Tikrit Journal of Pure Science 22 (2) 2017 *ISSN: 1813 – 1662 (Print)*
E-ISSN: 2415 – 1726 (On Line)

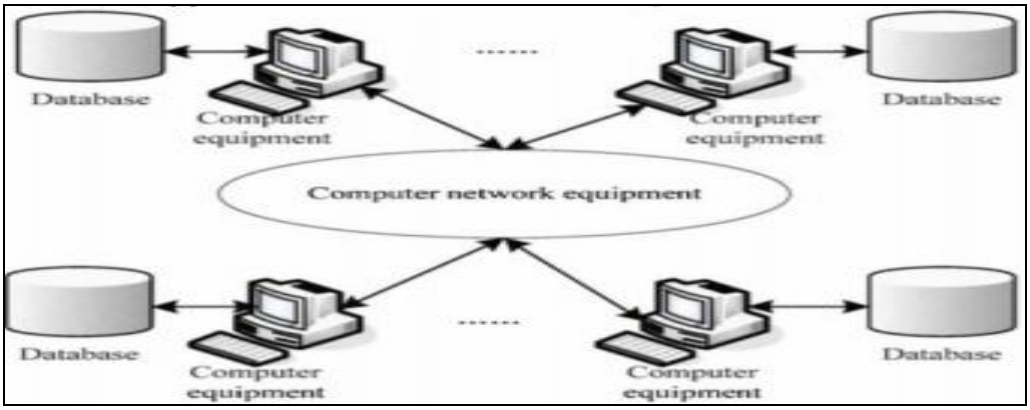
An Electronic System for Summer Training Students Distribution in Organizations with Comparative Study of Association Rule Algorithms

Anhar Khairualdeen Mohammed , Suhair Abid Dawood
MIS Department , Administration & Economic College , Mosul University , Mosul , Iraq

Abstract
The training of students is considered as one of the most promising forms of training to inform students with the reality of practical environment and what they require from serious and exact work. It may give the chance to other public sector organizations to be acquainted with the students' abilities and skills, in addition to the benefits of informing youths to join summer vocation. In order to solve the problem of students distribution to organizations and guarantee the equivalency between students desires and the capacity of governmental and privates offices, some algorithms were used to mine up data to uncover essential hidden relationships with huge data, & Distributed Database has been designed for summer training . The data mining were also used to set reports that may refer to the delicate number of students required for training according to the specializations in the four departments of the College of Administration and Economics (application environments) with the number of nominee students for training in these departments using (oracle 11g.).

Keyword: CDBMS, DBMS, Data Warehouse, Data mining, Apriori Algorithm

1. Introduction
The practicum center of University of Mosul responsible with the placement of students in the industry for the internship program, it is experiencing difficulty in matching organization's requirement with the student profile for several reasons .This situation could lead to a mismatched between organization requirement and students' background , Hence students will face problems in giving good service to the company On the other hand, companies too could be facing difficulties in training the students and assigning them with a project, so we built Database for an integrated Summer Training Students, with integrated computer system that handles all of all summary training students Information .
Distributed Database (DDB) technology emerged as merger of two technologies database technology and data communication technology. These systems have started to become the dominant data management tools for highly accessed data.
Distributed database (DDB) is a collection of multiple logically related database distributed over a computer network[1] as shown in figure(1), and a distributed database management system as a software system that manages a distributed database[2] while making the distribution transparent to the user. Consequently, an application can simultaneously access and modify the data in several databases in a network. Data may be replicated over a network using horizontal and vertical fragmentation similar to projection and selection operations in Structured Query Language (SQL).



Figure(1):Distributed Database

There are two main types of distributed databases is homogeneous database & is heterogeneous database
The proposed program database using Oracle will achieve integration in the data for all section and continuous updating of the query.
Oracle Database is the industry foundation for high performance, scalable, and optimized data warehousing. Oracle Expand data Database Machine is a complete hardware and software solution that delivers extreme performance and database consolidation for data warehousing.
A data warehouse is a relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but can include data from other sources [3]. In general, the data warehouse is maintained separately from operational databases of the organization for several reasons

[4].A data warehouse (DW) is a special database used for storing business-oriented information for future analysis and Decision-making [5].
we use data warehouse in order to provide integrated administrative information for planning and reporting purposes, This DW will be suitable to carry out statistical.

2. Columnar database management system (CDBMS)

A columnar database is a database management system (DBMS) as shown in figure (2) that stores data in columns instead of row [6]. Generally speaking, a row-orientated focus is preferable for online transaction processing (OLTP) systems and column-oriented focus is preferable for online analytical processing (OLAP) systems [7][8].

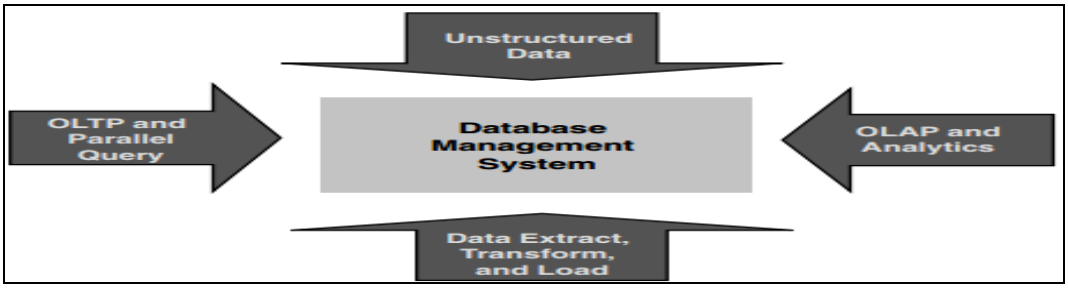


Figure (2): Database Management System Technology

(CDBMS)- Well-suited for data warehouses that have a large number of similar data items. A column-based relational database is exactly what its name suggests, a relational database that stores and retrieves data by column instead of by row as shown in figure(3) [9][6].

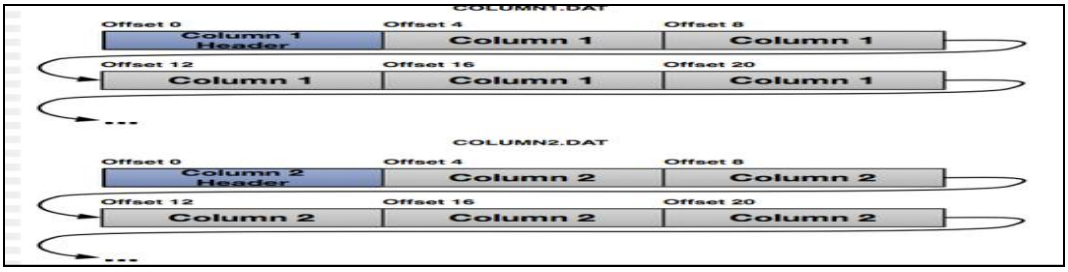


Figure (3): CDBMS

One of the main benefits of a columnar database is that data can be highly compressed. The compression permits columnar operations like MIN, MAX, SUM, COUNT and AVG to be performed very rapidly [10].

Another benefit is that because a column-based DBMSs is self-indexing, it uses less disk space than a relational database management system (RDBMS) containing the same data as shown in figure (4) [11].

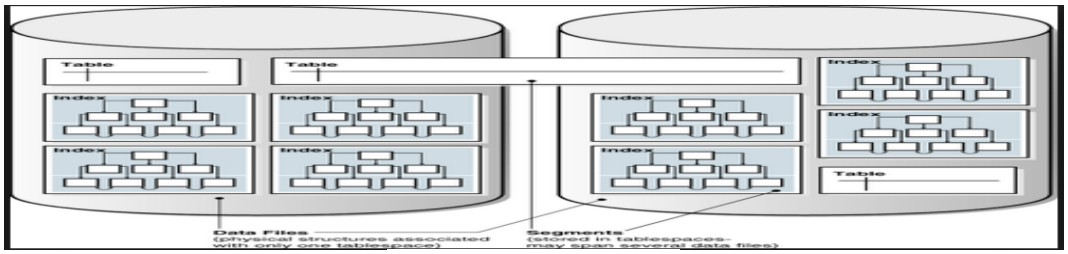


Figure (4): Self –Indexing

3. OLAP

Dr. E.F. Codd introduced the term OLAP (Online Analytical Processing) in 1993[12]. The objective of the OLAP is to facilitate solving data analysis problems and accurate decision-making.
(OLAP) applications and tools are those that are designed to ask "complex queries of large multidimensional collections of data" [13] to provide quick access to strategic information for the purposes of advanced analysis[14].
some additional operators that are more common in OLAP tools, called OLAP operators can be describes in[15] : 1) Drill: de-aggregates. 2) Roll: aggregates. 3) Slice & dice 4) Pivot

3.1 OLAP Technology in the Oracle Database

Oracle Database offers the industry's first and only embedded OLAP server. Oracle OLAP provides

native multidimensional storage and speed-of-thought response times when analyzing data across multiple dimensions. The database provides rich support for analytics such as time series calculations, forecasting, advanced aggregation with additive and no additive operators, and allocation operators. These capabilities make the Oracle database a complete analytical platform, capable of supporting the entire spectrum of business intelligence and advanced analytical applications [16].

