International Journal of Computer Applications Technology and Research
Volume 3– Issue 8, 519 - 528, 2014, ISSN: 2319–8656

Data mining Innovativeness of data give-and-take service station in Teradata client-server construction

K. Hepsiba
Department of CSE
Gokula Krishna College of Engineering,
Sullurpet, Nellore(Dt)
Andhra Pradesh, India.

Y. Madhusekhar
Department of CSE
Gokula Krishna College of Engineering,
Sullurpet, Nellore(Dt)
Andhra Pradesh, India.

Abstract: Teradata is a relational database management system that drives a company's data warehouse. Teradata provide the foundation to give a company the power to grow, to complete in today's dynamic marketplace, to achieve the goal of "transforming transactions into Relationships" and to evolve the business by getting answer to a new generation of questions. Teradata's scalability allows the system to grow as the business grows, from gigabytes to terabytes and beyond. Teradata's unique technology has been proven at customer sites across industries and around the world. Teradata is a large database server that accommodates multiple client applications making inquiries against it concurrently. Various client platforms access the database through a TCP-IP connection across an IBM mainframe channel connection. The ability to manage large amounts of data is accomplished using the concept of parallelism, where in many individual processors perform smaller tasks concurrently to accomplish an operation against a huge repository of data.

Keywords RDBMS, Data warehouse, Transformation, Scalability, Parallelism, shared-nothing, server-client architecture, leaner expansion.

1. INTRODUCTION

Teradata is a relational database management system which is especially designed for running very large commercial databases. Teradata uses the parallelism to manage terabytes of data. Teradata is a shared nothing architecture. Can start with teradata as small as gigabytes and grow large as volume of data increase. Teradata supports UNIX and Windows operating system. Teradata supports ANSI standard SQL. Teradata act as a database server for many client applications. Teradata supports the Network and mainframe connectivity. Fault tolerance at all levels of hardware and software. It has the data integrity and reliability.

1.1 Brief History

In 1979 Teradata corporation founded in Los Angeles, California Development begins on a massively parallel database computer.

In 1984 Teradata sells the first database computer DBC/1012 to wells Fargo Bank of California.

In 1989 Teradata and NCR partner on next generation of DBC.

In 1990 First Terabyte system installed and in production.

In 1991 NCR is acquired by AT&T.

In 1992 Teradata is merged into AT&T/NCR.

In 1995 Teradata version 2 for UNIX operating systems released.

1.2 Teradata

Teradata is a large database server that accommodates multiple client applications making inquiries against it concurrently. Various client platforms access the database through a TCP-IP connection across an IBM mainframe channel connection. The ability to manage large amounts of data is accomplished using the concept of parallelism, wherein many individual processors perform smaller tasks concurrently to accomplish an operation against a huge repository of data. To date, only parallel architectures can handle databases of this size.

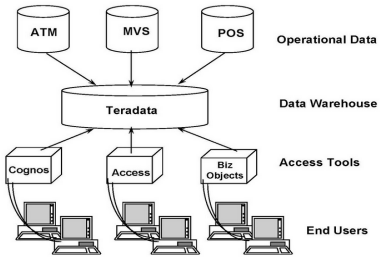

Figure:1.2 Tera data overview

• Designed to process large quantities of detail data.
• Ideal for data warehouse applications.
• Parallelism makes easy access to very large tables possible.
• Open architecture (system) – uses industry standard components.
• Performance increase is linear as components (nodes) are added.
• Runs as a database server to client applications.
• Runs on multiple hardware platforms.

2. TERA DATA MANAGEABILITY

One of the key benefits of Teradata is its manageability. The list of tasks that Teradata Database Administrators do not have to do is long, and illustrates why the Teradata system is so easy to manage and maintain compared to other databases. Things Teradata Database Administrators Never Have to Do Teradata DBAs never have to do the following tasks:

- Reorganize data or index space.
- Pre-allocate table/index space and format partitioning. While it is possible to have
- Partitioned indexes in Teradata, they are not required.
- Pre-prepare data for loading (convert, sort, split, etc.).
- Unload/reload data spaces due to expansion.

With Teradata, the data can be redistributed on the larger configuration with no offloading and reloading required. Write or run programs to split input source files into partitions for loading.

