

Developing an Agricultural Product Price Prediction Model using HADT Algorithm

S. Rajeswari, K. Suthendran



Abstract: *Big Data Predictive Analytics and Data mining are emerging recent research field to analyse the agricultural crop price. The applications and techniques of data mining as well as Big Data using agriculture data is considered in this paper. In particular, the farmers are more concern about estimating that how much profit they are about to expect for the chosen crop. As with many other sectors the amount of agriculture data are increasing on a daily source. In this work, agriculture crop price dataset of Virudhunagar District, Tamilnadu, India is considered and for the price prediction model based on data mining decision tree techniques. The main goal is to establish the new predictive model based on Hybrid Association rule-based Decision Tree algorithm (HADT). The outcome for the suggested HADT forecast model is heartening and precise to predict agricultural product prices than other current decision tree models.*

Keywords : *Crop price, Big data, Data mining, Classification, Association rule, Decision tree*

I. INTRODUCTION

Agriculture outcome adds value to an Indian economy in that its role is fundamentally instrumental in terms of market, aspect and product contributions. India's economy depends primarily on the growth of agricultural yields and their associated agro-industry [1]. Agricultural area accounts for 60 percentage of Tamilnadu in the total geographical area.

Agricultural sector performance mainly depends on natural forces such as spatio-temporal distribution of rainfall, temperature, climate, etc., resulting in any deviation of the monsoon from the normal pattern resulting in huge fluctuations in area and production. Annually, a crop yield comes in to play on domestic and global economies, and the prediction yield [2] contributes much in the food and agriculture industry.

Data mining classification methods are used to create a pioneering model for anticipating the corresponding market price of commodities. Price estimation in agriculture is mandatory to forecast the market price for the chosen commodities, and it is supportive for farmers to schedule their crop cultivation operations and thus they gain more

profit. Market price prediction is mandatory for both private and public sectors to plan and execute agricultural development programs to steady the commodities market price. With this price forecast, consumers can preplan their expenses based on their need. This ground breaking idea comes in to play for the betterment of farmers and customers, and also for agricultural planning, farming and market planning agricultural policies and schemes. Predictive modeling is a method to predict the results using data mining and probability.

Data prediction modeling contains four phases, viz., historical data analysis (descriptive), data preprocessing, data modeling, and performance estimation. Classification technique in data mining provides the best solution for the process of prediction. Analysis of decision tree techniques observes an autonomous (predictor) and dependent (target) connection between the data set characteristics. This method enables the time series estimation of information and discovers the fundamental impact between these factors.

Data analytics (DA) is the method of examining data sets to draw conclusions about the information contained therein, progressively using specialized technologies and software [3]. The forecast of crop yields was carried out by considering the experience of the farmer in a specific field and crop. However, as circumstances change very quickly every day, farmers are compelled to grow more and more plants.

Since this is the present scenario, many of them have insufficient understanding of the fresh plants and are not fully conscious of the advantages they gain from farming them. In addition, by knowing and predicting crop efficiency in a multitude of environmental circumstances, farm productivity can be improved.

As one who should understand how much he would expect for their plants, predicting prices is a very significant problem for many farmers. Over the previous few years, price prediction has been produced by assessing farmers experience on a specific crop and field. Suppose they have prior year information accessible in which distinct respective price projections are recorded and these recorded price projections are used to classify future price projections [4].

By using these predictive models, the government can do agricultural development planning to stabilize the respective product prices. The suggested scheme applies machine learning and forecast algorithms such as decision tree algorithms to define information patterns and then process them according to input circumstances.

Manuscript published on 30 December 2019.

* Correspondence Author (s)

S. Rajeswari*, Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnan koil - 626126, Tamilnadu, India. Email: rajeswari@klu.ac.in

K. Suthendran, Department of Information Technology, Kalasalingam Academy of Research and Education, Krishnan koil - 626126, Tamilnadu, India. Email: k.suthendran@klu.ac.in

© 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/>

This primary objective is used to suggest the best crops based on the circumstances of the environment. Also, establish the new crop price analysis system for Virudhunagar District, Tamilnadu, India. Based on the farmer location this scheme recommends the numbers of beneficial plants by using the available past data sets. This paper is structured as follows: the literature review required for the study work is described in Section II. Section III describes the Methods and Materials required for the study. Dataset description discussed on Section V. Section VI was discussed with Results and analysis. Section VII concludes the work with possible potential enhancement.

