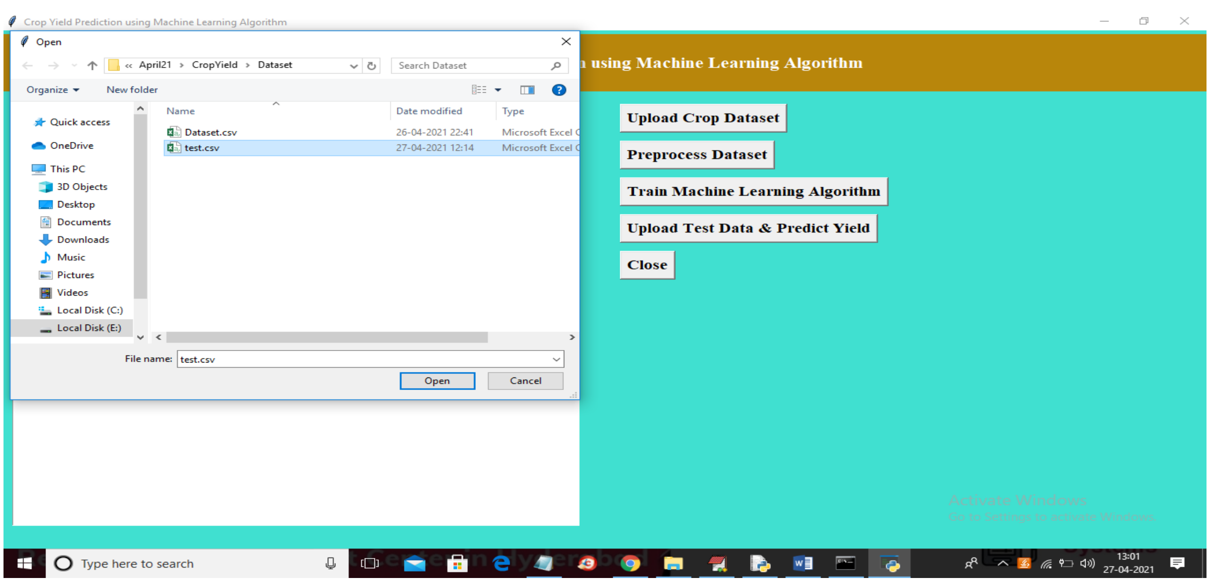

Fig.6 Results screenshot 6

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

In above screen selecting and uploading 'test.csv' file and then click on 'Open' button to load test data and then application will give below prediction result.

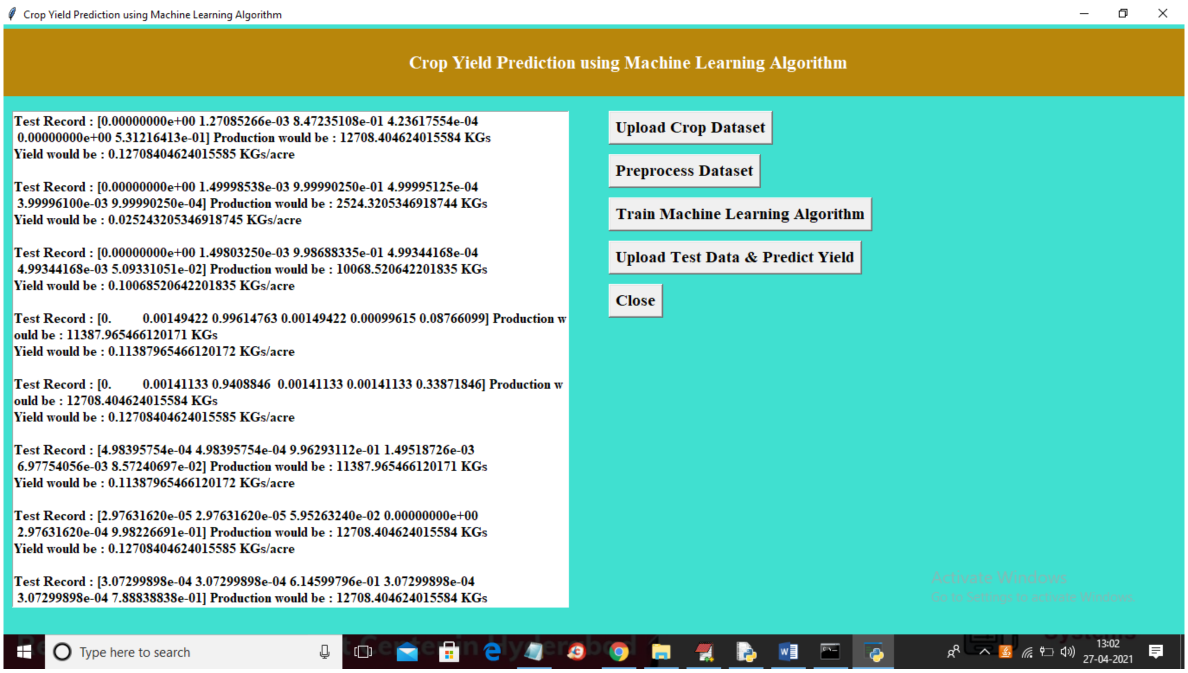

Fig.7 Results screenshot 7

In above screen each test record is separated with newline and in above screen in square bracket we can see test data values and after square bracket we can see predicted production and after that we can see predicted YIELD per acre. So, each test record and its prediction is separated with newline.