With Teradata, the workload for creating a table of 100 rows is the same as creating a table with 1,000,000,000 rows. Teradata DBAs know that if data doubles, the system can expand easily to accommodate it. Teradata provides huge cost advantages, especially when it comes to staffing Database Administrators. Customers tell us that their DBA staff requirements for administering non-Teradata databases are three to 10 times higher. How Other Databases Store Rows and Manage Data Even data distribution is not easy for most databases to do. Many databases use range distribution, which creates intensive maintenance tasks for the DBA. Others may use indexes as a way to select a small amount of data to return the answer to a query. They use them to avoid accessing the underlying tables if possible. The assumption is that the index will be smaller than the tables so they will take less time to read. Because they scan indexes and use only part of the data in the index to search for answers to a query, they can carry extra data in the indexes, duplicating data in the tables. This way they do not have to read the table at all in some cases. As you will see, this is not nearly as efficient as Teradata method of data storage and access.

Other DBAs have to ask themselves questions like:

- How should I partition the data?
- How large should I make the partitions?
- Where do I have data contention?
- How are the users accessing the data?

Many other databases require the DBAs to manually partition the data. They might place an entire table in a single partition. The disadvantage of this approach is it creates a bottleneck for all queries against that data. It is not the most efficient way to either store or access data rows. With other databases, adding, updating and deleting data affects manual data distribution schemes thereby reducing query performance and requiring reorganization. A Teradata system provides high performance because it distributes the data evenly across the AMPs for parallel processing. No partitioning or data re-organizations are needed. With Teradata, your DBA can spend more time with users developing strategic applications to beat your competition.

2.1 Scalability

"Linear scalability" means that as you add components to the system, the performance increase is linear. Adding components allows the system to accommodate increased workload without decreased throughput. Teradata was the first commercial database system to scale to and support a trillion bytes of data. The origin of the name Teradata is "tera-," which is derived from Greek and means "trillion".
The chart below lists the meaning of the prefixes: 10^3

Table 1:

Prefix	Exponent	Meaning
Kilo	10^3	1,000(thousand)

Mega	10^6	1,000,000(million)
Giga	10^9	1,000,000,000(billion)
Tera	10^{12}	1,000,000,000,000(trillion)
Peta	10^{15}	1,000,000,000,000,000(quadrillion)
Exa	10^{18}	1,000,000,000,000,000,000(quintillion)

Teradata's scalability provides investment protection for customer's growth and application development. Teradata is the only database that is truly scalable, and this extends to data loading with the use of parallel loading utilities. Teradata is scalable in multiple ways, including hardware, complexity, and concurrent users.

Hardware

Growth is a fundamental goal of business. A Teradata MPP system easily accommodates that growth whenever it happens. The Teradata Database runs on highly optimized NCR servers in the following configurations:

SMP - Symmetric multiprocessing platforms manage gigabytes of data to support an entry-level data warehousing system.

MPP - Massively parallel processing systems can manage hundreds of terabytes of data. You can start small with a couple of nodes, and later expand the system as your business grows. With Teradata, you can increase the size of your system without replacing:

Databases - When you expand your system, the data is automatically redistributed through the reconfiguration process, without manual interventions such as sorting, unloading and reloading, or partitioning.

Platforms - Teradata's modular structure allows you to add components to your existing system.

Data model - The physical and logical data models remain the same regardless of data volume.

Applications

Applications you develop for Teradata configurations will continue to work as the system grows, protecting your investment in application development.

Complexity

Teradata is adept at complex data models that satisfy the information needs throughout an enterprise. Teradata efficiently processes increasingly sophisticated business questions as users realize the value of the answers they are getting. It has the ability to perform large aggregations during query run time and can perform up to 64 joins in a single query.

Concurrent Users

As is proven in every benchmark Teradata performs, Teradata can handle the most concurrent users, who are often running multiple, complex queries. Teradata has the proven ability to handle from hundreds to thousands of users on the system simultaneously. Adding many concurrent users typically reduces system performance. However, adding more components can enable the system to accommodate the new users with equal or even better performance.

