2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N) | 979-8-3503-5681-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICAC2N63387.2024.10895212

Development of a Machine Learning-Based System for Optimizing Crop Recommendations

Vinay Tomar
Dept of Chemistry
Jiwaji University
Gwalior, India
chemistry.scholar2022@gmail.com

Gaurav Sharma
Technical Architect
Espire Infolabs
Gurgaon, India
sha.unchained@gmail.com

Rahul Rajkumar
Dept. of Food Technology
Dhanalakshmi Srinivasan college of Engineering
Coimbatore , India
rahulrajkumar0706@gmail.com

Akanksha Singh
Dept. of Agriculture
Mangalayatan University
Aligarh, India
akanksha.singh@mangalayatan.edu.in

Sarvendra Singh Dhakare
Dept. of Agriculture
J.S. University
Shikohabad, India
ssdhakare1986@gmail.com

Kamlesh kumar
Dept. of Agriculture
J.S university
Shikohabad, India
chaurasiya.kamlesh857@gmail.com

***Abstract*—In precision agriculture, crop recommendation systems play a crucial role in enhancing crop productivity. This research paper proposes a machine learning-based crop recommendation system that leverages climatic variables—such as temperature, humidity, and rainfall—as well as soil characteristics, including nitrogen, potassium, and pH levels. Utilizing a dataset that integrates soil parameters, climate data, and corresponding crop yield information, we aim to train and evaluate several machine learning algorithms to determine their efficacy in providing crop recommendations. By comparing the performance of these algorithms, our proposed system is designed to assist farmers and agricultural experts in selecting and managing crops more effectively, thereby improving overall crop yields and productivity**

***Index Terms*—Crop Prediction, Recommendation, Supervised Learning, Multilabel Classification, Machine Learning**

I. Introduction

Crop recommendation is a crucial aspect of precision agriculture that involves the identification of the best crop to grow can depend on various factors like the climate and soil properties. The goal of this process is to maximize crop yield and quality while minimizing environmental impact and resource utilization. Traditionally, crop recommendation was based on expert knowledge and experience, which often lacked consistency and relied heavily on intuition. With the advent of machine learning (ML) techniques, the process of crop recommendation has become more efficient and accurate. ML algorithms can analyze large datasets of agricultural information, including soil properties, weather data, and historical crop yields, to identify the optimal crop

to grow in a given region. The results of several studies have been promising and suggest that the development of crop recommendations is progressing well. The main aim of the study is to make a crop recommendation system using ML algorithms, specifically Decision Tree, Random Forest, Naive Bayes, and Support Vector Machines (SVM). This studyhave used a dataset consisting of soil parameters, climate data, and crop yield information to train and test our models. The system considers factors such as temperature and humidity, in addition to a number of other soil variables, when making a recommendation for crop growth. The essay is organized in the following order. The section which follows will look at previous work on crop recommendation systems and machine learning research. This studyshall detail the methodology utilized in developing the Crop recommendation system as well as the research that led to it. Afterwards, this studywill put together a summary of the research this studyperformed and give insight into what the different combinations of technology would look like. Finally, this studywill discuss the implications of our findings and conclude with potential future work. Crop recommendation systems based on ML techniques have the potential to revolutionize the way crops are grown and managed. These systems can provide farmers with accurate and personalized recommendations based on the specific conditions of their land, leading to improved crop yields, reduced resource utilization, and increased profitability. Such systems have been developed in a number of research employing a variety of ML techniques, such as Decision Trees, Random Forests, SVM. In recent years, researchers have also explored the use of Neural networks and convolutional neural networks are examples of deep learning techniques in crop recommendation systems. These techniques have shown promising results, particularly in image-based crop recognition systems, but they require large amounts of training data and computational resources.

