iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: VI Month of publication: June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6119

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

Data Analytics Model for Optimal Crop Management

Gaurav Sharma[1], Akhilesh Tiwari[2]
[1, 2]*Department of CSE & IT, Madhav Institute of Technology & Science, Gwalior (M.P), India*

Abstract: ***Agriculture is the largest domain of subsistence especially in India. A huge amount of the population depends on agriculture for their daily wages. Due to variation in bio-diversity, Huge amount of data generated every year. At the rural level, the knowledge of harvesting a crop at its most efficient level according to the various factors is much low. The alternatives for the poor environmental and nutritional state are not known. Due to this, the crop production rate is decreasing and so the yield. Considering parameters like Rainfall, Soil type, Soil PH, Soil Nutrient Value (P, N, K), etc. identification of parameters that causes low production of a certain crop is possible. There are certain scientifically and experimentally proven rules for every crop. Integrating those rules into a knowledge base and validating with the input defected parameter the designed model can be able to recommend the better and optimized crop production parameters. This intends to improve the harvested crop as well as for future crop planning management.***
Keywords: Data Mining, Data Analytics, Knowledge base, Crop Planning, Crop Management

I. INTRODUCTION

In India, a major part of the population depends on Agriculture. Approximately 70 percent of the total rural area of the country uses agriculture as their primary source of livelihood. Agricultural Lands of India reflects various behavior. Diversity in the geographical and ecological environment plays a vital role in overall crop production throughout the country. Increasing crop Production is always been a crucial task. Every crop has its certain features that hold some geographical, environmental and nutritional requirements.

In Agriculture, approximately 4 to 6 crop-related data generated every year, there is a huge amount of data generated quarterly as well as annually. The data generated includes annual crop production, Overall Yield, Fertilizer information, diseases, soil type, etc. For Agriculture, some Agri - meteorological data such as temperature, rainfall, humidity, etc. can also be used for prediction and rule generation purposes.

There are various methods and technologies available for various tasks in the agricultural field. Data mining is one of the most applied practices for achieving the specified intended objective. Data mining facilitates the exploration of intention - specific knowledge from the large raw heterogeneous data.

The data further can be analyzed for the prediction of overall production of crop that can be used by farmers for optimizing crop production. Knowledge generated after mining can also be used by respective government bodies for making plans and policies that intend to increase the overall crop production.

II. LITERATURE REVIEW.

A. Yogesh gandge and Sandhya perform a study on several data mining approaches for crop yield prediction. In the paper, there are various approaches applied to the different crop data for prediction.

B. A. T. M. Shakil Ahamed, Navid Tanzeem Mahmood, et. al. tested few data mining techniques for prediction of the annual yield of major crops in Bangladesh. In this paper, the clustering technique is used to predict the results. There are parameters such as temperature, humidity, minimum temperature, maximum temperature, average sunshine, Soil PH and salinity are used to predict the Annual crop production. K- means clustering is used by the author for recommending plant crops in the districts of Bangladesh.

C. Niketa Gandhi and Leisa J. Armstrong determine the applicability of data mining approaches that improve decision support in agriculture. In this paper, the author explains the applicability of Artificial Neural Network(ANN), Bayesian Network(BN) and Support Vector Machines(SVM) in the field of agriculture.

III. DESCRIPTION

Data Analytics is the sequence of techniques and operations performs on the large raw data intends to observe, analyze and extraction of the beneficial knowledge. These techniques include Preprocessing, actual processing and post-processing sequences of operations. Data Mining is a subset of Data Analytics. Below given Fig 1 Shows the Generalized analytical model.

