IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 06, Issue 05 (May. 2016), ||V3|| PP 15-18

A practical Approach of Pattern Interesting Measure

Prakashkumar Hitendrabhai Trivedi
Ph.D. Scholar: Shri J J T University, Rajasthan. India

Abstract: As discussed in the present research work, the function of pattern evaluation is utilized to evaluate the interestingness of discovered knowledge from the user's viewpoint. And the research work is the approach which is providing practical approach of interestingness measured. It is focused on advancement in pattern evaluation. Pattern got from the process of data mining Which is carried out from the large pool of data organized in Data Base or Data Warehouse.

Keywords: *Pattern Evaluation*

I. INTRODUCTION

Data mining system has a capacity for developing thousand of pattern. The study has objectives, like identify patterns and interestingness. The interesting pattern has certain capacity. The capacity of that pattern can be estimated by some practical approaches. For estimating the capacity of the study is using the data which have obtained from the education institute. The data are very bulky and the data mining process is functioning on it. For proving the capacity, there is certain hypothesis. The work will test the hypothesis and prove the interestingness.

II. LITERATURE REVIEW

Laube P, Kreveld M, Imfeld S(2004)In the work "Finding REMO-detecting relative motion patterns in geospatial lifeline" have analyzed that there is fundamental concept of moving point object data. It also instruct that they were available since the development of GPS and Radio transmitters.

Miller H.J., Han J.(2001), editor, Geographic Data Mining And Knowledge Discovery. The Geographic data mining performance have noted in the research work.

Romar Kay(2007) Distributed Mining of spatio-temporal event pattern in sensor network. The study based on sensor and provides analysis. There is data mining technique to discover frequent event pattern.

JiaweiHan, Dong Guozhu, Yin Yiwen(1999), Yiwen Yin, Efficient Mining of Partial Periodic Pattern in Time series database have presented the periodic patterns. Time factors prove to be important in this work. The work also provides algorithms for efficient mining of partial periodic patterns.

III HYPOTHESIS

The work will test the hypothesis and prove the interestingness. Hypothesis: (1) The pattern has capacity of easily understandable for a person. (2) Degree of certainty of the pattern can be measured on new or test data. It can be useful. It can be novel. A user has their works with the data and also has their functions with data mining. The user is exploring knowledge from large amount of data obtained from an educational institute. The present study obtains a target. The information obtained by a user will be regarded as a target. For obtaining target there should be a belief. A user has a hypothesis of interesting pattern that can produced knowledge. A pattern is interesting if it validates a hypothesis. Here user's belief is important so subject related interestingness is depending on beliefs of data. Belief has a relation with the data. The present study has a belief based on data.

IV RESEARCH METHODOLOGY

The work is based on scientific research. A practical approach have used from beginning to end of the work.

V PRACTICAL APPROACH

Utilization:

Source of data at initial stage:

Student data from MCMSR collage. Visnagar. SQL Server 2008 kept the data as a data warehouse repository. .mdf is a file name. dbo is table related extension.

Process of converting the file from SQL Server 2008 to Excel:

Right click on database, select task option and select export data. Click the next and choose the data source and choose the server name. Once again select database and press next button. At last choose the destination. Select Microsoft Excel from the dropdown box of the destination. Brows the path and view the Excel new file name

and press open button and next button. At the end select copy the data from one or more table or view. Then select next, once again go to next and press Finish.

1. Data: Large amount of data from the educational institute. Data obtained in .xls format.
2. Belief: A user's view point is a belief. The Belief is constructing on the basis of the present data current system. There will be three kinds of result "Pass", "Fail" and "Absent". A Pattern be Interesting when a users hypothesis is compound.
3. Using Tool: Tanagra 1.4, a data mining software.
4. Algorithms: To be used by the data mining tools (Tanagra 1.4) are

(1) Apriori Algorithm
(2) Clustering Algorithms
2.1 K-Means Algorithms
2.2 Expectation Maximum (EM) Algorithms

5. System to be used: Data mining, Data warehouse

The work utilizes educational institution based data and current area of the work is also related of it. The data, Target, Result, Pattern of the study, Algorithms and Tools to be used are also be related to it. The work is exploring the out put based on two systems (i) Data mining, and (ii) Data warehouse. The present study is exploring the pattern evaluation function. The function is used to evaluate the interestingness of knowledge on the basis of user's view point.

