Proceedings of the International Conference on Sustainable Computing and Smart Systems (ICSCSS 2023)
IEEE Xplore Part Number: CFP23DJ3-ART; ISBN: 979-8-3503-3360-2

Agricultural Crop and Fertilizer Recommendations based on Various Parameters

Mahalakshmi B,
Professor, Department of CSE,
K S Rangasamy College of Technology
Tiruchengode, India
mahalakshmi@ksrct.ac.in

Sakthivel V,
Scholar, Department of CSE,
K S Rangasamy College of Technology
Tiruchengode, India
sakthivelvishnu20016@gmail.com

Sumitha Devi B,
Scholar, Department of CSE,
K S Rangasamy College of Technology
Tiruchengode, India
sumithabalubs@gmail.com

Swetha S,
Scholar, Department of CSE,
K S Rangasamy College of Technology
Tiruchengode, India
swethasuresh626@gmail.com

2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) | 979-8-3503-3360-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSCSS57650.2023.10169320

Abstract— **Agricultural uncertainty lowers the Tamil Nadu's productivity because it is a coastal state. More production should be possible only with more people and land, but it is not possible. Farmers used to rely on word-of-mouth, but climate circumstances prevent them from doing so. Soil is one of the important factors in cultivating the crops. The new farmers are unaware about the crop which has to be cultivated in the particular season. To solve this problem, this system helps the farmer to know about the crop which has to be cultivated in the particular season to increase the production of the crops. Machine learning methods allow us to forecast the future and produce a precise model from the data. Agricultural issues like crop rotation, crop forecasting, water and fertilizer requirements, and crop protection can be solved. The need for an effective approach to simplify crop cultivation and the need to assist farmers in managing and producing their crops results from the environment's unstable climatic conditions. This could enable aspiring farmers to practice better agriculture. With the aid of data mining, a farmer may be given a system of suggestions to aid in crop production. Crops are suggested for use in such a strategy depending on their quantity and meteorological considerations. The analysis of the crop dataset resulted in crop recommendations based on season, soil type, water and productivity. Various techniques like Linear Regression, Support Vector Regression and Adaptive Network Based Fuzzy Inference System (ANFIS) are used in this system, which enable accuracy comparison and assist farmers in selecting the best crop under various situations.**

Keywords—Smart Farming, Linear Regression, Support Vector Regression (SVM), and Adaptive Neuro-Fuzzy Inference System (ANFIS).

I. INTRODUCTION

With a population ranking of 6, Tamil Nadu is the seventh-largest region in India. It is the top manufacturer of agricultural goods. The majority of people in Tamil Nadu were farmers. In this fictitious society, agriculture has a good tone. The main crop farmed in Tamil Nadu is rice. Other crops are also grown, such as sugarcane, cotton, coconuts, and groundnuts. Bio-fertilizers are successfully produced. Farming is a significant source of employment in many places. A nation's economy is greatly impacted by its agricultural sector. Agriculture farming is getting worse today as a result of changing environmental conditions. Agriculture is directly impacted by direct sunlight, moisture, variety of soil, rainfall, highest and lowest temperatures, the weather, pesticides, fertilizers, and other environmental variables. It is necessary to have knowledge about effective crop harvesting for agriculture to flourish. In India, the winter season runs from December to the second week of March. Between April and June third week is the summer time. From July to September, there is a monsoon or rainy season. Autumnal or post-monsoon season, lasting from October through November. It is important to evaluate the best crops to farm because of the variety of seasons and rainfall. The crop can be recommended using a machine learning technique. Generally, machine learning is the subfield of artificial intelligence, which describes the accuracy of the predicted outcome. Artificial Intelligence describes the human intelligence by processes by the machines.