The Oracle OLAP option includes the following features[17]:

- OLAP Cube Definition, Storage, and Querying
- OLAP API and Metadata
- OLAP Cube Materialized Views
- Analytic Workspaces
- SQL Access to OLAP Cubes

3.2 OLAP Guidelines

Dr. E.F. Codd created a list of guidelines and requirements as the basis for selecting OLAP systems [18][19] :

1- Basic Features(Multidimensional analysis , Consistent Performance, Fast response times for interactive queries, Drill-down and roll-up, Navigation in and out of details, Slice-and-dice or rotation, Multiple view modes, Easy Scalability, Time intelligence(year-to-date, fiscal period)

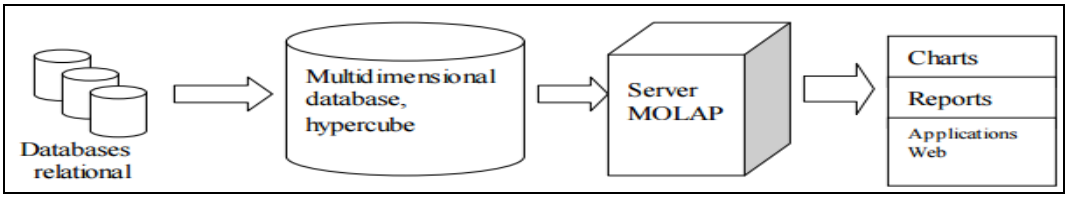
2-Advanced Features(Powerful Calculation, Cross-dimensional Calculations, Pre-Calculation or Pre-Consolidation, Drill-through across dimensions or details, Sophisticated presentation & displays, Collaborative decision making, Derived data values through formulas, Application of alert technology, Report generation with agent technology.

3.3 Types of OLAP servers

ROLAP versus, MOLAP versus, HOLAP

MOLAP (Multidimensional OLAP):

MOLAP systems are much faster in terms of data aggregation and in terms of queries, however, generates large volumes of data hedge. Response time the query is improved because of pre-aggregated summaries of such data and responses to queries are prepared before launching the application [20]. The MOLAP data store is built specifically to handle multidimensional queries as shown in figure(5) and offers fast, efficient, and manageable access to multidimensional data [21].



Figure(5): Architecture MOLAP

The features of MOLAP are [22][23]:

a) Store and manage warehouse data in multi-dimensional DBMS.

b) Array based storage structure.

c) Direct access to array data structure.

d) Excellent performance: MOLAP cubes are built for fast data retrieval, and are optimal for slicing and dicing operations.

e) Can perform complex calculations: All calculations have been regenerated when the cube is created. Hence, complex calculations are not only doable, but they return

4. Association rule algorithms

Association rules are used to find the frequent pattern, association or correlation in a transaction database. Association rule mining can be used in Basket Data Analysis, Educational Data mining, Classification, Clustering etc. The association Rule algorithm is Apriori, sampling, partitioning & Parallel Algorithm.[24].

4.1 Apriori Association Rule

The Apriori algorithm was first proposed by Agrawal [25] .It uses prior knowledge of frequent tools for association rule mining. The basic idea of the Apriori Algorithm is to generate frequent item set of a given dataset and then scan the dataset to check if their counts are really large the process is iterative and candidates of any pass are generated by joining frequent item set of the proceeding pass .Apriori is a confidence–based Association Rule Mining algorithm The confidence is simply accuracy to evaluate rules, produced by this algorithm .The rules are ranked according to the confidence value. if two or more rules share the same confidence then they are initially ordered using Their support and secondly the time of discovery.[26]

Support : for the association rule X→**y** is the percentage of transactions in the database that contains X U Y .

Confidence: For the association rule is X → y is the ratio of the number of transactions that contains X U Y to the number of transactions that contain X.

the generation of item sets & frequent Item sets where the minimum support count is 2.

To generate the association rule from frequent item set we use the following rule:

For each frequent item set L, find all nonempty subsets of L.

For each non-empty subset of L, write the association rule S → (L-S) if support count of L/support count of S>=Minimum Confidence.

The best rule from the item set L= {2, 3, 5} are calculated as follows

Consider the minimum support is a 2 & minimum confidence is 70% all nonempty subsets of {2,3 and 5} are {2,3} , {3,5}, {2}, {3}, {5}as shown in figure(6). [27].

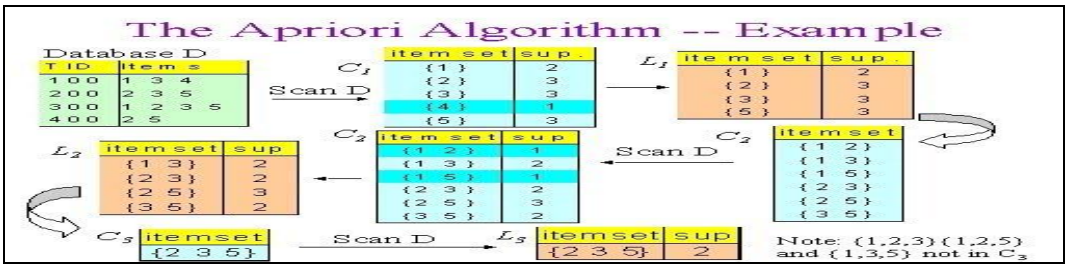


Figure (6) generation item sets & frequent item sets

The algorithm makes many searches in database to find frequent item sets where k-itemsets are used to generate k+1-itemsets. Each k-itemset must be greater than or equal to minimum support threshold to be frequency. Otherwise, it is called candidate item sets. In the first, the algorithm scan database to find frequency of 1-itemsets that contains only one item by counting each item in database. The frequency of 1-itemsets is used to find the item sets in 2-itemsets which in turn is used to find 3-itemsets and so on until there are not any more k-itemsets.[28].

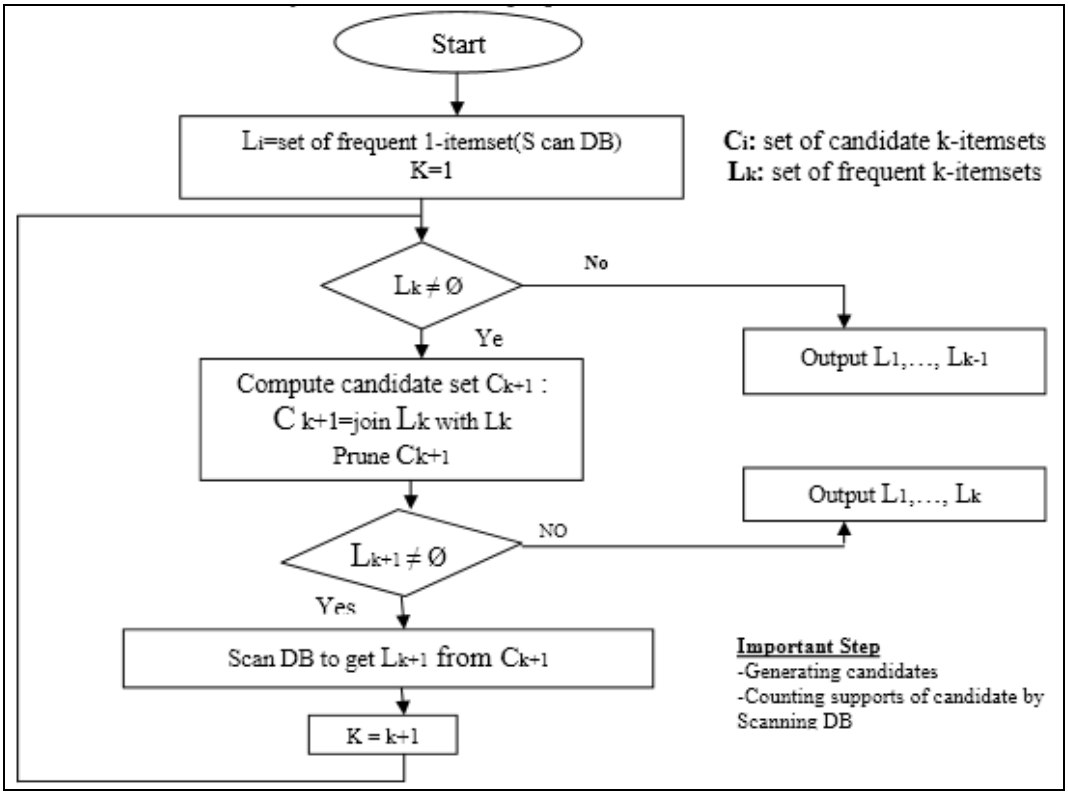


Figure (7) Flow Chart for Apriori Association Rule Algorithm [29]