II. RELATED WORKS

Agriculture comes in to play in the Indian economy, as 70% of the Indian families are exclusively reliant on this field. In 2015-16, Agriculture contributed nearly 17 percent of Gross Value in India. But after that, its contribution to Gross Value is not elevating. The food is on the top chart of primary needs for every human being, so the farmer's role is inevitable. Agricultural product price prediction supports the farmers to take beneficial decisions. This section discuss about various related works already done in data mining techniques using agriculture dataset. Most of the researchers focused on the similar problem like agricultural product commodity price prediction are as follows.

Huy Vuong Nguyen et al. developed a time bounded smart system for price estimation based on time series models. The scheme performs sequence of processes viz., Gathering transactions details from a website, pre-processing the collected unrefined information, and using an ARIMA model. Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) used as a performance metrics to assess the forecast precision strategy. The transaction details of PriceMe website is considered in this work. Moreover, the proposed model is compared with existing Moving Average (MA) model. The outcome discloses that the forecast details of the proposed model are time bounded when compared to MA model [5].

Sean McNally et al. proposed a Bitcoin price prediction method in. The price information is derived from the price index of Bitcoin. By implementing a Bayesian optimized recurrent neural network (RNN) and a Long Short Term Memory (LSTM) network, the task is accomplished with variable degrees of achievement. The LSTM achieves 52% maximum classification precision and 8% RMSE. As a comparison with the deep learning models, the common ARIMA model for time series prediction is introduced [6].

Manpreet Kaur et al. studied about different data mining techniques viz., K-Means, K-Nearest Neighbor, Artificial Neural Networks and Support Vector Machines and its crop price prediction capability. Estimating the crop price using available information is in high demand. Based on the study the author introduced a model that assists the farmers to take better decision on estimating the crop price with accuracy [7].

Ticlavilca et al. analyzed multiple methods for estimating the agricultural commodity prices. They have practiced a

machine learning approach based on Bayesian for multiple-time-ahead predictions. The performance this approach was weighted against existing manifold output model such as the Artificial Neural Network. The stability of the MVRVM and ANN's algorithm is analyzed based on bootstrapping concept [8].

Hemageetha and Nasira suggested an idea for vegetable price prediction based on Back-propagation Neural Network and Radial Base Neural Function Network. In particular, they have predicted the tomato price for Coimbatore [11]. Their MATLAB simulation outcome discloses that the RBF is competent than the neural network Back-propagation [9].

Luo et al. analyzed four models to predict the Beijing Xinfadi wholesale price for the *Lentinus edodes*. BP neural network model, genetic algorithm-based neural network model, RBF neural network model, and an embedded model are considered for prediction. Between 2003 and 2009, they have considered 84 records to train and test the system and this was performed on above four models. The observation outcome reveals that the BP neural network model produced poor outcome among the four. The genetic algorithm based model is more perfect than the RBF model. [10].

Nima Amjady designed an effective Fuzzy Neural Network approach for market price estimation. With a fresh hyper cubic training system, this fuzzy neural network is made up of inter-layer and feed-forward architecture. It estimates hourly market-clearing rates. The outcome analysis clears that this method outperforms the other neural networks viz., ARIMA time series, wavelet-ARIMA, MLP, and RBF [12].

Aman Vohra et al. planned a decision-building model to guess the agricultural product price. This model assists the farmers to obtain more benefits for their crop. This also enables the farmers to ensure distinct commodities prior prices. This model offers maximum options to the farmer so they can plan and estimate for different crops and finally they can go with beneficiary one [13].

Khalid et al. exhibited ARMA and wavelet based methods to calculate the market prices of grains. 300 observations were noted over three decades between 1983 and 2013. The accuracy of proposed models is calculated based on three distinct error tests. Outcome reveals that the wavelet forecasting technique offers best grain prices than ARMA[14].

Haoyang Wu et al. analyses day to day daily price estimation problems in different models viz., ARIMA model, PLS regression and neural networks. This model is suitable for large size of agricultural commodities. The findings acquired in the daily life of customers are precise and useful. Estimating the quantity of agriculture consumables comes in to play while deciding the price for commodity. Preventing real data uncertainty is a big challenge [15].

Predict the prices of agricultural products on the Virudhunagar District market using the Hybrid Association rule-based Decision Tree (HADT) algorithm in this research. The main aim is not to prove that above mentioned methods are best, but it is to show that they are applicable to daily life problems. Using HADT algorithm generate the monthly price prediction results for agricultural products. It enables the farming community make decisions and thus making profit is achievable.

