ISSN 2063-5346

CROP RECOMMENDATION SYSTEM FOR IDENTIFYING THE BETTER YIELDING CROPS USING HYBRID EVOLUTIONARY ALGORITHM SS-NMRA

Dhanya K[1], B.Jayanthi[2]

Article History: Received: 01.02.2023 Revised: 07.03.2023 Accepted: 10.04.2023

Abstract

The Indian economy is based on agriculture. In India, agriculture is a significant industry. The nation's 1.3 billion inhabitants are fed by agriculture, which occupies more than 60% of its land. Compared to conventional agriculture methods, precision farming is a new strategy. In order to help farmers to forecast crop before deciding to final crop projecting the proper crop utilizing data such as district, rainfall, temperature, and area. The novel aspect of the proposed approach is that it gives farmers instructions on how to increase crop yield while also recommending the most profitable crop for a given area. The proposed approach would help to maximize production while reducing the challenges farmers experience when selecting a crop. Proposed a hybrid The SS-NMRA evolutionary method is used to find the best characteristics to fit in the models. It is advised to choose a suitable crop while employing RF and SS-NMRA. The suggested approach achieves accuracy of 98%.

Keyword:- Crop Recommendation, Feature Selection, Sparrow Search, Crop Yield, NIA.

[1]Research Scholor,Department of Computer Science,Rathnavel Subramanian College of Arts and Science,Sulur. Mail id: dhanyakmail@gmail.com

[2]HOD, Department of Computer Science,Rathnavel Subramanian College of Arts and Science,Sulur. Mail id: jayanthi@rvsgroup.com

DOI:10.31838/ecb/2023.12.s1-B.446

1. Introduction

India is one of the top five agriculturally producing nations in the world. To advance India's agricultural output, it is crucial to support the economic welfare of farmers. Numerous crops are produced by Indian agriculture, each with its own needs, expenses, and climatic requirements. Indian agriculture contributes to the country's rich cultural diversity in addition to providing a means of subsistence. Crop profitability differs when taking into account all the cost elements, regional diversity, and socioeconomic conditions in different Indian states. The demand for a crop, the weather prediction, and the cost of cultivation, which includes the cost of seeds, fertilizer, manpower, and machinery, all affect how profitable it is to produce. About 55% of India's population relies on agriculture as their main source of income, making India one of the leading players in the global agricultural industry. India is the world's largest producer of milk, pulses, and spices. It also has the largest herd of cattle (buffaloes), the largest area planted to wheat, rice, and cotton. It is the second-largest producer of wheat, rice, cotton, sugar, farmed fish, fruit, vegetables, tea, and farmed vegetables. About half of the population of India is employed in agriculture, which has the second-largest agricultural land area in the world. As a result, farmers play a crucial role in the industry that produces food for us.

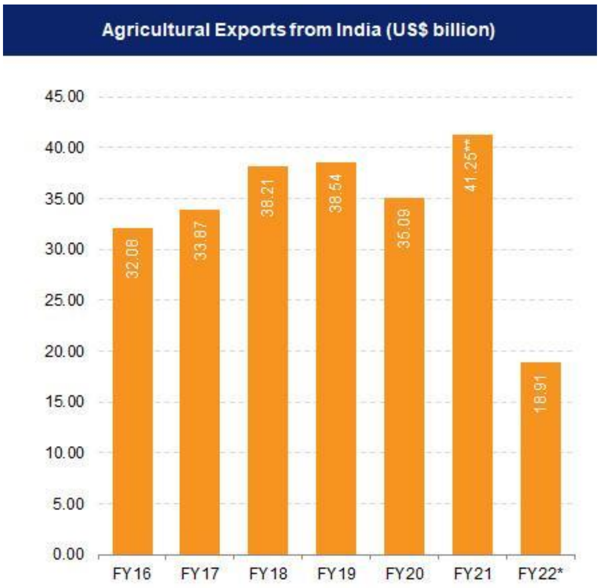

Figure: 1.1 Export Trends of Agriculture Products

By 2025, Inc42 projects that the Indian agricultural sector would grow to a value of US$24 billion as shown in fig.1.1. The sixth-largest food and grocery market in the world is in India, where 70% of sales are made through retail. According to First Advance Estimates for FY23 (Kharif alone), the nation's total production of food grain is predicted to be 149.92 million tonnes. India's rapid population growth is the primary force behind the industry's growth. This is further supported by the increasing income levels in rural and urban areas, which have boosted the demand for agricultural products across the country. Accordingly, the market is being encouraged by the increasing use of innovative technologies such as blockchain, artificial intelligence (AI), geographic information systems (GIS), drones, and remote sensing technologies, as well as the introduction of numerous e-farming applications. Due to increased investment in agricultural infrastructure, including irrigation systems, warehousing, and cold storage, the agriculture sector in India is anticipated to gain further pace in the coming years. Additionally, it's likely that the increased usage of genetically modified crops will increase Indian farmers' yields. The increase in the minimum support price and the concerted effort of scientists to obtain early maturing types of pulses are likely to make India self-sufficient in pulses within the next several years.