Description of data:

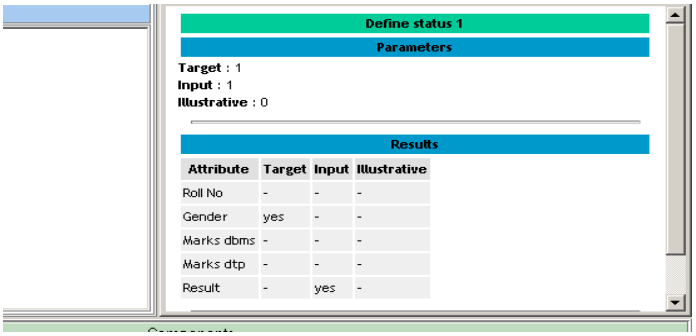
The roll no is a unique and there are five attribute including roll no. The large number of data to be taken from the MCMSR-College. Visnagar. The student of bca college and their marks of ADBMS and DTP Subject is taken. Above 28 marks will be awarded as "Pass". There will be result of "Fail" and "Absent". The mined data is giving the information of overall phenomena of the data. The target result of the mined process will be expected as: (1) how many attributes are used may be obtained. (2) How many result expression may be observed. (3) Gender vice result my be obtained. Sample Data of MCMSR-College. Visnagar

Roll No	Gender	Marks dbms	Marks dtp	Result	Roll No	Gender	Marks dbms	Marks dtp	Result
1	Male	42	57	pass	34	Male	64	60	pass
2	Female	22	18	fail	35	Female	64	56	pass
3	Male	42	52	pass	36	Male	69	62	pass
4	Male	65	52	pass	37	Female	33	36	pass
5	Female	32	30	pass	38	Female	63	58	pass
6	Female	11	51	fail	39	Female	45	54	pass
7	Female	54	52	pass	40	Male	67	53	pass
8	Male	33	35	pass	41	Female	66	53	pass
9	Male	55	53	pass	42	Male	61	53	pass
10	Female	28	33	fail	43	Female	58	53	pass
11	Female	0	0	Absent	44	Male	35	52	pass
12	Male	35	33	pass	45	Male	43	49	pass
13	Male	55	50	pass	46	Female	48	51	pass
14	Male	61	56	pass	47	Male	67	58	pass
15	Female	35	50	pass	48	Male	35	40	pass
16	Female	33	49	pass	49	Female	64	40	pass
17	Male	65	58	pass	50	Male	57	51	pass
18	Female	0	0	Absent	51	Female	38	44	pass
19	Female	63	54	pass	52	Female	67	54	pass
20	Male	65	52	pass	53	Female	38	44	pass
21	Female	65	60	pass	54	Female	64	50	pass
22	Female	66	56	pass	55	Male	10	54	fail

Table 1: Sample data of MCMSR College Visnagar.

The subsets of the data have taken from sybca students of MCMSR college. Visnager. Attributes of the data are Roll No, Marks of DBMS and DTP subject have selected for data mining purpose. The original data have taken in lager bulk there are two division of the student and more attributes have taken. There are marks of different subjects like D.B.M.S, D.T.P, STATESTIC AND OPERATION RESEARCH, C++, NETWORK etc. Tanagra 1.4.48 is using as data mining tool for displaying the data set. The data to be utilized is in .xls file format.

Tanagra 1.4.48 is using certain component. The data visualization, Feature selection, Spv learning, Statistics, Regression, Meta-spv learning, Nonparametric statistics, Factorial Analysis, Spv learning assessment, Instance selection, PLS, Scoring, Feature Construction, Clustering, Association.
Result is displaying as "Yes" in the data mining tool. The Screen shot as per below.