4.2 Predictive Apriori

In the case of Apriori , every so often we can find rules with higher confidence but low support on respective items of generating rules , sometimes , rules are produced with large support but low confidence [30] introduced this algorithm with the concept of " larger support has to trade against a higher confidence ". Predictive Apriori is also a confidence – based ARM algorithm. But rules ranked by this algorithm are sorted according to "expected predictive accuracy". This interestingness measure of predictive Apriori suits the requirement of a classification task [31] it tries to maximize expected accuracy of an association rule rather than confidence in Apriori .

Finding a unique association rule mining algorithm based on data characteristics

5. Experimental Results

5.1 Appendix (1) explains the flowchart about the summer training students system.

5.2 Designing database

Database Life Cycle as shown in figure(8)

ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

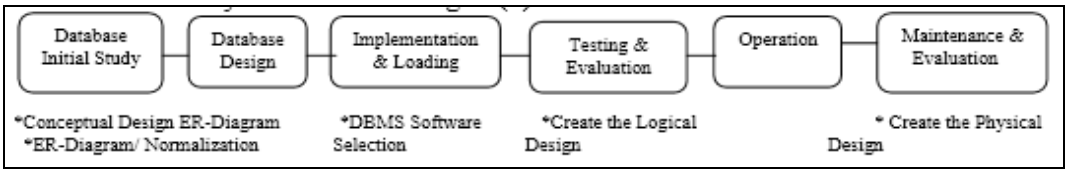


Figure (8): Database Life Cycle

A special database has been designed for summer training for the college of Administration and Economics including many tables, the tables as each entity will be a table in the database and special qualities of this entity will become the fields for this table, and identify the relationships between objects (entity) as the entity selection process should clarify the relationships that bind them. The Appendix (2) shows the E-R (Entity Relationship) gives the conceptual model of the world .

5.3 System Forms

1-Student: a student is regarded as one of the training operation elements (the most important element), for the prospects and attitudes he has and the relationship of that with the level of training and prior planning which is the responsibility of the Administration Departments to attain the expected benefits and the reflection of that on the training operation as a definite gain. A special Form has been designed for the student related to the table of the student existing in the database for summer training, and as explained in figure(9).

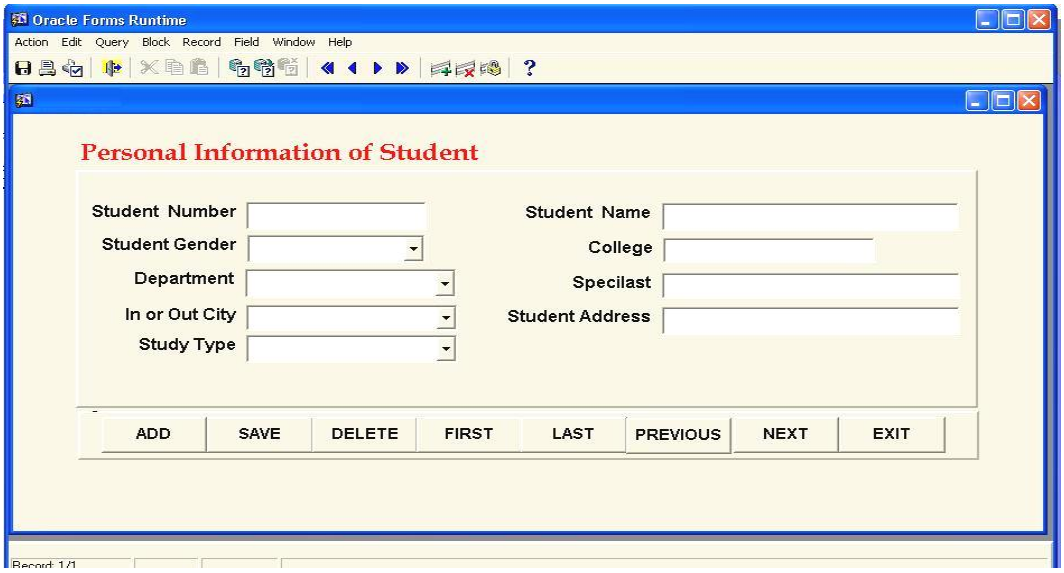


Figure (9) personal Information of Student

2- A member of Teaching Staff: A member of teaching staff forms a basic foundation in the field of summer training, starting from the following up of a student at the site of training and drawing the shape of work for him, together with the coordination of Practical supervisor at site and according to a prior prepared programmed in a detailed manner and a follow up at the field of training and stating the weak and strong points in them and ruling out the weak points. And as the university lecturer acquaints himself with the nature and policy of site work which he intends to supervise and by doing that his experience and reaction with the concerned people will increase, and in turn this is reflected on the development of expertise , and so, three forms have been developed ; a form including the writing down of tasks that a student carries out inside the organization and as shown in figure(10), and figure(11) includes the following up of university lecturer (scientific supervisor) and Practical supervisor (inside the organization) about the student through their giving the weekly merits for the student under training in that organization .

ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

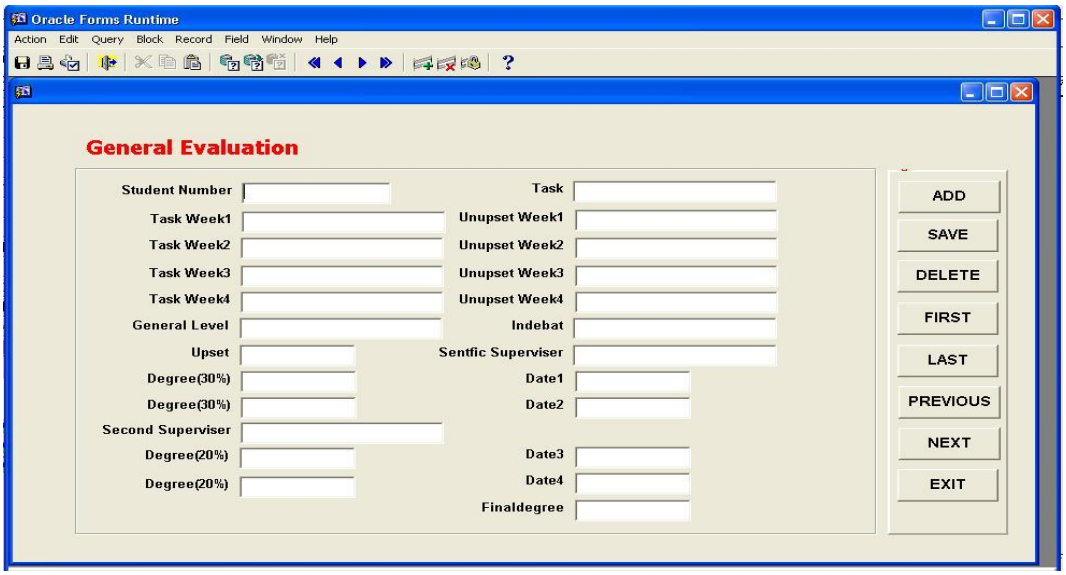


Figure (10) General Evaluation Form

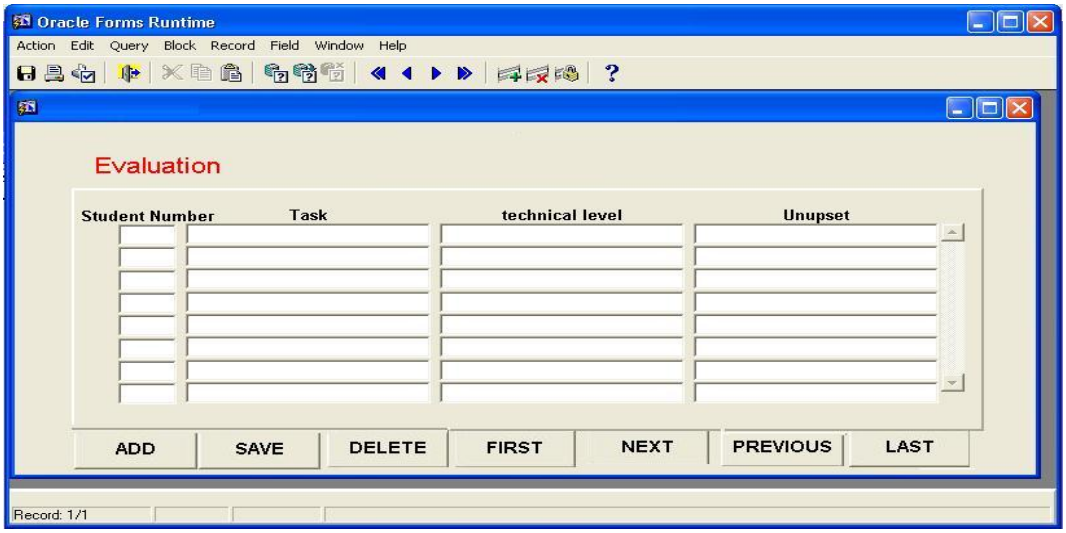


Figure (11) Evaluation Form

5.4 Selection

Data used in this study was obtained from some Departments of College of Administrations & economics. These data have been obtain by different reports among others registered students report , student lists based on city Reports this data include all 2012-2013 Bachelor in administration Department & MIS Department & Accounting department & Department of Banking and Financial Science. The initial data contain The performance profile gathered from a number of 194 students with 16 listed attributes which include company name , number, gender (female, male), Average, Dept/Accounting , Dept/ Banking, Dept/admin , Dept/MIS, Side, address, success stage , /first success stage /second, desire1, desire2, Class, The data contains various types of values either string or numeric value . The target is represented as the organization's name. The Organization name was grouped according to two Categories (Government, private) based on the discussion with the program coordinator; all 194 data are used in this study.

5.5 MOLAP Sample Application

One instruction summer training for university students to create tables containing information on the number of students, section, site training (proposed) the training to be sent to the training unit at college and it prepared reports which include the preparation of students for each department and each department where after approaching state departments to send the number that can be taken in for the purposes of training students and then prepare a plan by the summer training unit for the purpose of training students and then prepare a plan by the summer training unit for the distribution of students to the relevant departments and according to Specialization for each section.

The following are the most important reports that have been obtained using model MOLAP through oracle 11g.

We can show:

ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

- The Dimension which contains (Dept., Location, Time) that shows the hierarchies and levels of time dimension shown in Figure (12).
- The hierarchy levels and the histogram shown in Figure (13) which represents the total number of students expected trained to be in Government Organizations and private sector for each Dept. at the College of Administration and Economics.
- The number of students from each Department of the Departments expected to be trained in each of the government and private Organizations shown in Figure (14).

The installation of the dimensions of the data cube summer training, which includes three dimensions (time, department, Location) shown in Figure (15).

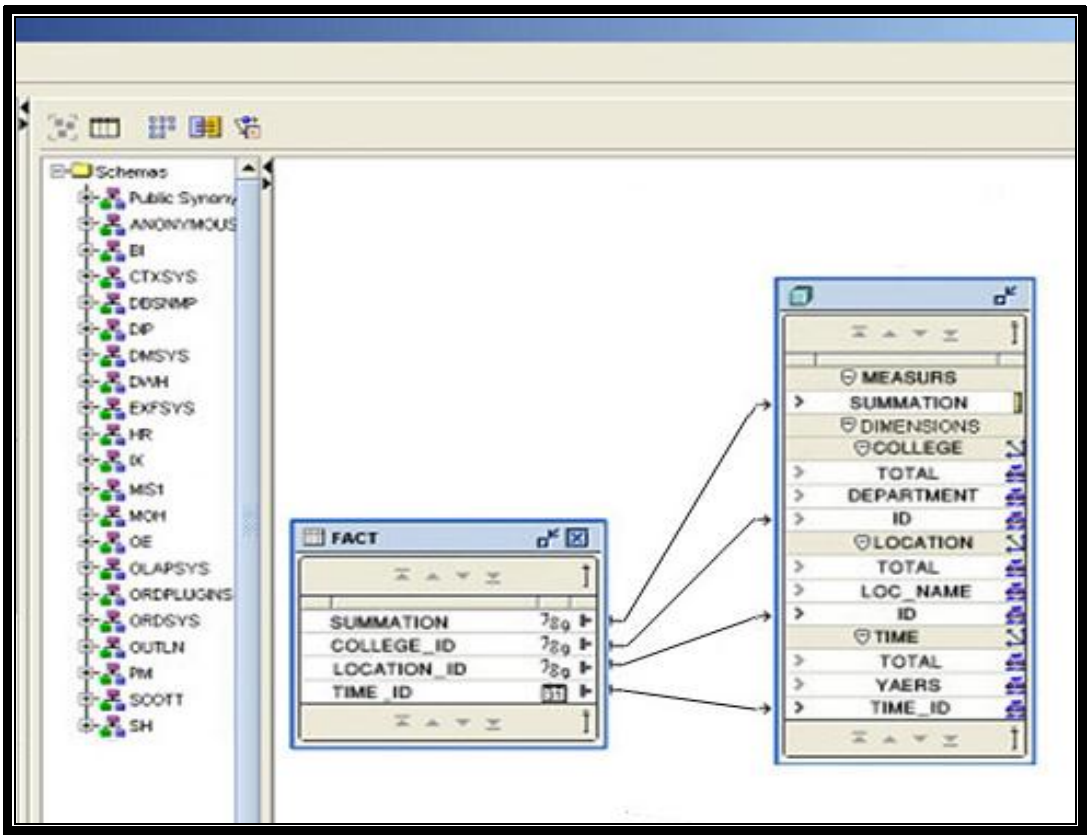


Figure (12) Fact Table

Figure (13) Total no. of students trained in organizations

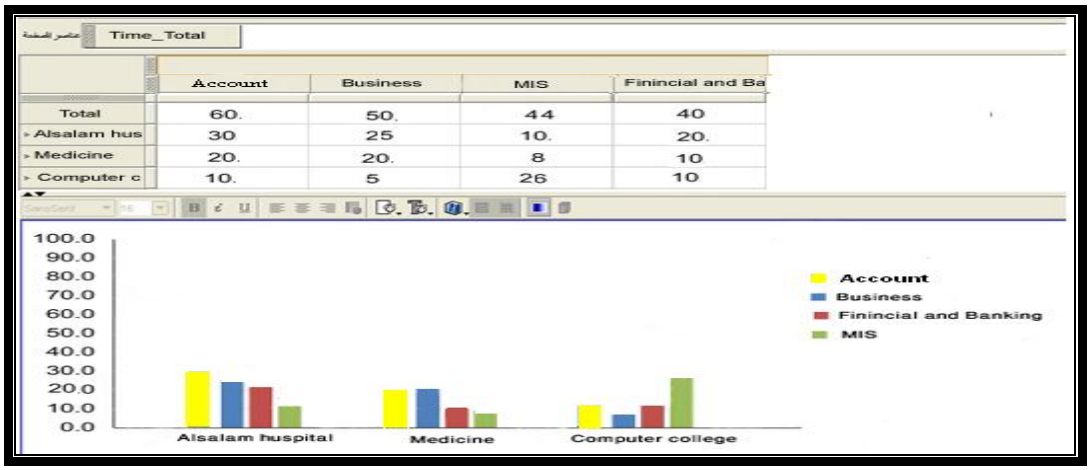
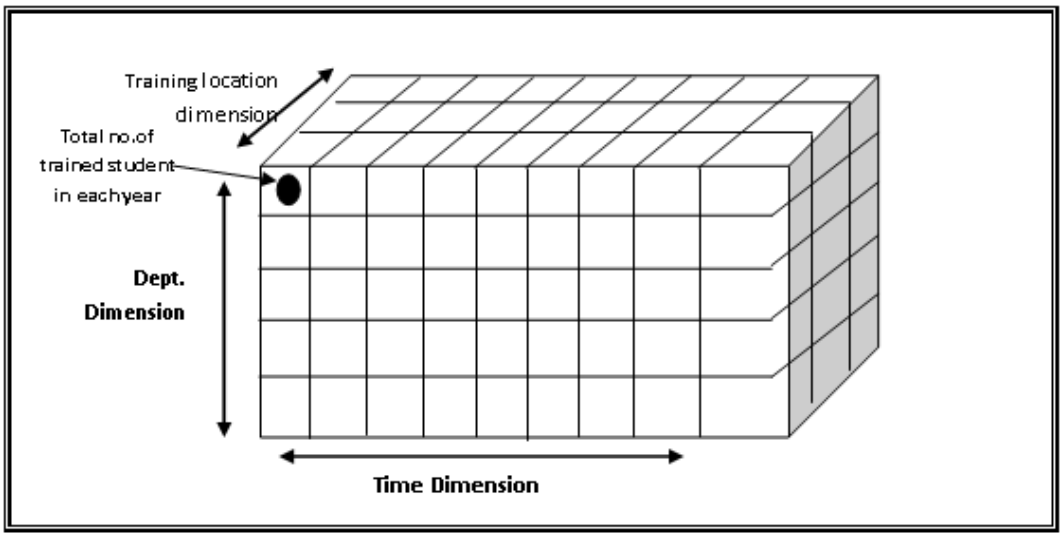


