Agriculture Crop Recommendation System using Machine learning

Sangeetha Allam
Department of Computer Science & Engineering
MLR Institute of Technology
Hyderabad, India
allamsangeetha@gmail.com

Bollimuntha Manjusha
Assistant professor
Department of computer science and Engineering
MLR Institute of Technology, Hyderabad
bollimuntha.manjusha@gmail.com

Asha Vuyyuru,
Assistant Professor,
Department of Computer Science and Engineering (AIML),
BVRIT HYDERABAD College of Engineering for Women,
Hyderabad.
vuyyuru.asha@gmail.com

P. Veeranjaneyulu,
Assistant Professor,
Department of CSE- (CyS, DS) and AI&DS,
VNR Vignana Jyothi Institute of Enginnering and Technology,
Hyderabad.
pagadalaveeru@gmail.com

K. Pushpa Rani
Department of Computer Science and Engineering
MLR Institute of Technology
Dundigal, Hyderabad-500043
rani536@mlrinstitutions.ac.in

Sushruta Mishra
Assistant Professor
Kalinga Institute of Industrial Technology, deemed to be University, India
sushruta.mishrafcs@kiit.ac.in

Abstract*— India's socioeconomic system is mostly dependent on agriculture. Farmers used to rely on word-of-mouth, but the current climate prevents them from doing so. The IT industry has grown in order to provide farmers with precise agricultural information, which has resulted in several advancements in the field of agriculture sciences. In the current climate, having access to contemporary technologies for agriculture is desirable.[1] Using machine learning methods, it is feasible to make predictions and create a precise model from the data. There are solutions for agricultural problems such crop rotation, crop forecasts, water and fertilizer needs, and crop protection. Due to the environment's unpredictable climatic elements, an effective method is required to expedite crop cultivation and support farmers in their production and management. It's possible that aspiring farmers can now engage in improved agriculture.[2] A farmer can receive a number of recommendations to aid with crop yield. Based on the properties of the soil, the timing of the sowing, and the location, farmers occasionally choose the incorrect crops. [3]Many as a result commit suicide, stop farming, and go to cities in search of employment. By taking into account all relevant aspects, including the planting season, soil, and geographic location, this research effort has developed a method to assist farmers in choosing crops. Following a careful review of crop data gathering, crop recommendations were given.

Keywords—KNN, SVM, TensorFlow, Numpy, Keras, Datasets, Hand Gestures

I. Introduction

. Agriculture is one of the most prevalent and highly recognised sectors in India. It is the most varied economic sector and essential to the growth of the entire nation.[4] More than 60% of the country's land is used for agriculture, which meets the needs of its 1.3 billion citizens. Thus, integrating contemporary agricultural methods is crucial. The farmers in our country will gain from this. Based on their understanding of a particular region, farmers had previously predicted crops and yields. They chose the crop that was more established or in vogue in the area specifically for the purpose of their property since they lacked sufficient knowledge regarding the quantity of soil nutrients like nitrogen, phosphate, and potassium on the land. [5]Crop rotation is not used currently, and the soil does not receive enough nutrients. This diminishes production, causes soil pollution (soil acidity), and harms the top layer of the soil. With the intention of enhancing the farmer's position while taking all of these factors into consideration, we produced the solution utilising machine learning. Machine learning is transforming the agricultural sector (ML). In the setting Machine learning, a subfield of artificial intelligence, has emerged as a result of huge data and powerful computers. to offer new options for data-intensive research in the multidisciplinary field of agricultural technology.[6] In the realm of agriculture, for instance, using this algorithms we can get good results suggestions offered by the algorithm regarding the best crop to plant on a specific plot of land based on climate extremes including temperature, humidity, and PH, as well as the type of soil.

These had already been made using the meteorological service, the official website, and V C Farm Manda. Sensors are used to gather temperature, pH, and humidity values from farms and other sources. [7]Support Vector Machine (SVM) and In order to find patterns in input data and analyse it in line with input requirements, machine learning predictive algorithms called Decision Tree are utilised. The technique recommends a specific crop be planted and specifies the amount of nutrients that should be provided to guarantee the crop's success[14]. A anticipated yield in q/acre, the required seed dose in kg/acre, and the crop's current market price are all displayed by the system[15].