979-8-3503-5681-6/24/$31.00©2024 IEEE

II. Literature Survey

Pandé et al. (2021) proposed a crop recommender system using machine learning, focusing on the benefits of providing personalized crop recommendations for farmers. The system considers environmental factors, offering practical solutions for improving agricultural yields and decision-making [1]. Asadollah et al. (2024) presented a novel ML approach to agricultural productivity optimization using remote sensing data across Europe. Their study utilized advanced data analytics, focusing on the potential of ML algorithms to enhance crop recommendations through improved environmental monitoring [2]. Chougule et al. (2018) examined crop suitability and fertilizer recommendations using data mining techniques. Their research emphasized how data-driven approaches can effectively match crops with appropriate fertilizers, improving yield predictions and fostering sustainable farming practices [3]. Gayatri et al. (2015) highlighted the role of IoT in providing smart agricultural solutions for farmers, with a focus on real-time data collection to optimize crop recommendations. This study stressed the integration of sensors and IoT to gather key environmental parameters [4]. Reddy et al. (2019) explored a crop recommendation system designed specifically for the Ramtek region. Their study integrated machine learning to maximize crop yield by considering local environmental conditions, such as temperature and soil type [5]. Senapaty et al. (2024) proposed a decision support system for crop recommendations using machine learning classification algorithms. This study emphasized how decision support systems can improve agricultural practices by aiding farmers in making informed choices regarding crop cultivation [6]. Lahza et al. (2023) focused on optimizing crop recommendations using novel machine learning techniques. Their approach incorporated various datasets, including soil properties and weather patterns, to fine-tune the crop prediction models for enhanced accuracy [7]. Musanase et al. (2023) explored a data-driven analysis for revolutionizing farming practices through machine learning-based crop and fertilizer recommendations. The study advocates for data-driven decision-making to improve crop yields and farming efficiency [8]. Lacasta et al. (2018) discussed an agricultural recommendation system aimed at crop protection. They focused on the intersection of crop protection and recommendation systems, demonstrating how machine learning can be used for both pest control and crop suitability [9].

Barvin and Sampradeepraj (2023) conducted a systematic review of crop recommendation systems based on environmental factors, proposing the use of Graph Convolutional Networks for more precise crop recommendations. Their review examined current challenges and the future potential of deep learning in agriculture [10]. Gosai et al. (2021) introduced a crop recommendation system utilizing machine learning, emphasizing the importance of environmental data such as soil quality and weather patterns. The system aimed to optimize crop choices based on local conditions [11]. Balakrishnan et al. (2023) discussed an agricultural crop recommendation system that integrates machine learning techniques. Their study aimed at improving crop productivity through the use of intelligent algorithms, offering farmers data-driven insights [12]. Malki (2023) reviewed the sustainable growth challenges faced by SMEs in agriculture, with a focus on how technology, including crop recommendation systems, could alleviate these challenges and contribute to sustainable agricultural practices [13]. Banerjee et al. (2020) proposed a fuzzy logic-based crop recommendation system, offering an alternative to traditional ML approaches. Their system aimed to account for uncertainties in environmental data, helping farmers make more adaptable crop choices [14]. Ujjainia et al. (2021) introduced an ensemble-based technique for crop recommendation, focusing on improving crop productivity. Their approach combined multiple machine learning algorithms to provide more reliable and precise crop recommendations [15]. S. M. Pande et al. (2021) developed a crop recommender system using ML algorithms to suggest suitable crops based on climate and soil type, improving farming productivity by providing tailored solutions for farmers (Pande et al., 2021). Similarly, D. Gosai et al. (2021) proposed a crop recommendation system that employs machine learning to predict suitable crops by analyzing parameters like soil quality and weather conditions, addressing the challenges of crop selection in different regions [16]. M. A. A. Tomh (2022) explored factors influencing design changes in the construction industry, providing insights into the broader application of machine learning in decision-making processes. Although focused on construction, the methodologies could inform decision-making in crop recommendations where environmental and operational changes occur frequently [17]. Gosai et al. (2021) developed a crop recommendation system using machine learning, significantly enhancing precision agriculture practices [20]. " P. A. and colleagues (2021) introduced an intelligent crop recommendation system at ICCMC 2021, showcasing advancements in machine learning applications for agriculture [19]." Kuraishi (2023) discussed the growing role of analytics, particularly big data and AI, in various industries, including agriculture. While the focus was on supply chain management, the principles of using large datasets and predictive analytics are applicable to crop recommendation systems, demonstrating the broader utility of these technologies [20]. Elbasi et al. (2023) proposed a crop prediction model using machine learning algorithms, emphasizing the importance of integrating diverse environmental variables to predict crop yields accurately. Their work highlighted how ML can be applied to predict agricultural outcomes, optimizing crop selection and improving yield forecasting [21]. Kaushik et al. (2024) developed a novel approach using ensemble learning and LSTM to tackle hate speech in smart environments, demonstrating enhanced detection capabilities [22]. Sikarwar et al. (2024) explore the integration of OpenCV and IoT for improving lane management in smart cities, highlighting the potential for enhanced traffic control [23]. Tanwar et al. (2024) introduce a CNN-based method for detecting brain hemorrhages, aiming to enhance healthcare in intelligent environments through accurate diagnostics [24]. Kaushik et al.

