Crop Guidance and Farmer's Friend – Smart Farming using Machine Learning

Tanvi Daware[1,*], *Pratiksha* Ramteke[1,**], *Uzma* Shaikh[1,***] and *Smita* Bharne[1,****]

[1]Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Navi Mumbai, India

Abstract: Agriculture is the centre of our country's social and economic development. But the majority of the farmers are still struggling for better agricultural facilities. The government is also helping farmers with their benefits with different government policies. Today's farmers are well aware of using the internet and smartphone for smart farming. There are expanding quantities of instructed individuals in the farming field and they began utilising cell phones in rustic areas. Thus, the main aim of this Crop Guidance and Farmer's Friend application is to assist the farmers with expanding their yield as well as to be aware of the current agriculture-based data used for smart farming. The proposed system is designed to track weather conditions at a specific location, such as temperature, humidity, and changes in the environment. This application diminishes the time and efforts of farmers and assists them with getting the day-to-day market cost for various harvests, fertilisers and vegetables without visiting the market. The data (value, climate, most recent rural technique) will be shipped off to the farmers, through this application. This application helps farmers a great deal and keeps them updated. It offers a livelihood for almost all of the population, contributing to countrywide income and gainful employment.

1 Introduction

Agriculture is the backbone of our Indian economy, which plays an important role in increasing our country's economic development. The main focus of this app is on the agricultural society because the farmers have the highest contribution to our country's GDP.[1] But the farmer still does not get the price for the harvest. This mainly happens due to improper watering or incorrect plant selection, or sometimes the crop yield is lower than expected. By analysing the soil and atmosphere in a given region, the best crop will produce a higher yield and the net yield of the crop can be predicted using Water level, distance depth, and soil ph. In the wake of innovation, numerous farms and agribusinesses are gradually receiving innovation into farm activities as it extensively affects usefulness and productivity.[2] Agribusinesses and other horticultural businesses, both administrative and non-legislative, are going to different programming and equipment computerised cultivating instruments, for example, ranch the board arrangements, Internet of things (IoT), and different types of precision agriculture. Current innovation in farming is something other than the utilisation of versatile applications or an equipment framework to robotize repetitive cycles and lessen reliance on human work.[5] This prediction will help farmers choose the appropriate crops for their farm based on soil type, temperature, humidity, water level, depth of spacing, soil pH, season, fertiliser, and months.[7] Rising technologies are used to improve the productivity of the crops by changing traditional farming to e-farming. About 70 per cent of primary and secondary activities are based on agriculture. Therefore, to improve agriculture, many farmers began to use new technologies and methods.[9] The key issue however to be resolved is cultivating precise crops at precise times. This could be done with the help of machine learning algorithms that are found to be an efficient methodology for predicting the appropriate crop.[11]

2 Literature Survey

In paper [1] a system has been designed based on sensors and IoT systems to monitor the crop field with the help of a sensor. The system is proposed for monitoring weather conditions in a particular place like temperature, humidity, and changes in the environment.

In paper [2] Machine Learning Algorithms are used and designed to make a system easier to compare prices. It achieved an accuracy rate of 95. The algorithm was created to predict modal prices at 14-day intervals across the major mandis in MP.

In paper [3] Time series and Regression algorithms are used, as a means to assess Moisture Content (MC) in their grain. It calculates Moisture Content (MC) with high accuracy, using Relative Humidity (RH) and Temperature (T) time series.

In paper [4] Random Forest Algorithm is used and a broader dataset results in a higher rate of accuracy. Random forest shows to be a superior prediction algorithm. To predict the yield of a particular state that uses some of the various factors which affect the manufacturing of crops.

* Corresponding author: tanvidaware17@gmail.com
** Corresponding author: pratikshasramteke@gmail.com
*** Corresponding author: ukshaikh210595@gmail.com
**** Corresponding author: smita.bharne@rait.ac.in

In the paper [5] authors used Image Processing, Machine Learning, the Internet of things and Artificial Intelligence. Comparative research is conducted between the advanced framework and the present framework. Survey of different methods for Crop selection, sowing of the crop, weed detection, and monitoring of the system thereby yielding productive output.