2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC) | 979-8-3503-7018-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ASSIC60049.2024.10507899

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. LITERATURE SURVEY

Ashwani Kumar Kushwaha explains methods for predicting crop yields and suggests an appropriate crop to raise farmer profits and industrial standards[16]. They collect massive volumes of data, or "big data," for this job in order to forecast crop yield. using the Hadoop framework and a farming algorithm. [8]As a result, crop quality can be increased and crop compatibility for a specific circumstance can be forecasted using repository data[17].

Crop production and rainfall forecasts are discussed by Girish L.[9]The effectiveness of several machine learning methods, including decision trees, liner regression, SVM, and KNN method, as well as a variety of machine learning strategies for crop production and rainfall prediction are examined in this paper. They find that SVM is the most reliable algorithm for predicting rainfall.

The many machine learning methods used to increase crop output are described by Rahul K. This paper discusses a variety of artificial intelligence techniques, such as large data analysis for algorithms for machine learning and precision agriculture.[10] They go over crop recommendation systems using KNN, ensemble-based models, neural networks, and other technologies in their explanations.

Most emerging economies place a significant emphasis on agriculture as a source of revenue.[13] The sector of modern agriculture is constantly growing in terms of farming practises and agricultural developments. It is challenging for farmers to adapt to the changing needs of the planet, businesses, customers, and other groups. Farmers deal with a variety of challenges.[11] I discuss industrial emissions and soil erosion as contributors to climate change. (ii) A shortage of essential minerals like potassium, nitrogen, and phosphorus can result in nutrient deficits in the soil, which can hinder crop development.[12] (iii) Farmers make the error of consistently growing the same crops without attempting new varieties.

III. PROPOSED WORK AND PROCEDURE WORK FOR MODEL DESIGN

A. Existing system:

Input criteria like soil fertility, climate, crop production, and geographic locations have all been considered in the current system. The accuracy of the model is low when compared to the suggested system. With the current system, many machine learning methods were applied.

B. Proposed system:

The proposed system would predict the ideal Based on the kind of soil, climate variables like temperature and humidity, soil PH, and rainfall, a crop can be grown on a specific piece of land. We will forecast healthy or unhealthy crops using CNN (Convolutional neural network). The accuracy of our suggested algorithm is 90%.

IV. MACHINE LEARNING ALGORITHMS

Artificial intelligence (AI) applications such as machine learning enable systems to autonomously learn from experience and progress without the need for programmatically explicit. The act of creating software that can access data and utilise it to learn for itself is the main goal of machine learning.

A. Supervised Learning

The machines are able to predict the output with the use of well-labelled training data—labelled data is input that has already been allocated to the proper output.

The proper input and output data are fundamentally provided to the machine learning model using the supervised learning approach. The aim of supervised learning algorithms is to relate the input variable (x) to the output variable by locating a mapping function or other similar mechanism (y).

B. Unsupervised Learning

As the name implies, unsupervised learning is a subset of machine learning in which models are not guided by training datasets. On the other hand, models examine the available data to elucidate buried patterns and insights. The cognitive process that takes place in the human brain when learning a new skill is strikingly comparable to this.

V. PROPOSED ALGORITHM

A. Decision Tree Algorithm

The decision tree is the most efficient and popular categorization and prediction technique. A decision tree is a tree structure made up of internal nodes that represent tests on attributes, leaf nodes (terminal nodes) that represent class labels, and branches that mimic flowcharts. A structure that resembles a tree is a decision tree.

B. Logistic Regression

There are only two valid classes as a result because the dependent variable is dichotomous in nature [18-19]. A dependent variable is, to put it simply, a binary variable with the values 1 or 0. where 1 denotes achievement/yes and 0 denotes failure/no.

C. Random Forest Algorithm

Random Forest algorithm is a supervised machine learning algorithm. This algorithm is popular for classification and Regression problems in machine learning.

