



# PREDICTION OF CROP YIELD AND FERTILIZER RECOMMENDATION USING MACHINE LEARNING ALGORITHMS

Devdatta A. Bondre  
Student, NICT Solutions & Research,  
Belagavi, Karnataka, India

Mr. Santosh Mahagaonkar  
Research Head, NICT Solutions & Research,  
Belagavi, Karnataka, India

**Abstract-** Machine learning is an emerging research field in crop yield analysis. Yield prediction is a very important issue in agriculture. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data. Machine learning techniques are the better choice for this purpose. Different Machine learning techniques are used and evaluated in agriculture for estimating the future year's crop production. This paper proposes and implements a system to predict crop yield from previous data. This is achieved by applying machine learning algorithms like Support Vector Machine and Random Forest on agriculture data and recommends fertilizer suitable for every particular crop. The paper focuses on creation of a prediction model which may be used for future prediction of crop yield. It presents a brief analysis of crop yield prediction using machine learning techniques.

**Keywords-** Agriculture, Yield Prediction, Machine learning, Support Vector Machine, Random Forest, Soil Nutrients.

## I. INTRODUCTION

From ancient period, agriculture is considered as the main and the foremost culture practiced in India. Ancient people cultivate the crops in their own land and so they have been accommodated to their needs. Since the invention of new innovative technologies and techniques in the agriculture field is slowly degrading. Due to these, abundant invention people are been concentrated on cultivating artificial products that is hybrid products where there leads to an unhealthy life. Nowadays, modern people don't have awareness about the cultivation of the crops in a right time and at a right place. Because of these cultivating techniques the seasonal climatic conditions are also being changed against the fundamental assets like soil, water and air which lead to insecurity of food.

The machine learning learns the algorithm based on the supervised, unsupervised, and Reinforcement learning each

has their importance and limitations. Supervised learning-the algorithm builds a mathematical model from a set of data that contains both the inputs and the desired outputs. Unsupervised learning-the algorithm builds a mathematical model from a set of data which contains only inputs and no desired output labels. Semi-supervised learning- algorithms develop mathematical models from incomplete training data, where a portion of the sample input doesn't have labels.

This paper aims to improve the yield of the crop in several ways and recommends fertilizer suitable for every particular crop.

## II. RELATED WORK

The literature has many reported works in the domain.

[1] **Shivnath Ghosh, et al. (2014)** In this paper machine learning system is divided into three steps, first sampling (Different soil with same number of properties with different parameters) second Back Propagation Algorithm and third Weight updating.

[2] **P.Vinciya, et al. (2016)** This paper mainly focused on analyzing the agriculture analysis of organic farming and inorganic farming, time cultivation of the plant, profit and loss of the data and analyzes the real estate business land in a specific area. This work goes for finding reasonable information models that accomplish a high precision and a high consensus as far as yield expectation abilities.

[3] **Zhihao Hong, et al. (2016)** This paper proposes an information driven approach on structure PA answers for gathering and information demonstrating frameworks. Soil dampness, a key factor in the yield development cycle, is chosen for instance to exhibit the viability of our information driven methodology. On the accumulation side, a responsive remote sensor hub is built up that expects to catch the elements of soil dampness utilizing soil dampness sensor. The prototyped gadget is tried on field soil to show its usefulness and the responsiveness of the sensors. On the information examination side, a one of a kind, site-explicit soil dampness expectation system is based over models produced by the AI procedures Support



Vector Machine and Relevance Vector Machine. The structure predicts soil dampness  $n$  days ahead dependent on a similar soil and natural characteristics that can be gathered by our sensor hub.

[4] **Sabri Arik, et al.(2016)** In this paper, we propose a method for predicting functional properties of soil samples from a number of measurable spatial and spectral features of those samples. The method used is based on Savitzky-Golay filter for pre-processing and a relatively recent evolution of single hidden-layer feed-forward network (SLFN) learning technique called extreme learning machine (ELM) for prediction.