A. Water

Water (H2O), an inorganic substance that is transparent, tasteless, odorless, and nearly colorless, is the main component of the hydrosphere on Earth and the fluids of all known living beings [1]. It is essential for all currently recognized forms of life even though it doesn't offer food, energy, or organic micronutrients [2]. There is water everywhere in the natural world. It creates aerosols that resemble fog as well as rain-like precipitation. Clouds are made up of suspended ice and water droplets. When crystalline ice is finely fractured, snow may form as a precipitation. Water that is gaseous is referred to as steam or water vapor. The oceans and seas, which make up around 71% of the earth's surface, contain the vast bulk (about 96.5% of the total volume of water on the planet). [3] Groundwater (1.7%), glaciers and ice caps in Antarctica and Greenland (1.7%), clouds (composed of ice and liquid water suspended in air), and precipitation (0.001%) of the sky all contain negligible amounts of water. Water and runoff often end up in the ocean. Water plays an important role in the world economy. More than 70% of the freshwater that people use is used for agriculture.

B. Smart Farming

The terms "digital agriculture," "smart farming," or "e-agricultural" refer to agricultural equipment that gather, store, analyse, and disseminate electronic data or information [4][20]. The "digital agricultural revolution" is the term used

by the Food and Agricultural Organization of the United Nations to describe the process of digitising agriculture. Precision farming is one element of digital agriculture. Pre-production to post-production of the agro-food value chain are all impacted by digital agriculture, as opposed to precision agriculture. As a result, precision agriculture and digital agriculture are increasingly recognised as including agricultural technology such as yield mapping, GPS guidance systems, and variable-rate applications. Traditional farming methods could change as a result of contemporary digital technology.

C. Yield

Let's talk about typical physical design techniques to increase yield, yield analysis. The yield is the proportion that can be produced [5] in proportion to the quantity that can be sold. The average production cycle lasts more than six weeks. Consistently high yield is important for a quicker time to benefit given such significant investments.

D. Machine Learning

The machine learning technique discusses both basic and complex machine learning principles. Computers are now capable of autonomously learning [6][18] from historical data. Machine learning uses a range of techniques to build mathematical models and predict the future based on historical data or expertise. A highly effective tool for predictive analytics, machine learning is now being used extensively in the agriculture industry as well as in fraud detection, healthcare, and business process automation. The yield of agricultural cultivation will enhance with the adoption of machine learning in farmland.

II. Literature Review

A. Weather-based crop management and the impact of data analytics

According to this study by A. Swarupa Rani et al., making decisions about a variety of topics pertaining to the agricultural area requires the use of data mining. The goal of data mining is to take information from a collection of already-existing data and transform [7][17] it into a format that people can understand for a future use. The climate in a particular agricultural location affects crop management since it has a significant influence on crop productivity. Real-time weather information can assist in achieving effective crop management [8]. By using information and communication technology, significant data extraction can be automated in an effort to compile data and identify trends. This reduces labour costs, making it easier to move data directly from electronic sources to a secure electronic system of documentation, eliminates the need for manual work, and raises yield and market value.

B. Using data mining classification techniques to analyze soil data

In this study, [9][19] make the argument that soil is a crucial aspect of agriculture. Utilizing data mining classification techniques, the objective of the effort is to predict soil type. Methods/Analysis: To forecast the kind of soil, JRip, J48, and Naive Bayes data mining classification algorithms are utilized. The choice of crop is one of the most crucial aspects of agricultural planning. Farmers' losses are reduced when correct knowledge about the best crop for their field and season is available to them. The rate at which a crop produces depends on a variety of variables, including soil characteristics, such as soil moisture, and geographical characteristics, such as slope. These qualities are taken from various datasets, which are then gathered and examined. When creating a prediction model, choosing the appropriate data sources is crucial because it affects the model's accuracy. By using these classifier algorithms, information may be extracted from soil data and two categories of soil, Red and Black, are taken into consideration. Findings: An overview of data mining and agricultural data mining is given in this article. This data can yield more precise conclusions using the JRip model, and the forecast's Kappa Statistics improved. Application /Improvement: Effective approaches may be developed to address Big Data problems by improving the precision of categorization of enormous soil data sets via data mining.