D. CNN Algorithm

CNNs are a type of artificial neural network used for deep learning as well as image recognition and classification. As a result, Deep Learning makes use of a CNN to recognize objects in a picture.

E. Support Vector Machine (SVM)

SVM is a technique for supervised machine learning. This is employed in both regression and classification analysis. Data is graphed in n-employ the linear SVM classifier in our research because we are dealing with linearly separable data.

IV. SYSTEM ARCHITECTURE DESIGN:

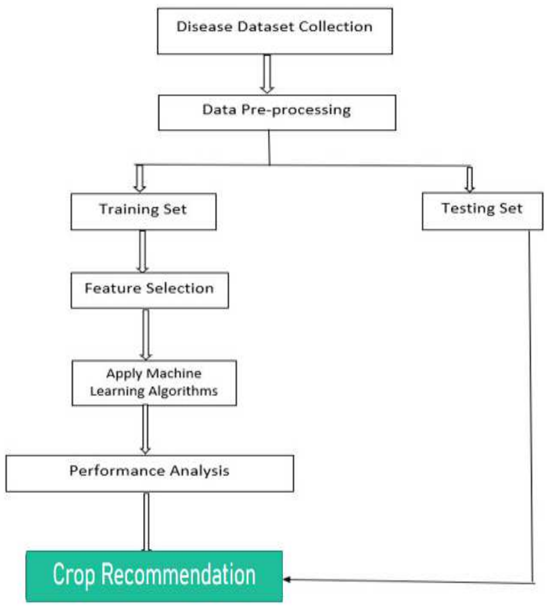

Fig 1. Flow chart of proposed Method

System architecture and data flow in analysis are depicted in Figure 1. The collection of data is the first step in the analysis process. Kaggle is used for collecting disease datasets. After that, data pre-processing is required because, for example, soil pH, P, K cannot be zero, null, or otherwise. These records have been pre-processed. Following pre-processing, data set preparation and model building for various crops using available datasets .figure 2 & 3 explains about dataset description of proposed method

The data is then split into training and testing datasets, with training datasets making up 80% and 20%, respectively, of the total data. The choice of features is now complete. By using only pertinent data and deleting irrelevant data from your model, feature selection is a technique for lowering the input variable. Then, for each disease, we used machine learning and deep learning algorithms. Following that, the performance of all algorithms is reviewed, and the best algorithm with the highest accuracy is chosen .figure 4,5 & 6 explains about signup page, log in page and farming description example of proposed method

h. This model will now recommend a crop based on the input factors.

1	temperatu	humidity	ph	rainfall	N	P	K	label	label1
2	20.87974	82.00274	6.502985	202.9355	0.7	0.1	0.8	rice	1
3	21.77046	80.31964	7.038096	226.6555	0.5	0.7	0.4	rice	1
4	23.00446	82.32076	7.840207	263.9642	0.7	0.6	0.1	rice	1
5	26.4911	80.15836	6.980401	242.864	0.8	0.1	0.7	rice	1
6	20.13017	81.60487	7.628473	262.7173	0.5	0.8	0.2	rice	1
7	23.05805	83.37012	7.073454	251.055	0.8	0.4	0.6	rice	1
8	22.70884	82.63941	5.700806	271.3249	0.2	0.5	0.8	rice	1
9	20.27774	82.89409	5.718627	241.9742	0.6	0.6	0.7	rice	1
10	24.51588	83.53522	6.685346	230.4462	0.7	0.6	0.3	rice	1
11	23.22397	83.03323	6.336254	221.2092	0.6	0.7	0.6	rice	1
12	26.52724	81.41754	5.386168	264.6149	0.1	0.6	0.5	rice	1
13	23.97898	81.45062	7.502834	250.0832	0.2	0.8	0.7	rice	1
14	26.8008	80.88685	5.108682	284.4365	0.1	0.8	0.8	rice	1
15	24.01498	82.05687	6.984354	185.2773	0.7	0.5	0.5	rice	1
16	25.66585	80.66385	6.94802	209.587	0.7	0.2	0.8	rice	1
17	24.28209	80.30026	7.042299	231.0863	0.1	0.8	0.7	rice	1
18	21.58712	82.78837	6.249051	276.6552	0.2	0.5	0.6	rice	1
19	23.79392	80.41818	6.97086	206.2612	0.8	0.5	0.8	rice	1
20	21.86525	80.1923	5.953933	224.555	0.1	0.8	0.7	rice	1
21	23.57944	83.5876	5.853932	291.2987	0.5	0.6	0.4	rice	1
22	21.32504	80.47476	6.442475	185.4975	0.1	0.6	0.5	rice	1
23	25.15746	83.11713	5.070176	231.3843	0.1	0.7	0.6	rice	1
24	21.94767	80.97384	6.012633	213.3561	0.6	0.5	0.2	rice	1
25	21.05254	82.6784	6.254028	233.1076	0.6	0.2	0.8	rice	1
26	23.48381	81.33265	7.375483	224.0581	0.8	0.8	0.7	rice	1