Figure (14): no. of students Trained from each Department



Data Cube Figure (15)

5.6 Apriori & Predictive Apriori Association rule Results

At this stage, we try to compare the two association rule algorithm in predicting the student placement in the organization, Apriori Association rule and Predictive Apriori Association rule, we need the algorithm where the Association rules consist of "Government" and "Private" , so we compare these results using these two Association rule algorithms.

Upon examining table (1), we found that Apriori Algorithm could generate patterns that are believed to be the factors that effect the matching processing process, the data has been grouped into two groups based on the organization category, example of pattern extracted are:

- If student are from the Accounting or from Banking or MIS Department and their Average between 66-70 and Sex=female and their place in the left side of Mosul then the students were placed in Alsalam Hospital in a Government Organization.
- If student are from the MIS Department and their Average between 76-80 and Sex=male and their place in the right side of Mosul then the students were placed at the college of computer science in a Government Organization.
- If student are from the Admin or from Banking or MIS Department and their Average between 71-75 and Sex=female and their place in the right side of Mosul then students were placed in Medicine in a Private Organization.

Table (1): rule generated based on Organization Category

Organization	Region	Criteria (Apriori)
Government	Alsalam Hospital	Major=Accounting & Banking & MIS Average=66-70 female=y SIDE=Left
Government	College of Computer Science	Major=MIS Dept. Average=76-80 male=y SIDE=right
Private	Medicine	Major=Admin &Banking & MIS Average=71-75 female=y SIDE=RIGHT

The Appendix (3) represent the result using Apriori Association Rule, this is used in the placement of students in the organization, As we increase the lower support bound, We get the refined rule as shown in these paragraphs, The rules were evaluated based on the confidence and support the best rule were chosen when confidence is 90% and the support also shows 10% good support.

Paragraph 2 in the Appendix (3), represented result using predictive Apriori association rule algorithm, this predictive accuracy is used to generate the Apriori association rule, the best rules accuracy start at 0.99329 and decrease to 0.62506.

6. Conclusion and future work

-we found that the Apriori Association rule algorithm performed best with confidence based ranking and predictive Apriori had performed better on accuracy based ranking.

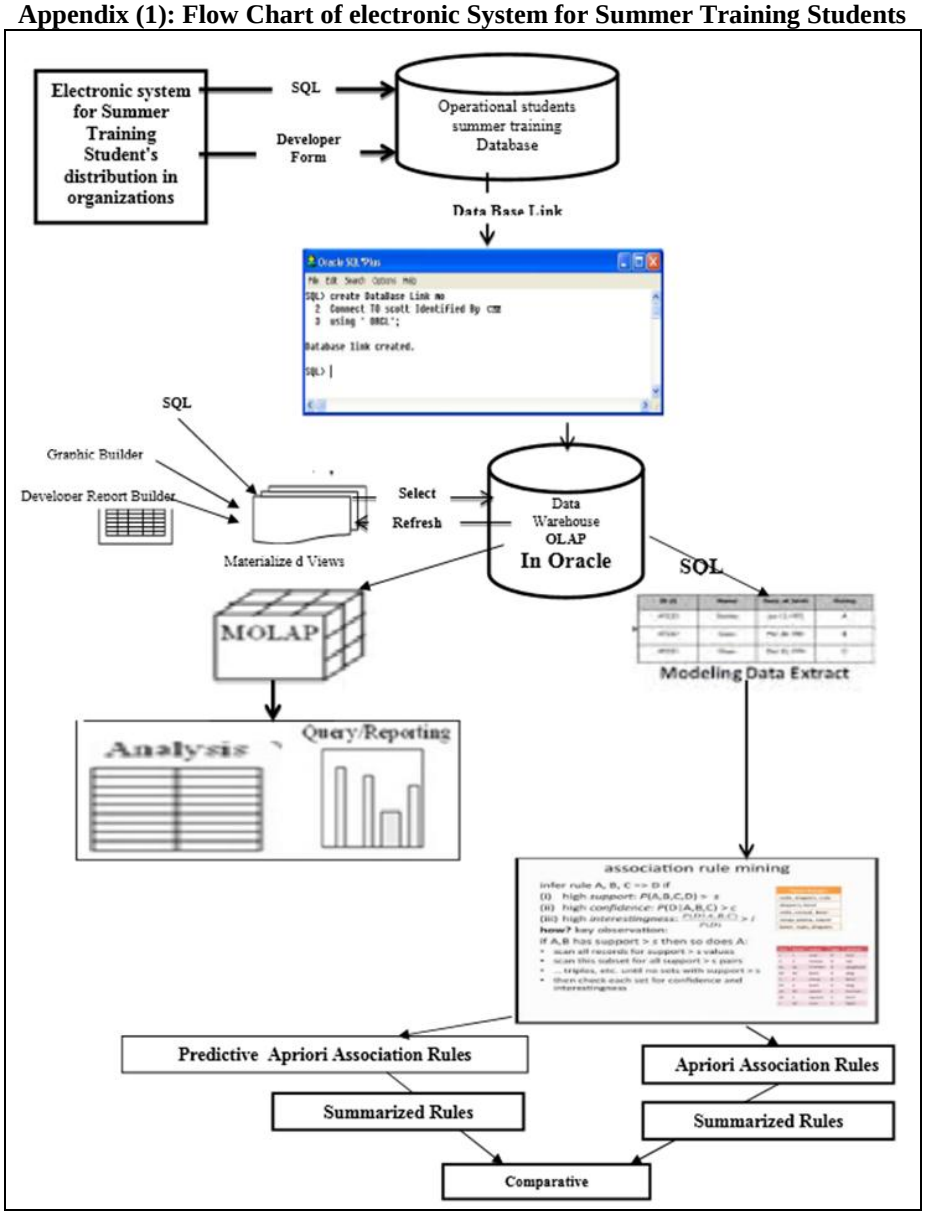
- The distributed database designed for summer training makes it possible to share data by multiple applications or users & reaching the saved data in the database.

Data warehouse provides summer training information, highly detailed reports (shows the results of inquiries in multiple formats through figures and charts), analyzes of value and quality, as derived data is formatted through construction processes (extraction, transform, and load) before they are loaded into the warehouse database.