(2024) present a method for road segmentation from aerial images, utilizing advanced machine learning techniques to improve smart transportation systems [25].

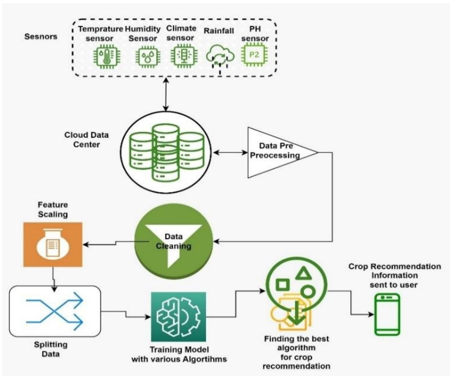

Fig. 1. Working Diagram of Proposed Model

III. PROPOSED METHOD

In this study, the dataset used in our study was obtained from the internet and consisted of soil parameters, weather data, and crop yield information for various crops grown in the region and performed initial data exploration to Gain a better understanding of the data would be beneficial.

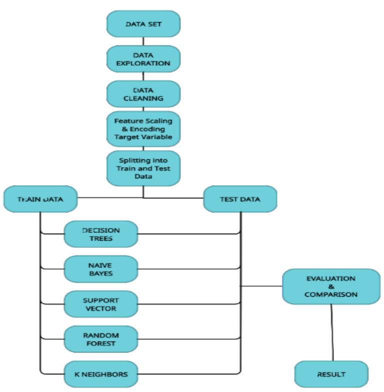

Fig. 2. Working Flowchart

Following this study have cleaned the data., which involved identifying and handling missing or incorrect values, handling categorical variables, and addressing outliers. This studyt hen scaled the independent features to prepare them for use in our machine-learning models and encoded the target variables if necessary.

This study divided the produced data into training and testing sets after finishing the preprocessing processes. Then, utilizing testing data, this study trained the data using several machine learning algorithms and assessed their effectiveness. In the end, this studychose the model that performed the best based on the evaluation measures, and this studyapplied this model to forecast new data. The study's aim was to. choose the most suitable machine learning method for the given dataset and challenge.

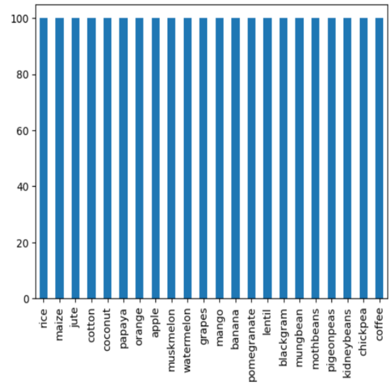

Fig. 3. Target Variable

A. *The Machine Learning algorithms are:*

1) Decision Tree Classifier: The Decision Tree Classifier is a method that decides how to divide data into various categories using a tree-like model. Recursively dividing the data into subsets according to its features, it then builds a tree with each internal node standing for a feature and each leaf node for a classification label. Applications where the objective is to classify data into several categories based on a set of input features frequently utilize the decision tree classifier.

