2021 5th International Conference on Computing Methodologies and Communication (ICCMC) | 978-1-6654-0360-3/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICCMC51019.2021.9418375

Intelligent Crop Recommendation System using Machine Learning

Priyadharshini A
Assistant Professor,
Information Science and Engineering
CMR Institute of Technology
Bengaluru, India
priyadharshini.a@cmrit.ac.in

Swapneel Chakraborty
B.E, Information Science and Engineering
CMR Institute of Technology
Bengaluru, India
swapneelchakraborty43@gmail.com

Aayush Kumar
B.E, Information Science and Engineering
CMR Institute of Technology
Bengaluru, India
aayushkumar2899@gmail.com

Omen Rajendra Pooniwala
B.E, Information Science and Engineering
CMR Institute of Technology
Bengaluru, India
1295omen@gmail.com

Abstract- Agriculture plays a vital role in the socio-economic fabric of India. Failure of farmers to decide on the best-suited crop for the land using traditional and non-scientific methods is a serious issue for a country where approximately 58 percent of the population is involved in farming. Sometimes farmers were failed to choose the right crops based on the soil conditions, sowing season, and geographical location. This results in suicide, quitting the agriculture field, moving towards urban areas for livelihood. To overcome this issue, this research work has proposed a system to assist the farmers in crop selection by considering all the factors like sowing season, soil, and geographical location. Furthermore, precision agriculture is being implemented with a modern agricultural technology and it is evolving in developing countries that concentrates on site-specific crop management.

Keywords- Agriculture, Crop Recommendation, Machine Learning

I. INTRODUCTION

Agriculture is one of the major sources of livelihood for about 58% of our nation's population [14] [15]. As per the 2016-17, Economic survey the average monthly income of a farmer in 17 states is Rs.1700/- which results in farmer suicides, diversion of agricultural land for non-agricultural purpose. Besides, 48% of farmers don't want their next generation to take care of their agriculture instead want to settle down in urban areas. The reason behind this is that the farmers often take wrong decision about the crop selection [9] for example selecting a crop that won't give much yield for the particular soil, planting in the wrong season, and so on. The farmer might have purchased the land from others so without previous experience the decision might have been taken. Wrong crop selection will always result in less yield. If the family is fully dependent on this income then it's very difficult to survive.

Both availability and accessibility of correct and up to date information hinder potential researchers from working on developing country case studies. With resources in our reach, a system has been proposed to address this problem by providing predictive insights on crop sustainability and recommendations based on machine learning models trained considering essential environmental and economic parameters.

In the proposed system the environmental parameters such as rainfall, temperature, and geographical location in terms of the state along with soil characteristics such as soil type, pH value, and nutrients concentration are being considered to recommend a suitable crop to the user. In addition to this, if the right crop is selected by the farmer then they will get the prediction about the yield also. The objective is to, 1. Build a robust model to give a correct and accurate prediction of crop sustainability in a given state for the particular soil type and climatic conditions. 2. Provide recommendation of the best suitable crops in the area so that the farmer does not incur any loss3. Provide profit analysis of various crops based on the previous year's data.

The proposed system is implemented using machine learning which is one of the applications of Artificial Intelligence that allows the systems to learn and evolve automatically without explicitly programmed by a programmer. Followed by that, the accuracy of the program will be improved without human intervention. Many researchers are researching this field to assist the farmers in the selection discussed as follows, to choose a suitable crop with its various factors like physical, environmental, and

economic factors were taken into consideration. Artificial Neural Network is taken into consideration to choose the crop with the high yield rate [1], before cultivation, the crops were ranked based on Decision Tree Learning-ID3 (Iterative Dichotomiser 3) and K-Nearest Neighbors Regression algorithms [9]. Crop features were analysed based on the random forest algorithm and BigML [10]. Machine learning algorithms were implemented to prevent the impacts from the water stress in plants and have given a set of decision rules used in plant's state prediction [11]. Machine learning techniques were used to predict the cost of crops and smart systems were used to provide real-time suggestions [5]. In this work [8], a survey has been made on several applications of machine learning algorithms in agricultural production systems. Further AI-enabled systems were used to provide recommendations concerning crop management. Deep learning techniques can be used to yield better in crop cultivation [12] [19]. In this paper [2] real-time monthly weather is taken into consideration to design an efficient yield forecasting mechanism. A non-parametric statistical model along with nonparametric regression methods was being used to implement the above-said forecasting mechanism.