From the literature survey, we analysed that the existing methods, the process of finding the soil type, identifying the leaf disease, and the suitable fertiliser had been all carried out manually. Even when the framework became digitised, it had certain issues such as predicting numerous fertilisers for a soil kind, plant sickness, or fertiliser. The existing systems lack efficiency due to limited algorithm-based prediction.

3 Proposed Methodology

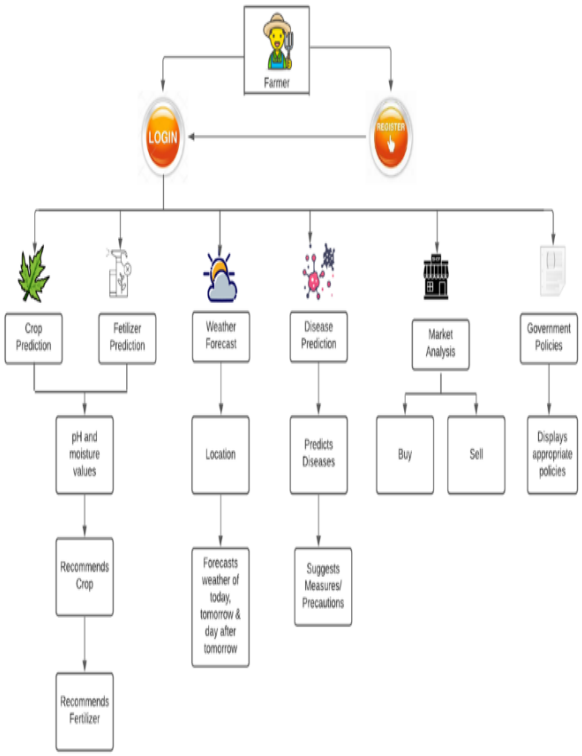


Fig. 1. Proposed System Architecture

Figure 1 shows the detailed system architecture of the proposed system. It portrays all the modules of the system and the functionalities in each module.