1.1. Indian Government Initiatives

The greatest exporter of agricultural goods worldwide is India. As a result of liberalization, it is one of the important growing themes in agricultural marketing. Due to deregulation, the number of agricultural exports is rising under WTO regulations, and this trend is anticipated to last for some time. Due to its favorable climatic conditions, low labour costs, low input unit costs, and low import material costs, India has a strong position in terms of farm exports. Exports of agricultural products are crucial to the growth of the industry. To analyze crops, digitize land records, and spray nutrients and insects, the Indian government plans to introduce Kisan Drones. A statewide network of 729 Krishi Vigyan Kendras has been established at the district level to ensure that newer technology, such as improved crop seeds, new breeds and strains of animals and fish, and improved production and protection measures, reach farmers. The Indian government has created the Digital Agriculture Mission for 2021–25 for agriculture initiatives based on cutting-edge technologies including artificial intelligence, blockchain, remote sensing and GIS technology, drones, robotics, and others. The Indian government would contribute Rs. 2,000 crore (US$ 306.29 million) for the computerization of the Primary Agricultural Credit Society (PACS) in order to ensure that cooperatives profit from digital technology. Additionally, it is broadening agricultural activities and expanding employment opportunities.

1.2. Problem

In Indian Agriculture, farm lands are used for cultivating multiple set of crops on varied seasons. But they find difficult to identify the crops best suitable for right time to increase their productivity and earnings. To overcome such situation crop recommendation systems build. Even though lots of recommendation system works now still there is a complication in determining the exact features for better productivity. Need to identify the better performing feature to the model is essential now.

1.3. Contribution

- Proposed a Hybrid SS-NMRA evolutionary model to identify the quality features for better yielding crops.
- Identify the features of outlier to normalize them for fit on the model.
- Handling the imbalanced features data thorough scaling and standardizing the feature sets.

1.4. Organization

The papers is divided into five sections: Section 2, which includes previous studies on models for crop yield and crop recommending, Section 3, which elaborates on the research work suggested, Section 4, which compares the results of the models, and Section 5, which wraps up the paper.

2. Literature Review

Agricultural information is derived from a variety of sources, including weather, soil, crop characteristics, and more. The performance of a machine learning model is improved by the use of feature selection algorithms to find significant features, remove irrelevant features, and identify features that matter. In (J & M, 2022), filter and wrapper approaches are combined to suggest a hybrid feature selection strategy. In order to create a crop recommendation model with greater accuracy and performance, the suggested method extracts the best features from soil characteristics, crop characteristics, and climatic parameters. All of the features in the dataset and the features derived from the suggested approach are taken into account when evaluating the model's efficacy. The evaluation metrics MSE, RMSE, MAE, and R 2 are used to validate the suggested feature selection approach. The chosen features are used to create the machine learning models, such as artificial neural networks and decision trees.

In (Aditya Shastry & H.A., 2021) to select the best features that affected the goal attribute (in this example, crop yield), a fitness function based on "Mutual Information" (MutInf) and "Root Mean Square Error" (RtMSE) was devised. A feature extraction process utilising "weighted principal component analysis" (wgt-PCA) was then used to these chosen features. Following feature extraction, the features were fed into a variety of machine learning models, including "Regression" (Reg), "Artificial Neural Networks" (ArtNN), "Adaptive Neuro Fuzzy Inference System" (ANFIS), "Ensemble of Trees" (EnT), and "Support Vector Regression" (SuVR). Trials on 3 benchmark and 8 real-world farming datasets showed that the developed hybrid feature selection and extraction technique outperformed FS and FeExt methods like Correlation Analysis (CA), Singular Valued Decomposition (SiVD), Genetic Algorithm (GA), and wgt-PCA on "benchmark" and "real-world" farming datasets with respect to Rsq2, RtMSE, and "mean absolute error" (MAE).