Screen Shot 1: Result of Mined Data using Tanagra 1.4.48

Brief introduction about the algorithms used by Tanagra 1.4.48 as data mining tools.
(1) Apriori Algorithm:
The research work uses classical Apriori Algorithms. The Apriori Algorithms is extracting the association rules. Mining based on association rule is depicted as J= { $p_1 p_2 \ldots p_n$ } taking as a set of n binary attributes identify as item. H= { $q_1 q_2 \ldots q_n$ } taking as a set of transaction identifying as database. H have unique transaction ID as roll no and taken and taken a subset of the item related to J. In the true sense the rule has a phenomena of form X=>Y where X, Y$\underline{C}$ I and X ∩ Y=Ø. The set of items X and Y are regarded antecedent (left hand side) and consequent (Right hand side) of the rule. To find interesting rules from the set of all possible rules. Constraints on various measures of significance and interest can be used. Constraints are related to support and confidence.
(2) Clustering Algorithms:
Clustering Algorithms can be used as (i) K-Means Algorithms and (ii) Expectation Maximum (EM) Algorithms. Clustering is the process of considering a set of object into groups. And so the objects in the same cluster are more similar. Clustering is a main work of data mining. Within this data mining there is statistical data analysis used in the fields, including machine learning, pattern reorganization etc.
Hypothesis conformation
-The pattern has capacity of easily understandable for a person.

3 attribute(s)
66 example(s)

Attribute	Category	Informations
Rno	Continue	-
Result	Discrete	3 values
Jender	Discrete	2 values

Screen Shot 2: Three attribute using Tanagra 1.4.48

-Degree of certainty of the pattern can be measured on new or test data. Depicted as per below.

Results

Attribute	Target	Input	Illustrative
Roll No	-	-	-
Gender	yes	-	-
Marks dbms	-	-	-
Marks dtp	-	-	-
Result	-	yes	-

Screen Shot 3: Degree of certainty of the pattern using Tanagra 1.4.48

-It can be useful. An authority takes decision on the basis of result.
-It can be novel.
The result is the advance and the information is novel.

VI. CONCLUSSION

Here the study introduced statistical implementation of data using data mining tool (Tanagra 1.4.48). The work has indicated case studies. At the end of the work which shows the result and proves the importance of pattern evolution using data mining and warehousing.

REFERENCES

[1] Ahmed, T., & Khan, R. (2012). Intelligent Dision Support System-A Framework. *Information And Knowledge Management , 2*, 6.
[2] Arya, S., Mount, D. M., Netanyahu, N. S., Silverman, R., & Wu, A. (1998). An Optimal Algorithm For Approximate Nearest Neighbour Fixed Dimensions. *ACM , 45* (6), 891-923.
[3] Chazelle, B. (1993). Cutting Hyperplanes For Divide And Conquere. *Discrete And Computational Geometry , 9* (2), 145-158.
[4] Cheng, R., Kalashnikow, D. V., & Prabhakar, S. (2004). Querying Imprecise Data In Moving Object Environments. *IEEE Transactions On Knowledge And Data Engineering* (16), 1112-1127.
[5] Erickson, J. (1999). New Lower Bounds For Convex Hull Problems In Odd Dimensions. *SIAM Journal Of Computing , 28*, 11981214.
[6] Erwig, M., & Sehneiber, M. (2002). Spatio-temporal Predicates. *IEEE Transection On Knowledge and Data Engineering* (14), 881-901.
[7] Fejes, Goodman, J. E., & O'Rourke, J. (2004). *Handbook Of Discrete And Computational Geometry.* Chapman & Hall/CRC.Frank, A. U., Raper, J. F., & Cheylan, J. P. (2001). *Life and Motion of Spatial Socio-Economic Units.* London: Taylor&Frances.