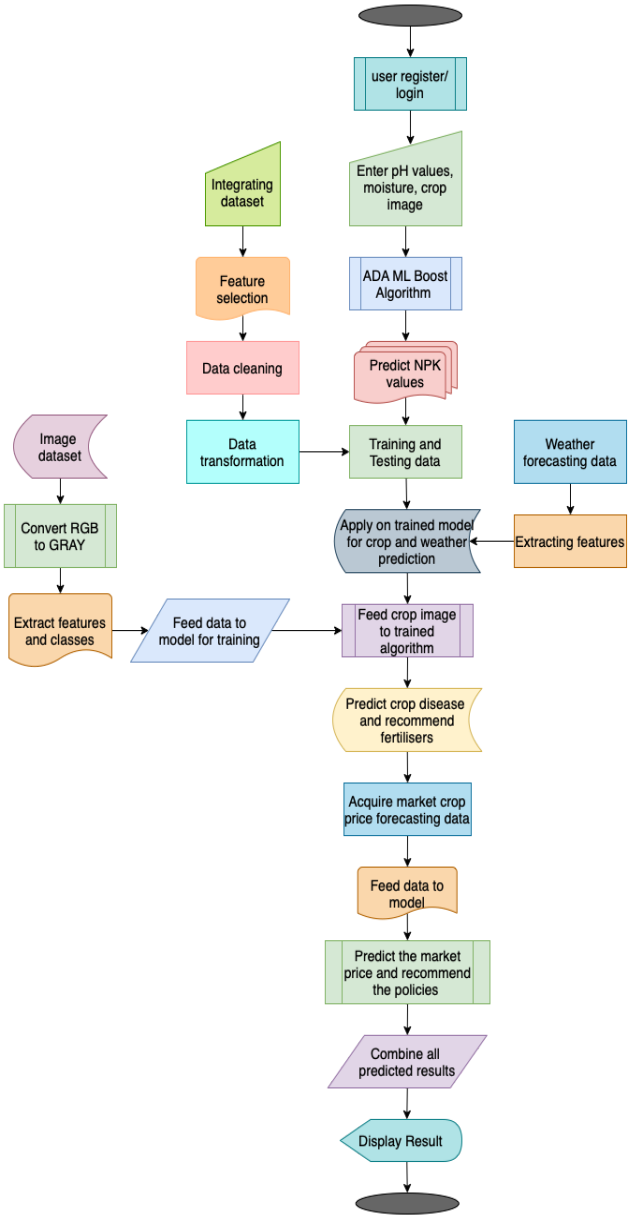


Fig. 2. Detailed flow of the system

Figure 2 shows the detailed flow of the proposed system. It portrays the control flow from user registration/login to displaying the results based on the inputs given by the user.

Composts and fertilisers are added to get the nutrient supplement levels to the proportion needed by the crop the farmer intends to develop, prompting loads of costly base work of developing soil.[3] Be that as it may, our model as of now predicts which harvest develops best given current conditions. Following are the descriptions of the modules included in the proposed system.

3.1 Crop Prediction

The system will use the Ada boost algorithm on moisture and PH values to predict the N-P-K values which are further used in the ensemble method technique which includes algorithms like Naive Bayes, logistic regression, ADA Boost, Random forest, Decision tree, and SVM for crop recommendation.

3.2 Fertiliser Prediction

Enter the nutrient contents of your soil and the crop you need to grow. The set of rules will inform which nutrient the soil has extra or lacks. Accordingly, it's going to supply pointers for getting fertilisers.

3.3 Government Policies

This provides various schemes offered by the Government of India in the field of agriculture. It also describes the eligibility criteria, benefits, procedure to apply, and documents required to take the advantage of a particular scheme.

3.4 Weather Prediction

Anticipating the real weather conditions in current scenarios and then accordingly providing information to the end-user to take the correct measures and also plan the development of the crop.

3.5 Market Price

There will be an analysis of the market to provide the user with the best deal. We will be supplying statistics of the huge gamers of the marketplace including Big Basket, scraping the retail charges from a genuine supply and permitting farmers to make their bargain and ship quotations.

3.6 Disease Prediction

Upload a picture of the leaves of your plant. The set of rules will inform the crop kind and whether or not it's far diseased or healthy. If it's diseased, it'll inform you of the purpose of the sickness and advocate to you the way to prevent the sickness accordingly.