Adaptive production systems in agriculture are required by climate variability, frequently necessitating substantial irreversible expenditures. (West, 2019) evaluate uncertainty regarding climate effects and the timing of investments for horticultural operations using a multi-objective evolutionary algorithm, and we derive the ideal periods to adapt using cultivar replacement approaches. We discover that choosing an investment based on conventional valuation techniques can lead to suboptimal outcomes and bad choices, potentially compromising adaptation efforts. We also demonstrate how timing the investment decision incorrectly might cause enterprise value to be negatively impacted by conflicting economic and climatic variables. Using a vector-based geographic information system on a farm where different orchard sections are susceptible to different rates of production, deterioration, and age, the use of the genetic algorithm solver is presented.

When deciding which crop to grow, farm managers must balance a number of competing goals. When determining yield potential, soil characteristics are crucial. Liming and fertilization are frequently employed to modify soils so that they meet the nutritional needs of the crops that will be grown there. An intriguing approach to minimizing the need for soil treatment, saving costs and potential environmental harms, is planting the crop that will best fit the soil features. Farmers typically search for investments that offer the highest potential returns with the fewest dangers. The crop-selection problem may be challenging to address with conventional tools, taking into account the objectives to be taken into account. In order to assist in the selection of an acceptable cultivation plan while simultaneously taking into account five crop options and five objectives, (Brunelli & Von Lücken, 2009) provides an approach based on multi-objective evolutionary algorithms.

(Whitmire et al., 2021) main contribution is the proof that ML, with feature selection, shows promise in crop yield prediction even on small datasets with a few features, and that reporting accuracy in R and R2 offers an appealing approach to compare results across different crops. Almost everyone's everyday lives depend on accurately predicting lucerne biomass and crop yield for livestock feed, and many characteristics of data from this domain, together with matching weather data, can be used to train machine learning models for yield prediction. Cross validation was used in the development of nearest neighbours, regression trees, support vector machines, neural networks, and Bayesian regression. A correlation-based method, the

ReliefF method, and a wrapper method were the feature selection techniques that were contrasted. We discovered that the correlation-based approach was the most accurate, and the features it identified were the Julian day of harvest, the number of days between the dates of sowing and harvest, the total solar radiation since the last harvest, and the total rainfall since the previous harvest.

3. Research Methodology

Here, proposed a hybrid Feature selection model for recommending crops to get better yields. The Proposed model is shown in fig. 3.1.

Final Dataset

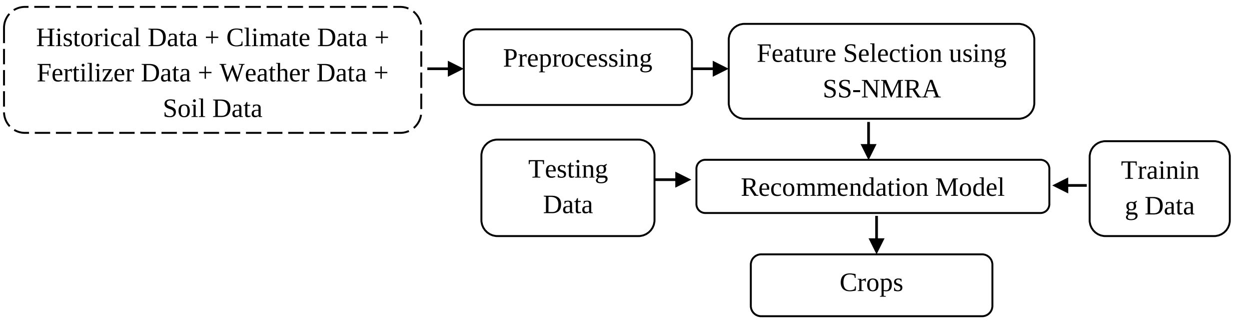

Figure: 3.1 Proposed Model

3.1. Dataset

Dataset collected from the available online sources such as data.world portal and http://data.icrisat.org/. Data on 20 important crops, including cereals, pulses, oilseeds, cotton, sugarcane, total fruits and vegetables, are included in the area, production yield dataset. Data has been gathered for crop and production by season, as well as a breakdown of fruits and vegetables by type. Data on soil, climate, and fertilizer were also included. Here 80% data used for training and 20% used for testing the model.

3.2. Preprocessing

The data collected from different sources and they are of different type. First and foremost the data has to be cleaned. Raw data may consist of null, missing values which will bring down accuracy of prediction models. Checked whether any NaN, Null data available on the dataset. No Null values found. Tried to identify the unique values of each features and found few features having unique entries upto 2.2k.