The farmers are being assisted towards crop selection based on soil characteristics, particular geographical location, sowing season, and environmental factors, by combining machine learning and data mining techniques [3]. The regression technique is used to analyse soil dataset [4]. In this paper [6], five different algorithms were used to suggest the crops for the underlying soil series. They are as follows, Support Vector Machine, Bagged Tree, Adaboost, Naive Bayes, and Artificial Neural Network. Further, the ensemble method is included to provide more accurate results. In another work [7], Precision agriculture is being discussed which suggests the right crops based on site-specific parameters, the precision agriculture is also used to detect pests in the coconut tress using drone [18]. The proposed recommendation system has used an ensemble model with a majority voting technique using CHAID, K-Nearest Neighbour, Naïve Bayes, and Random tree as learners.

II. PROPOSED MECHANISM

In our proposed research, both environmental and soil parameters are taken into account carefully. The reason behind this is a particular type of soil will support a crop whereas the weather conditions won't support that, such that the yield will suffer. The overall working of the proposed system is depicted in Fig. 1.

A. DATA ANALYSIS

This is an attempt to find the presence of any relationships between the various attributes present in the dataset.

Acquisition of Training Dataset:

Various datasets from government website [16] and Kaggle [17] are fed into the system. Datasets include i) Yield Dataset: This dataset contains yield for 16 major crops grown across all the states in kg per hectare. A yield of 0 indicates that the crop is not cultivated in the respective state. ii) Cost of Cultivation dataset: This data set provides the cost of cultivation for each crop in Rs. per hectare. iii) Modal price of crops: This dataset gives the average market prices for those crops over a period of two months. iv) Standard price of crops: This dataset gives the current market price of the crops in Rs per hectare. v) Soil nutrient content dataset: This dataset has five columns with the attributes in the order-State, Nitrogen content, Phosphorous content, Potassium content, and average ph. vi) Rainfall Temperature dataset: This dataset contains crops, max, and min rainfall, max and min temperature, max and min rainfall, and ph values.

Profit analysis is performed using the cost of cultivation, market price, standard price, and yield dataset. This is being performed as a first step to know how much impact profit can have on crop prediction. The profit is calculated for each crop grown in the state and assigns a -1 value for the states with 0 or no production of the given crop.

B DATA PREPROCESSING

This step includes replacing the null and 0 values for yield by -1 so that it does not affect the overall prediction. Further, the data-set has to be encoded so that it could be fed into the neural network. Data pre-processing is an important step as it helps in cleaning the data and making it suitable for use in machine learning algorithms. Most of the focus in pre-processing is to remove any outliers or erroneous data, as well as handling any missing values. The values in the dataset are in string format. To pass the input to the neural network, this should be converted into integer values. Further to reduce the amount of data going into the linear regression model the crops are being filtered based on the required nutrients and nutrients present in the soil. If the nutrient content of the soil is below that required by the crops, then the respective crop will be discarded, in this, the training time has been reduced a lot.

C TRAINING MODEL AND CROP RECOMMENDATION

After the pre-processing step, the data-set is used to train different machine learning models like neural network and linear regression to attain accuracy as high as possible.

