



# LOCATION BASED PREDICTION OF CROPS, ANALYSING THE YIELD AND MARKET DEMAND USING R-FOREST AND MLR ALGORITHMS

Prof. Mamatha A  
Assistant Professor  
Department of CS & E  
Sai Vidya Institute of Technology  
Bangalore, Karnataka, India

Anil Kumar DV, AC Shravya, Anusha S Petakar, Dheekshitha R  
Students  
Department of CS & E  
Sai Vidya Institute of Technology  
Bangalore, Karnataka, India

**Abstract - India being an agricultural country, its economy largely depends on the growth of agricultural produce and the agricultural products of its allied industry. In India, agriculture is heavily influenced by rainwater which is extremely unpredictable. Agricultural growth also depends on various soil types, namely Nitrogen, Phosphorus, Potassium, Crop rotation, Soil humidity, global warming and climate factors including temperature, precipitation, etc. India is now making rapid progress in technological development. Therefore, technology will prove to be beneficial to agriculture which will improve crop production leading to better yields for the farmer.**

**Keywords — random forest, decision tree, multi-line retrieval, Data Analytics.**

## I. INTRODUCTION

In India most of the population is in agriculture. Due to climate change, bottles are available to boost crop forecast in India. Finding the right levels of crop yield has been a difficult task. Various factors that directly affect plant growth, productivity, should be considered. estimation of crop yields is one of the key factors in agricultural production. Before seeds can be sown in their fields, farmers need information about the crop yield in order to produce more yields. In recent years, the use of technology in agriculture has increased and data analysis is one of the new entrants to the agricultural sector. Realizing the usefulness of big data analysis is the biggest obstacle to using big data in agriculture. Efforts are underway to understand how big data analytics can improve agricultural productivity. The proposed framework provides details of the

various data analysis methods used to estimate the crop yield, and shows the value.

## II. LITERATURE SURVEY

A. “Machine Learning approaches for crop yield prediction and nitrogen status, estimation in precision agriculture”, Anna Chlingaryan, Salah Sukkarieh, Brett Whelan.[1]

This review shows in particular that the M5-Prime Regression Tices trees are a good tool for multi-level crop forecasting and the Least Squads Support Machine is a promising tool for re-analyzing Nitrogen status. Second, the Fuzzy Cognitive Map (FCM) is used to model and represent expert knowledge in crop forecasting and crop management.

B. “Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods”, Michael D. Johnson, William W. Hsieh, Alex J. Cannon, Andrew Davidson.[2]

In this review, Multiple Linear Regression (MLR) and two non-linear Bayesian Neural Networks (BNN) and model-based recursive partitioning (MOB) are used to predict crop yields by combining MODIS-NDVI, MODIS-EVI and NOAA-NDVI as predictors. Estimates of fruit yields made using forecasts tested for a validated error index for the period 2000–2011.

C. “Crop Prediction System using Machine Learning”, Prof. D.S. Zingade, Omkar Buchade,



Nilesh Mehta, Shubham Ghodekar, Chandan Mehta.[3]

This review proposed a Smart Agriculture solution by looking at an agricultural field that could assist farmers in growing their produce on a large scale. Climate data obtained from IMD (Indian Ministry of Environment) such as temperature and precipitation and saturation of soil boundaries provides an understanding of which plants are suitable for cultivation in a particular area. It uses data analysis techniques to predict the most profitable crop in the current climate.

D. "A Survey on Crop Yield Prediction based on Agricultural Data", Dhivya B, Manjula, Siva Bharathi, Madhumathi.[4]

The review proposed that seasonal, economic and biodiversity factors contributed to crop production but unexpected changes to these factors resulted in significant losses for farmers. These risks can be eliminated when appropriate mathematical or mathematical methods are applied to data related to soil, climate and past yields. With the onset of data mining, crop yields can be predicted by obtaining useful information from these agricultural data that help farmers decide which crop they would like to plant in the next year which yields the most profit. This presents a survey on various algorithms used to predict crop yield.