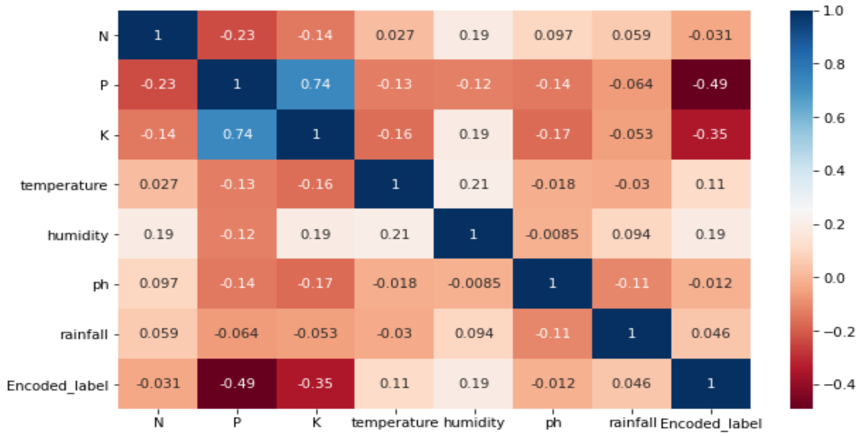

Figure: 3.2 Correlations on Features

Label Encoding performed on the targeted variables there are 22 classes found. LabelEncoder used for label the target output. Dropped duplicate values from dataset. Checked Correlation between the features and found that (as given in fig.3.2), the variable 'K' and 'P' have high positive correlation. Also identified outliers (as shown in fig. 3.3) on few features, those are normalized using Mean Normalization method.

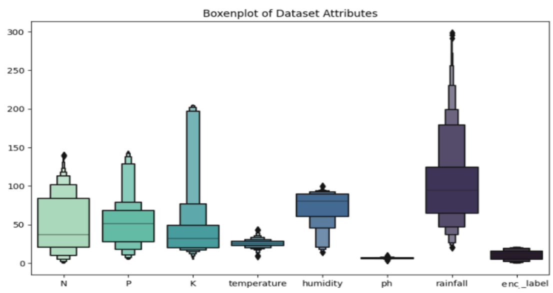

Figure: 3.3 Outlier Detection

3.3. Feature Selection using SS-NMAR

Identify the best performing Features is the essential section for a model. Using evolutionary algorithm can get ret of the issue of identifying the best performing feature set for the models. Here for the recommendation model there is a need to pick the optimal subset of features and simplify the model to speed up training so used new nature inspired algorithm Naked Mole Rat Algorithm with Sparrow Search.

In comparison to conventional heuristic search methods, Sparrow Search (SS) offers a quick convergence speed, powerful optimization capability, and a wider range of application scenarios. It nevertheless has certain flaws with existing swarm intelligence optimization algorithms, such as uneven initial population distribution, poor iteration-to-iteration convergence, propensity for early stagnation, and ease of falling into local optimums. Naked Mole-Rat algorithm (NMRA) has the ability to avoid local minima, has a rapid convergence rate, and finds the global minima with a high degree of consistency. The main cause of this is because these algorithms have the added benefits of quicker convergence, fewer possibilities of local optima stagnation, high difficulty, and considerable contribution compared to their competitors. The limitation of SS was overcome by NMRA [10].

Algorithm: SS-NMAR

Step 1: SS_{Pop}: Initialize, A random population matrix p of size n x m is produced for population sizes n and m and the number of features m. Population initialises the SS.

Step 2: SS_{fit}: Initialize, The best-fitting chromosomes are chosen to replicate the following set of solutions. For the following step, chromosomes with a higher fitness score or those chosen at random are chosen.

Step 3: In the crossover step, bits are switched inside the chosen parents to create children with different solutions. Calculated using, where parameter: 0k1 is a crossover probability. Number of bits b chosen from parent based optimal set.

Step 4: According to the "survival of fittest" theory, only those chromosomes that are particularly fit will be passed down to the next generation. Fitness as determined by SS

Step 5: Save the feature list.

3.4. Recommendation Model

A form of decision support system called a crop recommendation system aids farmers or other agricultural professionals in making data-driven decisions about which crops to grow based on elements like soil quality, weather, and market demand [7]. Major literatures used the base models of regression algorithms such as Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Support Vector Machine (SVM) [9]. Further the recommendation system developed with base models to understand the performance of the proposed hybrid SS-NMRA algorithm. Without any enhancement the models applied directly with SS-NMRA feature selection.

