

# Preliminary and Predictive Model for Finding Crop Product Patterns

Pooja Yadav, Praveen Yadav

Guest Assistant Professor, Department of Computer Science and Engineering

Indira Gandhi University, Meerpur, Rewari, Haryana, India

**Abstract-**The weka tool is designed to hold in the research of material learning technology to real survival dataset. In this paper we focused overall the application of weka in the agriculture department datasets so that we can recall statistics about the production of crop. this can lead to lessen the death rate of farmers as a result of increased harvest production and destitution factors. this paper deals with predictive method for finding crop product pattern.

**Keywords:** component; Season crop; Farmer suicide; crop destroy.

## I. INTRODUCTION

The universal accessibility of vast quantity of agriculture report has devote a pressing call for the research of data mining. Generating a set of rule mutually more suitable for agriculture database can be done applying different techniques of word mining. Association bulldoze mining is virtually important campaign to penetrate abide pattern, correlation, association or casual structure inserted item sets in the selling database or repositories.

To study support(s) and confidence(C) for association rule that satisfy the predefined minimum support and belief from if database.a-periori algorithm is the most popular move of faction menace for finding frequent item sets over transaction database. a-periori algorithm is very easily done to travail to find haunt patterns or begin database.

- Use of faction rule:

Association rules are widely used in various

Areas of agriculture such as soil type Including livestock, forestry and fishery spell sectors, telecommunication networks, market and risk management, inventory approach etc.

- Use of association rule:

Association rules are generally used in various

Areas of agriculture one as soil humorIncluding livestock, forestry and fishery supplant sectors, telecommunication networks, market and risk management, inventory approach etc.

- Implementation of A-periori algorithm:

The dominant objective of this paper is to bring about A-priori algorithm by WEKA simulator which suggest a platform to achieve algorithm on the story set. pre-process of datasets will improve the status of association rule. This algorithm mostly calculates frequent item apply, back and

minimum belief in a very rational manner. Apriori is the sooner association rule mining algorithm that pioneered the evaluate of support-based pruning to systematically act the high on the hog of candidate item sets. Here the database is used different type of attribute appreciate season, natural case, or human made cause, loss of hay etc. It extracts bodily frequents trends of datasets.

The function has particular focused on the association bully base A-priori algorithm in edict to court with maximum accessible data rule in order to have efficient result.

## II. METHODOLOGY

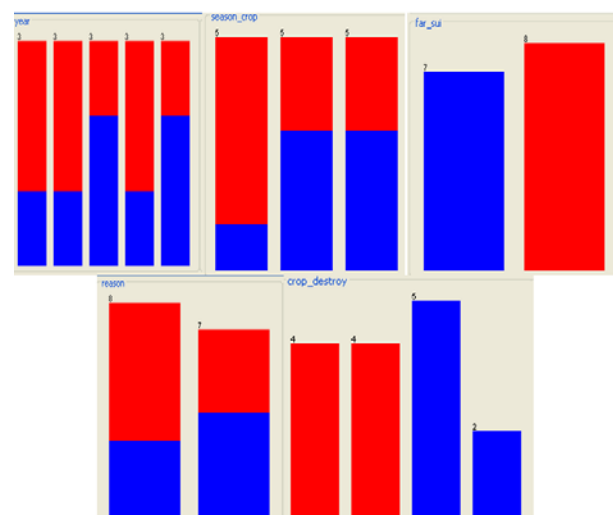
- Data sets:

Many reasons to destroy crops and production are present form the datasets. Only seven type of attribute are represent here, the following attribute are consider as nominal values.

Year	Nominal
Reason	Nominal
Season_crop	Nominal
Crop_destroy	Nominal
Farsui	Nominal

## III. VISULIZATION OF DATASETS

Graphically visualization of each attribute with its corresponding value.



- Performance of a-periori algorithm:

Part1: Frequent Item Set: The algorithm works as follows-For a given set of transaction (it is assume that we require minimum support of p %)

Step1: Scan all transactions and find all frequent items that have support p%.

Step2: Built potential sets of K-items from LK-1 by using pairs of item sets in LK-1 such that each pair has the first K-2 common items and the one reaming item from each of two item sets are combined two form A K item sets. The set of such potentially frequent K item seta is the candidate sets CK.

Step3: Scan all transactions and find all -item sets in CK that are frequent. The frequent set so obtained is LK.

Now terminate when no further frequent item sets are found,otherwise continue with step2.

Part2:Finding the rules: To find the association rules from the frequent item sets we take a large frequent item set, say P and find each non empty item set.

The rule A→ (P-A) is possible if it satisfied the confidence. Confidence of this rule is given by

- Support

The support  $supp(X)$  of an item set X is defined as the proportion of transactions in the data set which contain the item set.

$$Supp(X) = \text{no. of transactions which}$$

- Confidence

The confidence of a rule is defined:

$$Conf(X \rightarrow Y) = Supp(X)$$

The confidences contain the transaction which is used in set X and Y and contain the item set X transaction.