III. METHODOLOGY AND DATASET

For their economic growth, an agro-based nation relies on agriculture. As the country's population increases dependence on agriculture, it also improves and affects the country's subsequent economic growth. Prediction is a forward-looking statement [16]. Agricultural commodity pricing has become the hour's need for farmers. Although future occurrences are unsure, it is not feasible to predict precision. This paper includes a model of decision-making support that can be helpful in predicting prices for farmers.

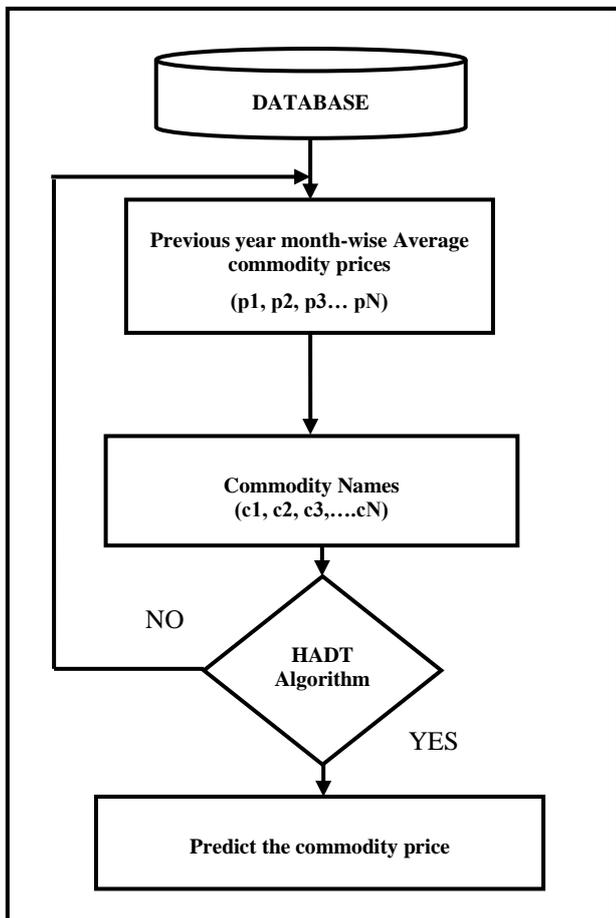


Fig. 1. Proposed Method Work flow

Figure 1. shows the farmers must enter the name of the commodity and the prior crop selling price. Based on past prices, the HADT algorithm will be able to provide average rates for a specific crop to help farmers make better choices and predict prices.

A. Dataset and Study Area

Many variables such as environment, supply and demand, etc., affect market prices of agricultural goods. It is more

complex to predict than business products. It is hard to correctly and timely obtain the information of the impact factor [17]. Thus, in this research, information on agricultural product prices gathered as experimental materials for Virudhunagar District-11 Blocks, Tamilnadu, India. The daily cost for the year 2016-2018 was gathered from the wholesale market of Virudhunagar District. The monthly price is gain to prepare for forecasting after it has been weighted. Dataset has 6 characteristics and 50,000 samples in total. The explanations for the input characteristics are shown in Table 1.

Table -I: Explanation Attributes

| Attributes | Explanations |
|------------------------|--|
| C. No | Commodity code |
| Year | Year |
| Month | Month |
| C.N | Commodity Names |
| Price | Commodity Price |
| Price prediction Class | Very High, Very Low, High, Medium, Low |



Fig. 2. Experiment Location

Figure 2. Shows the Virudhunagar District is the experiment place profile located in Tamilnadu, India. It is made up of 8 taluks and 11 blocks of development, i.e. Aruppukkottai, Kariapatti, Srivilliputtur, Rajapalayam, Watrap, Virudhunagar, Sivakasi, Sattur, Vembakkottai, Tiruchuli and Narikkudi. There are six municipalities in this district, 11 town panchayats and 600 villages. The headquarters of the district is Virudhunagar [18].

B. Proposed Methodology

Hybrid Association Decision Tree algorithm (HADT) focuses on producing the association rules that result in only one particular attribute.

The rules of association produced are called the rules of class association. By introducing a constraint to the attribute that must appear on the consequence of the rule, it differs from general association rule mining. This can be used to construct a template or classifier.