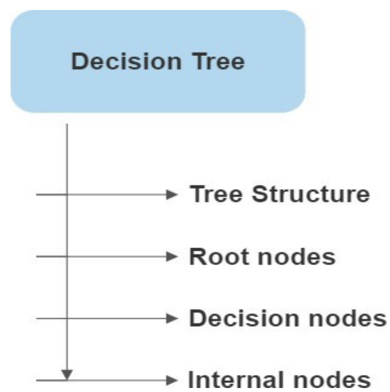

Fig. 4. Decision Tree Features

2) Naive Bayes Classifier: The The chance of a data point being part of the data is estimated using a machine learning method. to a specific class using Bayes' theorem. In order to determine the likelihood of each class based on the occurrence of each feature in the data, it works by assuming that the features of the data are independent of one another.

3) Support Vector Classifier: A machine learning algorithm called the Support Vector Classifier (SVC) identifies the most effective hyperplane for classifying data. It operates by maximising the distance between each class's closest points and the hyperplane. Applications where the objective is to divide data into two or more groups based on a set of input features frequently employ the SVC.

4) Random Forest Classifier: Several decision trees are created using the Random Forest Classifier, a machine learning

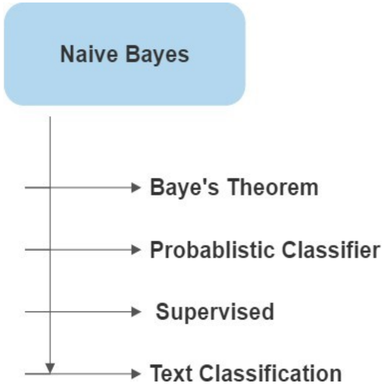

Fig. 5. Features of Naïve Bayes

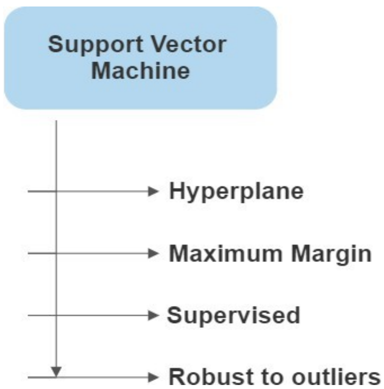

Fig. 6. SVM Features

technique, and their predictions are combined to get a final classification. It works by randomly selecting subsets of the data and features for each tree, and then aggregating the results of each tree to produce the final prediction. The Random Forest Classifier is often used in applications where the goal is to classify data into multiple categories based on a set of input features.

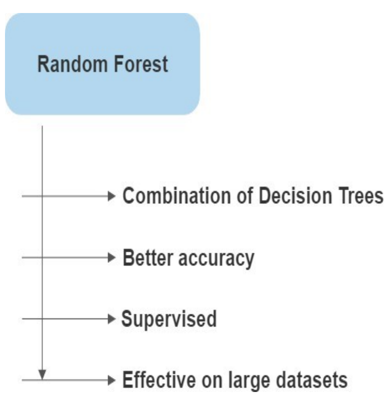

Fig. 7. Features of Random Forest

5) K-Neighbors Classifier: A machine learning technique known as the K-Neighbors Classifier makes predictions about the classification of a data point based on the classification of its k-nearest neighbors. It works by calculating the distance between each data point and its neighbors and then selecting the k closest neighbors To determine the classification data point. The K-Neighbors Classifier is used in applications where the goal is to classify data into two or more categories based on a set of input features.

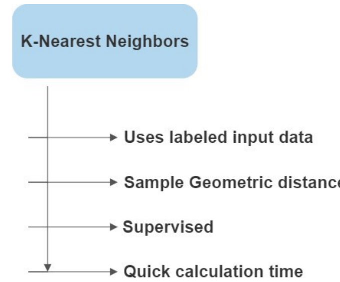

Fig. 8. Features of KNN

IV. Result & Analysis

The world of machine learning is rife with algorithms designed to extract insights from complex datasets. To assess the performance of these algorithms, accuracy is often used as a key metric. In this study, this studyanalyzed the accuracy of five classification algorithms on a given dataset, and the results are intriguing.