Fig 2. Dataset description 1

27	25.07564	80.52389	7.778915	257.0039	0.8	0.8	0.4	rice	1
28	26.35927	84.04404	6.2865	271.3586	0.8	0.5	0.3	rice	1
29	24.52923	80.54499	7.07096	260.2634	0.7	0.8	0.3	rice	1
30	20.77576	84.49774	6.244841	240.0811	0.8	0.6	0.8	rice	1
31	22.30157	80.64416	6.043305	197.9791	0.8	0.5	0.4	rice	1
32	21.44654	84.94376	5.824709	272.2017	0.6	0.6	0.8	rice	1
33	22.17932	80.33127	6.357389	200.0883	0.6	0.3	0.5	rice	1
34	24.52784	82.73686	6.364135	224.6757	0.1	0.6	0.8	rice	1
35	20.26708	81.63895	5.014507	270.4417	0.6	0.6	0.8	rice	1
36	25.73543	83.88266	6.149411	233.1321	0.1	0.7	0.5	rice	1
37	26.79534	82.14809	5.950661	193.3474	0.6	0.6	0.6	rice	1
38	26.75754	81.17734	5.96037	272.2999	0.8	0.5	0.4	rice	1
39	23.8633	83.15251	5.561399	285.2494	0.7	0.8	0.8	rice	1
40	21.01945	82.95222	7.416245	298.4018	0.5	0.1	0.6	rice	1
41	24.17299	83.72876	5.58337	257.0344	0.8	0.7	0.7	rice	1
42	22.78134	82.06719	6.43001	248.7183	0.4	0.8	0.7	rice	1
43	25.6298	83.52842	5.534878	209.9002	0.4	0.6	0.5	rice	1
44	25.59705	80.14509	6.903986	200.8349	0.3	0.5	0.5	rice	1
45	23.83067	84.8136	6.271479	298.5601	0.6	0.5	0.4	rice	1
46	26.31355	82.36699	7.224286	265.5356	0.3	0.7	0.6	rice	1
47	24.89728	80.52586	6.134287	183.6793	0.1	0.5	0.7	rice	1
48	24.95878	84.47963	5.206373	196.956	0.8	0.5	0.4	rice	1
49	23.24114	84.59202	7.782051	233.0453	0.5	0.7	0.3	rice	1
50	21.66628	80.70961	7.062779	210.8142	0.6	0.3	0.8	rice	1

