EasyChair Preprint

№ 2413

Krishak Sahayata: Prediction of Best Crop Yield

Gresha Bhatia, Urjita Bedekar, Simran Bhagwandasani,
Rahul Bhatia and Pranit Naik

January 18, 2020

KRISHAK SAHAYATA : PREDICTION OF BEST CROP YIELD

Dr. Mrs. Gresha S Bhatia, Urjita Bedekar, Simran Bhagwandasani, Rahul Bhatia and Pranit Naik

Vivekanand Education Society's Institute of Technology,
Mumbai, India

Abstract. Agriculture is one of the most critical occupations practiced in our country. It is an economic sector that plays an important role in the overall development of the country. Thus, the modernization of agriculture is significant and thus will lead the farmers of our country towards profit. Earlier sowing of crops was performed by considering the farmer's knowledge in a particular field and crop. However, as the weather conditions change very rapidly, farmers cultivate more and more crops that do not give an expected yield. Being this as the current situation, many of them do not have enough knowledge about the new crops and are not entirely aware of the benefits they get while farming them. Also, farm productivity can be increased by understanding and forecasting crop performance in a variety of environmental conditions. Thus, the proposed system takes the location of the user as an input. From the location, the nutrients of the soil, such as nitrogen, phosphorous, potassium is obtained. The processing part also takes into consideration two more datasets i.e. one obtained from the weather department, forecasting the weather expected in the current year, and the other data being static. This static data is the crop production and data related to the demands of various crops obtained from various government websites. The proposed system applies Machine Learning algorithms to identify the pattern among data and then process it as per input conditions which in turn, will propose the best feasible crops according to given environmental conditions. Thus, this system will only require the location of the user, and it will suggest several profitable crops providing a choice directly to the farmer about which crop to cultivate. Since historical data is also taken into account, the output is expected to be much more accurate.

Keywords. Machine Learning, Yield Prediction, Indian Agriculture.

I. INTRODUCTION

Agriculture in India has a full-size history. More than 60 % of the land in the country is used for agriculture to cater to the needs of 1.2 billion people. Nowadays, India is ranked second worldwide in farm output. Agriculture and other sectors like forestry and fisheries accounted for 16.6 percent of GDP in 2009 [1][2]. The monetary contribution of agriculture to India's GDP is regularly declining. The production of crops relies on different factors like climatic, geographical, and organic elements. Accurate statistics about the character of an ancient yield of the crop is a critical modeling input, which is useful to farmers and authorities organization for decision-making techniques in establishing the right policies associated with subsequent manufacturing. The advances in computing and information storage have mainly provided at the maximum of information. The project has been to extract expertise from these statistics mining that may be able to bridge the understanding of the facts to the crop yield estimation. This task aimed at statistics mining strategies and follow them to the various variables consisting inside the database to set up if significant relationships may be discovered and the usage of fuzzy common sense to discover the circumstance of crops on a diverse situation of climatic conditions.

II. LITERATURE SURVEY

Literature survey is the most important part which guided us in our research. We had a look at previous papers aiming same objectives which we have and on the basis of study we plan to improve the limitations, challenges faced and various drawbacks analysed from those papers. Referring to the work done by Kunal Teeda, Nandini Vallabhaneni and Dr. T.Sridevi, [3] we understood that climatic analysis based on location, rainfall, soil type, yield capacity, irrigation techniques is a fundamental prerequisite that needs to be carefully implemented to plan better farming structures for a satisfactory yield. Farming risks have greatly reduced with the advent of technologies. In order to predict the weather parameters, empirical approaches like regression, artificial neural network, fuzzy logic and institution approach of statistics have been employed. For prediction of rainfall, clustering and classification methods of data mining have been used. The work done by D.S Zingade [4] is spectacularly presented. They have considered significant parameters mainly soil contains, water level and temperature. For the soil measure they have considered nitrogen, phosphorus and potassium along with the pH value of the soil. Taking Zinc and Sulphur nitrate into consideration along with these can help to improvise the system. Moreover, we have chosen to perform classification using random forest classifier which overfits the linear regression used above, since the parameters are quite significant in number. Analysing the Farmer's handbook published by Desai Fruits, [5] we have an understanding of choosing the sufficient parameters required for selecting the type of crop suitable for the cultivation. Project by P Priya, published in International Journal of engineering science and research technology have performed the prediction of crop yield Focusing only Tamil Nadu district[6]. We Intend to build a system that satisfies majority places of possible farming in India. Random Forest Algorithm has worked with great accuracy amongst all the other algorithms.