E. "Agricultural Crop Predictor and advisor using ANN for Smartphones" Giritharan Ravichandran, Koteeshwari R S.[5]

The program uses the Artificial Neural Network, which is another very effective tool for modeling and forecasting. There are various parameters that determine the production output. these parameters are considered to be the input of the program and based on the illusion of this input, the desired result must be produced. The parameters include pH, phosphate, potassium, nitrogen, depth, temperature, precipitation. Also, the proposed plan proposes some fertilizers that can improve the product.

### III. EXISTING SYSTEM

Many NGOs assist the rural people by building information about the farm through the network. They provide information such as Pesticides, Compost, Seed. Documentation of different technologies used in cultivation in different countries. They also run camps in various villages. There are many books and websites that cover additional information on livestock such as weather, machinery, fertilizers, seeds, pesticides and more.

### IV. PROBLEM STATEMENT

Implementing the system, which uses the R-Forest and MLR algorithm to predict the most beneficial yields and favorite

fertilizers to be used in current weather and soil conditions such as input and monitoring data from the weather department and static data from various government websites.

### V. PROPOSED SYSTEM

Agricultural product manufacturers are concerned about changes in consumer demand and demand change, in addition to climate and climate change that are contributing to the growth of agricultural products. In this system a prediction of the yield success rate and demand from the Linear Regression Technique, Multiple Linear Regression and the Random Forest.

### VI. SYSTEM DESIGN

System design is the process of defining architecture, modules, communication platforms, and program data to satisfy specific requirements. Designing systems can be seen as the application of product development strategies.

#### A. Block Diagram

The block diagram is a special, high-flowchart used in engineering. Its structure provides a high-level overview of the major system components, key process participants, and key performance relationships. The focus is on input and output of the program.

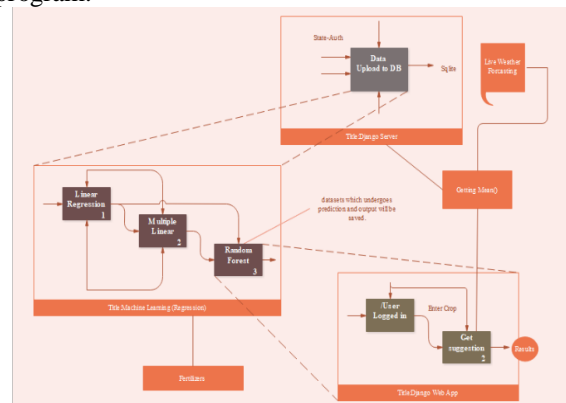


Fig. 1. Block Diagram of the proposed System.

The block diagram.1 Fig.1 shows the overall performance of the Time series forecast for the demand for agricultural products using a web application. It contains the following:

- Recalibration: regression models have been used to predict continuous value. Predicting the crop yield given by the characteristics of the crop such as the name of the plant, soil, climate, etc. It is one of the classic examples of Regression. It's a guarded road. Linear Regression: Linear Regression is a machine learning algorithm based on supervised learning. He does the postponing work. The regression models determine the value of the target prediction based on independent variables. Where there



is only one input element (x) and there is a need to use vertical alignment, this is called vertical line correction.

- Direct multiplication of several lines: Multiple Linear Correction effort
- Random Forest: Random forest is a trade name for clustering decision trees. In the Random Forest, there is a collection of deciduous trees (called "Forests"). To categorize something new based on characteristics, each tree gives the tree "votes" in that category. The forest prefers the most demarcated subdivisions (above all the trees in the forest). Random forests are a combination of tree prognosticators that each tree depends on the values of the already available offline sampling alone and the same distribution across all forest trees
- Django Server: Python's advanced web framework that enables faster and more secure web development. Designed by experienced developers, they take care of most of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It is a free and open source, has a prosperous and efficient community, great documentation, and many free support and payment options.

### B. Sequence Diagram

Sequence diagram shows the interaction of an object in a chronological order. It shows the objects and categories involved in the situation and the sequence of messages exchanged between the elements needed to achieve the status quo.