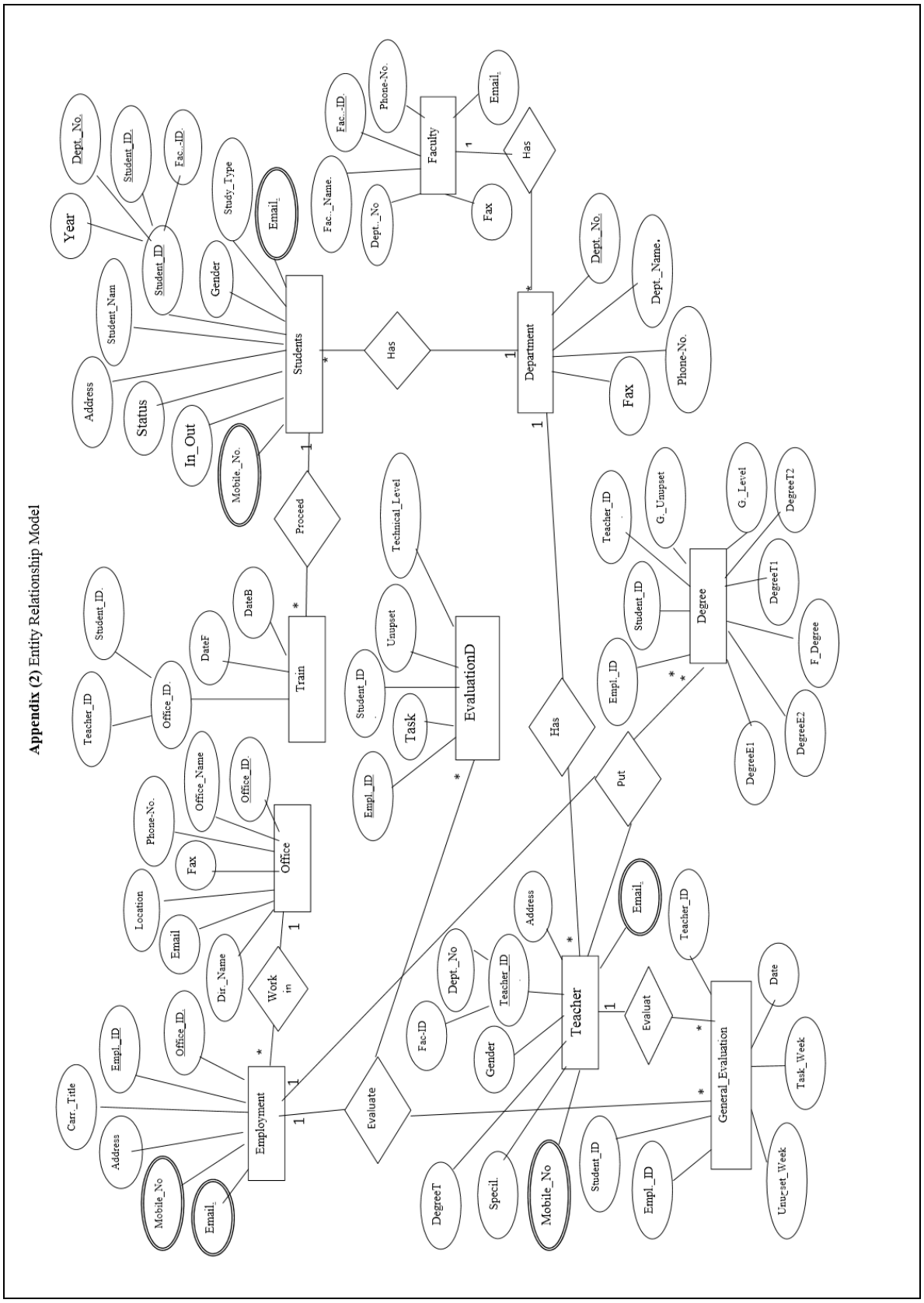
As Future Work includes:

- Other categories OLAP can be applied such as ROLAP…etc.
- Other Association Rule Algorithms can be applied in Distribution of the student placement in organizations.

Appendix (1): Flow Chart of electronic System for Summer Training Students



ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)



Appendix (2) Entity Relationship Model

Appendix(3): Apriori Association Rule & Predictive Apriori algorithm Rule

1. Apriori Association Rule

Attributes: 16
company name
number
female
male
Average
Dept/accounting
Dept/banking
Dept/admin
Dept/MIS
SIDE
address
SUCCES STAGE/first
SUCCES
STAGE/second
desier1
desire2
class

1.1 Apriori
=======
1.Minimum support: 0.15 (29 instances)
2.Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Best rules found:

1. male=y 77 ==> class=government 77 conf:(1)
2. SIDE=LEFT 70 ==> class=government 70 conf:(1)
3. female=y SIDE=LEFT 68 ==> class=government 68 conf:(1)
4. desier1=Alsalam Hospital 65 ==> class=government 65 conf:(1)
5. desire2=college of computer Science57 ==> class=government 57 conf:(1)
6. company name=Alsalam Hospital 51 ==> class=government 51 conf:(1)
7. company name=medicine 51 ==> class=private 51 conf:(1)
8. company name=college of computer Science51 ==> class=government 51 conf:(1)
9. Average=66-70 51 ==> class=government 51 conf:(1)
10. Average=71-75 51 ==> class=private 51 conf:(1)
11. Average=76-80 51 ==> class=government 51 conf:(1)
12. company name=Alsalam Hospital female=y 51 ==> class=government 51 conf:(1)
13. company name=Alsalam Hospital Average=66-70 51 ==> class=government 51 conf:(1)
14. company name=Alsalam Hospital SIDE=LEFT 51 ==> class=government 51 conf:(1)
15. company name=Alsalam Hospital desier1=Alsalam Hospital 51 ==> class=government 51 conf:(1)
16. company name=Alsalam Hospital dept/accounting=y 23==> class=government 23 conf:(1)
17. company name=Alsalam Hospital dept/money=y 20==> class=government 20 conf:(1)
18. company name=Alsalam Hospital dept/mis=y 8==> class=government 8 conf:(1)
19. company name=medicine female=y 51 ==> class=private 51 conf:(1)
20. company name=medicine Average=71-75 51 ==> class=private 51 conf:(1)
21. company name=medicine SIDE=right 51 ==> class=private 51 conf(1)
22. company name=medicine dept/Admin=y 30==> class=private 30 conf:(1)
23. company name=medicine dept/mis=y 17==> class=private 17 conf:(1)
24. company name=medicine dept/money=y 4==> class=private 4 conf:(1)
25. company name=college of computer Science male=y 51 ==> class=government 51 conf:(1)
26. company name=college of computer Science Average=76-80 51 ==> class=government 51 conf:(1)
27. company name=college of computer Science Side=right 51 ==> class=government 51 conf:(1)
28. company name=college of computer Science desire2=college of computer Science 51 ==> class=government 51 conf:(1)
29. company name=college of computer Science Dept/mis =y 45 ==> class=government 45 conf:(1)
30. company name=college of computer Science dept/admin=y 6 ==> class=government 6 conf:(1)

1.2 Apriori
=======
1.Minimum support: 0.25 (49 instances)
2.Minimum metric <confidence>: 0.9
Number of cycles performed: 15

Generated sets of large item sets:

Size of set of large item sets L(1): 16

Size of set of large itemsets L(2): 26

Size of set of large item sets L(3): 24

Size of set of large item sets L(4): 11

Size of set of large item sets L(5): 2

The same Best rules found in minimum support 0.15:

Attributes: 1 **company name** **number** **female** **male** **Average** **Dept/accounting** **Dept/banking** **Dept/admin** **Dept/Mis** **SIDE** **address** **SUCCES STAGE/first** **SUCCES** **STAGE/second** **desier1** **desire2** **class**	**1.3 Apriori** **=======** **1.Minimum support: 0.3 (59 instances)** **2.Minimum metric <confidence>: 0.9** **Number of cycles performed: 14** **Generated sets of large item sets:** **Size of set of large item sets L(1): 6** **Size of set of large item sets L(2): 1** **Best rules found:** **1. male=y 77 ➔ class=government 77 conf (1)** **2. SIDE=LEFT 70 ➔ class=government 70 conf (1)** **3. female=y SIDE=LEFT 68 ➔ class=government 68 conf (1)** **4. desier1=Alsalam Hospital 65 ➔ class=government 65 conf (1)**	**1.4 Apriori** **=======** **1.Minimum support:0.4 and 0.5 and 0.6 and 0.7 and 0.8 and 0.9** **2.Minimum metric <confidence>: 0.9** **No large item sets and rules found**
	2. Predictive Apriori algorithm Rule	
Attributes: 16 **company name** **number** **female** **male Average** **Dept/accounting** **Dept/Bank** **Dept/admin** **Dept/MIS** **SIDE address** **SUCCES STAGE/first** **SUCCES** **STAGE/second** **desier1** **desire2 class**	**1. male=y 77 ==> class=government 77 acc:(0.99329)** **2. SIDE=LEFT 70 ==> class=government 70 acc:(0.99295)** **3. desier1=Alsalam Hospital 65 ==> class=government 65 acc:(0.99264)** **4. desire2=college of computer Science57 ==> class=government 57 acc:(0.99199)** **5. company name=Alsalam Hospital 51 ==> class=government 51 acc:(0.99132)** **6. company name=medicine 51 ==> class=private 51 acc:(0.99132)** **7. company name=college of computer Science51 ==> class=government 51 acc:(0.99132)** **8. Average=66-70 51 ==> class=government 51 acc:(0.99132)** **9. Average=71-75 51 ==> class=private 51 acc:(0.99132)** **10. Average=76-80 51 ==> class=government 51 acc:(0.99132)** **11. female=y SIDE=right 51 ==> class=private 51 acc:(0.99132)** **12. address=hay alkdsea 41 ==> class=government 41 acc:(0.9896)** **13. SIDE=right desire2=medicine 38 ==> class=private 38 acc:(0.98884)** **14. company name=Cement Plant 27 ==> class=government 27 acc:(0.98415)** **15. Average=50-100 27 ==> class=government 27 acc:(0.98415)** **16. female=y desier1=medicine 26 ==> class=private 26 acc:(0.98348)** **17. address=hay alsuger 23 ==> class=government 23 acc:(0.98103)** **18. desier1=medicine desire2=medicine 17 ==> class=private 17 acc:(0.97302)** **19. company name=hos.alknsm 16 ==> class=government 16 acc:(0.97101)** **20. Average=50-55 16 ==> class=government 16 acc:(0.97101)** **21. desier1=General Hospital 15 ==> class=private 15 acc:(0.96872)** **22. female=y address=TALKEF 13 ==> class=private 13 acc:(0.963)** **23. mis=y desier1=Cement Plant 13 ==> class=government 13 acc:(0.963)** **24. SIDE=right address=TALKEF 13 ==> class=private 13 acc:(0.963)** **25. SIDE=right address=FLFEL 13 ==> class=private 13 acc:(0.963)** **26. address=hay almuthna 12 ==> class=government 12 acc:(0.95939)** **27. address=ALHDBAA 12 ==> class=private 12 acc:(0.95939)** **28. accounting=y desier1=medicine 12 ==> class=private 12 acc:(0.95939)** **29. address=FLFEL desire2=medicine 12 ==> class=private 12 acc:(0.95939)** **30. desier1=hos.alknsm 11 ==> class=government 11 acc:(0.95511)** **31. desire2=factory algazzel alnaseeg 11 ==> class=government 11 acc:(0.95511)** **32. address=TALKEF desire2=medicine 10 ==> class=private 10 acc:(0.94998)** **33. SUCCES STAGE/second=SECOND desire2=hos.alknsm 10 ==> class=goverment 10 acc:(0.94998)** **34. admin=y SIDE=right SUCCES STAGE/first=FIRST 10 ==> class=private 10 acc:(0.94998)** **35. address=hay albreed 9 ==> class=government 9 acc:(0.94375)** **36. admin=y desire2=hos.alknsm 9 ==> class=government 9 acc:(0.94375)** **37. address=BESAn 8 ==> class=private 8 acc:(0.93606)** **38. desire2=college admin & economic 8 ==> class=government 8 acc:(0.93606)** **39. accounting=y SUCCES STAGE/second=SECOND desier1=Cement Plant 8 ==> class=government 8 acc:(0.93606)** **40. address=TALKEF SUCCES STAGE/first=FIRST 7 ==> class=private 7 acc:(0.9264)** **41. address=FLFEL SUCCES STAGE/second=SECOND 7 ==> class=private 7 acc:(0.9264)** **42. admin=y SUCCES STAGE/second=SECOND desire2=medicine 7 ==> class=private 7 acc:(0.9264)** **43. SUCCES STAGE/second=SECOND desier1=Cement Plant 34 ==> class=government 32 acc:(0.91892)** **44. desire2=office alshaa nignoa 6 ==> class=government 6 acc:(0.91397)** **45. admin=y address=FLFEL 6 ==> class=private 6 acc:(0.91397)** **46. admin=y address=badoush SUCCES STAGE/first=FIRST 6 ==> class=government 6 acc:(0.91397)** **47. desire2=hos.alknsm 18 ==> class=government 17 acc:(0.90963)**	

	48. accounting=y SIDE=right SUCCES STAGE/first=FIRST 17 ==> class=private 16 acc:(0.90397) **49. desier1=Cement Plant 49 ==> class=government 45 acc:(0.90199)** **50. accounting=y address=TALKEF 5 ==> class=private 5 acc:(0.89753)** **51. accounting=y address=FLFEL 5 ==> class=private 5 acc:(0.89753)** **52. accounting=y address=badoush SUCCES STAGE/second=SECOND 5 ==> class=government 5 acc:(0.89753)** **53. admin=y SIDE=right desier1=medicine 5 ==> class=private 5 acc:(0.89753)** **54. admin=y SUCCES STAGE/first=FIRST desier1=medicine 5 ==> class=private 5 acc:(0.89753)** **55. address=TALKEF 14 ==> class=private 13 acc:(0.88292)** **56. address=FLFEL 14 ==> class=private 13 acc:(0.88292)** **57. address=hay alarabe 4 ==> class=government 4 acc:(0.87499)** **58. address=hay almasarf 4 ==> class=government 4 acc:(0.87499)** **59. address=hay alwahda 4 ==> class=government 4 acc:(0.87499)** **60. address=hay albaker 4 ==> class=government 4 acc:(0.87499)** **61. female=y desier1=Cement Plant 4 ==> class=private 4 acc:(0.87499)** **62. SIDE=right address=badoush 4 ==> class=private 4 acc:(0.87499)** **63. SIDE=right desire2=Alsalam Hospital 4 ==> class=private 4 acc:(0.87499)** **64. desier1=medicine desire2=Alsalam Hospital 4 ==> class=private 4 acc:(0.87499)** **65. money=y 44 ==> class=government 39 acc:(0.86957)** **66. admin=y SIDE=right 20 ==> class=private 18 acc:(0.86448)** **67. mis=y SUCCES STAGE/first=FIRST 20 ==> class=government 18 acc:(0.86448)** **68. mis=y 46 ==> class=government 40 acc:(0.85417)** **69. SUCCES STAGE/second=SECOND desire2=medicine 18 ==> class=private 16 acc:(0.85069)** **70. address=badoush 30 ==> class=government 26 acc:(0.84376)** **71. address=hay almshaq 3 ==> class=government 3 acc:(0.84255)** **72. address=hay alfalah 3 ==> class=government 3 acc:(0.84255)** **73. address=hay domeez 3 ==> class=government 3 acc:(0.84255)** **74. address=hay althreer 3 ==> class=government 3 acc:(0.84255)** **75. desier1=QAIM hospital 3 ==> class=government 3 acc:(0.84255)** **76. desier1=Technical Institute of Nineveh 3 ==> class=private 3 acc:(0.84255)** **77. desire2=bank alrasheed/alzhoor 3 ==> class=government 3 acc:(0.84255)** **78. mis=y desire2=Alsalam Hospital 3 ==> class=private 3 acc:(0.84255)** **79. admin=y desire2=medicine 19 ==> class=private 16 acc:(0.80956)** **80. accounting=y SIDE=right 27 ==> class=private 22 acc:(0.79311)** **81. number=qqq 2 ==> class=private 2 acc:(0.79259)** **82. address=hay alzhoor 2 ==> class=government 2 acc:(0.79259)** **83. desier1=office alshaa nignoa 2 ==> class=government 2 acc:(0.79259)** **84. desire2=college medicine 2 ==> class=government 2 acc:(0.79259)** **85. desire2=Technical Institute of Nineveh 2 ==> class=private 2 acc:(0.79259)** **86. admin=y desier1=college medicine 2 ==> class=government 2 acc:(0.79259)** **87. SUCCES STAGE/second=SECOND desire2=Alsalam Hospital 2 ==> class=private 2 acc:(0.79259)** **88. female=y money=y address=badoush 2 ==> class=private 2 acc:(0.79259)** **89. female=y address=badoush SUCCES STAGE/second=SECOND 2 ==> class=private 2 acc:(0.79259)** **90. desire2=medicine 49 ==> class=private 38 acc:(0.76471)** **91. SUCCES STAGE/second=SECOND 87 ==> class=government 67 acc:(0.76405)** **92. admin=y desier1=medicine 6 ==> class=private 5 acc:(0.75355)** **93. female=y address=badoush 5 ==> class=private 4 acc:(0.71721)** **94. SUCCES STAGE/first=FIRST 108 ==> class=government 77 acc:(0.70909)** **95. SIDE=right desier1=medicine 41 ==> class=private 26 acc:(0.6279)** **96. admin=y 49 ==> class=government 31 acc:(0.62745)** **97. SIDE=right SUCCES STAGE/first=FIRST 49 ==> class=private 31 acc:(0.62745)** **98. desire2=Alsalam Hospital 6 ==> class=private 4 acc:(0.62506)** **99. mis=y SUCCES STAGE/second=SECOND desier1=medicine 6 ==> class=private 4 acc:(0.62506)**

