

DATA MINING IN ORGANIZATION

Marathe Dagadu Mitharam And A. D. Bhosale

Commerce & Management Department R. C. Patel A.C.S. College, Shirpur
Department of Commerce G. S. Science, Arts & Commerce College, Khamgaon

Abstract: Data mining automates the detection of relevant patterns in a dataset. In organization data mining techniques uses e.g. classification of employees, classification products, classification of department etc. Discovery and analysis of some patterns in organization using Association rule, clusters.

e.g.

$$\text{Quality of Product} = \frac{XUY.\text{count}}{N}$$

X=products, Y= defects and N= Number of operation

Keyword: Data Mining, Association Rule, Cluster

i) INTRODUCTION

Data mining means to retrieve, to capture and to accept data from datasets. Suppose there is no use of data mining in the organization then actual data not get. In the mining process we use some kinds of techniques just, Classification, association rule, cluster, etc. Means Supervised learning and Unsupervised learning are use. In this organization some dataset are use for experimental. Some mining techniques are,

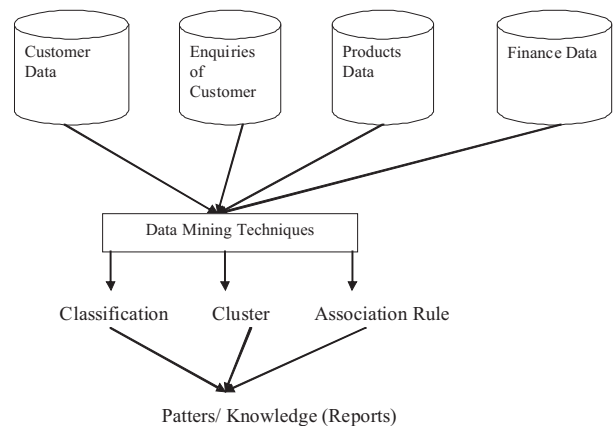
Association Rule- Association rules are an important class of regularities in data. Mining of association rules is a fundamental data mining task. Association rule discovery and statistical correlation analysis can find groups of items or pages that are commonly accessed or purchased together.

Decision Tree- Decision tree learning is one of the most widely used techniques for classification. Its classification accuracy is competitive with other learning methods, and it is very efficient.

Naïve Bayesian Classification- Supervised learning can be naturally studied from a probabilistic point of view. The task of classification can be regarded as estimating the class posterior probabilities.

Cluster- Clustering is the process of organizing data instances into groups whose members are similar in some way. A cluster is therefore a collection of data instances which are “similar” to each other and are “dissimilar” to data instances in other clusters.

2) EXPERIMENT:-



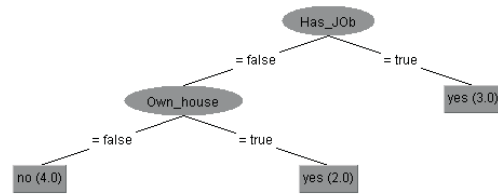
Visit “Priyadarshani Sahakari Soot Girani Ltd TANDE Tal- Shirpur Dist- Dhule (Maharashtra)” is leading commercial manufacturing organization. Organization is to provide the maximum work environment, maximum profit of framers in this district.

Creates some questionnaire and arrange some interviews in Feb and March 2013. In this questionnaire create questions like this, how mine the data, how to generate various reports in company, how to done classification of data, how done same objects of data, etc...

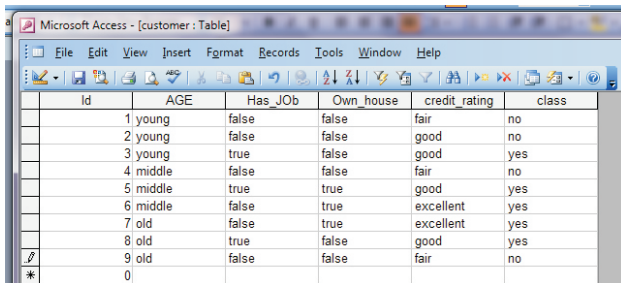
Suppose collected dataset is e.g.
A loan application data set

ID	Age	Has_job	Own_house	Credit_rating	Class
1	young	false	false	fair	No
2	young	false	false	good	No
3	young	true	false	good	Yes
4	middle	false	false	fair	No
5	middle	true	true	good	Yes
6	middle	false	true	excellent	Yes
7	old	false	true	excellent	Yes
8	old	true	false	good	Yes
9	old	false	false	fair	No