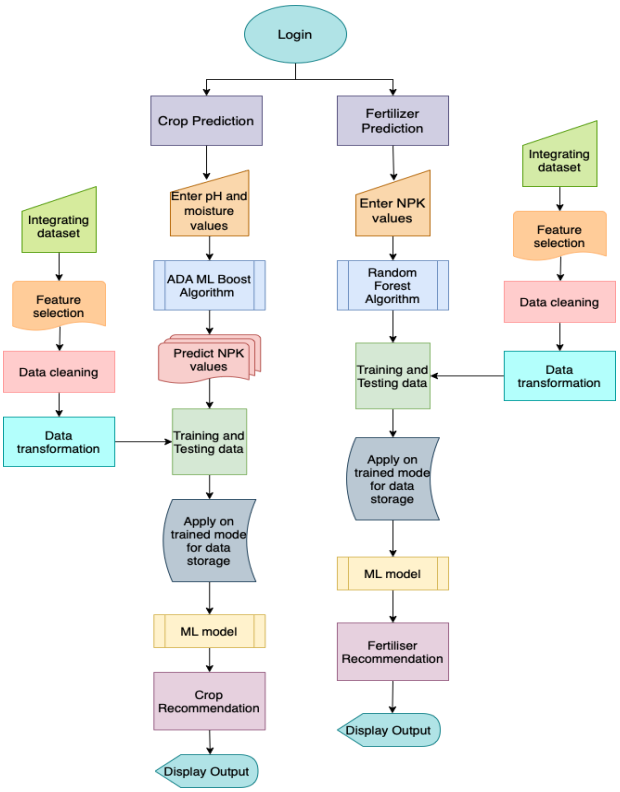


Fig. 3. Flowchart of Crop and Fertiliser Prediction

The given Fig. 3 shows the system module that consists of a portal that will predict the crop and fertiliser suitable for a particular user. When the user enters moisture and pH value for the recommendation of the crop, the ADA Boost Classifier will fit the Decision tree adjusting the incorrect weight such that subsequent classifiers focus on difficult cases. Later it will predict the NPK values and then we will integrate, clean, and transform the dataset which is followed by training and testing the data. The next stage is to apply the trained models which in turn will predict crops based on the ML models and finally display the name of the crop. Similarly, for the fertiliser prediction, our model will take inputs of NPK values. On integrating the dataset, cleaning, and transforming the data, we will train and test the data and apply the trained data to the ML model and display the name of the fertilisers.

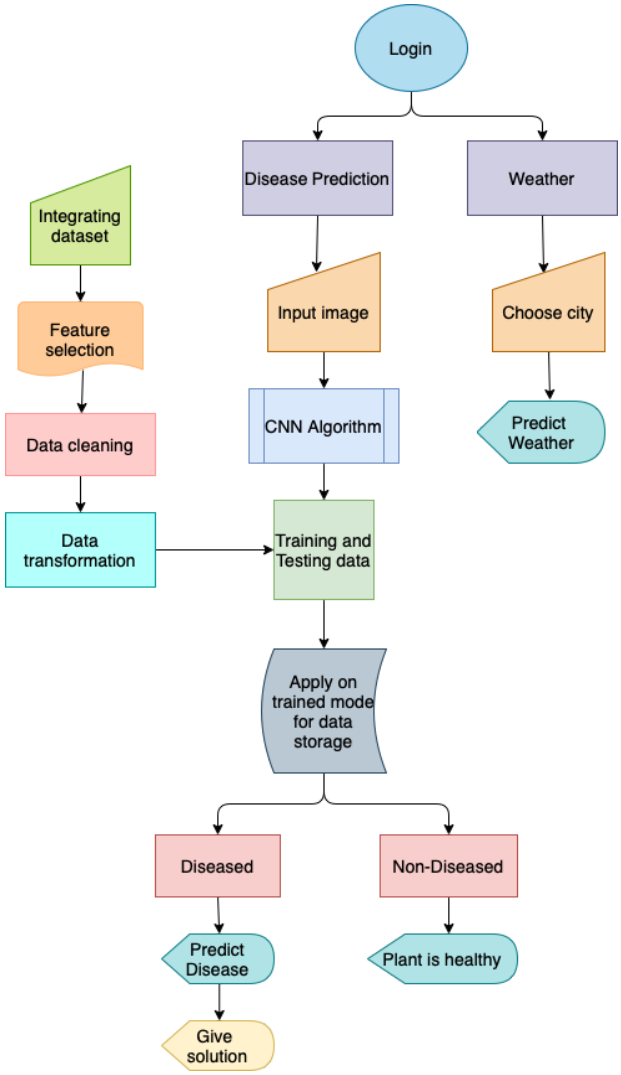


Fig. 4. Flowchart of Weather and Disease Prediction Module.

Figure 4 describes the flowchart for the plant disease detection and weather prediction modules. For the prediction of disease, the user has to give an image as an input for the system, using the CNN algorithm, and then we will integrate, clean, and transform the dataset which

is followed by training and testing of data. The next stage is to apply the trained models which in turn will predict whether the plant is diseased or not based on the ML models. For the prediction of weather, the user has to choose a particular city that will finally predict the weather pattern.

