

Agriculture Price Prediction Using Data Mining

Veeresh Kadlimatti, S V Saboji



Abstract: Agriculture is the main source and backbone of Indian Economy and plays a vital role in individual life. In the total Gross Domestic Product (GDP) agriculture nearly contributes sixteen percent and for increasing foreign exchange it contributes nearly ten percent to the total country exports. As the population continuously increasing and to manage the livelihood of the country there's requires a proper utilization and management of agriculture products. Data mining is a better technique and best choice in predicting the accurate prices of the agriculture based on previous data. In this work various data mining algorithms are applied on the dataset to predict the future prices of the agriculture products. **Keywords:** Agriculture, Data-Mining, Prediction.

Keywords: Agriculture, Data Mining, Prediction, Regression

I. INTRODUCTION

Agriculture takes a tremendous role in the development of India and it is the foundation of Indian economy. Agriculture still contributes more to the GDP of India compared to other several emerging sectors such as BPOs and IT sectors. Agriculture considered as more dominant sector. Agriculture increases the foreign exchange by contributing more to the exports of India. The above factors along with some more impact on cultivation of crop. For the reason farmers need accurate advice to predict the crop price and productivity of crop, a concerted examination should be carried out to achieve desired results accurately. Yield is an important agricultural issue. Large amount of data can be gathered from Indian agriculture sector. Knowledge acquired from data is highly useful for many purposes.

Data Mining is the process for analysing the large dataset, which include several steps called analysing, classification, clustering of the data. This process is necessary in analysing the raw and random data to process the further steps in predicting the price.

The work focuses on analysing such agricultural data using techniques of data mining and gaining the knowledge obtained from the result. The overall comparison of results obtained from several algorithms will help in selecting the more appropriate algorithm for agricultural data.

II. RELATED WORK

From the research is found that applying Data Mining for agriculture data can solve many complex problems avoiding interaction of human. Several researches carried out work are:

The paper [1] focus on that yield prediction system evaluation using the techniques of Data Mining on several datasets related to agriculture. In [2] they include the hybrid model to make better the agriculture production by using techniques of data-mining. [3] Various methods which are applied to determine the water needed for crops and effective rainfall for the Vijayapur district in Karnataka for improving the yield of the crops. [4] Focuses in regression analysis for forecasting the crop production using data mining. In [5] they have demonstrated to estimate the crop yield, choose the most excellent crop. Several algorithms are used to estimate the yield and choosing the most excellent and accurate crop. [6] Shows the system, which makes use of several techniques of data mining to predict the various category of soil data. In [7] major classification algorithms were discussed in detail for agriculture data. [8] Uses the spatial-temporal mining for the climate and weather dataset for better understanding. [9] focuses on the applications of several association rule mining techniques for the Rajasthan state rice crop yield. [10] various studies related to the wheat crop and generating rules using decision tree for crop yield. This research [11] predict the rice yield in Bangladesh region. [12] Gives the focus on estimating and predicting the future yield for cultivation of tea. [13] Describes several approaches for predicting the weather such as ANN, fuzzy logic and regression. [14] Presents a brief comparative study of various papers that deal with various techniques used to figure out the crop yield. [15] Analyses the soil data of Salem District using techniques of data-mining. [16] Describes various techniques data-mining to predict several category of soil data. The [17] techniques used for agricultural data are discussed. [18] Focuses on study which examines the data visualization application to find the similarity between the rice crop and climate factors. This work [19] experiments with different data mining issues like classification, clustering and association to study the given data using weka package. The present [20] work focus on several applications of data mining to predict production of crops such as Maize, Wheat and Rice. [21] A survey has done to outline several techniques of data mining that are used for analysing agriculture data. This paper [22] examines the agriculture data and provide the solution to maximize the crop production using optimal parameters. A multidimensional model [23] for the analysis of soil properties is built. An analysis of it is performed using OLAP operations. Paper [24] uses decision tree for analysis of for gross output value of agriculture for more than 30 cities of China. [25] aims to analyse the agricultural data of India using data mining algorithms.

Manuscript published on 30 August 2019.