The performance improved by this algorithm, the problem of scalability still exists. In addition, the databases can be spread across separate places in today's Internet environments. It has the following two components of algorithms are [19]:

First part: Rule creation

- (1) It first searches the training data set and discovers all the frequent item set. Then selects the values of the attributes in support descending order, F-list, and scans the training data to build a frequent item set again.
- (2) Sort the rules in a compact tree structure. It investigates the possibility of exchanging rules and thus saves room as the rules that have common frequencies share the route portion.
- (3) Put the rules in order. Use general and high trust rules to prune more specific and lower trust rules. Select only rules that have been corrected positively. Select a subset of rules based on the coverage of the database

Second part: Multiple rules generated using classification algorithm.

For prediction existing C5.0: ADT classifier and association rule mining algorithms are used. These two algorithms are suitable to process the big amount data with better performance. Here, the fundamentals of C5.0: ADT decision tree algorithms are included in the MapReduce concept and the decision tree rules are generated. Decision tree rules are taken into the proposed Association rule mining algorithm for design the new MapReduce framework functions for generate the candidate itemset derived from the support, confidence count. The derived frequent itemset counts for selected crops prediction patterns are generated [20].

Pseudo code for HADT Algorithm

Let D denotes the dataset classification rules and T represents the training data. The aim of this model is to choose a set of highly predictive rules in D to cover the training data T. The classifier formed as: $\langle d_1, d_2, \dots, d_i, \text{default class} \rangle$ where $d_j \in D$, d_m, d_n if $m < n$. Default class is the default label used when none of the rules can classify an instance.

Rule Generation

Input: Relevant features are selected from the dataset

Output: Crop price values

Map Phase

Step 1: Read the characteristics in the data set of the crop price

Step 2: Let C be a classification of a set S of substance into grouping as C1, C2, ... , Cr. Let p_i exist chance that

an entity in S is in group of C_i . The entropy of S is describe as

$$\text{Entropy}(S) = - \sum_{i=1}^m P_i \log_2(P_i)$$

Step 3: For any known element A, believe the set V(A) of feasible values of A. For any $v \in V(A)$, let S_v exist the set of all elements of S includes value v for the element A. The information gain of A with high opinion to S is known by

$$\text{Info.Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} * \text{Entropy}(S_v)$$

Step 4: Compute the Gain Ratio using Entropy and Information gain

$$\text{Gain Ratio}(A) = \text{Entropy}(D) - \text{Info.Gain}(S, A)$$

Reduce Phase

Step 1: Discover the highest gain ratio attribute value

Step 2: If there is no more attribute, the rule is generated

Pattern Generation

Input: C5.0: ADT classification Rules prediction result for the Dataset

Output: Generate the rules for Crop price prediction value

Map Phase:

Step 1: Convert the transaction data from the classification output

Step 2: For each transaction split the data into the k-itemset

Step 3: Find the partial frequency and value of k-itemsets in the dataset

Step 4: Frequent item set are converted into rule format

Reduce Phase:

Step 1: Calculate the elements value of pairs occurrences in the k - itemset

Step 2: Split occurrences and find partial frequent k - itemsets with values

Step 3: Until the frequent itemset found the process repeatedly works

Step 4: Generated association rules are taken into for prediction process

From the above, Pseudo code perform the task to find the frequent itemset for the crop price prediction values using the new novel proposed Hybrid Association rule-based Decision Tree algorithm (HADT) algorithm. It generates many rules for similar set of features. Our work is aimed to eliminate redundancy in the item set and in turn it reduces the redundant time operation of algorithm.

Figure 3 shows the work flow of Hybrid Association Decision Tree Algorithm for calculating the frequent item sets using MapReduce function. The map key used in the Mapper function represents the v-itemsets where v is the number of items used to form the key. Firstly, the C5.0 ADT classification is applied on the dataset attributes shown in Table. 1 for generate the rules for prediction.



These rules are converted in to transactional database. During Mapping process, the minimum support count and maximum confidence count is calculated. In Reduce phase, pruning function is performed to identify the frequent itemsets. The support count and confidence value is calculated for all identified by frequent itemsets. The frequent items are rated based on support count and confidence value and corresponding patterns are generated for the same [21].