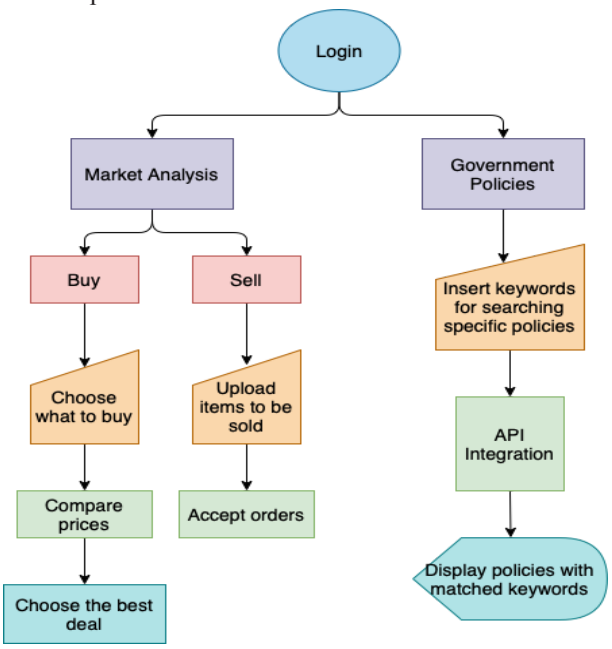


Fig. 5. Flowchart of Market Analysis and Government Policies

Fig.5 describes the flowchart for the market analysis and government policies modules. Our system allows users to buy and sell products and also compare prices to choose the best deals. This provides information regarding various government policies of India related to agriculture.

4 System Details

Machine learning is an important support tool for crop expectations, including support options for plant development and what to do during the plant development phase. The following algorithms are applied to the various features:

4.1 Decision Tree

The decision tree creates characterization or fallback models within the level of a tree structure. It divides the dataset into a moderate subset and gradually evolves a decision tree. It deals with all kinds of downright and mathematical information.

4.2 Naive Bayes

Naive Bayes is used to construct classifiers: models that assign to drawback cases and labels described as vectors of entity values wherever the class labels are drawn from a finite set. It takes into consideration that the worth of a particular feature is insignificant to the real worth of alternative characteristics given class variables.

4.3 Support Vector Machine

SVM makes use of a subset of training factors in the decision function (referred to as guide vectors), so it's also memory efficient.

4.4 Random Forest Algorithm

The Clustered Random Forest (CRF) methodology is used to identify intruders into the network by dividing the entire network into character systems so that glowing comments can be made on communications initiated within the network. It examines every node in the network and creates log files that are given as input for the proposed technique.

$$Gini\ Index = 1 - \sum_{i=1}^{n} (Pi)^2$$

4.5 ADA Boost

It is used as an ensemble method and it adapts a series to differently weighted training data. It starts with predicting the original data set and giving equal weight to each observation.

$$H(x) = sign\ (\alpha_t h_t(x))$$

$h_t(x)$ is the output of weak classifier for input x.

Where $\alpha_t = 0.5 * \ln((1 — E)/E)$ (where E = error rate)

4.6 Neural Network

A neural network is a computational model with network architecture and specific parameters that can be used to modify it to perform specific tasks.

4.7 Dataset

Crop Prediction data is collected from the Kaggle website for the prediction of crops and fertilisers. It has 22 different crops as labels and 7 features - Nitrogen, Phosphorus, Potassium, Temperature, Humidity, Rainfall and Moisture. The input dataset consists of around 3000 data values of crops (rice, maize, orange, moth beans etc) for different PH values and moisture content in the soil.

	N	P	K	ph	Moisture
51	76	60	39	6.766240	51
1656	17	16	14	6.625539	27
245	35	64	78	7.496645	49
357	34	59	18	5.744118	44
267	41	69	82	6.715587	47
...	...	...	...	...	...
643	2	47	15	6.401456	37
131	63	43	19	6.641906	44
1663	26	11	11	7.609348	25
130	87	35	25	6.178056	56
1137	35	18	26	5.279389	35

660 rows × 5 columns

Fig. 6. Dataset for crop prediction module

	Crop Type	Nitrogen	Potassium	Phosphorous	Fertilizer Name
0	Maize	37	0	0	Urea
1	Sugarcane	12	0	36	DAP
2	Cotton	7	9	30	14-35-14
3	Tobacco	22	0	20	28-28
4	Paddy	35	0	0	Urea
...	...	...	...	...	...
94	Pulses	24	0	19	28-28
95	Tobacco	4	17	17	10-26-26
96	Wheat	39	0	0	Urea
97	Millets	15	0	41	DAP
98	Sugarcane	12	0	10	20-20

99 rows × 5 columns

Fig. 7. Dataset for fertiliser prediction

Based on the given pH and moisture content, we can predict the NPK values to recommend suitable crops and using NPK value and crop type we can predict the suitable fertilisers.