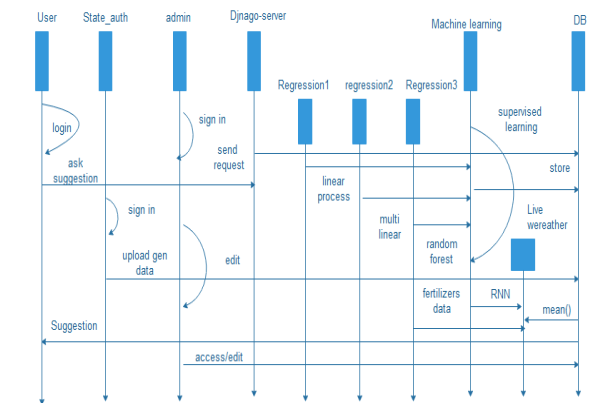


Fig. 2. Sequence Diagram of the proposed system.

In the proposed scheme:

- The user must first register and sign in.
- Email can be used for registration.
- The user enters the fruit he or she wants to plant with the region (such as land details), asking for suggestions.
- The state authority must first register and sign in.
- The state authority unloads everything needed.

- The state authority will upload three machine learning results obtained through the three control algorithms (row, clustering and random forest algorithm) database.
- Management must be logged in and have the authority to edit the database.
- The live weather report is available through the google API and is stored in a database. It also contains the temperature and rainfall of all locations and is stored in a database.
- The user where the correct crop application shows the percentage of forms on the web app screen that if the proposed plant is suitable for cultivation.

### C. Use Case Diagram

Use case diagrams that are a way to capture the effectiveness of programs and requirements in UML diagrams. It draws the flexibility of a live program. Contains use case and character. It represents a different functionality of a program, object, package or a class.

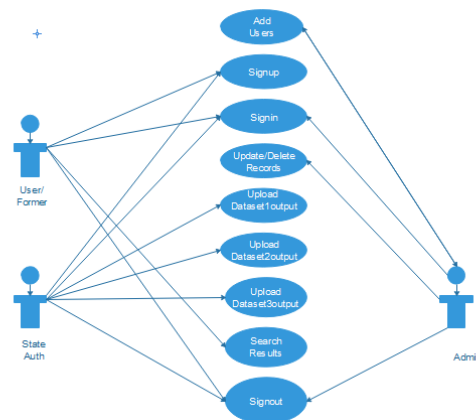


Fig. 3. Use case Diagram of the proposed system.

- In the proposed scheme:
- User can subscribe, login, search results and logout.
- The State Authorities will log in, download the three Algorithm results and log out.
- Administrators can add new users, log in and update or delete records or save and exit.
- Later, the result will be generated on the web app result screen.



## VII. METHODOLOGY

As a first step in crop forecasting, the system displays results using only the data. In addition to plant data, other related features such as meteorological data, estimation of preferences for analysis of SNS data. The framework is designed to reflect this future learning potential. The selection of targeted agricultural products varies depending on the sampling times. The most representative products are selected and evaluated on a periodic basis.

The new learning algorithms algorithm used and demonstrated good predictive performance somewhere in the future. Although understanding the accuracy of the predicted value at a specific time point. The predecessor who is trained by past data and provides a measure of success over a period of time, that trade in agricultural products seems to have seasonal variations, e.g. Climate movements are seasonal and show one-year operations at regular intervals. total, the basic predictor can see these features-respectively.