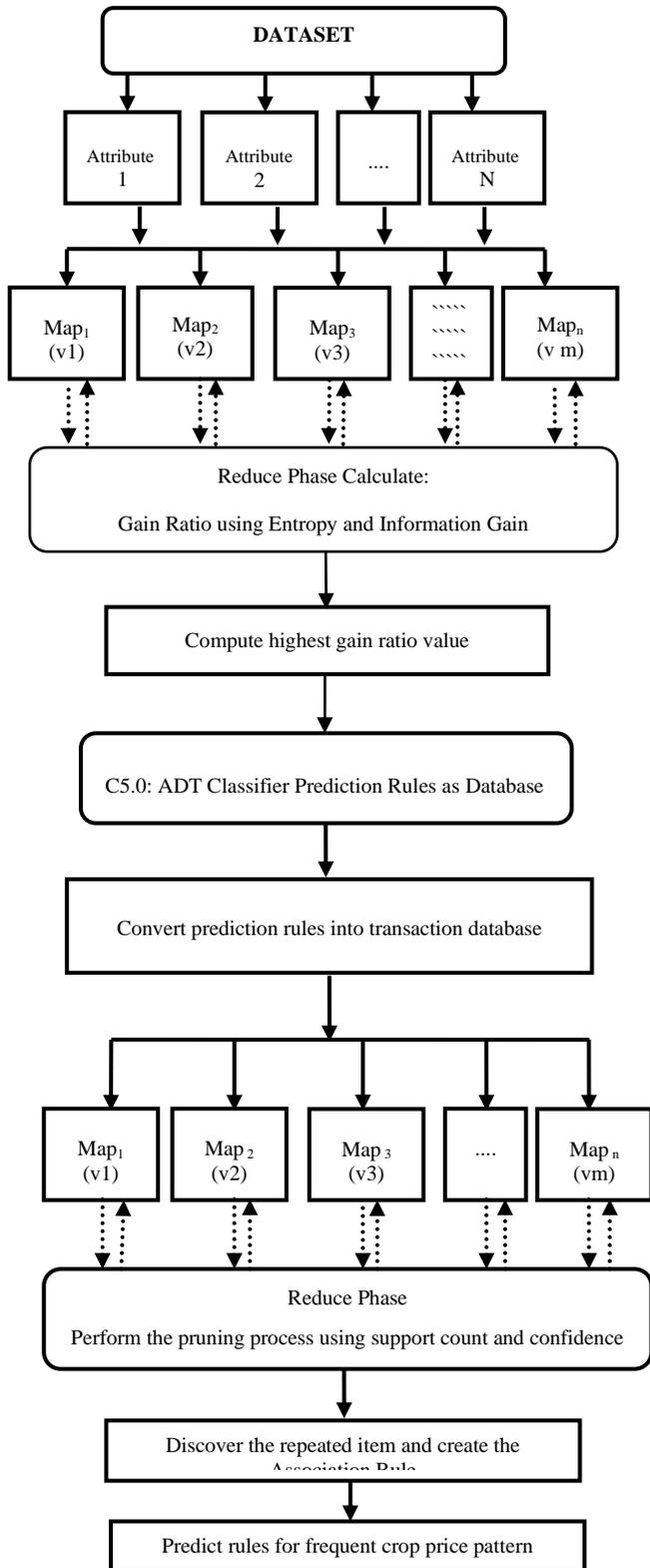


Fig. 3.HADT Algorithm Work flow

The proposed algorithm main objective is used to move towards to take away redundancy by means of higher rank closed frequent item set. So, the amount of generated rule set to develop the feature and effectiveness exclusive of several losses [22]. The advantages of Hybrid Association rule-based Decision Tree algorithm (HADT) algorithms are as follows:

- Training the data is very capable apart from the volume of data set
- High dimensionality training sets can be dealt with effectively and assumptions on dependent and free attributes are not required
- Classification process was finished in fast way
- Humans can easily understand the Generated rules
- It works in three stages, creates a bunch of association rules using training set with positive support and confidence threshold as candidate rules
- For pruning methodology introduce the over fitting within the discovered rules
- Classification techniques are used to predict the test information and measure the precision of the classifier model

IV. DISCUSSION OF RESULT

In this work, R programming language is used to process the agricultural data. It is open source software and used to processes the statistical data. The integrated R and Hadoop environment provides RHadoop, RmR, RHdfs packages. The proposed system model was developed using this package to handle big data [23].

The motivation of this proposed model is to recommend the right crop seed for sowing based on the available crop price previous year data. The accuracy calculation should be a very essential component for classification algorithm. It will testify the past active classification algorithms are moreover “Right” or else “Wrong” [24].