4. Result and Analysis

Proposed models performance is evaluated using precision, accuracy, f1-score and recall. The model which scores good accuracy applied for prediction. To compare the performance PSO is used as Base Feature Selection (FS) method.

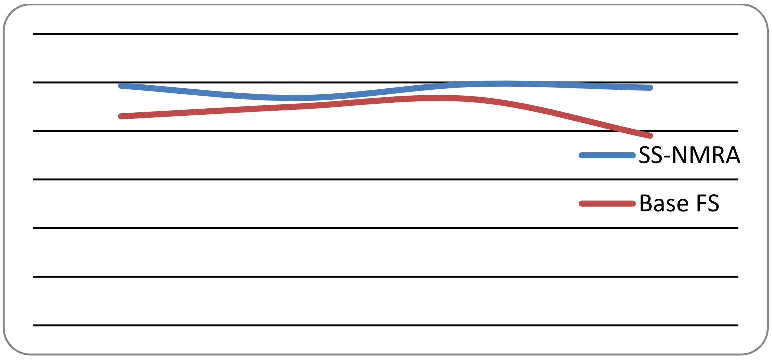

Figure: 4.1 F1-Score

Fig: 4.1 shows that compare to other base feature selection algorithms, the proposed hybrid SS-NMRA algorithm performs better with base models. Particularly with Random forest yields a high f1-score of 0.93.

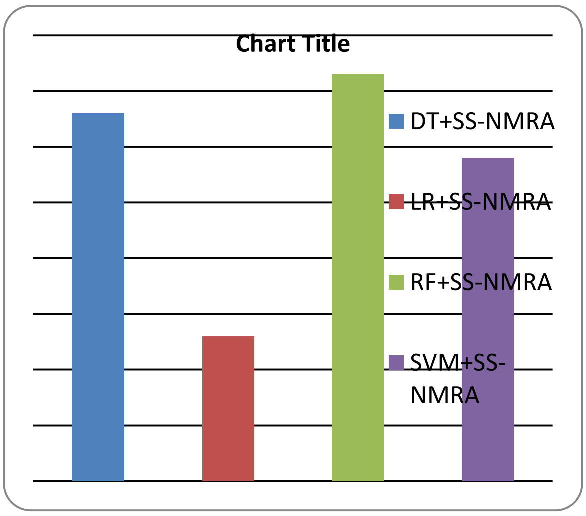

Figure: 4.2 Accuracy

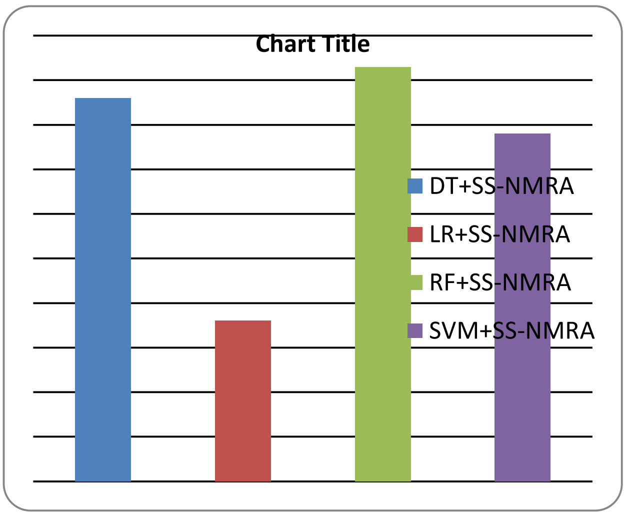

Figure: 4.3 Recall

Fig: 4.2 shows that compare to other base feature selection algorithms, the proposed hybrid SS-NMRA algorithm performs better with base models. Particularly with Random forest yields a high accuracy of 0.99.

Fig: 4.3 shows that compare to other base feature selection algorithms, the proposed hybrid SS-NMRA algorithm performs better with base models. Particularly with Random forest yields a high recall of 0.98.

RF+SS-NMRA used further for the crop recommendation as derived best suitable crops both on the single and combinational crops (as shown in fig: 4.4). The model will deliver result which gives a better suggestion to farmers to choice of crops even they opt for more than one crops mean those combinational crops.