In order to predict the yield further to predict the success rate at a specific location in the future, the generation-and filtering method using the basic estimation combinations is used. Basic predictors of a specific time point are selected for their accuracy, and are arranged in chronological order. The basic idea is as follows. A random forest is used to generate a network. The length of the word 'n' insertion is set as 20. The forecast period for k varies from 1 to 53 for recognition of the next one year forecast. With each word 'k' ranging from 1 to 53, the direct process is trained 90% of training data from last 312 weeks Training is set to start at 300 or 200 epochs in the Random Forest, with good results a minimum of 30 guaranteed minimum loss was selected among the trained ANN weights. As a result, the set of the basic 1590 prediction  $\{P_i(n, k) | T \text{ in } [1, 30], n = 20, 'k' \text{ in } [1, 53]\}$  received. Accuracy of validation data is less than 0.01 for all baseline predictors, which means they show good performance of 10% validation data in the past. Prediction expert is designed as a weather for basic predictors. For each first 'k' the predicate 'Pa (n, k)' in  $[1, 2, \dots, 30]$  is available. Sort basic predictors in order of their validity in the validation data. Group the best predictors from each 'k' from 1 to  $k_{max} = 53$ , and make a sequence of predictors in sequence k:  $S_f(1) = [p(n, 1), p(n, 2), \dots, p(n, k_{max})]$  Similarly, the sequence created by the second best character is shown by Sf (2), Sf (3), respectively. This way, find Sf (i) (i =  $[1, 2, \dots, 30]$ ).

The predicted sequence values from the previous spj sequence sequence (Past Sequence 'j') are denoted by S f (i) (spj). In many fields including trading on agricultural products, the price range has a value similar to or more important than the actual price in a particular area, since the primary purpose of forecasting is to avoid risk caused by fluctuations in success rate, not profit. In the present section, the proposed method

shows that the predicted future yield has the same characteristics of the real-time past, a detailed analysis is left for future research. Through the analysis of the probability distribution of real and predicted time, both appear to have the same trend in the mathematical sense.

## VIII. RESULT

The most important point is that when analyzing predictions based on time, there is a need to pay attention also to mathematical features such as probability distributions in addition to accuracy and accuracy in specific areas. This may sound strange, because it is common to assume that when the precision of the value for each point increases, the accuracy of the predicted time series should increase. However, this understanding does not always apply to targeted programs including social and / or human entities, such as trade, economics, communications, local community communication, and so on. The reason is that, first of all, the accuracy of future predictions is low in nature in such systems. To say more precisely, precise forecasting at some future location is difficult or in many cases impossible. In addition to searching for great accuracy and precision over a period of time, there is a need to pay attention to the numerical transformation method, which can be measured in a mathematical way. Although one of the methods is to analyze the probability distribution as described, there are many other ways to verify the consistency between the original and predicted future series. In many fields including trading on agricultural produce, the price range has a value similar to or more important than the price itself in a given area, since the primary purpose of forecasting is to avoid risk caused by price fluctuations, not profit. In the present section, our approach suggests that the estimated future output captures the data of past years and almost the same real-time signals, detailed analyzes are left for future research.