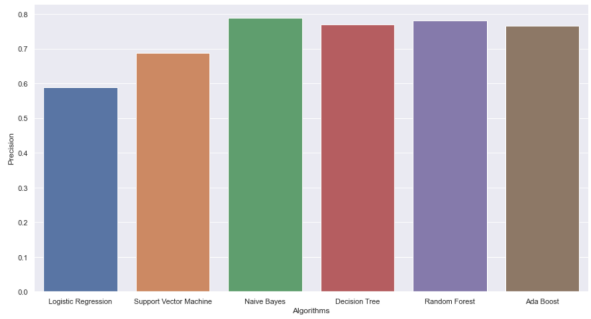
Fig. 8. Dataset for Disease Prediction

For the disease prediction module, the dataset is collected from the Kaggle website, the system takes the image of the infected leaf as input and predicts the disease remedy. Input dataset consists of 2000 different images of healthy and infected leaves of plants like tomatoes, potatoes, bell pepper etc.

4.8 Mathematical formulae

1. Precision: Precision is the ability to correctly identify the number of positive predictions.

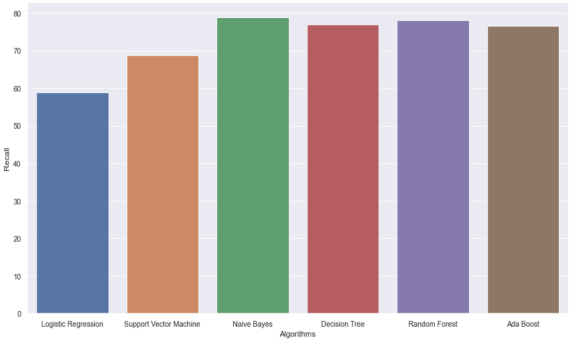
$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$



Average Precision = 0.72467

2. Recall: It finds all positive results from the confusion matrix.

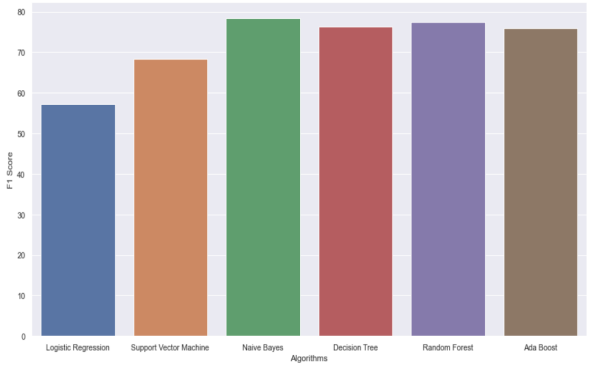
$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$



Average Recall = 73.015

3. F1 Score: It finds a balance between precision and recall.

$$\text{F1 score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$



Average F1 score = 72.2867

4. Accuracy: It is calculated by dividing correct predictions by total predictions.

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$$

5 Results

In particular, our contributions are as follows:

• Predicts the suitable crop and fertiliser with maximum accuracy.

• Detects the disease on a plant with the help of an image of its leaves.

• Provides the latest updated government schemes.

• Provides a marketplace so as to compare the prices and sell or buy the required item. Here, we propose a methodology for crop and fertiliser prediction using the ensemble method and the disease will be predicted using the CNN algorithm. The fetching of weather and government schemes will be done by integrating the APIs.

This model is trained using a dataset of around 100 epochs and 3000 values compiled with sparse categorical cross-entropy as the loss function for crop and fertiliser prediction.