2.2 Unconditional Parallelism

Teradata provides exceptional performance using parallelism to achieve a single answer faster than a non- parallel system. Parallelism uses multiple processors working together to accomplish a task quickly. An example of parallelism can be seen at an amusement park, as guests stand in line for an attraction such as a roller coaster. As the line approaches the boarding platform, it typically will split into multiple, parallel lines. That way, groups of people can step into their seats simultaneously. The line moves faster than if the guests step onto the attraction one at a time. At the biggest amusement parks, the parallel loading of the rides becomes essential to their successful operation. Parallelism is evident throughout a Teradata system, from the architecture to data loading to complex request processing. Teradata processes requests in parallel without mandatory query tuning. Teradata's parallelism does not depend on limited data quantity, column range constraints, or specialized data models -- Teradata has "unconditional parallelism".

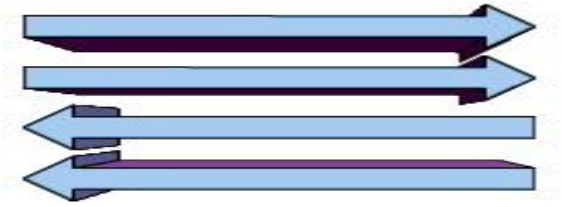

Figure2.2: unconditional parallelism

2.3 Ability To Model The Business

A data warehouse built on a business model contains information from across the enterprise. Individual departments can use their own assumptions and views of the data for analysis, yet these varying perspectives have a common basis for a "single version of the truth".With

Teradata's centrally located. logical architecture, companies can get a cohesive view of their operations across functional areas to:

- Find out which divisions share customers.
- Track products throughout the supply chain, from initial manufacture, to inventory, to sale, to delivery, to maintenance, to customer satisfaction. Analyze relationships between results of different departments.
- Determine if a customer on the phone has used the company's website.
- Vary levels of service based on a customer's profitability.

You get consistent answers from the different viewpoints above using a single business model, not functional models for different departments. In a functional model, data is organized according to what is done with it. But what happens if users later want to do some analysis that has never been done before? When a system is optimized for one department's function, the other departments' needs (and future needs) may not be met.

A Teradata system allows the data to represent a business model, with data organized according to what it represents, not how it is accessed, so it is easy to understand. The data model should be designed with regard to usage and be the same regardless of data volume. With Teradata as the enterprise data warehouse, users can ask new questions of the data that were never anticipated, throughout the business cycle and even through changes in the business environment. A key Teradata strength is its ability to model the customer's business.

Teradata's business models are truly normalized, avoiding the costly star schema and snowflake implementations that many other database vendors use. Teradata can do Star Schema and other types of relational modeling, but Third Normal Form is the methodology Teradata recommends to customers. Teradata's competitors typically implement Star Schema or Snowflake models either because they are implementing a set of known queries in a transaction processing environment, or because their architecture limits them to that type of model. Normalization is the process of reducing a complex data structure into a simple, stable one. Generally this process involves removing redundant attributes, keys, and relationships from the conceptual data model. Teradata supports normalized logical models because Teradata is able to perform 64 table joins and large aggregations during queries.

Figure 2.3 : Ability to model the business

3. TERADATA COMPONENTS

1) Parsing Engine 2) BYNET 3) AMP 4) VDISKS

3.1 Parsing Engine

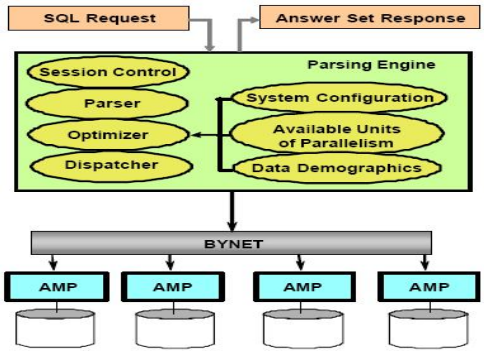

Fig 3.1. P.E Architecture

The Parsing Engine is responsible for:

• Managing individual sessions (up to 120 sessions per PE)

• Parsing and optimizing your SQL requests

• Building query plans with the parallel-aware, intelligent Optimizer

• Dispatching the optimized plan to the AMPs

• Sending the answer set response back to their questing client.