C. Based on spatial temporal analysis of image time series, spiking neural networks for crop yield estimation

To make a recommendation [10] for first SNN computational model for estimating agricultural productivity using normalized difference index image time series is presented. It discusses the construction and evaluation of a methodological framework that trains an SNN to predict agricultural production in real time using spatially accumulated time series of data from a 250-m-resolution moderate resolution imaging spectrometer and historical crop yield data. This also looks at the optimum number of attributes needed to improve the results of our experimental data collecting. Reduction in the relative error and lower crop yield prediction accuracy are the Neural Network's main drawbacks. Similar to this, supervised learning algorithms failed to recognize the nonlinear relationship between input and output variables, which presented a challenge during the selection, grading, or sorting of fruits.

D. Data mining is used in a smart farming system

A "smart farming system" [11] should be employed to boost agricultural output by utilizing high-tech farming techniques. without the need for human participation. The report provides a summary of current smart farming software. The proposed system makes use of data mining methods and information acquired from satellite data, the Internet, and soil test results entered into the existing databases [12]. By monitoring crop growth stages, with proper water use, along with the selection of fertilizer to use according to crop stage, as well as the pesticide to use to protect crops from diseases and insect attack, it uses clustering algorithms in an elegant way to make decisions based on awareness of weather changes. By effectively controlling agricultural activities, this technology may raise field output. The largest user of water worldwide is agriculture, which uses around 70% of all water used for irrigation.

III. Existing Approach

Crop management, crop yield, and crop productivity output are just a few of the severe concerns that farmers face [13][14]. Data in the agricultural sector is growing daily. A system is required to have clear analyses of agricultural data

and to extract or use important information from the growing data. Data must first be learned in order to yield insights. One of the recent works [15][16] says the crop yield is predicted by using only one parameter like climate changes or soil type which cannot give accurate prediction It observed that at least 50% of the farm production gets wasted because of unawareness usage of fertilizer for the recommended crops.

IV. Proposed Approach

Machine learning (ML) approaches are currently being employed in many industries to provide useful and effective solutions. The agricultural production may be predicted using a variety of ML techniques [20] based on classification, clustering, and neural networks. In this work, a system that identifies weather quality, soil forecasts the best crop to cultivate using techniques like and Linear Regression SVR and ANFIS. This system, take weather, soil and agricultural data into account when developing the algorithm. Also, this method makes fertilizer recommendations based on the expected crop. The test results show that this method effectively forecasts crop selection and production, which is very helpful to farmers. ANFIS combines the learning capacity and relational structure of artificial neural networks with the fuzzy logic decision-making process. As with artificial neural networks, ANFIS implements learning with samples using a train data set. In this manner, the most ideal ANFIS structure for resolving the associated issue is obtained. The discovered structure is put through a testing process to evaluate how it affects samples that it has never seen before. The ANFIS model meets the requirements as seen by the lower error levels. The inability to explain the weight values obtained is one of the artificial neural network's most important drawbacks. The fuzzy inference mechanism that is part of the ANFIS framework eliminates this drawback.

V. Methodology

A. Input Data

Mining entails obtaining information from the data collection. It seeks to provide farmers with reliable findings. It takes significant data from the massive data gathering and extracts it. It is a step in the process of databases discovering knowledge. In addition to the KDD technique, Machine Learning has recently evolved as a way to manage large volumes of data and incorporates high speed computation.

A crop dataset classifies as:

- Season
- Crop Name
- Water Content
- Area
- Production (in tons)
- Crop Year

B. Dataset Preprocessing

Data preparation is a required configuration to improve the performance of the proposed model. Some irrelevant and less dependent attributes make up the data set. To make the most of the processing time, this system deletes several properties. This applies the classifier using some missing variables.

C. Feature Engineering

The same datasets cannot be used by the machine learning algorithm. To provide the appropriate input for the particular algorithm and increase the classifier accuracy, the dataset must be converted into a related algorithm. This system, employ several feature extraction methods. Class characteristics now have alphabetic values instead of their previous numerical values. This system creates performance class clusters that fall within the appropriate range.