[5] **Vaneesbeer Singh, et al. (2017)** This work presents an approach which uses different Machine Learning techniques in order to predict the category of the yield based on macro-nutrients and micro-nutrients status in dataset. The dataset considered for the crop yield prediction was obtained from Krishi Bhawan (Talab-Tillo) Jammu. The parameters present in the data are Macro-Nutrients (ph, Oc, Ec, N, P, K, S) and Micro Nutrients (Zn, Fe, Mn, Cu) present in samples collected from different regions of Jammu District. After analysis Machine learning algorithms are applied to predict the category of yield. The category, thus predicted will specify the yield of crops. The problem of predicting the crop yield is formulated as Classification where different classifier algorithms are used.

[6] **E.Manjula et al.(2017)** This paper chooses Nitrogen, Phosphorus, Potassium, Calcium, Magnesium, Sulphur, Iron, Zinc, and so forth, nutrients for investigating the soil supplements utilizing Naïve Bayes, Decision Tree and hybrid approach of Naïve Bayes and Decision Tree. The performance of the classification algorithms are compared based on accuracy and execution time.

[7] **Rohit Kumar Rajak et al. (2017)** This method is characterized by a soil database collected from the farm, crop provided by agriculture experts, achievement of parameters such as soil through soil testing lab dataset. The data from soil testing lab dataset given to recommendation system it will use the collect data and do ensemble model with majority voting technique using support vector machine and ANN as learners to recommend a crop for a site specific parameter with high accuracy and efficiency.

**Problem Definition:** The works done till now only concentrated on crop prediction using different soil properties and Data Mining Techniques. Fertilizer Recommendation is not taken into consideration. So, it is necessary to develop crop yield prediction and fertilizer recommendation system which predicts crop yield based on soil nutrients crop yield data and recommend fertilizer for selected crop based on different datasets like fertilizer data, location data and crop yield data.

### III. PROPOSED WORK

The aim of proposed system is to help farmers to cultivate crop for better yield. The crops selected in this work are based on important crops from selected location. The selected crops are Rice, Jowar, Wheat, Soyabean, and Sunflower, Cotton, Sugarcane, Tobacco, Onion, Dry Chili etc. The dataset of crop yield is collected from last 5 years from different sources.

There are 3 steps in proposed work.

#### 1) **Soil Classification:**

Soil classification can be done using soil nutrients data. Two Machine learning algorithms used for soil classification are Random Forest and Support Vector Machine. The two algorithms will classify, and display confusion matrix, Precision, Recall, f1-score and average values, and at the end accuracy in percentage as output.

#### 2) **Crop Yield Prediction:**

Crop Yield Prediction can be done using crop yield data, nutrients and location data. These inputs are passed to Random Forest and Support Vector Machine algorithms. These algorithms will predict crop based on present inputs.

#### 3) **Fertilizer Recommendation:**

Fertilizer Recommendation can be done using fertilizer data, crop and location data. In this part suitable crops and required fertilizer for each crop is recommended.

- Third Party applications are used to display Weather information, Temperature information as well as Humidity, Atmospheric Pressure and overall description.