* Correspondence Author (s)

Veeresh Kadlimatti, Basaveshwar Engineering College, Bagalkot., Karnataka, India. (Email: veeresh.kadlimatti@gmail.com)

S V Saboji, Basaveshwar Engineering College, Bagalkot., Karnataka, India. (Email: sabojishivakumar@gmail.com)

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

III. ISSUES AND CHALLENGES

From the survey identified several issues and challenges which are classified as:

1. Price: The price of the agricultural data is not constant and need to handle varying prices of crop.
2. Huge data: The dataset for analysis and prediction contains large amount of data.
3. Dataset: There is no available dataset need to build the dataset with daily prices of crop.
4. Operations: In upcoming years many more algorithms and operations are available for data mining.

IV. PROBLEM STATEMENT

The Raw and random data instances containing values like crop name, year, and pro- duction rate per hectare collected from the statistical department is converted into structural format (.arff File Format) using MS-Excel, then the dataset containing the previous year’s crop prices is supplied to Weka tool as a input which in turn creates a training model where the test set can be checked onto training set. The prices of agricultural products (such as vegetables and fruits) can be predicted using Decision Tree algorithm by cross validation process.

V. PROPOSED MODEL & RESULT

Agriculture is the primary source of livelihood for about 60 per cent of India's population. It contributes nearly sixteen percent to the total GDP of India. The system analysis and predict the cost of agricultural data with high accuracy.

The following diagram depicts the process model for analyses and prediction of agricultural crop price using data mining:

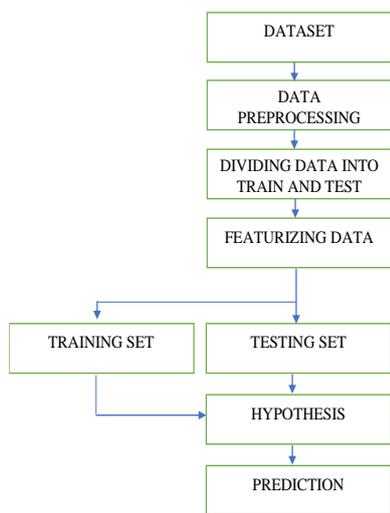


Fig.1. Flow of work

A. Dataset

Agriculture crop prices are influenced by many various factors of agriculture parameters which are climate condition, demand for crop and supply. The data include the daily base data of several commodity of several markets of Karnataka state. Collecting data with all other parameters is more difficult task hence considered only the price of various crop prices. From Karnataka State Krishi department the data are

collected.

Table I: Dataset Sample

Date	Commodity	Price	Market
2009/01/01	BAJRA	1050	BENGALURU
2009/01/01	COTTON	2500	VIJAYAPURA
2009/01/01	GROUNDNUT	2240	GADAG
2009/01/01	JOWAR	710	GADAG
2009/01/01	MAIZE	731	BAGALKOT
2009/01/01	ONION	780	HUBBALLI
2009/01/01	POTATO	600	HUBBALLI
2009/01/01	RAGI	950	BENGALURU
2009/01/01	SOYABEEN	1751	HUBBALLI
2009/01/01	WHEAT	1180	GADAG

- Price is the target variable,
- Date is the time series variable,
- Market and Commodity are categorical variables.

B. Data Pre-processing

The raw and random data collected from the Krishi department need to be pre-processed in order to make data suitable for future analysis.

C. Data Featurization

Featurization include conversion of categorical into numerical data using techniques of data mining.

D. Hypothesis

It is a pre-trained model with 70% data which we call hypothesis.

E. Predicting the price

The crop price is taken as input with the attributes including year, month, day, commodity name, market. By applying algorithm, the price of the crop thus be predicted using dataset.

VI. CONCLUSION

Application of data mining techniques in agriculture is an evolving field of research and there is a lot of work to be done. So, the application of computer science in agriculture would help the farmers to easily analyze and predict the prices of the agriculture products. Our system describes some data mining techniques on agriculture data to analyze and predict the prices of future agricultural products. From the survey of various papers, it can be possible to predict prices of crop with higher accuracy.

REFERENCES

1. Shruti Mishra, Priyanka Paygude, et al., “Use of Data Mining in Crop Yield Prediction”, IEEE Xplore Compliant - Part Number:CFP18J06-ART, ISBN:978-1-5386-0807-4; DVD Part Number:CFP18J06DVD, ISBN:978-1-5386-0806-7.
2. Sachin Chawhan, Pooja Ghodichor, et al., “Survey on Data Mining for Increasing Agriculture Productivity”, Volume 5 Issue III, March 2017 IC Value: 45.98 ISSN: 2321-9653.
3. Abishek.B, Akash Eswar M, et al., “Prediction of Effective Rainfall and Crop Water Needs using Data Mining Techniques”, IEEE, 2017.