D. Linear Regression Module

Using supervised learning, linear regression is a machine learning algorithm. Regression tends to make use of independent factors to model a goal posterior probability. Its primary goal is to render the correlation between factors and prediction more transparent. Regression models depend on the number of explanatory variables they incorporate as well as the nature of the correlations between the variables that are dependent and independent. Crop prediction makes use of linear regression to analyse and forecast the behaviour of a particular variable.

The objective of linear regression is to identify the most appropriate line that can forecast the outcome of the dependent variable (Y) based on the values of the independent variable(s) (X). The line with the right fit is the one that minimize the variance between the anticipated and actual values of Y. The primary objective is to make the correlation between factors and prediction more transparent. Both the number of explanatory factors included in regression models and the nature of the correlations between the dependent and independent variables are crucial. In order to assess and forecast the behaviour of a certain variable, linear regression is employed in crop prediction.

Both basic (one independent variable) and complex (more than one independent variable) regression models can be utilised using linear regression. It is frequently used to make predictions and examine correlations between variables in many different sectors, including economics, finance, marketing, and social sciences.

E. Support Vector Regression

In a space with high dimensions, the Support Vector Regression (SVR) regression method locates a hyper plane, also known as a decision boundary, that optimizes the difference between the predicted values and the actual ones. The support vectors that are closest to the hyper plane are used to create it. The following are the essential steps in how SVR operates:

- The data must first be split into groups to be trained and tested as the first step in the procedure. The model's performance is assessed using the set for testing after it has been taught using the training set.
- According to the properties of the data, a function called the kernel is selected. SVR frequently employs kernels with either linear or polynomial and radial basis functions. (RBF).

- The SVR model is trained using the training data collection. The ideal hyperplane that minimises the error between the predicted and actual numbers while capturing the training data is found by the model during training.
- The model's parameters are adjusted to enhance performance. Although the kernel parameter (gamma) governs the smoothness of the model, the regularization parameter (C) controls the trade-off between the complexity of the model and the amount of error permitted in the training data. Finally, the trained model is deployed to make predictions on new data.

F. Adaptive Neuro-Fuzzy Interference System (ANFIS)

ANFIS is a hybrid system that can conduct learning and inference tasks using both fuzzy logic and artificial neural networks (ANNs). Many applications, such as prediction, classification, and control, are possible using ANFIS models. The following steps can be used to breakdown how ANFIS functions:

- The input data are first fuzzified in ANFIS. The input variables are given membership functions to do this. The degree to which the input data are members of a specific fuzzy set is represented by the membership functions.
- The creation of fuzzy rules is the following phase. The fuzzy operators AND, OR are used to combine the fuzzy sets of the input variables to achieve this. A fuzzy rule is made up of any combination of fuzzy sets.
- Each fuzzy rule's output is calculated using a consequence function in this stage. The output variable is mapped by the consequence function from the input variables' fuzzy sets. Each rule produces a fuzzy value as its result.
- The output must be defuzzified as the last step. To do this, a defuzzification approach is used to combine the fuzzy values that were generated from the rule evaluation stage. The output of the ANFIS model is represented by the defuzzified value.

The gradient descent and least-squares approaches are combined to train the ANFIS model. The membership functions' and consequence functions' parameters are changed during training in order to reduce the discrepancy between expected and actual output values. The analysis was done utilising studies that dealt with soil management, water management, crop management, and livestock management. In order to select the ideal time for harvest and anticipate yield, this assesses the ripeness of the fruits. Various issues and approaches that are faced in the agricultural industry, particularly in the detection of illnesses, in the fields of image processing and machine learning

In order for ANFIS to function, the input data must first be fuzzified, followed by the generation of fuzzy rules, evaluation of the rules, defuzzification of the output, and training of the model to maximize performance.