3.2 BYNET (Banyan Network) Architecture

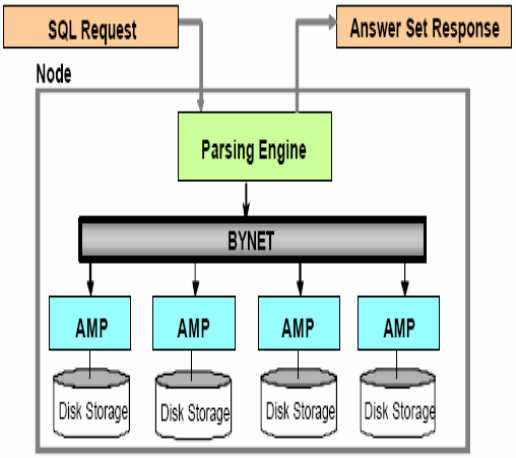

Fig.3.2 BYNET Architecture

• Automatic load balancing of message traffic.

• Automatic reconfiguration after fault detection.

The Bynet Connects All The Amps On the System

• Between nodes, the BYNET hardware carries broadcast and point-to-point Communications. (1 AMP….1AMP)

• On a node, BYNET software and PDE together control which AMPs receive a multicast communication (1….more, 1….many.)(1….more AMPs)

Bynet Features

- Enables multiple SMP nodes (MPP) to communicate.
- Automatic load balancing of message traffic.
- Automatic reconfiguration after fault detection.
- Fully operational dual BYNETs provide fault tolerance.
- Scalable bandwidth as nodes is added.

3.3 Access Module Processor (AMP)

Amp is called as Heart of teradata and every AMP will consist of its own virtual disk (VDISK). It retrieves data and updates the data on its own virtual disks.

AMPs are responsible for:

• Storing and retrieving rows to and from disks
• Lock management (lock/unlock)
• Sorting rows and aggregating columns
• Join processing
• Output conversion and formatting (ANSI, ASCII, EBCDIC)
• Creating answer sets for clients
• Disk space management and Accounting
• Recovery processing (ARC, LOCKS, JOURNAL, FALLBACK...)

4. FAULT TOLERANCES

Fallback:

A fallback table is a duplicate copy of a primary table. Each row in a fallback table is stored on an AMP different from the one to which the primary row hashes. This reduces the likelihood of loss of data due to simultaneous losses of the 2 AMPs or their associated disk storage.

AMP Clusters:

Clustering is a means of logically grouping AMPs to minimize (or eliminate) data loss that might occur from losing an AMP. Note that AMP clusters are used only for fallback data.

4.1 Cliques

The clique is a feature of multimode systems that physically groups nodes together by multiport access to common disk array units. A clique is the mechanism that supports the migration of vprocs under PDE following a node failure. If a node in a clique fails, then AMP and PE vprocs migrate to other nodes in the clique and continue to operate while recovery occurs on their home node. PEs for channel-attached hardware cannot migrate because they are dependent on the hardware that is physically attached to the node to which they are assigned. PEs for LAN-attached connections do migrate when a node failure occurs, as do all AMPs.

4.2 Hot Standby Nodes

The Hot Standby Node feature allows spare nodes to be incorporated into the production environment so that the Teradata Database can take advantage of the presence of the spare nodes to improve availability. A hot standby node is a node that:

- Is a member of a clique
- Does not normally participate in the production
- Can be brought into the production to compensate for the loss of a node in the clique

Configuring a hot standby node can eliminate the system-wide performance degradation associated with the loss of a single node in a single clique. When a node fails, the Hot Standby

Node feature migrates all AMP and PE vprocs on the failed node to other nodes in the system, including the node that you have designated as the hot standby. The hot standby node becomes a production node. When the failed node returns to service, it becomes the new hot standby node.

5. RAID PROTECTION

RAID 1 (Mirroring)

- Each physical disk in the array has an exact copy in the same array.
- The array controller can read from either disk and write to both.
- When one disk of the pair fails, there is no change in performance.
- Mirroring reduces available disk space by 50%.
- Array controller reconstructs failed disks quickly.
- good performance with disk failures
- higher cost in terms of disk space

RAID 5 (Parity)

- Data and parity are striped across a rank of 4 disks.
- If a disk fails, any missing block may be reconstructed using the other three disks.
- Parity reduces available disk space by 25% in a 4-disk rank.
- Array controller reconstruction of failed disks takes lesser than RAID1.
- Reduced performance with disk failure
- Lower cost in terms of disk space

6. TEMPORARY TABLES

Global Temporary Tables:

Global temporary tables are tables that exist only for the duration of the SQL session in which they are used. The contents of these tables are private to the session, and the system automatically drops the table at the end of that session. However, the system saves the global temporary table definition permanently in the data dictionary. In addition, global temporary tables allow the database administrator to define a template in the schema, which a user can reference for their exclusive use during a session.