Fig 3. Dataset description 2

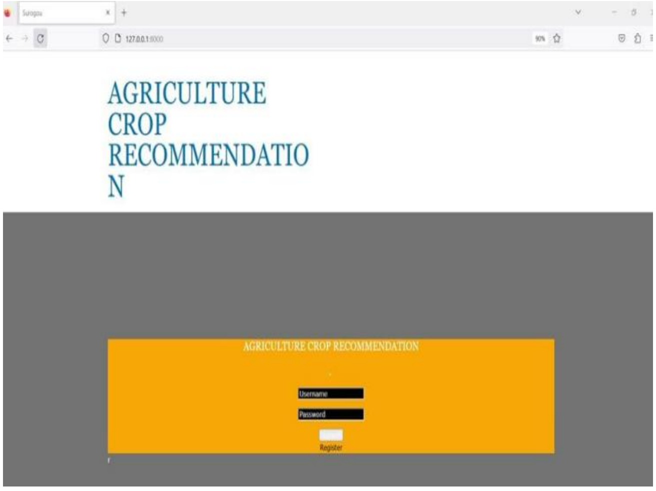

Fig 4. Signup page

Fig 5. Login page

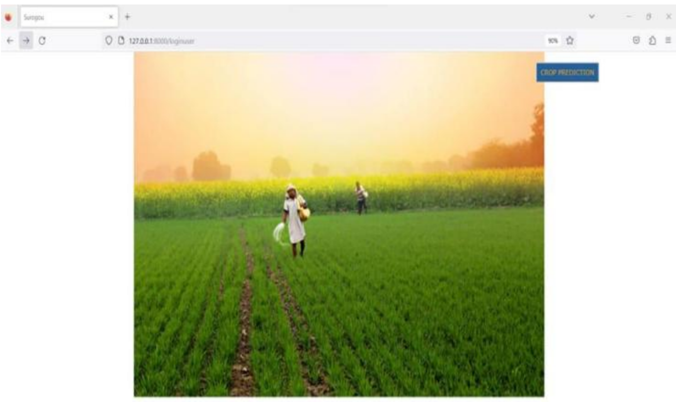

Fig 6. Farming description example

Fig 7. Different crop descriptions

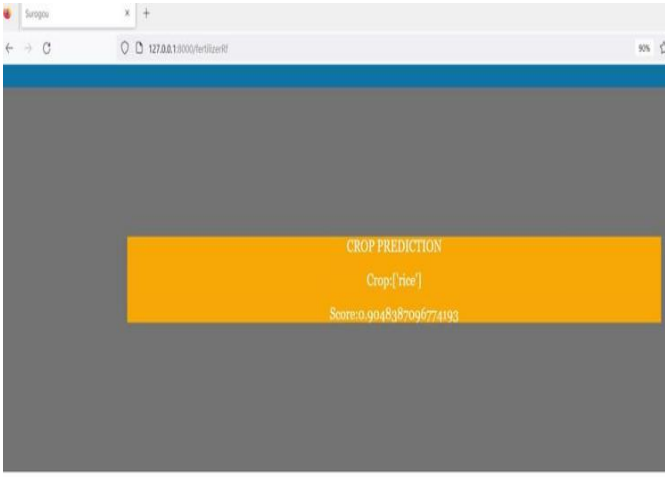

Fig 8. Prediction of CROP and Accuracy measurement

Figure 7 shows about prediction of crop and accuracy measurement.

VI. CONCLUSION

Our farmers run the risk of choosing the wrong crops for cultivation because they are currently not using technology and analysis efficiently, which would also lower their profitability. We developed a farmer-friendly system using ML that forecasts which crop is most suited to a particular plot of land and also offers crop information in order to lessen these kinds of losses. Farmers will be more likely to choose crops wisely as a result, and the agricultural industry will benefit from novel ideas.

REFERENCES

[1]. F1ilippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S.U., Bishop, T.F.A., 2019a. An approach to forecast grain crop yield using multilayered, multi-farm data sets and machine learning. Precis. Agric. 1–15. https://doi. org/10.1007/s11119-018-09628-4.

[2]. Sahoo, S., Mishra, S., Panda, B., Bhoi, A. K., & Barsocchi, P. (2023). An Augmented Modulated Deep Learning Based Intelligent Predictive Model for Brain Tumor Detection Using GAN Ensemble. Sensors, 23(15), 6930.

[3]. M.Sen, K. Sharma, S. Mishra, A. Alkhayyat and V. Sharma, "Designing a Smart and Intelligent Ecosystem for Autistic Children," 2023 4th International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2023, pp. 1-5, doi: 10.1109/ICIEM59379.2023.10166057.