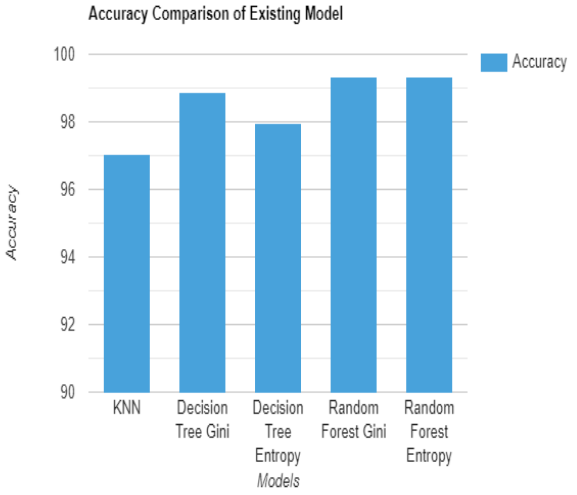


Fig. 9. Accuracy comparison of Existing Model

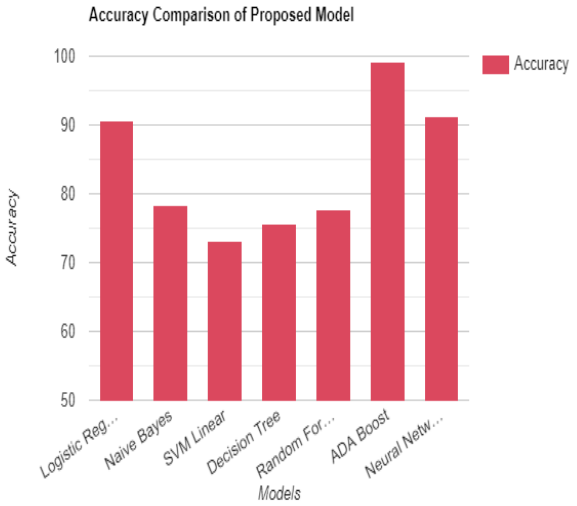


Fig. 10 Accuracy Comparison of Proposed Model for Crop Prediction

Table 1. Accuracy Comparison for Crop Prediction

Sr No.	Machine Learning Model	Accuracy
1	Logistic Regression	60.61%
2	SVM Linear	78.3%
3	ADA Boost	73.18%
4	Decision Tree	75.61%
5	Random Forest	77.73%
6	Naive Bayes	79.09%
7	Neural Network	61.18%

Table 1 summarises the model's accuracy of the crop prediction module. The highest accuracy obtained is 79.09% for the Naive Bayes Model.

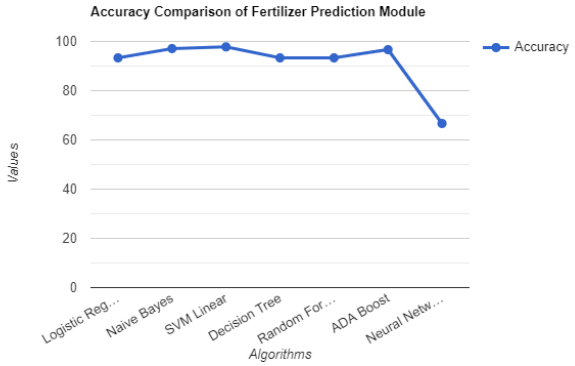


Fig. 11. Accuracy Comparison of Proposed Model for Fertiliser Prediction

Table 2. Accuracy Table for Fertiliser Prediction

Sr No.	Machine Learning Model	Accuracy
1	Logistic Regression	93.33%
2	SVM Linear	97.1%
3	ADA Boost	97.8%
4	Decision Tree	93.33%
5	Random Forest	93.33%
6	Naive Bayes	96.67%
7	Neural Network	66.67%

Table 2 summarises the model's accuracy of the fertiliser prediction module. The highest accuracy obtained is 97.8% for ADA Boost Model.

In the weather prediction module, the user has to enter a city name, which in turn will predict the weather pattern for the subsequent days.

As mentioned earlier, for the disease prediction module, we have taken the images dataset with 2000 images and the system got an average accuracy of 71.20% with the help of convolutional neural networks.