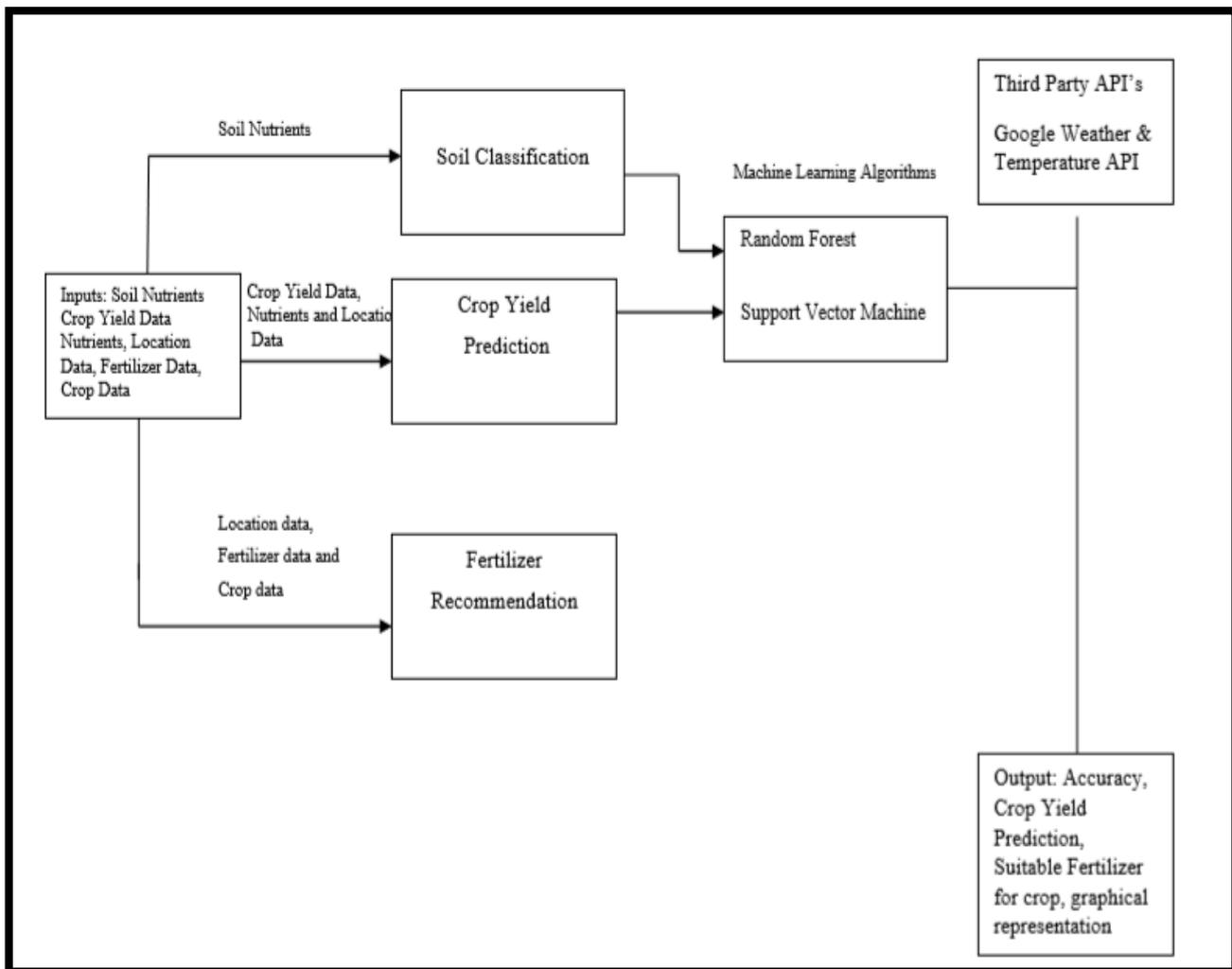


Fig 1. System Architecture

### Support Vector Machine

SVM develops a hyperplane or set of hyper planes in a high-or boundless dimensional space, which can be utilized for characterization, relapse, or different errands. Naturally, a great partition is accomplished by the hyperplane that has the biggest separation to the closest preparing information purpose of any class, since by and large the bigger the edge the lower the speculation blunder of the classifier. The computational burden have to be reasonable, the mappings are utilized by the SVM plan to guarantee the tiny items will be figured as far as the variable in the first degree, for that a bit capacity  $k(x, y)$  chose to get the ideal computational time.

#### Advantages

- 1) SVM calculation has a regularization parameter, which stays away from over-fitting.
- 2) SVM calculation utilizes the portion trap, so you can construct master learning about the issue.

### Random Forest

Random forest is a supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithm of the same type. Random Forest algorithm can be used for classification and regression problems.

#### Advantages

- 1) The random forest algorithm is not biased, since, there are multiple trees and each tree is trained on a subset of data.
- 2) Random Forest algorithm is stable if a new data point is introduced in the dataset the overall algorithm is not affected.