TABLE I
ACCURACY

S.NO	Algorithm	Accuracy
1	Decision tree	0.988636
2	Navie Bayes	0.995455
3	SVM	0.968182
4	Random forest	0.993182
5	KNN	0.956818

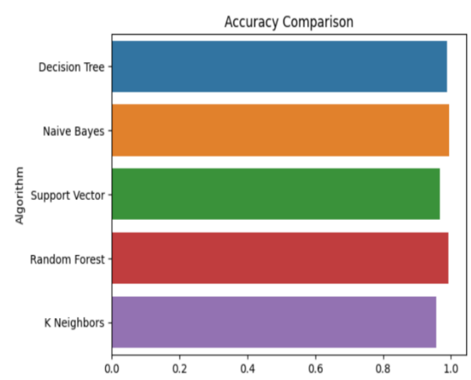

Fig. 9. Graphical Representation

At the top of the leaderboard is the Naive Bayes algorithm with an accuracy of 0.995455. This indicates that the model correctly classified almost 99% of the data, a remarkable feat indeed. The Naive Bayes algorithm's success lies in its

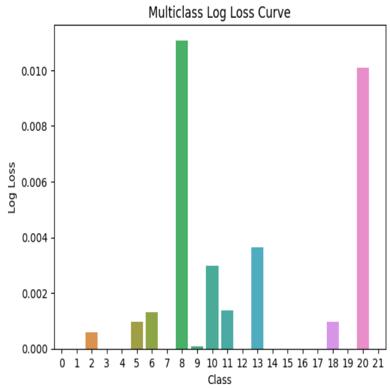

Fig. 10. MLL Chart of Naive Bayes

ability to model complex dependencies between features while maintaining simplicity and efficiency.

A close second is the Random Forest algorithm with an accuracy of 0.993182. This ensemble method combines multiple decision trees to achieve higher accuracy than individual trees. In this case, the Random Forest algorithm managed to outperform the Decision Tree algorithm, which achieved an accuracy of 0.988636, highlighting the importance of combining multiple models for better accuracy. The Support Vector algorithm, on the other hand, achieved an accuracy of 0.968182, which was lower than both the Naive Bayes and Random Forest algorithms. This could be due to the algorithm's sensitivity to the kernel function used to map the data into a higher-dimensional space. The K Neighbors algorithm had the lowest accuracy of 0.956818, indicating that it was the least effective at classifying the data compared to the other algorithms.

V. CONCLUSION

We suggest a machine learning based crop recommendation system based on performance on a dataset comprising soil characteristics, meteorological data, and crop yield data. Our results showed that the Naïve Bayes algorithm outperformed the other models, achieving an accuracy score of 0.9954 on the test set. With an accuracy score of 0.9931, Random Forest also did well, while Decision Trees, SVM, and other algorithms scored reasonably well. Our suggested crop recommendation method can aid farmers and agricultural professionals in selecting and managing crops in a way that will ultimately improve crop yields and productivity. Our results demonstrate that machine learning algorithms can analyze large datasets of agricultural information to identify the optimal crop to grow in a region. To enhance the effectiveness of crop recommendation systems, however, there are a number of topics that could be researched in the future. To give more thorough information on environmental conditions, it may be possible to combine multiple data sources, such as remote sensing and satellite photography. Future work will explore the neural network and other deep learning techniques. improve the accuracy and robustness of crop recommendation systems. As a result, it is advised to continue using the Naive Bayes and Random Forest algorithms with the provided dataset because they both showed the highest accuracy levels. When choosing the best algorithm for a certain task, it's crucial to keep in mind that additional aspects including the algorithm's complexity, interpretability, and computational needs should also be taken into account. Overall, the findings of our study show the potential of machine learning algorithms in creating useful and effective crop recommendation systems that can help farmers choose and manage crops based on data.