Compared with the other existing techniques, the new novel classification precision technique was suggested. The precision was calculated on the basis of the total number of predictions divided by the total number of data. In this whole collection of information, 25% of the information is allocated to the training stage and the remaining allocated to the test phase [25]. The precision is measured by,

$$Accuracy = \left(\frac{TP + TN}{TP + FN + FP + TN} \right) \quad (1)$$

Where, TP denotes True Positive for identified the total number of correctly classified positive instances, TN represents True Negative for identified the total number of correctly classified negative instances, FP denotes False Positive for identified the total number of incorrectly classified positive instances, and FN represents False Negative for identified the total number of incorrectly classified negative instances. The performance of the algorithm is calculated using the above formula (1).

The performance of the proposed algorithm is measured and compared with the existing algorithms for data mining. The performance is measured by accuracy, rate of error, and time of execution. The comparisons result for the prediction models are shown in Table 2. From that table it was observed that HADT classifier algorithm achieves high precision compared to the other existing algorithms with fewer amounts of error rate and execution time.

Table - II: Results of Prediction Models Comparison

| Algorithms used | Accuracy (%) | Error Rate (%) | Execution Time (Seconds) |
|----------------------|--------------|----------------|--------------------------|
| C4.5 | 73% | 28% | 15 Secs |
| C5.0 | 82% | 18% | 11 Secs |
| C5.0: ADT classifier | 85% | 15% | 10 Secs |
| HADT classifier | 90% | 10% | 6 Secs |

This result facilitates the virudhunagar district farmer get suggestions for deciding the best crop for sowing based on crop price predictions value. Also, clear that Hybrid Association rule-based Decision Tree algorithm (HADT) algorithm achieves better performance than other existing algorithm. So, the proposed approach is suitable for large datasets.

V. CONCLUSION AND FUTURE DIRECTIONS

In accounting for economic fluctuations, commodity prices have long played a significant role. Therefore, forecasting shifts in commodity prices are a significant element for policymakers looking forward. The increasing use of future markets has raised the issue of how much data these prices integrate into spot prices about future movements. In this paper, using the Hybrid Association rule-based Decision Tree algorithm (HADT) algorithms created and assessed the intelligent scheme for short-term price forecast.

The suggested scheme also has some constraints in relation to the characteristics described above. Implementing price forecasting in the suggested scheme continues step-by-step, leading to time-consuming procedures. The primary objective in the future is to enhance approach synchronization and efficiency by using an automation module. In addition, we strive to combine the HADT model with Fuzzy techniques to further enhance the precision of the predictions.

ACKNOWLEDGMENT

The First Author is grateful to the Kalasalingam Academy of Research and Education Management for offering research fellowship and for helping laboratory equipment during this study thanks to the National Cyber Defence Research Centre.

REFERENCES

- Chapshou Luo, Qingfeng Wei, Liying Zhou, Jungeng Zhang and R. Suien Sun "Prediction of vegetable price based on Neural Network and Genetic Algorithm". IFIP AICT 346, pp. 672-681, 2011