[4]. Ghosh, S., & Mishra, S. (2022, November). Intelligent Virtual Ambulance Model using Predictive Learning. In 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC) (pp. 1-5). IEEE.

[5]. Saha, L., Tripathy, H. K., Nayak, S. R., Bhoi, A. K., & Barsocchi, P. (2021). Amalgamation of customer relationship management and data analytics in different business sectors—A systematic literature review. Sustainability, 13(9), 5279.

[6]. Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S., 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Comput. Electron. Agric. 153, 213–225. https://doi.org/10.1016/J.COMPAG.2018.07.016.

[7]. Son, L. H., Tripathy, H. K., Acharya, B. R., Kumar, R., & Chatterjee, J. M. (2019). Machine learning on big data: A developmental approach on societal applications. Big Data Processing Using Spark in Cloud, 143-165.

[8]. Cakir, Y., Kirci, M., Gunes, E.O., 2014. Yield prediction of wheat in south-east region of Turkey by using artificial neural networks. In: 2014 The 3rd International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2014. https://doi.org/10.1109/AgroGeoinformatics.2014.6910609.

[9]. Charoen-Ung, P., Mittrapiyanuruk, P., 2019. Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning, pp. 33–42. https://doi.org/10.1007/978-3-319-93692-5_4.

[10]. Allam Balaram, Manda Silparaj, Rajender Gajula, " Detection of malaria parasite in thick blood smears using deep learning," Materials Today: Proceedings, Vol, 64, Part 1, pp. 511–516, 2022, DOI: 10.1016/j.matpr.2022.04.1012.

[11]. T. Sakthivel, Allam Balaram,"The impact of mobility models on geographic routing in multi-hop wireless networks and extensions - a survey," International Journal of Computer Networks and Applications, Vol. 8, Issue No. 5, pp. 634–670, 2021. DOI: 10.22247/ijcna/2021/209993.

[12]. S. Spandana, M. Sathiya, K. Krishnan and A. Kiran, "Defect Exposure in Vegetables and Fruits using Machine Learning," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICCCI56745.2023.10128504.

[13]. A. Kiran, G. R. Sakthidharan, D. D. Priya, B. U. Mahesh, K. P. Kumar and K. P. Kumar, "Voice Controlled Home Automation System using Google Assistants," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICCCI56745.2023.10128217.

[14]. Kumar, M.G.V.; N, V.; Čepová, L.; Raja, M.A.M.; Balaram, A.; Elangovan, M. Evaluation of the Quality of Practical Teaching of Agricultural Higher Vocational Courses Based on BP Neural Network. Appl. Sci. 2023, 13, 1180. https://doi.org/10.3390/app13021180.

[15]. Neeraja Koppul, Koppula Srinivas Rao, Shaik Abdul Nabi, Allam Balaram," A Novel Optimized Recurrent Network-Based Automatic System for Speech Emotion Identification," Wireless Personal Communications128(3), pp. 2217–2243, 2023, DOI: 10.1007/s11277-022-10040-5.

[16]. S. Sai Kumar, Anumalareethika Reddy, B. Sivarama Krishna, J. Nageswara Rao And Ajmeera Kiran, "Privacy Preserving With Modified Grey Wolf Optimization Over Big Data Using Optimal K Anonymization Approach" Journal Of Interconnection Networks, Https://Doi.Org/10.1142/S0219265921410395, 2022.[Scopus Indexed Journal]

[17]. D Shanthi, N Swapna, Ajmeera Kiran, A Anoosha",Ensemble Approach Of GP,ACOT,PSO,And SNN For Predicting Software Reliability", International Journal Of Engineering Systems Modelling And Simulation , 2022.

[18]. Chaudhury, P., & Tripathy, H. K. (2017). An empirical study on attribute selection of student performance prediction model. International Journal of Learning Technology, 12(3), 241-252.

[19]. Tripathy, H. K., Tripathy, B. K., & Das, P. K. (2008). A knowledge based approach using fuzzy inference rules for vowel recognition. J. Convergence Inf. Technol., 3(1), 51-56