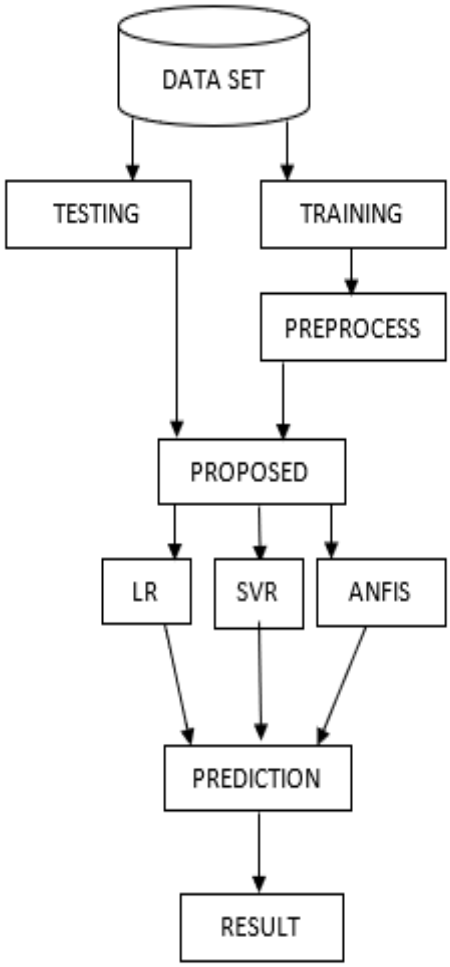

Fig. 1. Overall Flow Diagram

VI. RESULT ANALYSIS

A new trend towards heuristic-based ANFIS training algorithms for greater performance has been noticed. It appears to be proposed in derivative and heuristic-based hybrid algorithms at the same time. To assist researchers, the heuristic and hybrid methodologies used in ANFIS training are evaluated within the context of this work. According to various meteorological circumstances, ANFIS machine learning technique offer the best level of accuracy for crop production prediction. It comprises of how many decision trees are there. Each tree illustrates the dependencies of a single variable and combines all of the relationships in a proposed model.

TABLE I. CLASSIFICATION MODELS IN CROP RECOMMENDATIONS

Table 1: Outcomes of several classification models in crop recommendations		
Models	**Accuracy (%)**	**Execution Time (ms)**
LR	89.0	80
SVR	77.0	35
ANFIS	93.0	10

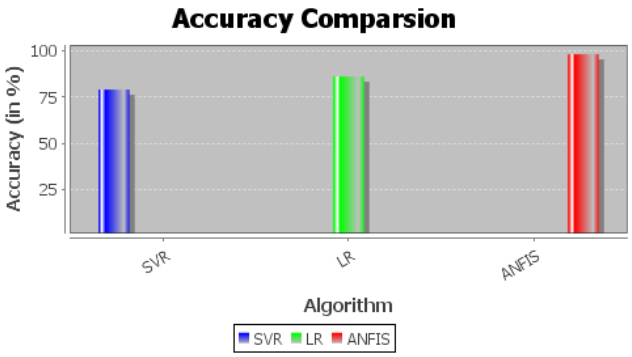

Fig. 2. Accuracy comparison of the classifier

VII. CONCLUSION

Research on crop management's significance is broad. Farmers require the aid of contemporary technologies to grow their crops. Accurate crop estimates may be quickly shared with agriculturalists. Techniques for machine learning have been applied to evaluate the parameters in agriculture. Farmers who consider variables like productivity and season may receive more individualised and pertinent advice, enabling them to produce a large quantity of goods. The ANFIS algorithm that offers the highest degree of accuracy in numerous comparisons of different algorithms like Linear Regression and Support Vector Regression.

From the standpoint of this research, incorporating image processing will be employed to find illnesses, which may enhance the model's accuracy.

VIII. REFERENCES

[1] Rajeswari and K. Arunesh, "Analyzing Soil Data Using Data Mining Classification Techniques," Indian Journal of Science and Technology, Vol. 9, 2016.

[2] A.Swarupa Rani, "The Impact of Data Analytics in Crop Management depending on Weather Conditions," International Journal of Engineering Technology Science and Research, Vol. 4, Issue no. 5, 2017.

[3] S. P. Raja, B. Sawicka, Z. Stamenkovic and G. Mariammal, "Crop Prediction Based on Characteristics of the Agricultural Environment Using Various Feature Selection Techniques and Classifiers," in IEEE Access, vol. 10, pp. 23625-23641, 2022, doi: 10.1109/ACCESS.2022.3154350.