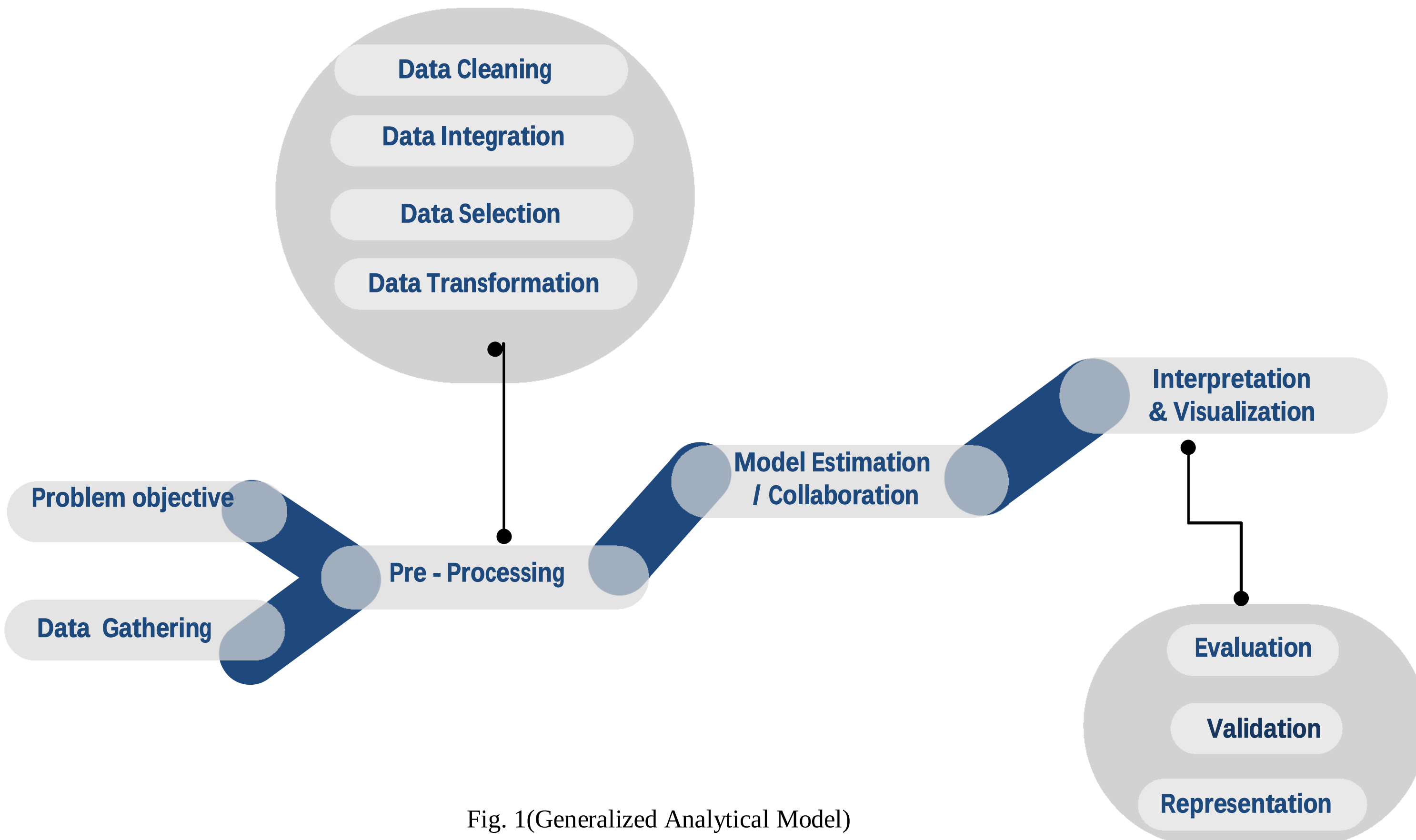

Fig. 1(Generalized Analytical Model)

The generalized analytical model contains steps such as Problem objective, Data Gathering, Preprocessing, Model estimation or model design collaborated with other required technologies and interpretation and visualization. The last step includes the evaluation, validation of results. It also includes the proper representation of the knowledge generated by the analytical model.

In Agriculture, there are varieties of data generated throughout the year. The data stored digitally is required a lot more preprocessing than other types of data. Agriculture-related data may contain missing as well as duplicate values. Before proceeding for actual mining the data is gone through 4 subphases of preprocessing like Data cleaning, Integration, selection of appropriate values and transformation in the processable form. The generalized model is capable to determine basic knowledge from the raw data. GAM (Generalized Analytical Model) provides model estimation according to the specific intention behind the knowledge extraction. The model can be integrated with current or upcoming data science techniques and approaches.

IV. RESEARCH METHODOLOGY

A. Data Set

or testing of this model raw agricultural data is taken. Other agricultural data can also be used as an input in this model. The data should be authentic and verified by any governmental organization or an agricultural research institute. The authentic data generates more accurate results and those results can lead towards the efficient generation of rules for optimized farming.

Parameters included are given below

1) Rainfall.
2) Soil Type.
3) Soil PH Value.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

4) Soil Nutrient Value (N, P, K).
5) Area Harvested
6) Crop Production
7) Crop Yield

B. *Theoretical Framework*

The Model proposed here works in 5 phases. Each phase has its own processes and subprocess includes.

The very first phase of the model includes preprocessing related tasks. Fig 2 shows the theoretical model for achieving optimized production.

1) *Preprocessing Phase:* The very first phase includes the proper cleaning of data, removal of missing and uncertain values.In this stage, the data from multiple sources are integrated into a single dataset with their corresponding values.For example in data set includes the crop-related data for a particular year and other data set includes the information of rainfall, temperature, and humidity related data of a certain area in a particular year.Then both of these data values merge together according to their respective fields.
2) *Classification Phase:* In the classification phase, the threshold is decided according to the production rate. According to the threshold, the data is classified into two classes i.e., High Production class and Low production class.The are of concern now is Low production class, contains the crop-related data for which production is below of certain applied threshold.
3) *Feature Assessment Phase:* After classification, low production class processed for every feature assessment.The detailed assessment of every feature related to a particular crop provides specific and detailed information regarding that crop.
4) *Collaboration Phase:* In the current phase, the knowledge base associated with the model comes into the role. The knowledge base already designed for every crop. It holds scientifically verified values for harvesting a crop. The output of Feature assessment now evaluated and validated with the knowledge base, and the result generated reports the parameters which are responsible for low production.
5) *Recommendation Phase:* The final phase recommends the calculated values for the parameters responsible for low productivity. The recommended values can be used for the current harvested crop as well as for future crop planning.