- Y. S. Son, R. Baldick, K. Lee, and S. Siddiqi, "Short-term electricity market auction game analysis: Uniform and pay-as-bid pricing," IEEE Trans. Power Syst., 19(4), pp. 1990–1998, 2004
- A. M. Gonzalez, A. M. San Roque, and J. G. Gonzalez, "Modeling and forecasting electricity prices with input/output hidden Markov models," IEEE Trans. Power Syst., 20(1), pp. 13–24, 2005
- Kaur, Manpreet, Heena Gulati, and Harish Kundra. "Data mining in Agriculture on crop price prediction: Techniques and Applications." International Journal of Computer Applications, 99(12), pp.1-3, 2014
- Huy Vuong Nguyen, M.Asif Naeem, Nuttanan Wichitaksorn, and Russel Pears. "A smart system for short-term price prediction using time series models", Computers and Electrical Engineering, 76, pp. 339–352, 2019
- Sean McNally. 2016. Predicting the price of Bitcoin using Machine Learning. Ph.D. Dissertation. Dublin, National College of Ireland.
- Manpreet Kaur, Heena Gulati, and Harish Kundra, "Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications", International Journal of Computer Applications, 99(12), pp. 0975 – 8887, 2014
- Ticlavilca, A. M., Dillon M. Feuz and Mac McKee. "Forecasting Agricultural Commodity Prices Using Multivariate Bayesian Machine Learning Regression." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, 2010
- N. Hemageetha, and G. M. Nasira, "Radial basis function model for vegetable price prediction", 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, IEEE Xplore digital library, pp.424-428, 2013
- Changshou Luo, Qingfeng Wei, Liying Zhou, Junfeng Zhang, and Sufen Sun, "Prediction of Vegetable Price Based on Neural Network and Genetic Algorithm", IFIP International Federation for Information Processing 2011, pp. 672–681, 2011
- Nasira, G. M, and Hemageetha, N. "Vegetable price prediction using data mining classification technique", International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012), pp.99-102, 2012
- Amjady, N, and Keynia, F. "Electricity Market Price Spike Analysis by a Hybrid Data Model and Feature Selection Technique". Electr. Power Syst. Res., 80, pp. 318–327, 2010
- Aman Vohra, Nitin Pandey, and S.K. Khatri, "Decision Making Support System for Prediction of Prices in Agricultural Commodity", 2019 Amity International Conference on Artificial Intelligence (AICAI), IEEE Xplore digital library, pp. 345-348,
- M. Khalid, Mariam Sultana, and Faheem Zaidi, "Prediction of Agriculture Commodities Price Returns Using ARMA and Wavelet", Journal of Natural Sciences Research, 4(23), pp. 5-10, 2014
- Haoyang Wu, Huaili Wu, Minfeng Zhu, Weifeng Chen, and Wei Chen, "A new method of large-scale short-term forecasting of agricultural commodity prices: illustrated by the case of agricultural markets in Beijing", Journal of Big Data, 4(1), pp. 1-22, 2017
- Rajeswari S, Suthendran K, and Rajakumar K, A Smart Agricultural Model by Integrating IoT, Mobile and Cloud-Based Big Data Analytics, International Journal of Pure and Applied Mathematics, 118 (8), pp. 365-370, 2018
- Agro Marketing website for data collection - <https://agmarknet.gov.in/>
- Virudhunagar District Profile: <http://kvkpk.org/District%20Profile.pdf>
- Lin X. (2014). MR-Apriori: Association Rules algorithm based on MapReduce, 2014 IEEE 5th International Conference on Software Engineering and Service Science, pp. 141-144.
- Rajeswari S, and Suthendran K. (2019). C5.0: Advanced Decision Tree (ADT) classification model for agricultural data analysis on cloud, Computers and Electronics in Agriculture, 156, pp.530-539.
- Chandanana A.K., and Shuklab M.K. (2015). Removal of Duplicate Rules for Association Rule Mining from Multilevel Dataset, Procedia Computer Science, 45, pp. 143 – 149.
- Yang Junrui., Xu Lisha., and He Hongde. (2012). A Classification Algorithm Based on Association Rule Mining, 2012 International Conference on Computer Science and Service System, IEEE Computer Society, pp. 2056 - 2059.



23. Han J. and Kamber M. (2012). Data Mining: Concepts and Techniques”, third edition, San Francisco: Morgan Kaufmann Publishers.
24. Rajeswari S, and Suthendran K. (2018) Chi-Square MapReduce Model for Agricultural Data, Journal of Cyber Security and Mobility, 7(1), pp. 13-24.
25. Ming-Yen Lin, Pei-Yu Lee, and Sue-Chen Hsueh. (2012). Apriori-based Frequent Itemset Mining Algorithms on MapReduce, ACM, ICUIMC’12, pp. 1-8, doi-10.1145/2184751.2184842.

AUTHORS PROFILE



S. Rajeswari received her B.Com (Computer Applications) from Madurai Kamaraj University in 2012; Master of Computer Applications and M.Phil (Computer Science) from Madurai Kamaraj University in 2015 and 2016 respectively. Now, she is a Research Scholar in the Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India. Her current research areas include Big Data, Predictive Analytics, and Data mining.



Suthendran Kannan received his B.E. Electronics and Communication Engineering from Madurai Kamaraj University in 2002; his M.E. Communication Systems from Anna University in 2006 and his Ph.D Electronics and Communication Engineering from Kalasalingam University in 2015. He was a Research and Development Engineer at Matrixview Technologies Private Limited, Chennai for a couple of years. He is now the Head, Cyber Forensics Research Laboratory and Associate Professor in Information Technology, Kalasalingam Academy of Research and Education. His current research interests include Cyber Security, Communication System, Signal Processing, Image Processing, etc.