## IV. EXPERIMENTAL RESULTS

The complete system is designed using Python. Different datasets like crop, crop yield dataset, Location, soil and crop nutrients, fertilizer datasets are gathered from other sources like agricultural books, agricultural websites. The



Figures (Fig 2) shows soil classification using Random Forest algorithm and Support Vector Machine. The output of these algorithms shows confusion matrix as summary of

algorithms different parameters like Precision, Recall averages and accuracy in percentage.

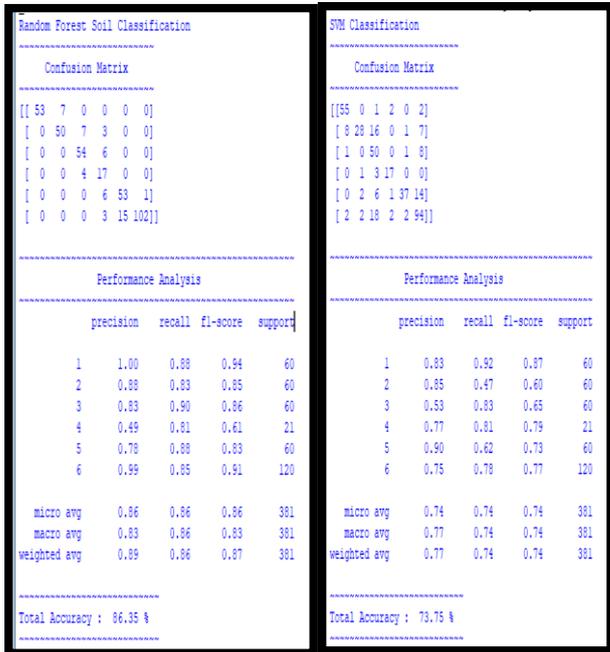


Fig 2. RF and SVM Classification

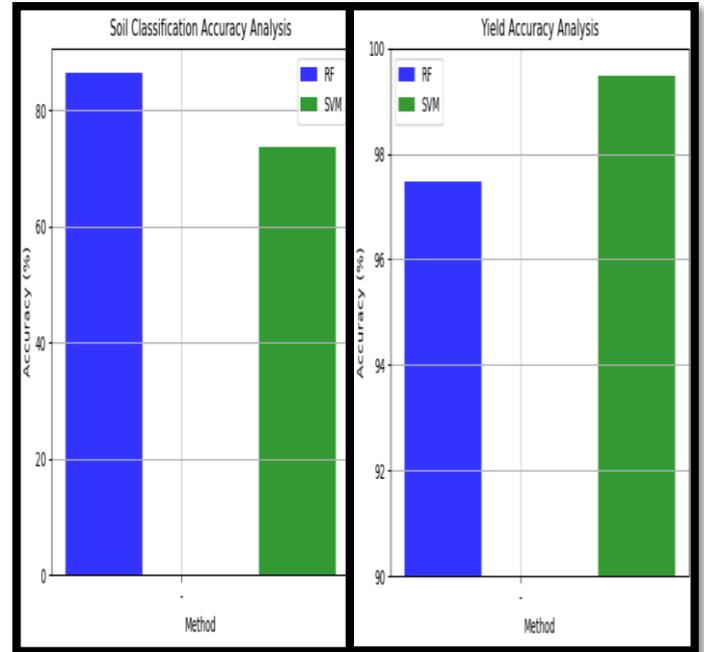


Fig 3. Soil Classification and Crop Yield Analysis

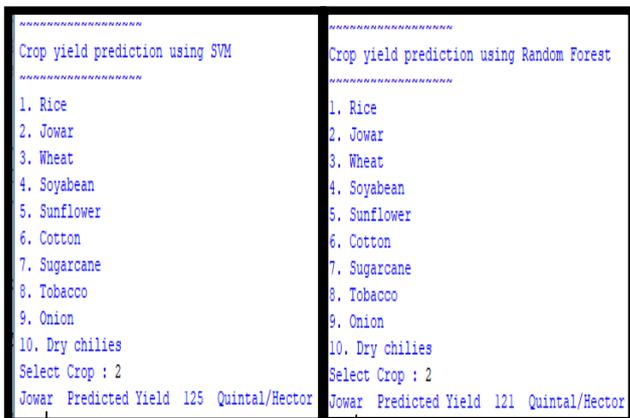


Fig 4. Yield Prediction using RF and SVM

The Figure 3 shows Soil Classification and Crop Yield analysis by graphical representations. The accuracy of Random Forest is 86.35% and Support Vector Machine is 73.75% so Random Forest algorithm is good for Soil Classification. The accuracy of SVM algorithm is 99.47% for yield prediction and RF accuracy is 97.48%. So, for crop yield prediction SVM algorithm is good.