References

1. S. K. Verma, Fragmentation Techniques for Distribution Database: A Review, International Journal of Innovative Computer Science & Engineering Volume 3 Issue 2; (2016),PP. 47-50

2. Karanvir Kaur, Ravi Kant Sahu, Query Optimization in Distribution Database, International Journal of Advanced Research in Computer Science and software Engineering,Volume 6,(2016),No.,PP679-684

3. O. Help, Center, Database Data Warehousing Guide,(2014), Oracle and/or its affiliates, p.**50**

4. W.Mundane, M.Hussein, M.Moukhtar, and F.M.Camino, Intelligent Data Compression Approach in Multidimensional Data Warehouse, International Journal on Computer Science and Engineering, Vol. 3, (2011), No. 2, pp. 951- 960.

5. L.Sapir and S.Armin, A Methodology for the Design of a Fuzzy Data Warehouse, 4th International

IEEE Conference "Intelligent Systems**",** (2008), IEEE.
6. D. Abadi, P. Boncz, S. Harizopoulos, S. Idreos and S. Madden, "The Design and Implementation of Modern Column-Oriented Database Systems", Foundations and Trends R in Databases, Vol. 5, (2014),No. 3 ,PP.197–280
7. S.Srinivasan, " DATA MODELING Applying Relevant Data Modeling Techniques to MPP Columnar Stores - A Point of View", A Sapient Case Study, (2014), Sapient Corporation,P.30 .
8. A. Bhatia , S. Patil, "Column Oriented DBMS an Approach", International Journal of Computer Science & Communication Networks, Vol. 1,(2011), No.2, PP.111-116.
9. http://birtanalytics.actuate.com
10. W. Mcknight, Best Practices in the Use of COLUMNAR DATABASES - How to select workloads for columnar database based on the benefits provided, Calpont,(2011),P.6.
11. G. MATEI,"Column-Oriented Databases, an Alternative for Analytical Environment", Database Systems Journal Vol. I, (2010), No. 2, PP.3-16
12. J. Celko's, Analytics and OLAP in SQL", Morgan Kaufmann Publishers is an imprint of Elsevier, (2006), PP.58-62.
13. A. Khalil, Advanced Database Systems, The American University in Cairo, (2004), Department of Computer Science, p.1.
14. P.Rob, C.Coronel, K.Crockett, Database Systems Design, Implementation &
Management, International Edition,(2008), Edition Cengage Learning EMEA, London, British, P.757.
15. K.Frank , H. A. Dehne, T. Eavis, B. Liang, Compressing Data Cube in Parallel OLAP
Systems. . Data Science Journal, Vol. 6,(2007), PP. 184-197 .
16. B. Jongmuenwai, P. Chinjoho and S.h Prachai, Data Warehouse use Geographic Information System for Ornamentals, Proceedings of the International Multi Conference of Engineers and Computer Scientists(IMECS), Vol I, (2016), Hong Kong, PP 157-162
17. M. Ahluwalia, P. Avril, C, Berger, M. Bird, C. Bruse, et.cl., Oracle Database Licensing Information, 11g Release 2 (11.2),
18. http://www.1keydata.com/datawarehousing/molap-rolap.html, accessed on 3 August (2009).
19. A. Berson, S.j.Smith , "Data Warehousing, Data Mining & OLAP", 1st Edition, (2008), McGraw-Hill, PP. 205-252.
20. M. Wilkie and A. Zaima, " cubes by design: ROLAP and HOLAP solutions using SAS and Teradata ", (2009),P.4 http://www.teradata.com/tdmo/v08n03/Tech2Tech/AppliedSolutions/CubesByDesign.aspx accessed .
21. A. Weinberger, M. Ender, "The Power of Hybrid OLAP in a Multidimensional World", (2000), SAS Institute Inc., Cary, NC, P.44.
22. A. Parekh, Introduction on Data Warehouse with OLTP and OLAP, International Journal Of Engineering And Computer Science, Volume2, (2013), No.8,PP 2569-2573.
23. K. Jayapriya., E. Girija, R.Uma M.C.A, A Technical Review on On-Line Analytical Processing (OLAP), International Journal of Computer Science & Communication Networks, Vol.5, (2011), NO.5, PP.315-324
24. S. B. Aher, L.M. Lobo, Data Mining in Educational System uses WEKA, IJCA Proceedings of International conference on Emerging Technology Trends, Foundation of computer science, (2011),New York, USA,P.20
25. B.Marvin, the Impact of Data Imputation Methodologies for Knowledge Discovery, Doctor of Business Administration, (2008),Morehead State University,P.2
26. M.M. Mazid, A. B.Shawkat A. Kevin, S. Tickle, "Finding a Unique Association Rule Mining Algorithm Based on Data Characteristics",5th International Conference on Electrical and Computer Engineering ICECE , (2008), PP.22-23
27. S.B. Aher, L.M. Lobo, A Comparative study of Association Rule algorithm for course Recommender system in E-learning , International Journal of Computer Applications, Volume 39, (2012),No.1, PP0975-8887
28. M. Mohammed & A. Mohammed, AN IMPROVED APRIORI ALGORITHM FOR ASSOCIATION RULES, International Journal on Natural Language Computing (IJNLC) 0,Vol. 3, (2014),No.1.P.22
29. N.C.Dhande, IDENTIFICATION BEYOND PASSWORD AN ALGORITHMIC APPROACH TO HUMAN MACHINE INTERFACE**,** IJAIT**,** Vol. 5,(2013). No. 2**,** PP.72.
30. T. Scheffer, Finding Association Rules that trade support Optimally Against confidence, in proceedings of the 5th European conference on principles and practice of knowledge Discovery in databases (PKDD'01), (2001),Freiburg, Germany: Springer – Verlag, PP.424-435 .
31. S. Mather ,M. Hall, and E. Frank., using classification to evaluate the output of confidence based association rule mining, Lecture notes in Artificial Intelligence, Advances in artificial intelligence – AI, Berlin, Springer, Vol. 3339, (2004), PP.538-549 .

Tikrit Journal of Pure Science 22 (2) 2017 *ISSN: 1813 – 1662 (Print)*
E-ISSN: 2415 – 1726 (On Line)

نظام الكتروني لتوزيع طلبة التدريب الصيفي على المنظمات مع حالة مقارنة لخوارزميات قواعد الارتباط

أنـهار خيرالدين محمد ، سـهير عبد داؤود

قسم نظم المعلومات الإدارية ، كلية الإدارة والاقتصاد ، جامعة الموصل ، الموصل ، العراق

الملخص

يُعد تدريب الطلبة خلال الإجازة الصيفية أحد أشكال التدريب الهادفة إلى تعريف الطلاب بواقع البيئة العملية وما تحتاج إليه من جد وانضباط ، كما أنه من جهة أخرى يتيح الفرصة لمؤسسات القطاع العام والخاص للتعرف على إمكانيات الطلاب ومهاراتهم ، هذا بالإضافة إلى فوائده فيما يتعلق بتوجيه الشباب نحو الاستفادة من الإجازة الصيفية .ولحل مشكلة توزيع الطلاب على المنظمات ولضمان التوافقية بين رغبات الطلبة واستيعاب الدوائر الحكومية والأهلية تم استخدام بعض من خوارزميات تنقيب البيانات (Aprior, Predictive) لاكتشاف علاقات هامة مخفية في مجموعة البيانات الضخمة، وتم تصميم قاعدة بيانات موزعة بكل مايتعلق بالتدريب الصيفي فضلاً عن استخدام مستودع البيانات (DW) لغرض إعداد تقارير يتضح من خلالها أعداد الطلبة المطلوب تدريبهم في المنظمات وحسب التخصصات لأربعة اقسام من كلية الإدارة والاقتصاد (بيئة التطبيق) مع أعداد الطلبة المرشحين للتدريب في كل قسم باستخدام برنامج (Oracle 11g.) .