Volatile Temporary Tables:

A volatile temporary table resides in memory but does not survive across a system restart. If a user needs a temporary table for a single use only, they should define a volatile temporary table. Using volatile temporary tables improves performance even more than using global temporary tables, since the system does not store the definitions of volatile temporary tables in the data dictionary. Moreover, users require no privilege to access volatile temporary tables.

The maximum number of temporary tables is provided in the following table:

global temporary tables (materialized) 2000 per session.
volatile tables 1000 per session.

6.1 Primary Keys And Primary Indexes

Indexes are conceptually different from *keys:*

- A *PK* is a relational modeling convention which uniquely identifies each row.
- A *PI* is a Teradata convention which determines how the rows are stored and accessed.

Primary Key	Primary Index
Logical concept of data modeling	Physical mechanism for access and storage
Teradata doesn't need to recognize	Each table must have exactly one
No limit on column numbers	64-column limit
Documented in data model (Optional in CREATE TABLE)	Defined in CREATE TABLE statement
Must be unique	May be unique or non-unique
Uniquely identifies each row	Used to place and locate each row on an AMP
Values should not change	Values may be changed (Del+ Ins)
May not be NULL—requires a value	May be NULL
Does not imply an access path	Defines most efficient access path
Chosen for logical correctness	Chosen for physical performance

- A significant percentage of tables may use the same columns for both the PK and the PI.
- A well-designed database will use a PI that is different from the PK for some tables.

Fig6.1. Difference between pk and Primary indiex

7. TERADATA WAREHOUSE MINER

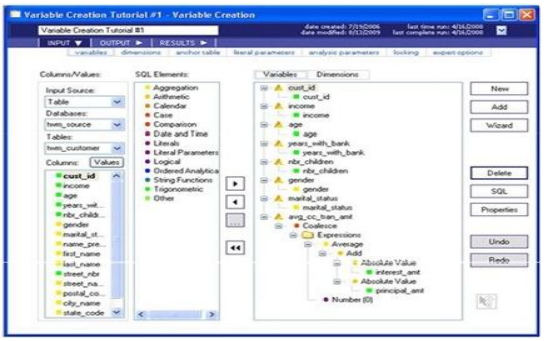

Fig 7. Warehouse miner main screen

There are three windows on the main screen, the largest of which is for viewing and editing analysis forms. On the right

is the Project explorer window where open projects and the analyses they contain are displayed in tree view. Underneath both of these areas is the execution status window. Directly over the analysis work area is a toolbar with icons for primary functions and over that is a series of menus topics, including file, view, project, tools, window and help. In the sample screen above, the open connection icon has been selected to connect to data source DBC twm, and the add new analysis icon has been selected to select Data Explorer from the Descriptive Statistics category. Now looking at the data Explorer input from covering most of the main screen, selectors can be seen on the left side of the form for selecting databases, tables and columns, and on the right area to drag selected columns into.(the arrow buttons in the middle can also be used to select and de-select columns.)

Over the selectors are tabs for INPUT, OUTPUT and RESULTS, with sub-tabs that depend on the type of analysis. After the parameters for an analysis have been specified, the analysis can be executed by clicking the run button above, by right clicking on the project or analysis in the project work area and selecting run, or by pressing the F5 key on the keyboard. The status of the execution will be displayed in the execution status window below. When execution is complete, the results tab will be enabled, and upon selection, the resulting data, graphs and generated SQL can be viewed.