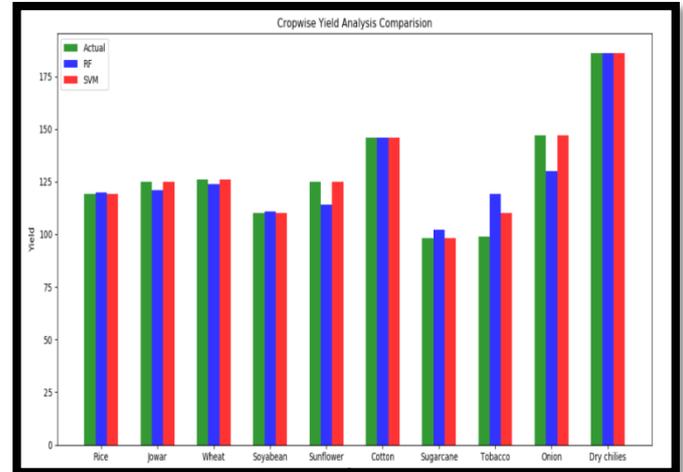


Fig 5. Crop Wise Yield Analysis

The Figure 4 shows Yield Predicted by both algorithms like if Jowar is selected then Random Forest predicts the “Jowar Predicted Yield 121 Q/H” and SVM predicts “Jowar Predicted Yield is 125Q/H”.

The Figure 5 Shows Crop wise yield analysis where each crop yield is taken into consideration.



Crops	Actual	RF	SVM	RF_Err	SVM_Err	RF_Ac	SVM_Ac
Rice	119	120	119	0.41841	0	99.5815	100
Jowar	125	121	125	1.62602	0	98.3739	100
Wheat	126	124	126	0.8	0	99.2	100
Soybeans	110	111	110	0.45249	0	99.5475	100
Sunflower	125	114	125	4.60251	0	95.3974	100
Cotton	146	146	146	0	0	100	100
Sugarcane	98	102	98	2.0	0	98	100
Tobacco	99	119	110	9.17431	5.26315	90.825	94.7368
Onion	147	130	147	6.13718	0	93.862	100
Dry chilli	186	186	186	0	0	100	100
<b>Avg Accuracy</b>						<b>97.48</b>	<b>99.47</b>

Table 1. Crop Yield Analysis

```

Location : Belgaum
Temperature (in degree unit) = 23.0
atmospheric pressure (in hPa unit) = 1009
humidity (in percentage) = 88
description = broken clouds

Suitable Crops and Required Fertilizer for Location Belgaum
1 Rice
    Amonium Sulphate
    Single Superphosphate
2 Jowar
    Amonium Sulphate
3 Corn
    Amonium Sulphate
    Groundrock Phosphate
4 Wheat
    Amonium Sulphate
    Groundrock Phosphate
5 Cowpeas
    Amonium Sulphate
    Groundrock Phosphate
6 Soyabean
    Amonium Sulphate
    Groundrock Phosphate
7 Peanut
    Amonium Sulphate
    Single Superphosphate
8 Sunflower
    Amonium Sulphate
9 Cotton
    Amonium Sulphate
    Groundrock Phosphate
10 Sugarcane
    Amonium Sulphate
    Single Superphosphate
11 Chilli
    Amonium Sulphate
    Single Superphosphate
12 Onion
    Amonium Sulphate
    
```

Fig 6. Fertilizer Recommendation

In fertilizer recommendation (Figure 5) suitable crops and required fertilizer for each particular crop is recommended. Also first it displays weather information of selected location first which is displayed by third party API's.