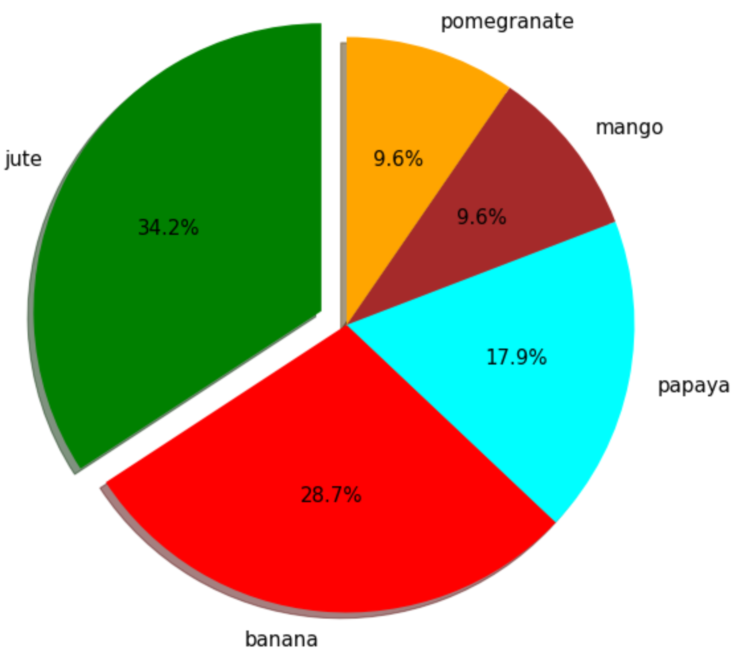

Figure: 4.4 Crop Recommendations

5. Conclusion

Prior to choosing to cultivate their ultimate crop, the farmer can anticipate the crop yield by selecting the proper crop utilizing criteria such as district, rainfall, temperature, and area. Proposed approach can help the farmer and give them important insights. In this research, the feature selection and classifier algorithms are inspired by natural processes. To improve the accuracy of these models merged various datasets found available online such as from government websites and competition forums. The proposed hybrid SS-NMRA algorithm score high f1-score 0.93 compare other base Feature selection models when applied with base classifiers. With that attained a result for the recommendation system with single and multi crop suggestions to attain better yield and profit.

Reference:-

1. J, M. and M, I. (2022) "Hybrid filter and wrapper methods based feature selection for crop recommendation," *2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC)* [Preprint]. Available at: https://doi.org/10.1109/icesic53714.2022.9783542.

2. Aditya Shastry, K. and H.A., S. (2021) "A modified genetic algorithm and weighted principal component analysis based feature selection and extraction strategy in Agriculture," Knowledge-Based Systems, 232, p. 107460. Available at: https://doi.org/10.1016/j.knosys.2021.107460.
3. West, J. (2019) "Multi-criteria evolutionary algorithm optimization for horticulture crop management," Agricultural Systems, 173, pp. 469–481. Available at: https://doi.org/10.1016/j.agsy.2019.03.016.
4. Raja, S.P. et al. (2022) "Crop prediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers," IEEE Access, 10, pp. 23625–23641. Available at: https://doi.org/10.1109/access.2022.3154350.
5. Brunelli, R. and Von Lücken, C. (2009) "Optimal crops selection using multiobjective evolutionary algorithms," AI Magazine, 30(2), p. 96. Available at: https://doi.org/10.1609/aimag.v30i2.2212.
6. Whitmire, C.D. et al. (2021) "Using machine learning and feature selection for alfalfa yield prediction," AI, 2(1), pp. 71–88. Available at: https://doi.org/10.3390/ai2010006.
7. Prasad Mane et al. (2023) "Crop recommendation using support vector machine (SVM) classifier," International Journal of Advanced Research in Science, Communication and Technology, pp. 505–508. Available at: https://doi.org/10.48175/ijarsct-8875.
8. Modi, D. et al. (2021) "Crop recommendation using machine learning algorithm," 2021 5th International Conference on Information Systems and Computer Networks (ISCON) [Preprint]. Available at: https://doi.org/10.1109/iscon52037.2021.9702392.
9. Gambhir, M.S. et al. (2023) "Crop recommendation system using machine learning," International Journal for Research in Applied Science and Engineering Technology, 11(2), pp. 640–644. Available at: https://doi.org/10.22214/ijraset.2023.49097.
10. Shilaskar, S. et al. (2022) "Artificial intelligence based crop recommendation and Plant Leaf Disease Detection System," 2022 3rd International Conference for Emerging Technology (INCET) [Preprint]. Available at: https://doi.org/10.1109/incet54531.2022.9824002.