III. CHALLENGES IN EXISTING SYSTEMS

IoT based Plants : For constant analysis of the fields and crop growth, many projects have obtained to install IoT devices for data gathering, which are way costlier and require regular maintenance.[7]

Web Interface : Many projects have their UI based on python which are hosted on web server, our Mobile application is very much user friendly and we intend to bring the information available in the locale languages which are region specific.[8]

Parameters Under Consideration : The existing projects have a limited number of parameters considered for training. We intend to consider majority of parameters which are accountable for the plants growth. [9]

IV. PROPOSED PROBLEM STATEMENT

Accurate yield estimation would be essential for farmers to sow the crops accordingly. Machine Learning Systems are being widely used in building decision support tools for the contemporary farming system to improve yield production while reducing operational cost and environmental impact. The capability of Machine Learning algorithm to access tons of data and mold them into relatable format helps in predicting the future results when a particular test case is provided to the trained system, thereby it would provide a great support in predicting the estimated yield of a particular crop provided with sowing area, time, climatic conditions of that particular crop. Our system proposes to use an algorithm which will work accurately even when the number of parameters under consideration are more.

V. PROPOSED SYSTEM

Since other systems developed are using Data Mining techniques to predict the optimal crop yield, our system aims to provide an accurate yet effective result based on more number of parameters dependent on climatic conditions to provide maximum profit, reduced risks and more than expected yield to the farmers. The proposed solution is represented in figure 1 and is described as follows :

Input : The prediction of the crop is dependent on various factors such as weather and past crop production to predict the crop correctly. The location of the user is taken as an input, which helps to get the weather conditions like humidity, temperature.

Data Acquisition : Depending on user location, the system mines the weather conditions like humidity, the temperature in the respective area from the other datasets. In a similar approach, earlier crop yield is extracted from previous data. This would give an excellent idea about the demands of crops in the location

Output : The random forest algorithm creates a forest with a number of trees. The most suitable crop is predicted by the system using the Random Forest Algorithm, and the user is provided with the best-suited crop for a given set of climatic conditions.

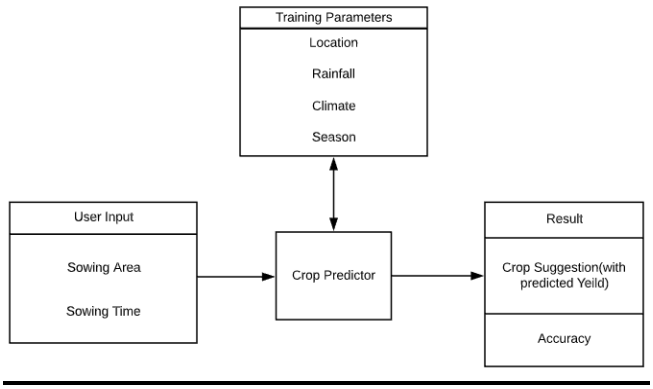

Fig .1. BLOCK DIAGRAM OF FARMER'S GUIDE

Sowing area and sowing time will be entered by user as parameters. Our classification algorithm will be analyzing climatic conditions such as rainfall based on the location and sowing time. Algorithm will then suggest best crop to plant for getting the maximum yield and this can be used by farmers to acquire maximum profit as observed in figure 2.

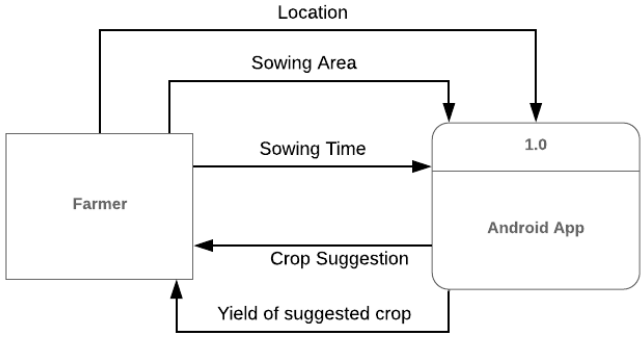

Fig .2. DATA FLOW DIAGRAM

The farmer provides location access to the android application and also provides the inputs such as the sowing area and expected sowing time. The Android Application then suggests the most suitable crop along with the expected yield.