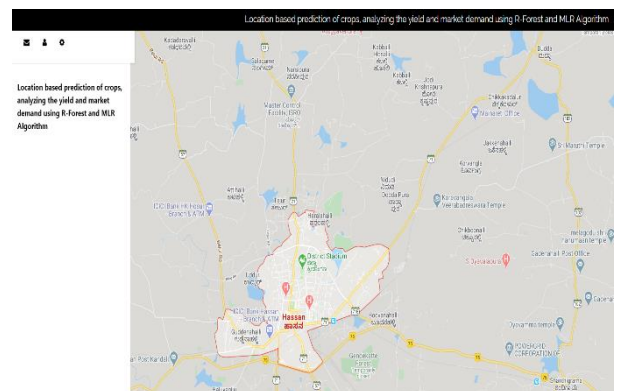


Fig. 4. Map of Hassan



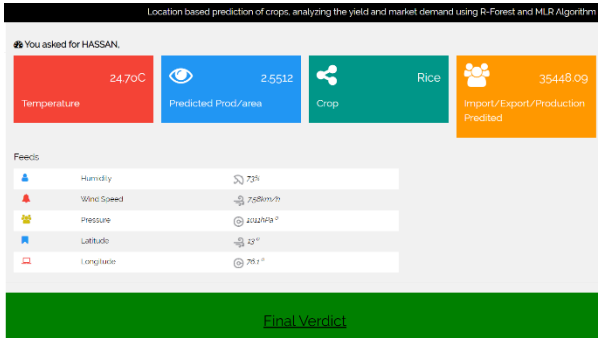


Fig. 5. Weather conditions of Hassan

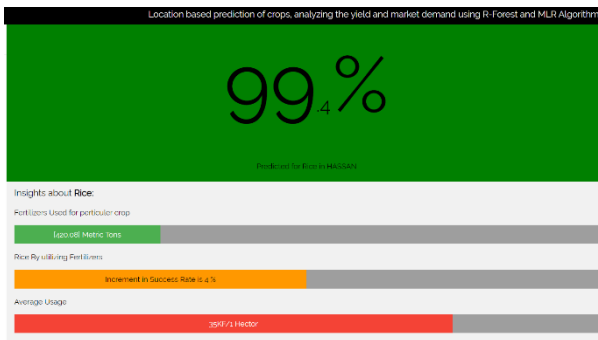


Fig. 6. Success rate for the entered crop

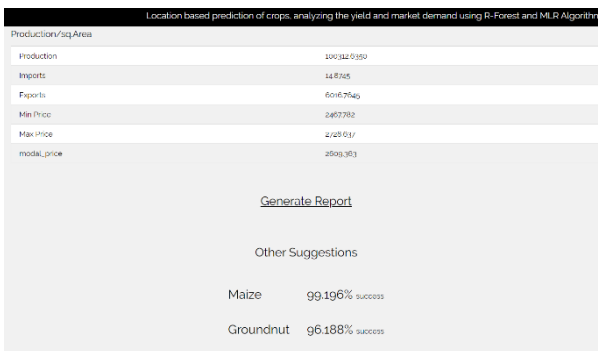


Fig. 7. Predicted production with other crop suggestions

When a farmer / user logs into a system or website and enters the area where he or she wishes to grow a crop and the name of the plant he / she is willing to grow and give us prediction, the server automatically locates the area as shown in figure 4 on the map and prediction.

First, the server looks at the local climate and predicts the production of each location as shown in figure 5 based on the previous year's dataset and also predicts the user's import and export information.

By taking all these factors into account the system predicts the success rate of the applied crop and weather conditions as shown in figure 6, and provides insight into the amount of fertilizer to be used to produce the additional

product. This time also shows how few other crops have a high success rate in the area with recorded weather conditions.

Finally the farmer is given the opportunity to get a complete report as shown in note 7, a forecast made with regard to the production of the crop at a specific location mentioned in his registered post name by simply clicking on the Generate Report option.

## IX. CONCLUSION

There are many programs available that use various methods of data mining and machine learning to help farmers make decisions before seeds are sown. But there is still a problem when farmers focus on just one crop and ignore all the other crops that can bring them more money. This program focuses on these issues and gives farmers an understanding of the different crops that can be planted to achieve high production levels.

## X. FUTURE WORK

Information on soil development, climate and the past year can be taken into account as well as productive plants that can be planted in the proper environment. All possible crops are calculated to help the farmer decide which crop to grow. Under this program, high-yielding crop types are covered, and farmers can find a crop that may not have been planted.

## ACKNOWLEDGEMENT

We extend our deepest thanks to the management, principals, director, head of department and team of Computer Science and Engineering of Sai Vidya Institute of Technology, Rajanukunte, Washington for providing the opportunity to do research and are very grateful for their inspiration, encouragement, guidance, suggestions and constant support for the finalization. work.

## XI. REFERENCES

- [1] Anna Chlingaryan, Salah Sukkarieh, Brett Whelan, "Machine Learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture" ELSEVIER May 2018.
- [2] Michael D. Johnson, William W. Hsieh, Alex J. Cannon, Andrew Davidson, "Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods" ELSEVIER Nov 2015.
- [3] Prof. D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, Chandan Mehta.,I.S, "Crop



Prediction System using Machine Learning”, IJAERD  
Dec 2017.

- [4] Dhivya B, Manjula, Siva Bharathi, Madhumathi, “A Survey on Crop Yield Prediction based on Agricultural Data”, International Conference in Modern Science and Engineering, March 2017.
- [5] Giritharan Ravichandran, Koteeshwari R S, “Agricultural Crop Predictor and advisor using ANN for Smartphones”, IEEE 2016.
- [6] <https://data.gov.in>