7.1 Exploring Data with a Data Explorer Analysis

Parameterized a Data Explorer analysis as follows

- Input source: MultiTable
- Available Databases: the databases where the demonstration data was installed.
- Available Tables:
  - TWM_CHECKING_ACCT
  - TWM_CREDIT_ACCT
  - TWM_CUSTOMER
  - TWM_SAVINGS_ACCT
- Analyses to Perform
  - Values: Enabled
  - Compute unique values: Enabled
  - Statistics: Enabled
  - Frequency :Enabled
  - Histogram : Enabled

Output Values analyses output table : twm_values

Statistics analyses output table : twm_stats

Frequency analyses output table : twm_freq

Histogram analyses output table : twm_hist

Run the analysis, and when it completes, click on the results tab.

Data

By clicking on data and then load, each of the four tables produced can be viewed by selecting the desired table in the pull-down selector.

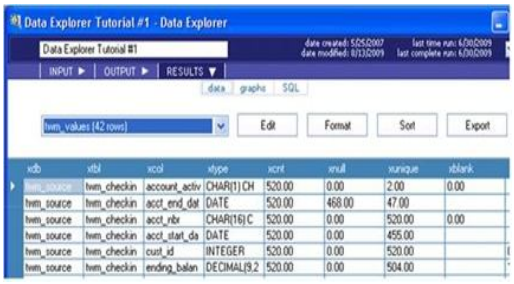

Fig 7.1 Data Explorer

Graph

The following is a snapshot of the icon displayed when the graph tab is selected.

Fig 7.2 : Graph menu

By clicking anywhere in this picture the subsequent display of the actual graph object is displayed.

Fig 7.3: Graph

Clicking on the city_name thumbnail graph leads to the following display, while clicking on the bar for san Diego adds the drill down box to the displayed. By clicking on the drill down button the customers in san Diego can be displayed.

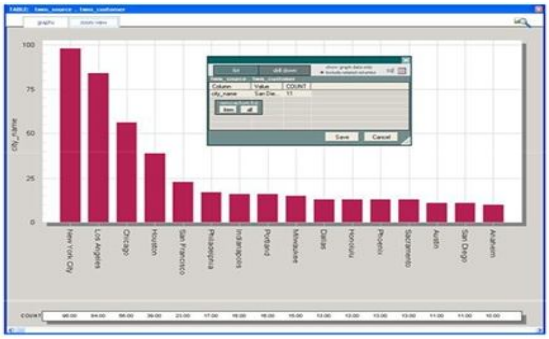

Fig 7.4.city name thumbnail graph

Creating an Analytic Data Set the following depicts an example of creating an analytic data set using the variable creation analysis. Following this depiction are step-by-step instruction for defining the variables creating in this example

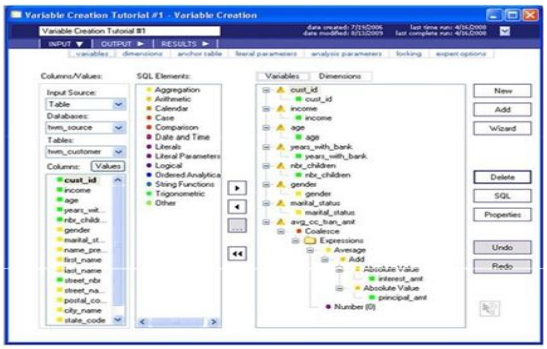

Fig 7.5: variable creation

Parameterize the above variable creation analysis as follows

1. Select TWM_CUSTOMER as the available Table.

2. Create seven variables by double-clicking on the following columns.

TWM_CUSTOMER.cust_id
TWM_CUSTOMER.inncome
TWM_CUSTOMER.age
WM_CUSTOMER.years_with_blank
TWM_CUSTOMER.nbr_children
TWM_CUSTOMER.gender
TWM_CUSTOMER.marital_status

3. Select TWM_CREDIT_TRAIN as the Available Table.

4. Create a variable by clicking on the new button and build up an expression as follows.

5. Drag an Add SQL Element over the variable, and then drag the following two columns over the empty arguments.

TWM_CREDIT_TRAN.insert_amt
TWM_CREDIT_TRAN.principal_amt

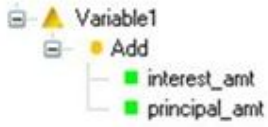

Fig 7.6 : Add(arthimetic)

Because there may be negative values, drag and drop an Absolute value (Arithmetic) SQL Element over both interest_amt and principal_amt.