VI. EVALUATION MEASURES

Accuracy : Accuracy can be defined as the degree to which the result of a measurement, calculation, or specification conforms to the correct value or a standard. The Accuracy for a system is defined with the formula given as : (TP / TN) / (P + N).[10]

Error Rate : Error rate refers to the frequency of errors occurred, defined as "the ratio of total number of data units in error to the total number of data units transmitted." The Error Rate can be calculated by using the formula : (FP+FN)/(P+N) or (1 - Accuracy).[11]

Sensitivity : Sensitivity measures the proportion of actual positives that are correctly identified as such. The sensitivity of a system can be defined using the formula : (TP)/(P)[12].

Precision : Precision or positive predictive value, is the fraction of relevant instances among the retrieved instances. The precision of a system can be calculated using the formula : (TP)/(TP+FP)[13].

VII. OUTPUTS OBTAINED

The parameters that we have selected to evaluate our system include accuracy, error rate and precision. Since the system is being evaluated on parameters like location, rainfall, season and climate, the quality of being precise is about 87.88% for the system developed. Having an error rate of 12.12% in a dataset of 500 trees .

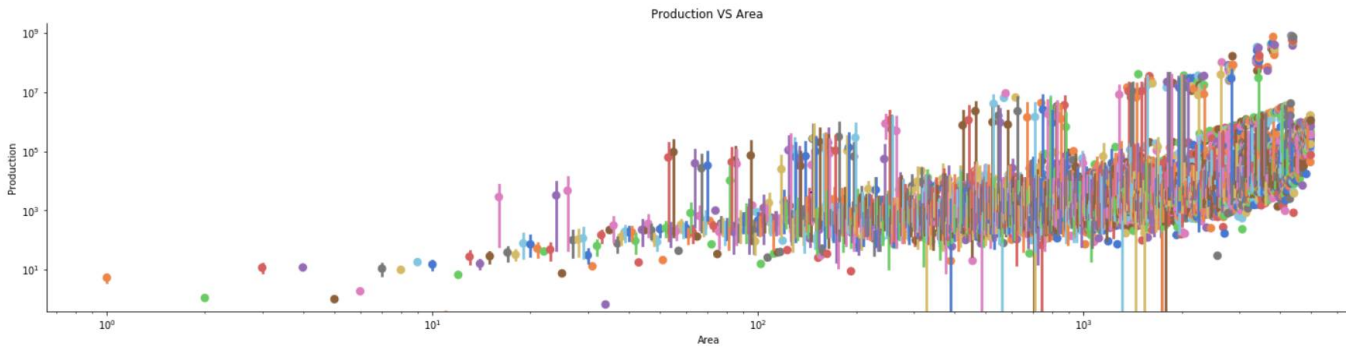

Fig. 3. Plot for Yield vs Area(Insight from data)

VIII. CONCLUSION AND FUTURE WORK

The proposed system takes the data related to weather, and past year production and suggests which are the most suitable crops that can be cultivated in environmental conditions of various parts of India. As the system lists out the best-possible crop, this would help the farmer in the decision making of which crop is appropriate to cultivate. Since this system takes into consideration past production, it will help the farmer get insight into the demand of crops cultivated at the stipulated time in the market. The

main aim of this research paper is to provide a methodology so that it can perform descriptive analytics on crop yield production in an effective manner. The future scope of the project is to collect soil nutrients for every particular piece of land and combine the trained system and data about soil nutrients to get a cumulative output on the basis of all these factors.

IX. REFERENCES

[1]http://www.expo2015.org/magazine/en/economy/agriculture-remains-central-to-the-world-economy.html

[2] https://en.wikipedia.org/wiki/Agriculture_in_India

[3]Analysis of Weather Attributes to Predict Crops for the Season Using Data Mining,Kunal Teeda1 Nandini Vallabhaneni2 Dr.T.Sridevi3

[4]Machine Learning based Crop Prediction System Using Multi Linear Regression, Prof D.S Zingade

[5]https://www.manage.gov.in/publications/farmerbook.pdf

[6]PREDICTING YIELD OF THE CROP USING MACHINE LEARNING ALGORITHM,P.Priya, U.Muthaiah & M.Balamurugan

[7]https://www.iofficecorp.com/blog/cost-of-iot-sensors

[8]https://www.quora.com/How-many-Indian-farmers-use-the-internet

[9]https://www.manage.gov.in/publications/farmerbook.pdf

[10]https://www.researchgate.net/post/How_can_I_calculate_the_accuracy

[11]https://classeval.wordpress.com/introduction/basic-evaluation-measures/

[12]http://gim.unmc.edu/dxtests/reviewof.htm

[13]http://gim.unmc.edu/dxtests/reviewof.htm