VI. FUTURE SCOPE

As in this paper we have discussed about the best training algorithm which predicts the crop recommendation .As in the future by using this algorithm we can find the best crop which suits the soil according to the conditions of different parameters and also they can be a Web application for crop recommendation so that it predicts the best crop according to the different conditions

REFERENCES

[1] AISHWARYA, K. ROHILLA and K. SHAURYA, ”Crop Recommender System Using Machine Learning Approach,” 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351.

[2] S. B. H. S. Asadollah, A. Jodar-Abellan, and M. Á. Pardo, "Optimizing machine learning for agricultural productivity: A novel approach with RScv and remote sensing data over Europe," Agricultural Systems, vol. 218. Elsevier BV, p. 103955, Jun. 2024. doi: 10.1016/j.agsy.2024.103955.

[3] A. Chougule, V. K. Jha, and D. Mukhopadhyay, "Crop Suitability and Fertilizers Recommendation Using Data Mining Techniques," Advances in Intelligent Systems and Computing. Springer Singapore, pp. 205–213, Jul. 10, 2018. doi: 10.1007/978-981-13-0224-4_19.

[4] M. K. Gayatri, J. Jayasakthi and G. S. Anandha Mala, ”Providing Smart Agricultural solutions to farmers for better yielding using IoT,” 2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR), Chennai, India, 2015, pp. 40-43, doi: 10.1109/TIAR.2015.7358528.

[5] D. A. Reddy, B. Dadore, and A. Watekar, "Crop Recommendation System to Maximize Crop Yield in Ramtek region using Machine Learning," International Journal of Scientific Research in Science and Technology. Technoscience Academy, pp. 485–489, Feb. 15, 2019. doi: 10.32628/ijsrst196172.

[6] M. K. Senapaty, A. Ray, and N. Padhy, "A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms," Agriculture, vol. 14, no. 8. MDPI AG, p. 1256, Jul. 30, 2024. doi: 10.3390/agriculture14081256.

[7] H. Lahza et al., "Optimization of Crop Recommendations Using Novel Machine Learning Techniques," Sustainability, vol. 15, no. 11. MDPI AG, p. 8836, May 30, 2023. doi: 10.3390/su15118836.

[8] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, and I. Kabandana, "Data-Driven Analysis and Machine Learning-Based Crop and Fertilizer Recommendation System for Revolutionizing Farming Practices," Agriculture, vol. 13, no. 11. MDPI AG, p. 2141, Nov. 13, 2023. doi: 10.3390/agriculture13112141.

[9] J. Lacasta, F. J. Lopez-Pellicer, B. Espejo-García, J. Nogueras-Iso, and F. J. Zarazaga-Soria, "Agricultural recommendation system for crop protection," Computers and Electronics in Agriculture, vol. 152. Elsevier BV, pp. 82–89, Sep. 2018. doi: 10.1016/j.compag.2018.06.049.

[10] P. Ayesha Barvin and T. Sampradeepraj, "Crop Recommendation Systems Based on Soil and Environmental Factors Using Graph Convolution Neural Network: A Systematic Literature Review," ECSA 2023. MDPI, Nov. 15, 2023. doi: 10.3390/ecsa-10-16010.

[11] D. Gosai, C. Raval, R. Nayak, H. Jayswal, and A. Patel, “Crop Recommendation System using Machine Learning,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology. Technoscience Academy, pp. 558–569, Jun. 01, 2021. doi: 10.32628.

[12] D. Balakrishnan, A. P. Kumar, K. Sai Kiran Reddy, R. R. Kumar, K. Aadith and S. Madhan, ”Agricultural Crop Recommendation System,” 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205756.

[13] M. Al Malki, “A Review of Sustainable Growth challenges faced by Small and Medium Enterprises,” International Journal for Global Academic & Scientific Research, vol. 2, no. 1. International Consortium of Academic Professionals for Scientific Research, pp. 53–67, Mar. 30, 2023. doi: 10.55938/ijgasr.v2i1.39.