**V. CONCLUSION AND FUTURE SCOPE**

The prediction of crop yield based on location and proper implementation of algorithms have proved that the higher crop yield can be achieved. From above work I conclude that for soil classification Random Forest is good with accuracy 86.35% compare to Support Vector Machine. For crop yield prediction Support Vector Machine is good with

accuracy 99.47% compare to Random Forest algorithm. The work can be extended further to add following functionality. Mobile application can be build to help farmers by uploading image of farms. Crop diseases detection using image processing in which user get pesticides based on disease images. Implement Smart Irrigation System for farms to get higher yield.

**VI. ACKNOWLEDGEMENT**

I would like to express my special thanks of gratitude to my guide Mr. Santosh Mahagaonkar, who gave me the opportunity to do this wonderful research work on the topic



Prediction of Crop Yield and Fertilizer Recommendation using Machine Learning Algorithms which also helped me in doing lots of research and I came to know many things I am really thankful to him and even my friends.

#### VI. REFERENCES

- [1] Rushika G., Juilee K, Pooja M, Sachee N, and Priya R.L.(2018). Prediction of Crop Yield using Machine Learning, Issue 02 IRJET (pg 2337-2339).
- [2] Ruchita T, Shreya B, Prasanna D, and Anagha C (2017). Crop Yield Prediction using Big Data Analytics, Volume 6, Issue 11, IJCMS.
- [3] Monali P, Santosh K, Vishwakarma, and Ashok V (2015). Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach, ICCICN 2015.
- [4] Mrs. N. Hemegeetha, and Dr. G.M. Nasira (2016). Analysis of Soil Condition based on pH value using Classification Techniques, IOSR-JCE, Issue 6, (pp.50-54).
- [5] E. Manjula, and S. Djodiltachoumy (2017). Data Mining Technique to Analyze Soil Nutrients based on Hybrid Classification. IJARCS.
- [6] Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare (2017). Crop Recommendation System to maximize Crop yield using Machine Learning. Issue 12 IRJET.
- [7] S. Veenadhari, Dr. Bharat Misra, Dr. CD singh (2014). Machine Learning Approach for Forecasting Crop Yield based on Climatic Parameters. ICCCI-2014
- [8] Sadia A, Abu Talha K, Mahrin Mahia, Wasit A, and Rashedur M.R.(2018). Analysis of Soil Properties and Climatic Data To Predict Crop Yields and Cluster Different Agricultural Regions of Bangladesh, IEEE ICIS 2018 (pp.80-85).
- [9] Arun K, Navin K, and Vishal V (2018). Efficient Crop Yield Prediction using Machine Learning, (pp.3151-3159), IRJET.
- [10] Subhadra M, Debahuti M, Gour, H. S.(2016). Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper, Vol 9(38), DOI:10.17485/ijst/2016/v9i38/95032 IJST.
- [11] Mrs. Prajakta P.B., Yogesh S. P, and Dinesh D.P. (2017). Improved Crop Yield Prediction using Neural Network, IJARIE, (pg.3094-3101).
- [12] Dhivya B H, Manjula R, Siva Bharathi S, and Madhumathi R (2017). A Survey on Crop Yield Prediction based on Agricultural Data, DOI:10.15680/IJRSET.2017.0603053, IJRSET.
- [13] Omkar B, Nilesh M, Shubham G, Chandan M, and D.S. Zingade (2017), Crop Prediction System using Machine Learning, Volume 4, Special Issue 5, IJAERD.
- [14] Sneha N, Dr. Jharna Majumdar (2017). Big Data Application in Agriculture to Maximize the Rice Yield Crop Production using Data Mining Technique DOI: 10.15680/IJRCCCE.2017. 0505045, IJRCCCE.
- [15] S.Bhanumathi, M.Vineeth and N.Rohit (2019). Crop Yield Prediction and Efficient use of Fertilizers, (pg.0769-0773), ICCSP.
- [16] Niketa G, Leisa A.J., Owaiz Petkar, and Amiya K.T.(2016). Rice Crop Yield Prediction in India using Support Vector Machines 978-1-5090-2033-1/16/\$31.00 ©2016 IEEE.