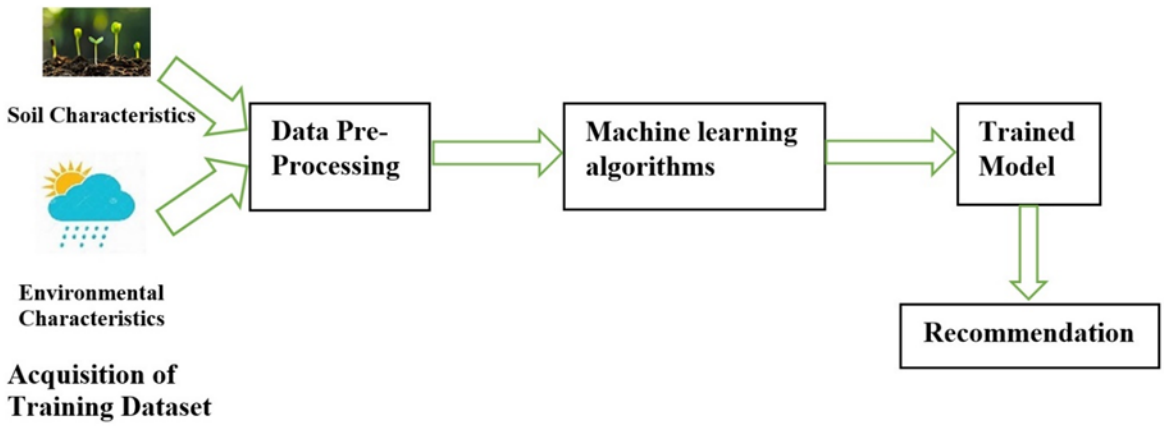

Fig 1: Crop Recommendation System

LINEAR REGRESSION

Linear regression fits a straight-line between rainfall, temperature, pH, and production which would return a y-pred value for each crop.

In the end, the crops are sorted based on the y-orvalue returned by the linear regression model using quick sort giving the crop with the best score first in the list.

NEURAL NETWORK

Our implementation of the neural network is facilitated with the help of the Keras module. A sequential model is implemented with 3 input layers and 15 output layers which gives the sustainability of each 15 crops given the input in terms of state, month, and soil.

III RESULTS AND PERFORMANCE ANALYSIS

The proposed system is implemented along with two algorithms namely, Linear regression and Neural network using Pandas, Numpy, Tensorflow, Keras and Sickitlearn libraries, tools, and Python as the programming language. Whereas both the algorithms are based on supervised learning. The results are analysed and are being compared with crop recommendation using K Nearest Neighbour, K Nearest Neighbour with K Nearest Neighbour with cross validation, Decision Tree, Naive Bayes and Support Vector Machine [13] in terms of accuracy.

The overall system is divided into three modules:

i) Profit analysis
ii) Crop recommender
iii) Crop Sustainability predictor

A PROFIT ANALYSIS

As shown in Fig.2 after the analysis we get the profit data for each crop grown in all the states. This provides a clear insight on which crop to be selected.

```
In [15]: profit_data
```

Out[15]:

	state	crop	profit
0	Andhra Pradesh	Rice	6.385184e+04
1	Andhra Pradesh	Jowar	1.097407e+04
2	Andhra Pradesh	Bajra	7.414478e+03
3	Andhra Pradesh	Maize	3.136984e+04
4	Andhra Pradesh	Ragi	5.636376e+03
5	Andhra Pradesh	Wheat	1.000000e+00
6	Andhra Pradesh	Barley	-1.000000e+00
7	Andhra Pradesh	Gram	5.058972e+03
8	Andhra Pradesh	Tur	1.000000e+00
9	Andhra Pradesh	Groundnut	1.017747e+04
10	Andhra Pradesh	Mustard	1.000000e+00
11	Andhra Pradesh	Soyabean	7.632153e+03
12	Andhra Pradesh	Sunflower	9.739718e+03
13	Andhra Pradesh	Cotton	1.000000e+00
14	Andhra Pradesh	Jute	-1.000000e+00
15	Andhra Pradesh	Mesta	1.609931e+03
16	Andhra Pradesh	Sugarcane	8.931305e+05
17	Arunachal Pradesh	Rice	9.995866e+02
18	Arunachal Pradesh	Jowar	-1.000000e+00

Fig. 2: Profit on crops per state

B CROP RECOMMENDER

The Crop recommendation model enlists the crops in an order where the first crop will have the highest productivity followed by the remaining in the list as depicted in Fig.4. Regression model output and input to the predictor is depicted in Fig. 3 and Fig.5.

C CROP SUSTAINABILITY PREDICTOR

The prediction value of sustainability for each crop has been observed given the three inputs. Hence from these values, one can get a clear idea of which crop will give better yield and is depicted in Fig. 6. The accuracy of proposed system and various machine learning algorithms in terms of crop prediction is given

in table 1 and the comparison is depicted in Fig. 7 respectively.