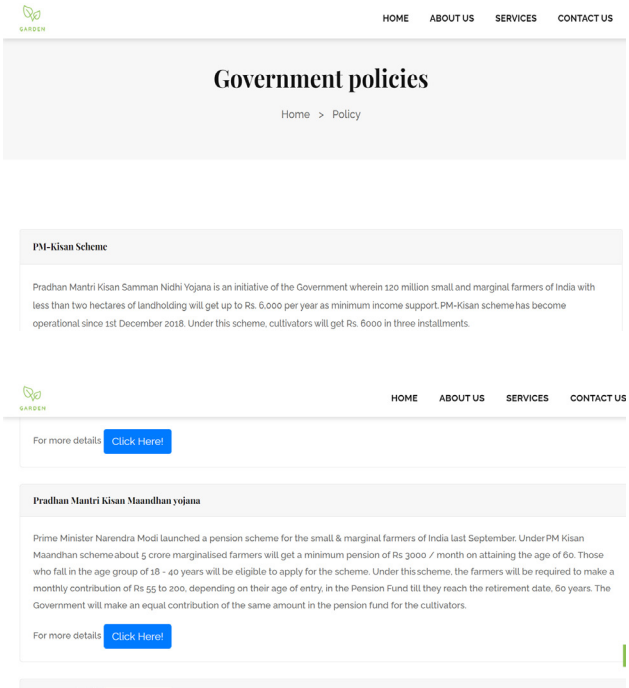


Fig. 12 Government Policies

The government policies will help farmers to access various government policies according to their needs and it will redirect them to the government portal.

6 Conclusion

The comparative study of different machine learning algorithms helps us understand which is best suited for the prediction of crops and fertilisers which will increase the yield. Hence Crop Guidance and Farmers Friend make farming easier by providing proper guidance to the farmers about irrigation, fertilisers, weather conditions, soil quality, and suggesting alternative crops and seeds. Since we are using the ensemble method, we got the highest accuracy of 79.09% for the Naive Bayes Algorithm for crop prediction and 97.8% for ADA Boost in Fertilizer Prediction and 71.2% for disease prediction using CNN.

References

1. V. R. V. S. S. T. Sooraj Krishna, Sreevijay, *Farmer friend*, IEEE, **10**, pp. 9–19 (2021)
2. M. Dr Anitha Govindraj, *Crop price prediction using machine learning*, IEEE, **6**, pp. 14–20 (2020)
3. C. B. W. W. S. Ghionea, *Algorithms to predict the moisture content of the grain*, IEEE, **1**, pp. 1–7 (2020)
4. P. M. R.Vidhya, *Predicting yield of crop using ml algorithm*, IEEE, **9**, pp. 1–13 (2020)
5. S. K. M. G. S. Prince Samuel S., K. Malarvizhi, *Machine learning and internet of things based agriculture*, IEEE, **2**, pp. 3–11 (2020)
6. S.D.M, Subramaniya Raman M.K, *E- Farming: A Breakthrough for Farmers, IJERT, **9** (2020)*
7. V. J. Monika Jadhav, *Android application for farmers*, IEEE, **6**, pp. 1–3 (2019)
8. P. K. M. B. Karthikeyan S.Sainavinraj M., Praveen Kumar A., *Farmer's friend*, IEEE, **6**, pp. 1–9 (2019)
9. G. A. Anirudha Vachaspati Vempati, *Agricultural problems and technology-based sustainable solutions*, IEEE, **5**, pp. 6–13 (2020)
10. R. C. N.Vanitha, N.Archana, *Agriculture analysis using data mining and machine learning techniques*, IEEE, **3**, pp. 8–15 (2019)
11. K. P. H. Trupti Bhange, Swati Shekapure, *Prediction of crop yield and suitable crop*, IEEE, **8**, pp. 1–5 (2019)
12. W. G. Parul Sharma, Yash Paul Singh Berwal, *Krishimitr (farmer's friend)*, IEEE, **4**, pp. 893–899 (2018)