International Journal of Research in Engineering, Science and Management
Volume-3, Issue-5, May-2020
www.ijresm.com | ISSN (Online): 2581-5792

Agriculture Crop Prediction Using Machine Learning Algorithms

N. P. Ganapathi[1*], K. Sudarshan[2], Ananthakrishna Bhatta[3]
[1,3]*Student, Department of Computer Science & Engineering, Srinivas Institute of Technology, Mangaluru, India*
[2]*Associate Professor, Department of Computer Science & Engineering, Srinivas Institute of Technology, Mangaluru, India*
*Corresponding author: sudhinganapathi@gmail.com

***Abstract*: India is generally an agricultural country. Agriculture is the single most important provider to the Indian economy. Agriculture crop production depends on the season, organic, and monetary cause. The prognostication of agricultural yield is challenging and pleasing task for every nation. Nowadays, Farmers are hostile to produce the yield because of erratic climatic changes and scarcity of water resource. The consequence of cultivating crops which are not suitable to the current conditions not only affect economic structure but also affect future crop growth. The main objective is collecting agricultural data which can be stored and analyzed for useful crop yield forecasting. To predict the crop yield with the help of data mining technique, advanced methods can be introduced to predict crop yield and it also helps the farmer to choose the most suitable crop, thereby improving the value and gain of the farming area. The increase in crop productivity ultimately enhances country's stature in all aspect.**

***Keywords*: Prediction, Dataset collection, Support Vector Machine, Random Forest.**

1. Introduction

Agriculture is one of the important industrial sectors in India and the country's economy is highly dependent on it for rural sustainability. Due to some factors like climate changes, unpredicted rainfall, decrease of water level, use of pesticides excessively etc. The level of agriculture in India is decreased. To know the level of production we performed descriptive analytics on the agriculture data. The main objective of this research work is to provide a methodology so that it can perform descriptive analytics on crop yield production in an effective manner. Although, some studies revealed statistical information about the agriculture in India, few studies have investigated crop prediction based on the historic climatic and production data. The system developed is a supervised based model. And it will work as mixed approach it means classification technique as well as regression technique. In this project the crop yield classification will perform to categorize on the basis of yield productivity and class labels will be low, mid, and high and range of productivity will be defined and regression will be performed to get the actual crop yield estimated cost.

2. Problem Statement

The production of agriculture is affected by several climate factors. Like as metrological parameters (Humidity, wind speed, temperature, and moisture), precipitation parameters (rainfall, region wise rainfall, irrigation etc.), and soil parameters (PH, organic carbon, phosphorus, fiber etc.). And due to continuously change in climate condition everything is messed. In India farmers still follow the traditional technology which they adopted from their ancestors. But the problem is that in earlier times climate was very healthy everything happened on time. But now most of the things have been changed due to global warming and many other factors. The main problem with agriculture in India is lack of rainfall in seasonal time. Humidity is also necessary for crops but it has been excessive, it also converts as drawback. In this research work, we have proposed a system which is based on descriptive analytics. By which farmers can know what happened in past time and what is going to happen. So here we collect several data from agriculture production, rainfall and soil data and prepare their respective datasets.

3. Methodology

Prediction of agricultural crops involves an approach to analyze large data set. When implementing a more accurate prediction model it might not be sufficient to just consider one or two parameters. Data about weather, irrigation, and yield from several other sources (e.g. meteorological station and irrigation-plan records) for past few decades are collected and analyzed to produce an output which has the highest productivity of each grains in their respective geographical conditions. Simultaneously, the data about weather, soil information, Rainfall, Land area etc. are recorded. From these records using SVM and Random Forest algorithms system can evaluate the perfect crop for the current geographical conditions [5]. Support vector machines (SVM) is set of supervised learning strategies used for classification, and regression. It's a classification technique [2]. Random forests is an ensemble learning algorithm. The basic premise of the algorithm is that building a small decision tree with few features is a

computationally cheap process [1]. Our main aim is to predict best suitable crop which gives better yield and better profit for farmers.

A. Dataset Collection

In this phase, we collect data from various sources and prepare datasets. And these datasets are used for analytics (descriptive and diagnostic). There are several online abstracts sources such as Raithamithra.karnataka.gov.in and Data.gov.in [2]. We will use annual abstracts about a crop for at least ten years' period. Collecting previous crop history data from places like Mangalore, Kodagu, Kasaragod, Mysore, Davangere, Hassan, Shivamogga, Chikkamagalur which belongs to Karnataka State. Collecting data related to crops like Coconut, Cardamom, Coffee, Areca nut, Ginger, Tea, Paddy, ground nut, Black gram, Cashew, Pepper are the crops which are commonly grown in these regions. We also collect data related to Rainfall. Humidity, Soil type, Irrigation type, Previous Yields, Location, Price, Year, type of crop, Crop diseases and its symptoms.

B. Learners Used in the Model

1) Support Vector Machine

Support vector machines (SVM) is set of supervised learning strategies used for classification, and regression. It's a classification technique. During this algorithmic rule, we have a tendency to plot every information item as some extent in n-dimensional house (where n is variety of options you have) with the worth of every feature being the worth of a selected coordinate. A Support Vector Machine (SVM) is discriminative classifier correctly bounded by a separating hyperplane. In alternative words, given labeled coaching information (supervised learning), the algorithmic rule outputs associate degree best hyperplane that categorizes new examples. Support vector simple machine (SVM) may be a set of supervised learning strategies used for classification, regression and helps in better prediction [1] [2]. Accuracy of Agricultural crop prediction using Support Vector Machine (SVM) algorithm is shown in the accuracy graph obtained using our system.