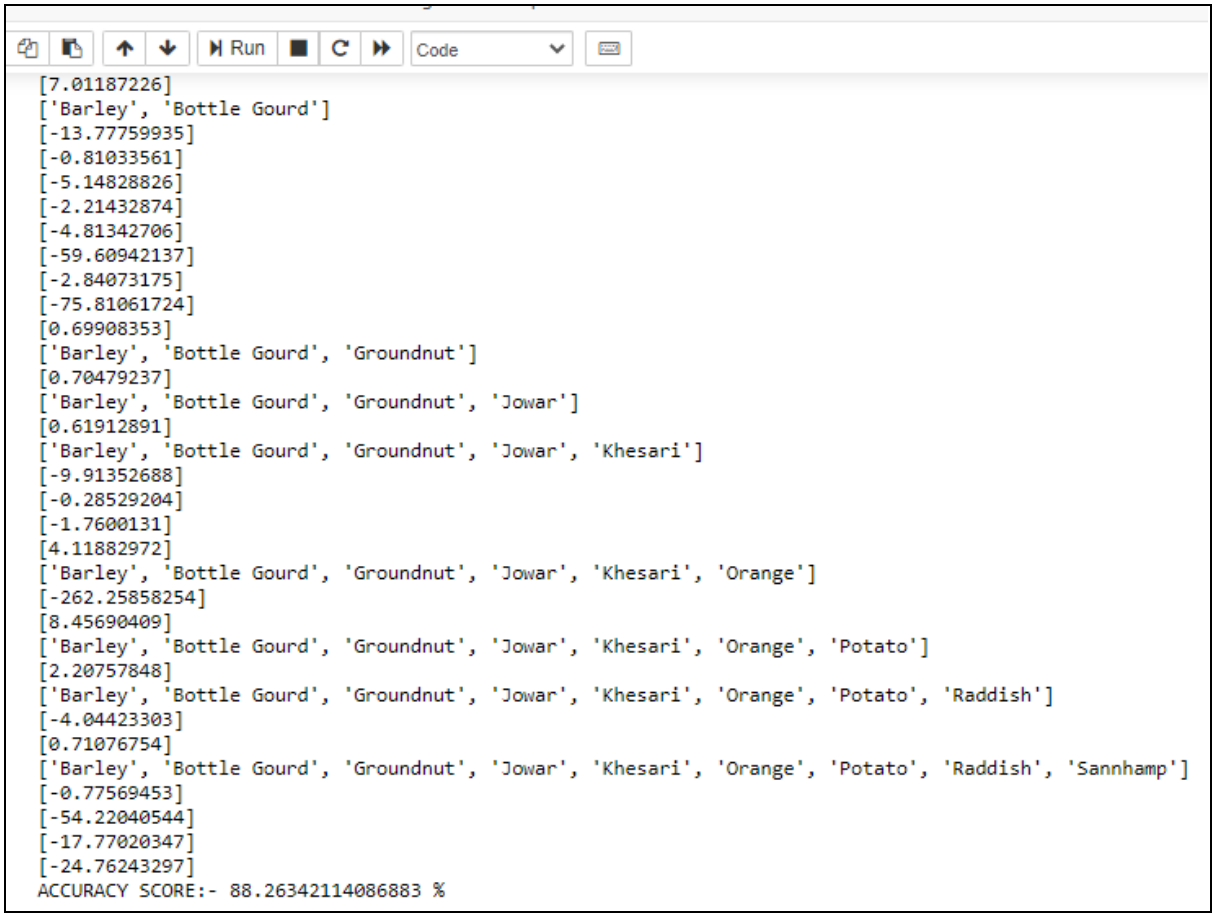

Fig.3: Regression model output

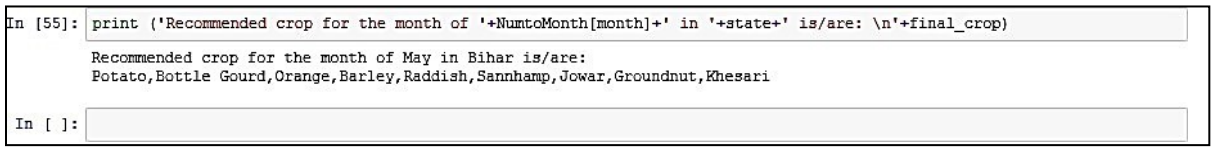

Fig. 4: Crop recommendation

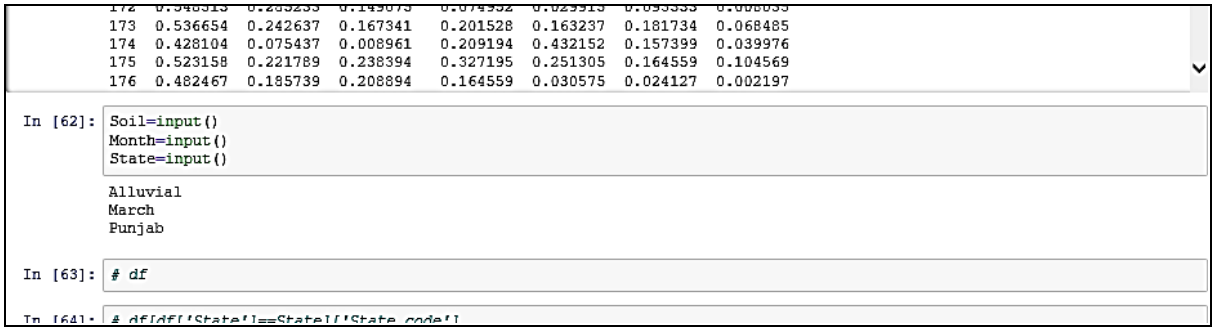

Fig.5: Input for the predictor

```
In [68]: pred = model.predict_proba(Choices)
         df2 = pd.DataFrame(pred, columns=["Rice","Wheat","Cotton","Sugarcane","Tea","Coffee","Cashew","Rubber","Coconut","O
         print(df2)
         df2.shape

              Rice    Wheat    Cotton  Sugarcane       Tea    Coffee    Cashew  \
         0  0.495499  0.2027  0.236318   0.184892  0.043014  0.035415  0.004605