4. Pallavi V. Jirapure, Prarthana A. Deshkar, "Qualitative data analysis using Regression method for Agricultural data", IEEE Sponsored World Conference on Futuristic Trends in Research and Innovation for Social Welfare (WCFTR'2016).
5. R.Sujatha, P.Isakki Devi, "A Study on Crop Yield Forecasting Using Classification Techniques", IEEE, 2016.
6. Monali Paul, Santosh K. Vishwakarma, et al., "Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach", 2015 International Conference on Computational Intelligence and Communication Networks.
7. Anshal Savla, Parul Dhawan, et al., "Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture", IEEE Sponsored 2nd International Conference on Innovations in Information, Embedded and Communication systems (ICIIECS)2015.
8. Chowdari K.K, Dr. Girisha R, et al., "A study of Rainfall over India Using Data Mining", International Conference on Emerging Research in Electronics, Computer Science and Technology – 2015.
9. Niketa Gandhi, Leisa J. Armstrong, "Assessing impact of seasonal rainfall on rice crop yield of Rajasthan", India using Association Rule Mining, 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India.
10. Rupinder Singh, Gurpreet Singh, "Wheat Crop Yield Assessment Using Decision Tree Algorithms", International Journal of Advanced Research in Computer Science, 8 (5), May/June 2017, 1809-1817.
11. Umid Kumar Dey, Abdullah Hasan Masud, et al., "Rice Yield Prediction Model Using Data Mining", International Conference on Electrical, Computer and Communication Engineering (ECCE), February 16-18, 2017, Cox's Bazar, Bangladesh.
12. Rupanjali D. Baruah, R.M. Bhagat, et al., "Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam India", 2016 International Conference on Information Technology 13. Kunal Teeda, Nandini Vallabhaneni, et al., Comparative Analysis of Data Mining Models for Crop Yield by Using Rainfall and Soil Attributes, IEEE Xplore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2.
13. Kunal Teeda, Nandini Vallabhaneni, et al., "Comparative Analysis of Data Mining Models for Crop Yield by Using Rainfall and Soil Attributes", IEEE Xplore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2.
14. Dr.A. Senthil Kumar, P.Arun, "A Survey on Agriculture Analysis for Crop Yield Prediction Using Data Mining Techniques", IOSR Journal of Computer Engineering (IOSR-JCE) e- ISSN: 2278-0661,p-ISSN: 2278-8727 PP 23-25.
15. N. Hema Geetha, G.M. Nasira, "Analysis of the Soil Data Using Classification Techniques for Agricultural Purpose", Volume-4, Issue-6 E-ISSN: 2347-2693 2016, IJCSE.
16. Supriya D M, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach", Vol. 5, Issue 5, May 2017 IJIRCCCE.
17. V. Rajeswari, K. Arunesh, "Analysing Soil Data using Data Mining Classification Techniques", Indian Journal of Science and Technology, Vol 9(19), DOI: 10.17485/ijst/2016/v9i19/93873, May 2016.
18. Niketa Gandhi, Leisa J. Armstrong, "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques", 2016 IEEE International Conference on Advances in Computer Applications (ICACA).
19. Anusha A. Shettar, Shanmukhappa A. Angadi, "Efficient Data Mining Algorithms For Agriculture Data", IJRTER-2016.
20. Suvridha Jambekar, Shikha Nema, et al., "Application of Data Mining Techniques for Prediction of Crop Production in India", IJIACS ISSN 2347 – 8616 Volume 7, Issue 4 April 2018.
21. Ms. Saima Humma, Prof. Anil Kumar Mishra, "Review on Analysis of Agriculture Data Using Data Mining Techniques", Volume 6 Issue IV, April 2018.
22. Suvridha Jambekar, Shikha Nema, et al., "Analysis of agriculture data using data mining techniques: application of big data", Majumdar et al. J Big Data (2017).
23. Constanta Zoie Radulescu, A Multidimensional Data Model for Analysis of Agricultural Soil Characteristics, 2008 First International Conference.
24. GAO Yi-yang, Ren Nan-ping, "Data Mining and Analysis Of Our Agriculture Based On The Decision Tree", 2009 ISECS.
25. Vinayak A. Bharadi , Prachi P. Abhyankar, et al., "Analysis And Prediction In Agricultural Data Using Data Mining Techniques", Special Issue 7-ICEMTE March 2017, IJRISE.