6. Take the average of this expression, by dragging and dropping an average (Aggregation) on top of the Add.

7. Because this analysis may generate many NULL values by joining TWM_CUSTOMER to TWM_CREDIT_TRAN, drag a coalesce (case) on top of the Average.

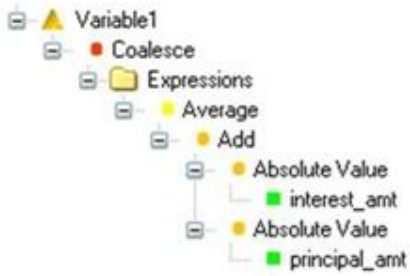

Fig 7.7 : Coalesce (case)

8. Drag and drop a number (Literal) 0 into the expression folder and rename it from variable to avg_cc_tran_amt to complete the variable.

9. Goto INPUT anchor Table and select TWM_CUSTOMER as the anchor table as seen below

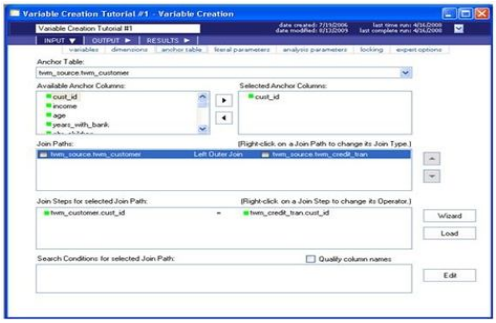

Fig 7.8: INPUT>Anchor Table : select TWM_CUSTOMER

10. Specify the join path from TWM_CUSTOMER to TWM_CREDIT_TRAIN by clicking on the Wizard button and specifying that they be joined on the column “cust_id”.

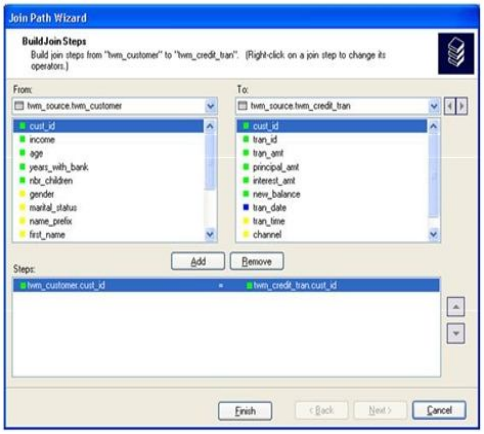

Fig 7.9 : Join Path Wizard

11. Go to OUTPUT storage, and select Store the tabular output of this analysis in the database.

Specify that a table should be created named twm_tutorials_vcl.

Creating and Scoring a Decision Tree model

Building a Decision Tree Model

The following depicts a tutorial example of creating a decision tree model. In this example a standard Gain Ratio tree was built to predict card ownership, based on 20 numeric and categorical input variables. Notice theat the tree initially built contained 100 nodes but was pruned back to only 11, couning the root node. This yielded not only a relatively simple tree structure, but also Model Accuracy of 95.72% on this training data.

Parameterize aDecision Tree as follows.

- Available Tables : twm_customer_analysis
- Dependent variable : ccacct
- Independent variables:
  - Income,age
  - Years_with_bank, nbr_children
  - Gender, marital_status
  - City_name, state_code
  - Female, single
  - Married, separated
  - Avg_ck_bal,avg_ck_tran_cnt
  - Avg_sv_tran_amt,avg_sv_tran_cnt
- Tree splitting : Gain Ratio
- Minimum Split count : 2
- Maximum Nodes :1000
- Maximum Depth :10
- Bin numeric variables :Disabled
- Pruning Method : Gain Ratio
- Include Lift Table : Enabled
- Response value : 1

Run the analysis and click onresults when it completes. For this example, the decision tree analysis generate the following pages.