Tree view

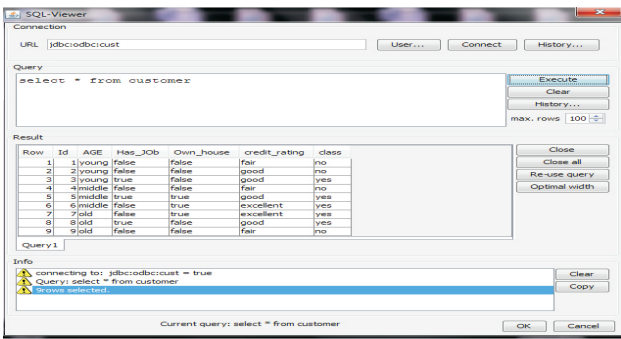


In Ms_Access File

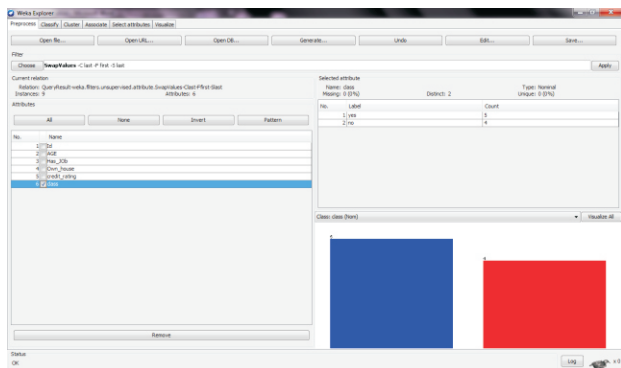


The WEKA package will be used for the analysis of collected data. Based on analysis, interpretation will be made to reach the meaningful conclusions. To perform operation in collected data in WEKA

Connect database using “jdbc:odbc:cust” in Weka software



Classification “Yes” and “No” Class type



4) RESULT

==== Run information ====

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: QueryResult

-weka.filters.unsupervised.attribute.SwapValues-Clas-

Ffirst-Slast

Instances:9

Attributes:6

- Id
- AGE
- Has_JOB
- Own_house
- credit_rating
- class

Test mode:evaluate on training data

==== Classifier model (full training set) ====

J48 pruned tree

Has_JOB = false
 | Own_house = false: no (4.0)
 | Own_house = true: yes (2.0)
 Has_JOB = true: yes (3.0)
 Number of Leaves : 3
 Size of the tree : 5

Time taken to build model: 0seconds

==== Evaluation on training set ====

==== Summary ====

Correctly Classified Instances	9	100 %
Incorrectly Classified Instances	0	0 %
Kappa statistic	1	
Mean absolute error	0	
Root mean squared error	0	
Relative absolute error	0 %	
Root relative squared error	0 %	
Total Number of Instances	9	

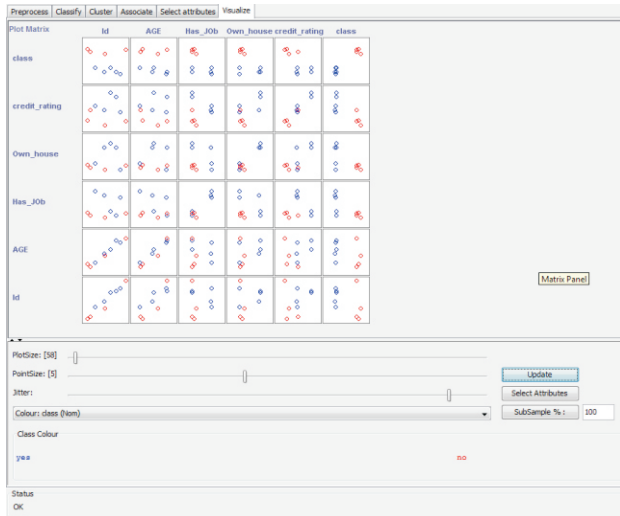
==== Detailed Accuracy By Class ====

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	yes
1	0	1	1	1	1	no
Weighted Avg.	1	0	1	1	1	1

==== Confusion Matrix ====

```
a b <-- classified as
50 | a=yes
04 | b=no
```

Visualize Output:-



Visualize All-



5) CONCLUSION:-

This paper has attempted to for the purpose of Computer (Datamining techniques) usages in organizations. The proposed methods were successfully tested on collected data. In the organization computer is useful and impact also create in the organization. The results which were obtained after the analysis were satisfactory and contained valuable information about the organization.

6) REFERNCES:-

1)	The Impact of Office Automation on the Organization	Margrethe H.Olson	25 Issue 11, Nov 1982
2)	Impact of computers in organizations	Thomas L. Whisler	1970 Year
3)	Priyadarshani Shoot Gini	Shirpur	
4)	Web Mining	Bing liu	Springer
5)	WEKA	Manual	
6)	1) www.au.af.mil/au/cadre/aspj/airchronicles/aureview/.../pribus.html .		
7)	2) Association Rule in Web Usage Mining- D.M. Marathe		