              Rubber   Coconut  Oilseed      Ragi     Maize  Groundnut    Millet  \
         0  0.002973  0.005234  0.06702  0.087935  0.13789   0.075435  0.125787

              Barley
         0  0.043136

Out[68]: (1, 15)
```

Fig.6: Crop Sustainability prediction values

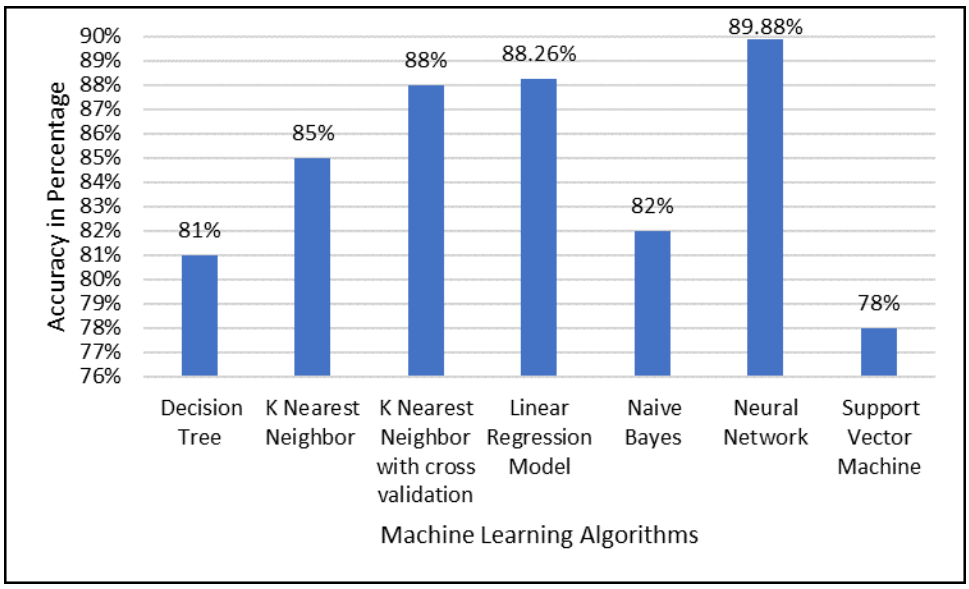

Fig. 7: Comparison of proposed system with other machine learning algorithms

TABLE 1: CROP RECOMMENDATION ACCURACY

S.No	Algorithms	Accuracy
1	Decision Tree	81%
2	K Nearest Neighbour	85%
3	K Nearest Neighbour with cross validation	88%
4	Linear Regression Model	88.26%
5	Naive Bayes	82%
6	Neural Network	89.88%
7	Support Vector Machine	78%