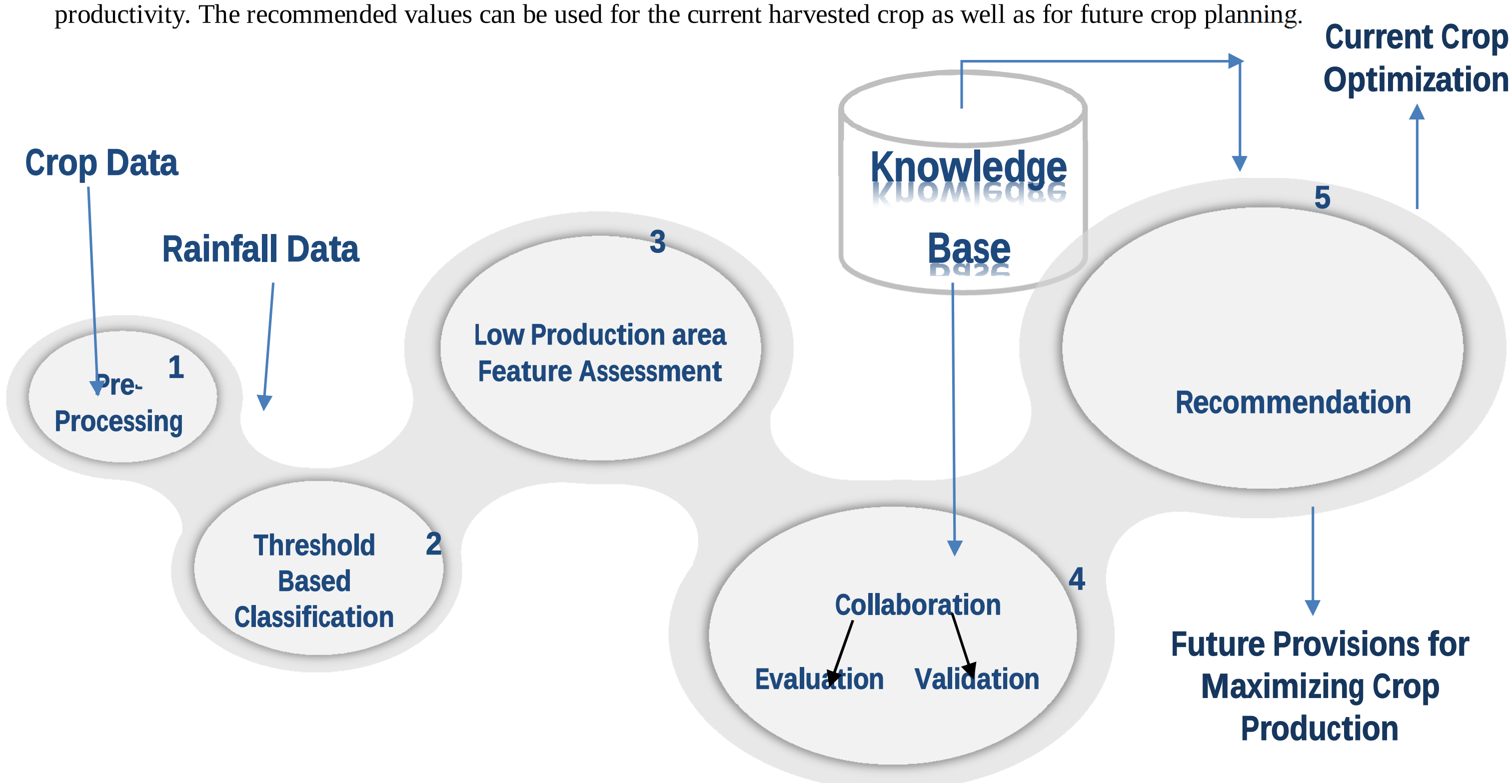

Fig 2 (Analytical Model for Crop Planning)

V. FUTURE WORK

It can easily be observed from the paper that there are yet many issues in the field of agriculture that are required to be addressed. The above-proposed model can also be integrated with various machine learning technologies to provide much more efficient results. The real-time data can be tested on the above model. By considering other parameters associated with the crop can also be integrated into the raw data set for better crop planning.

VI. CONCLUSION

The study shows that the data analytical model can provide a better result for the farmers. Not only the Current harvested crop production can be improved but the crop planning for the upcoming crop can also be managed properly. Considering all the parameters mentioned could be able to determine the lacking factors and so after, a better recommendation can be suggested to reduce the loss caused by the particular factor. The recommendation is authentic and trusted because of the validation itself compared by the scientifically and experimentally proven results of various research institutes.

REFERENCES

[1] Y. Gandge and Sandhya, "A study on various data mining techniques for crop yield prediction," 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, 2017, pp. 420-423. doi: 10.1109/ICEECCOT.2017.8284541

[2] A. T. M. S. Ahamed et al., "Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh," 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Takamatsu, 2015, pp. 1-6. doi: 10.1109/SNPD.2015.7176185

[3] N. Gandhi and L. J. Armstrong, "A review of the application of data mining techniques for decision making in agriculture," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, 2016, pp. 1-6. doi: 10.1109/IC3I.2016.7917925