Decision Tree Report

Table 1: Decision tree Report

Total observations	747
Nodes before pruning	33
Nodes after pruning	11
Model accuracy	95.72 %

Variables

Table 2: dependent variables

Dependent variables
ccacct

Table 3:Independent variables

Independent variable
Income
Ckacct
Avg_sv_bal
Avg_sv_tran_cnt

Confusion matrix

Table 4: confusion matrix

	Actual non-response	Actual response	Correct	Incorrect
Predict 0	340 / 45.52 %	0 / 0.00%	340 / 45.52%	0 / 0.00%
Predict 1	32 / 4.28%	375 / 50.20%	375 /50.20%	32 /4.28%

Cumulative lift table

Table 5: cumulative lift table

Table 5: Cumulative Lift Table

Decile	Count	Response	Response (%)	Captured Response (%)	Lift	Cumulative Response	Cumulative Response (%)	Cumulative Captured Response (%)	Cumulative Lift
1	5.00	5.00	100.00	1.33	1.99	5.00	100.00	1.33	1.99
2	0.00	0.00	0.00	0.00	0.00	5.00	100.00	1.33	1.99
3	0.00	0.00	0.00	0.00	0.00	5.00	100.00	1.33	1.99
4	0.00	0.00	0.00	0.00	0.00	5.00	100.00	1.33	1.99
5	0.00	0.00	0.00	0.00	0.00	5.00	100.00	1.33	1.99
6	402.00	370.00	92.04	98.67	1.83	375.00	92.14	100.00	1.84
7	0.00	0.00	0.00	0.00	0.00	375.00	92.14	100.00	1.84
8	0.00	0.00	0.00	0.00	0.00	375.00	92.14	100.00	1.84
9	0.00	0.00	0.00	0.00	0.00	375.00	92.14	100.00	1.84
10	340.00	0.00	0.00	0.00	0.00	375.00	50.20	100.00	1.00

Graphs

By default the tree browser is displayed as follows.

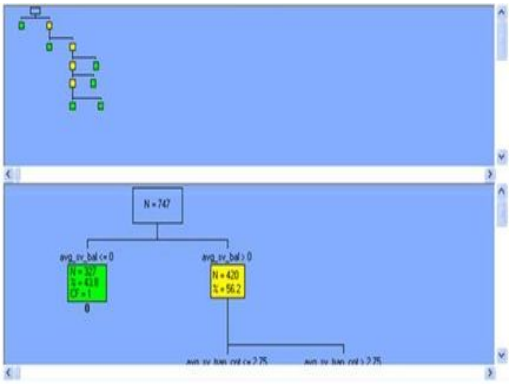

Fig 8: Tree Browser

Select the text tree tab to view the rules in textual format.

Fig 9 : Text tree tab

Additionally, you can click on lift chart to view the lift table graphically.

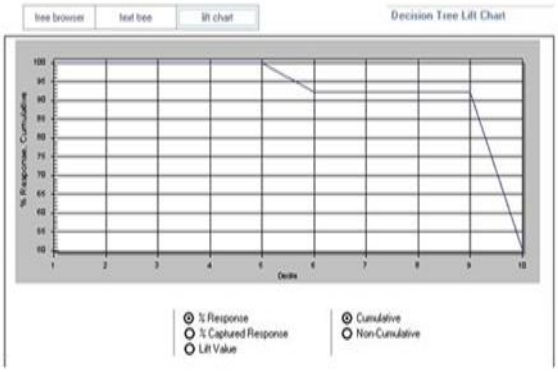

Fig 10: lift chart tab

8. CONCLUSION

Teradata is the forthcoming of the Data Mining. In upcoming the whole world we start using Teradata Database. Now it is expensive, works are successful on to reduce its cost. So, it will reach to small business people also. It is firm occupied environment will change organization goals.

9. ACKNOWLEDGMENTS

Our thanks to K Vasanth Kumar, Assistant Professor in NBKR Institute of Science and Technology, Vidyanagar, P RAJESH KUMAR, Assistant Professor in Siddhartha institute of Engineering and Technology College, Puttur, Andhra Pradesh for his guidance in regards of this paper and T GOPINATH Pursing Master of Computer Applications in JNTU, ANTHAPUR for his support in completing this paper.

10. REFERENCES

[1] https://www.teradata.com/.../TeradataData-Mining-Services-eb1719/
[2] http://www.teradatatech.com/?p=103
[3] http://www.teradata.com/businessneeds/data-mining-and- analytics/
[4] http://www.teradata.com/products-andservices/teradata-warehouseminer/?ICID=Ptwm
[5] http://decisionfirst.files.wordpress.com/2013/09/sap-acquires-kxen.pdf
[6] www.teradata.com/brochures/TeradataRapid-Insight-Service-eb6161