- Lift

The lift of a rule is defined as:

$$Lift(X \rightarrow Y) = \frac{Supp(X \cup Y)}{Supp(Y) * Supp(X)}$$

- Conviction

The conviction of a rule is defined as:

$$Conv(X \rightarrow Y) = \frac{1 - Supp(Y)}{1 - Conv(X \rightarrow Y)}$$

The conviction of the rule  $X \Rightarrow Y$  can be interpreted as the ratio of the expected frequency that X occurs without Y

(that is to say, the frequency that the rule makes an incorrect prediction) if X and Y were independent divided by the observed frequency of incorrect predictions.

Association rule mining	Evaluation criteria	value
A-periori algorithm	Min. Support(s)	0.15
	confidence (c)	0.9
	Number of instances	15

#### IV. SIMULATION RESULTS

```

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0
Relation: crop_data
Instances: 15
Attributes: 5
    year
    reason
    season_crop
    crop_destroy
    far_sui

=== Associator model (full training set) ===

Apriori
*****

Minimum support: 0.15 (2 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Generated sets of large itemsets:

Size of set of large itemsets L(1): 16
Size of set of large itemsets L(2): 39
Size of set of large itemsets L(3): 21
Size of set of large itemsets L(4): 3

Best rules found:

1. crop_destroy=51-75 5 ==> far_sui=yes 5   conf:(1)
2. crop_destroy=0-25 4 ==> far_sui=no 4   conf:(1)
3. crop_destroy=26-50 4 ==> far_sui=no 4   conf:(1)
4. reason=natural crop_destroy=0-25 3 ==> far_sui=no 3   conf:(1)
5. reason=natural far_sui=yes 3 ==> crop_destroy=51-75 3   conf:(1)
6. reason=natural crop_destroy=51-75 3 ==> far_sui=yes 3   conf:(1)
7. season_crop=rabi crop_destroy=0-25 3 ==> far_sui=no 3   conf:(1)
8. season_crop=khrib far_sui=yes 3 ==> crop_destroy=51-75 3   conf:(1)
9. season_crop=khrib crop_destroy=51-75 3 ==> far_sui=yes 3   conf:(1)
10. crop_destroy=76-100 2 ==> reason=unnatural 2   conf:(1)
    
```

#### V. CONCLUSION

Useful information from the large database has been extracted in the form of the association rules. There are many algorithm have been developed to extract the association rules from the large database. A-priori algorithm is the most popular algorithm to extract the association rules from the database.

For survival in this world it's the basic need to use and to store the information means to prepare a proper database or dataset to analyze. Using and storing the database is not an

issue, but find the relevant pattern or to analyze the meaningful dataset and solve the problem to taking decision about the production of crops and difficult situation of farmers.

#### REFERENCES

- [1]. S. Aruna, Dr S.P. Rajagopalan and L.V. Nandakishore, "An Empirical Comparison Of Supervised learning algorithms in Disease ion", Vol.1, August 2011.
- [2]. Leonard Gordon, Using Classification and Regression Trees (CART) in SAS Enterprise Miner For Applications in Public Health. Paper 089-2013, 2013
- [3]. Vikas Chaurasia, Saurabh Pal, "Early Prediction of Heart Diseases Using Data Mining Techniques", Vol.1, 2013.
- [4]. Thenmozhi, P. Deepika, M. Meiyappasamy, "Different Data Mining Techniques Involved in Heart Disease Prediction: A Survey", Volume 3, ISSN No. 2277-8179, September 2014.
- [5]. Sushilkumar Kalmegh, "Analysis of WEKA Data Mining Algorithm REPTree, Simple Cart and RandomTree for Classification of Indian News", , Vol. 2 Issue 2, February 2015.
- [6]. T. Miranda Lakshmi, A. Martin, R. Mumtaj Begum, Dr. V. Prasanna Venkatesan, "An Analysis on Performance of Decision Tree Algorithms using Student's Qualitative Data", IJ. Modern Education and Computer Science, 2013, 5, 18-27.
- [7]. Jyoti Rohilla, Preeti Gulia, "Analysis of Data Mining Techniques for Diagnosing Heart Disease", Volume 5, Issue 7, July 2015.
- [8]. Hlaudi Daniel Masethe, Mosima Anna Masethe, "Prediction of Heart Disease using Classification Algorithms", Vol II WCECS 2014, 22-24 October, 2014, San Francisco, USA.
- [9]. Nidhi Bhatla, Kiran Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques" International Journal of Engineering and Technology Vol.1 issue 8 2012.
- [10]. Chaitrali S. Dangare and Sulabha S. Apte, "Improved Study of Heart Disease Prediction Using Data Mining Classification Techniques", International Journal Of Computer Applications, Vol. 47, No. 10, pp. 0975-888, 2012.
- [11]. Chaitrali S. Dangare and Sulabha S. Apte, "Improved Study of Heart Disease Prediction Using Data Mining Classification Techniques", International Journal Of Computer Applications, Vol. 47, No. 10, pp. 0975-888, 2012.
- [12]. Atul Kumar Pandey, Prabhat Pandey, K.L. Jaiswal and Ashok Kumar Sen, "A Heart Disease Prediction Model using Decision Tree", IOSR Journal of Computer Engineering, Vol. 12, Issue.6, (Jul. – Aug. 2013), pp. 83 – 86.
- [13]. Tina R. Patil, Mrs. S.S. Sherekar, "Performance Analysis of Naïve Bayes and J48 Classification algorithm for Data Classification", International Journal Of Computer Science and Applications, Vol. 6, No.2, Apr 2013.