[14] G. Banerjee, U. Sarkar, and I. Ghosh, “A Fuzzy Logic-Based Crop Recommendation System,” Advances in Intelligent Systems and Computing. Springer Singapore, pp. 57–69, Nov. 24, 2020. doi: 10.1007/978-981-15-7834-2_6.

[15] S. Ujjainia, P. Gautam, and S. Veenadhari, “A Crop Recommendation System to Improve Crop Productivity using Ensemble Technique,” International Journal of Innovative Technology and Exploring Engineering, vol. 10, no. 4. Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP, pp. 102–105, Feb. 28, 2021. doi: 10.35940/ijitee.d8507.0210421.

[16] S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA, and K. SHAURYA, “Crop Recommender System Using Machine Learning Approach,” 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, Apr. 08, 2021. doi: 10.1109/iccmc51019.2021.9418351.

[17] M. A. A. Tomh, “Investigating the Factors Affecting the Design Change In Construction Industry In UAE,” International Journal for Global Academic & Scientific Research, vol. 1, no. 4. International Consortium of Academic Professionals for Scientific Research, pp. 60–73, Dec. 31, 2022. doi: 10.55938/ijgasr.v1i4.32.

[18] D. Gosai, C. Raval, R. Nayak, H. Jayswal, and A. Patel, “Crop Recommendation System using Machine Learning,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology. Technoscience Academy, pp. 558–569, Jun. 01, 2021. doi: 10.32628/cseit2173129.

[19] P. A, S. Chakraborty, A. Kumar, and O. R. Pooniwala, “Intelligent Crop Recommendation System using Machine Learning,” 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, Apr. 08, 2021. doi: 10.1109/iccmc51019.2021.9418375.

[20] Z. Mohammad Kuraishi, “Role of Analytics in Supply Chain Management Industry in Lithuania: Big Data Analytics & AI,” International Journal for Global Academic & Scientific Research, vol. 2, no. 4. International Consortium of Academic Professionals for Scientific Research, pp. 44–53, Dec. 31, 2023. doi: 10.55938/ijgasr.v2i4.65.

[21] . Elbasi et al., “Crop Prediction Model Using Machine Learning Algorithms,” Applied Sciences, vol. 13, no. 16. MDPI AG, p. 9288, Aug. 16, 2023. doi: 10.3390/app13169288.

[22] P. Kaushik, R. Rohilla, P. Walia, S. Shankar, M. M. Kaushik and T. S. Gupta, ”Confronting Hate Speech in SMART Environments: An Approach that Uses Ensemble Learning and LSTM,” 2023 International Conference on Smart Devices (ICSD), Dehradun, India, 2024, pp. 1-6, doi: 10.1109/ICSD60021.2024.10751441.

[23] S. S. Sikarwar, N. Chawla, P. Agarwal, M. F. Afzal, S. P. S. Rathore and N. Valecha, ”Improving Lane Management: Leveraging OpenCV and IoT to Empower Smart Environments,” 2023 International Conference on Smart Devices (ICSD), Dehradun, India, 2024, pp. 1-5, doi: 10.1109/ICSD60021.2024.10751495.

[24] S. Tanwar, N. Choudhary, V. Chaudhary, A. Dahiya, P. Kaushik and R. Rathore, ”Advancing Healthcare: CNN-Based Brain Hemorrhage Detection in Intelligent Environments,” 2023 International Conference on Smart Devices (ICSD), Dehradun, India, 2024, pp. 1-6, doi: 10.1109/ICSD60021.2024.10751290.

[25] M. M. Kaushik, M. Kakkar, M. Yadav, P. Kaushik, S. Jha and V. Jaitwal, ”Smart Road Segmentation from Aerial Images,” 2023 International Conference on Smart Devices (ICSD), Dehradun, India, 2024, pp. 1-6, doi: 10.1109/ICSD60021.2024.10751072.