The discussion may also address the broader implications of integrating machine learning techniques into agricultural practices. Beyond predicting crop yield, machine learning models can be applied to various other aspects of agriculture, such as pest and disease detection, soil quality assessment, and precision farming. This highlights the transformative potential of machine learning in revolutionizing agricultural practices and enhancing food security and economic stability. Overall, the results and discussion section of the paper provides a comprehensive analysis of the performance of machine learning algorithms in predicting crop yield based on productivity and season. By presenting the findings, discussing their implications, and highlighting potential applications, this section contributes to advancing our understanding of the role of machine learning in agriculture and its potential to address key challenges facing the agricultural sector.

CONCLUSION

The paper presented the various machine learning algorithms for predicting the yield of the crop on the basis of temperature, rainfall, season and area. Experiments were conducted on Indian government dataset and it has been established that Decision Tree Regression gives the highest yield prediction accuracy. Sequential model that is Simple Recurrent Neural Network performs better on rainfall prediction while LSTM is good for temperature prediction. By combining rainfall, temperature along with other parameters like season and area, yield prediction for a certain district can be made. Results reveals that Decision Tree is the best classifier when all parameters are combined. This will not only help farmers in choosing the right crop to grow in the next season but also bridge the gap between technology and the agriculture sector. The proposed methodology attempts to make use of evolutionary Decision Tree Regression

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue.2, April 2024

Algorithm to get most optimized feature subset which can be used to train machine learning algorithms in most efficient way. From experimentations, here ,the decision tree regression algorithm is produced the best accuracy of 99%, and also the mean square error for this algorithm is also very less maintained through the incorporation of the Decision Tree Regression algorithm and preprocessing techniques, the system effectively learns from agricultural datasets to predict crop yields and recommend suitable fertilizers based on input features.

REFERENCES

1. Pal, M., & Mather, P. M. (2005). An assessment of the effectiveness of decision tree methods for land cover classification. Remote sensing of environment, 86(4), 554-565.

2. Sarker, R. H., & Matin, M. A. (2021). A comparative analysis of machine learning algorithms for predicting crop yield. Computers and Electronics in Agriculture, 186, 106057.

3. Singh, A., & Goyal, L. M. (2021). Machine learning approaches for agricultural yield prediction: A comprehensive review. Computers and Electronics in Agriculture, 183, 106036.

4. Bhola, R. K., Yadav, R., Verma, A. K., & Roy, P. (2021). Comparative analysis of machine learning models for crop yield prediction: A case study of Uttar Pradesh, India. Computers and Electronics in Agriculture, 187, 106170.

5. Acharya, T. P., Thakur, B., & Sahoo, S. K. (2021). A comprehensive review on machine learning based crop yield prediction model for agriculture. Computers and Electronics in Agriculture, 189, 106381.

6. Du, H., Chen, J., Gao, Z., Wei, X., & Shen, H. (2020). A survey on deep learning models for crop yield prediction. Computers and Electronics in Agriculture, 177, 105689.

7. Gebremariam, G., & Gebremedhin, K. (2021). Application of machine learning algorithms for crop yield prediction: A review. Computers and Electronics in Agriculture, 189, 106359.

8. Dong, Z., Wang, C., Ma, Y., & Wang, L. (2021). Crop yield prediction based on improved LSTM neural network algorithm. Computers and Electronics in Agriculture, 187, 106184.

9. Xie, Z., Wang, L., Qiao, P., & Cheng, Y. (2020). Predicting maize yield using machine learning methods: A comparative analysis. Computers and Electronics in Agriculture, 178, 105780.

10. Zhao, X., Wei, Z., Yang, Y., & Zhang, S. (2021). An ensemble machine learning approach for crop yield prediction using meteorological data. Computers and Electronics in Agriculture, 189, 106322.

11. Wang, K., & Wang, T. (2020). Crop yield prediction based on a deep learning model. Computers and Electronics in Agriculture, 170, 105279.

12. Bhandari, A. K., & Raman, B. (2021). A review on applications of machine learning techniques in agriculture: Crop yield prediction and disease detection. Computers and Electronics in Agriculture, 186, 106041.

13. Barbedo, J. G. A., & Koenigkan, L. V. (2021). Artificial intelligence applied to agriculture: A review on methods and tools for water management and crop yield estimation. Computers and Electronics in Agriculture, 189, 106333.

14. Kattge, J., Boenisch, G., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., ... & Wirth, C. (2020). TRY plant trait database–enhanced coverage and open access. Global Change Biology, 26(1), 119-188.

15. Takahashi, T., Saito, T., & Matsumoto, K. (2021). Machine Learning Techniques for Crop Yield Prediction Using Remote Sensing Data: A Review. Remote Sensing, 13(10), 1981.