IV CONCLUSION

The proposed system helps the farmers to choose the right crop by providing insights that ordinary farmers don't keep track of thereby decreasing the chances of crop failure and increasing productivity. It also prevents them from incurring losses. In the future, it has been planned to incorporate a web interface as well as a mobile app to provide the recommendations of crop cultivation to the farmers can be accessed by millions of farmers across the country.

V REFERENCES

[1] R. Kumar, M. P. Singh, P. Kumar, and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique", 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy, and Materials (ICSTM), Chennai, 2015, pp. 138-145, DOI: 10.1109/ICSTM.2015.7225403.

[2] H. Lee and A. Moon, "Development of yield prediction system based on real-time agricultural meteorological information", 16th International Conference on Advanced Communication Technology, Pyeongchang, 2014, pp. 1292-1295, DOI: 10.1109/ICACT.2014.6779168.

[3] T.R. Lekhaa, "Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques", International Journal of Modern Trends in Engineering and Science (IJMTES), Volume 03, Issue10, 2016.

[4] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, "Soil Data Analysis Using Classification

Techniques and Soil Attribute Prediction", International Journal of ComputerScience Issues, Volume 9, Issue 3,2012.
[5] S. R. Rajeswari, ParthKhunteta, Subham Kumar, Amrit Raj Singh, Vaibhav Pandey, "Smart Farming Prediction using Machine Learning", International Journal of Innovative Technology and Exploring Engineering, 2019, Volume-08, Issue-07.
[6] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms", Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune,India, 2018, pp. 1-6, DOI: 10.1109/ICCUBEA.2018.8697349.
[7] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika and J. Nisha, "Crop recommendation system for precision agriculture," Eighth International Conference on Advanced Computing (ICoAC), 2017, pp. 32-36, DOI: 10.1109/ICoAC.2017.7951740.
[8] Konstantinos G. Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson and Dionysis Bochtis, "Machine Learning in Agriculture: A Review", Article on sensors, 2018, pp 1-29, doi:10.3390/s18082674.
[9] M. T. Shakoor, K. Rahman, S. N. Rayta and A. Chakrabarty, "Agricultural production output prediction using Supervised Machine Learning techniques", 1st International Conference on Next Generation Computing Applications (NextComp), Mauritius, 2017, pp. 182-187, DOI: 10.1109/NEXTCOMP.2017.8016196.
[10] Marion Olubunmi Adebiyi, Roseline Oluwaseun Ogundokun and Aneoghena Amarachi Abokhai, "Machine Learning-Based Predictive Farmland Optimization and Crop Monitoring System", open-access article distributed Volume 2020, pp 1-12, https://doi.org/10.1155/2020/9428281
[11] Dimitriadis Savvas and Christos Goumopoulos. "Applying machine learning to extract new knowledge in precision agriculture applications", 2008 Panhellenic Conference on Informatics, IEEE, (2008), pp:100-104.
[12] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data", IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 5, pp. 778-782, May 2017, DOI: 10.1109/LGRS.2017.2681128.
[13] Kevin Tom Thomas, Varsha S, Merin Mary Saji, Lisha Varghese, Er. Jinu Thomas, "Crop Prediction Using Machine Learning", International Journal of Future Generation Communication and Networking Vol. 13, No. 3, (2020), pp. 1896–1901.
[14] www.india.gov.in/topics/agriculture
[15] www.ibef.org/industry/agriculture-india.aspx
[16] https://data.gov.in/g
[17] https://www.kaggle.com/notebook
[18] Chandy, Abraham. "Pest Infestation Identification in Coconut Trees Using Deep Learning." Journal of Artificial Intelligence 1, no. 01 (2019): 10-18.
[19] Vijayakumar, T., and Mr R. Vinothkanna. "Mellowness detection of dragon fruit using deep learning strategy." Journal of Innovative Image Processing (JIIP) 2, no. 01 (2020): 35-43.