[4] Lontsi Saadio Cedric, Wilfried Yves Hamilton Adoni, Rubby Aworka, Jérémie Thouakesseh Zoueu, Franck Kalala Mutombo, Moez Krichen and Charles Lebon Mberi Kimpolo, "Crops yield prediction based on machine learning models: Case of West African countries," Smart Agricultural Technology, Vol. 2, 2022.

[5] Anna Chlingaryana and Salah Sukkarieh, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture," Computers and Electronics in Agriculture in Elsevier, Vol. 151, pp: 61-69, 2018.

[6] Ahmed Awad , Luo Wan a and Mustafa El-Rawy, "Proper predictions of the water fate in agricultural lands: Indispensable condition for better crop water requirements estimate," Ain Shams Engineering Journal in Elsevier, Vol. 12, pp. 2435-2442, 2021.

[7] Rubby Aworkaa and Lontsi Saadio Cedric, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries," Smart Agriculture Technology Elsevier, Vol. 02, 2022. [Art.no: 100048]

[8] Attaluri, Soumya & Batcha, Nowshath & Raheem, Mafas "Crop Plantation Recommendation using Feature Extraction and Machine Learning Techniques," Vol. 04, pp. 1-4, 2020.

[9] Dhruvi Gosai1, Chintal Raval and Rikin Nayak "Crop Recommendation System using Machine Learning" International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol. 7, Issue No.03, 2021.

[10] M. Torres, R. Howitt and L. Rodrigues, "Analyzing rainfall effects on agricultural income: why timing matters," EconomiA, Vol.20, Issue No. 1, pp. 1–14, 2019.

[11] F. Abbas, H. Afzaal, A.A. Farooque and S. Tang, "Crop yield prediction through proximal sensing and machine learning algorithms," Agronomy, Vol. 10, Issue No. 7, 2020.

[12] S. Khaki, L. Wang, "Crop yield prediction using deep neural networks", Frontier Plant Science. Vol. 10, 2019.

[13] M. Shahhosseini, G. Hu, I. Huber and S. Archontoulis, "Coupling machine learning and crop modeling improves crop yield prediction" US corn belt, Sci. Rep. Vol. 11, 2021.

[14] D. Paudel, H. Boogaard, A. de Wit, S. Janssen, S. Osinga, C. Pylianidis and I.N. Athanasiadis, "Machine learning for large-scale crop yield forecasting," Agriculture Systems, Vol. 187, 2021.

[15] J.F. McEldowney, "Chapter 22 - climate change and the law", T.M. Letcher (Ed.), The Impacts of Climate Change, Elsevier, Vol. 2021, pp. 503–519, 2021.

[16] S. Aziz, M.M. Dowling, H. Hammami, A. Piepenbrink, "Machine Learning in Finance: A Topic Modeling Approach, SSRN, pp. 1-44, 2019.

[17] D.M. Camacho, K.M. Collins, R.K. Powers, J.C. Costello and J.J. Collins, "Next generation machine learning for biological networks," Cell Vol. 173, Issue No. 7, pp. 1581– 1592, 2018.

[18] X. Xu, P. Gao, X. Zhu, W. Guo, J. Ding, C. Li, M. Zhu and X. Wu, "Design of an integrated climatic assessment indicator (ICAI) for wheat production: a case study," Jiangsu Province, China, Ecol. Indic. Vol. 101 (2019), pp. 943–953, 2019.

[19] Anjana, Aishwarya Kedlaya K, Aysha Sana, B Apoorva Bhat, Sharath Kumar and Nagaraj Bhat, "An efficient algorithm for predicting crop using historical data and pattern matching technique," Global Transitions Proceedings, Vol. 2, Issue No. 2, Pages 294-298, 2021.

[20] Karaboga, D and Kaya, E, "Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey," Artificial Intelligence Review, Vol. 52, pp. 2263–2293, 2019.