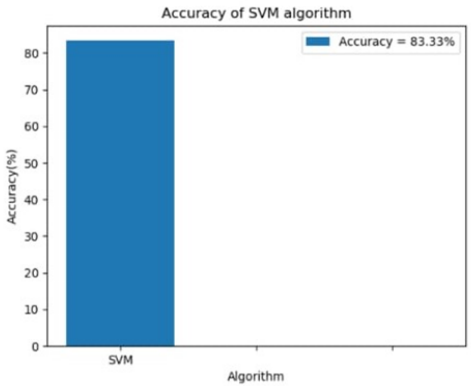

Fig. 1. Accuracy of SVM algorithm

2) Random Forest

Random forest is an ensemble learning algorithm. The basic premise of the algorithm is that building a small decision tree with few features is a computationally cheap process. If we can build many small, weak decision trees in parallel we can combine the trees to form a single, strong learner by averaging or taking the majority vote. In practice, random forests are often found to be the most accurate learning algorithms to date. The random forest algorithm uses the bagging technique for building an ensemble of decision trees. Bagging is known to reduce the variance of the algorithm. In traditional bagging with decision trees, the constituent decision trees may end up to be very correlated because the same features will tend to be used repeatedly to split the bootstrap samples. By restricting each split-test to a small, random sample of features, we can decrease the correlation between trees in the ensemble. Furthermore, by restricting the features that we consider at each node, we can learn each tree much faster, and therefore, can learn more decision trees in a given amount of time. Thus, not only can we build many more trees using the randomized tree learning algorithm, but these trees will also be less correlated. For these reasons, random forests tend to have excellent performance [1]. Accuracy of Agricultural crop prediction using Random Forest (RF) algorithm is shown in the accuracy graph obtained using our system.

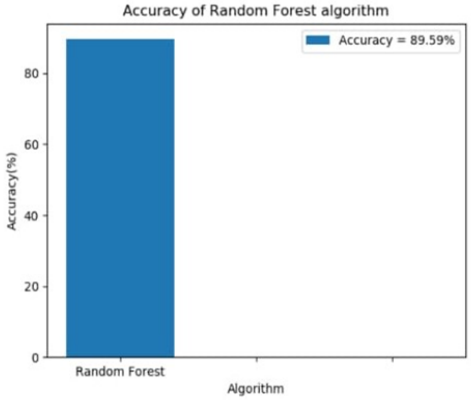

Fig. 2. Accuracy of RF algorithm

4. Proposed System

The proposed system takes into consideration the data related to rainfall. humidity, soil type, irrigation type, previous yields, location, price of crop, past year production, type of crop, crop diseases and its symptoms and suggests which are the best profitable crops which can be cultivated in the appropriate environmental condition. Our system helps in avoiding the use of sensors and reduce safeguards future profitability unnecessary cost. Thus, our system results in efficient usage of time and cost. A key aspect of Crop Prediction is to identify a suitable crop quickly and suggest the farmer as to which crop to grow. Our system helps in gathering all necessary information and giving a model of output which not only increases current economic gain but also safeguards future profitability.

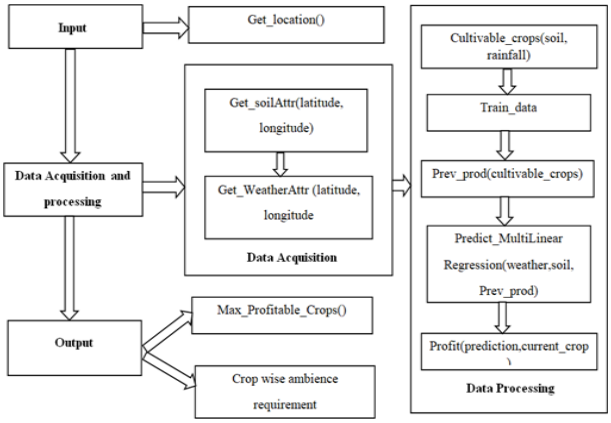

Fig. 3. System architecture of agricultural crop prediction using machine learning algorithms

5. Conclusion

The developed system is employed to gain knowledge about the crops that can be deployed to make efficient and useful harvesting. The developed system is not only cost-effective but also takes less time to execute and renders result in quick time. Although the system uses SVM and RF algorithm for making the prediction, in the back-end, we can further implement more accurate machine learning algorithm and use this system for better crop prediction. In future, by using latest technology, the system can be made more accurate and efficiency of the system can also be increased. The system can be implemented using more advanced machine learning algorithms for making better and stable predictions.

References

[1] Rohit Kumar Rajak and Ankit Pawar, "Crop recommendation system to maximize crop yield using machine learning technique", *International Research Journal of Engineering and Technology*, Volume 4, Issue 12 Dec. 2017.

[2] Arun Kumar and Naveen Kumar "Efficient Crop Yield Prediction Using Machine Learning Algorithm", *International Research Journal of Engineering and Technology*, Volume 5, Issue 6, June 2018.

[3] Shailesh S, "Crop Analysis and Profit Prediction using Data Mining Techniques", *IRJET*, March, 2019.

[4] S. R. Rajeswari, "Smart Farming Prediction Using Machine Learning", *International Journal of Innovative Technology and Exploring Engineering*, May, 2019.

[5] P. Priya "Predicting Yield of the Crop Using Machine Learning Algorithm" *IJESRT*, April, 2018.