Data Mining: Approach Towards The Accuracy Using "Teradata"!

Shubhangi Pharande
Department of MCA
NBNSSOCS,Sinhgad
Institute
Pune, Maharashtra
India

Simantini Nalawade
Department of MCA
NBNSSOCS,Sinhgad
Institute
Pune, Maharashtra
India

Ajay Nalawade
Cognizant Technology
Solutions
Pune, Maharashtra
India

Abstract: Data mining refers to the process or method that extracts or mines interesting knowledge or patterns from large amounts of data. Thus, the result of the natural evolution of information technology can be viewed by data mining. Data mining also involves assimilation, more willingly than a simple transformation. Data mining also have techniques from multiple directions such as database technology, statistics, machine learning, high-performance computing, pattern identification, neural networks, data revelation, information extraction as well as image and signal processing and spatial data analysis.
The large number of database systems having query and transaction processing eventually and naturally led to the need for data analysis and understanding. Hence, due to this increasing necessity data mining started its development. The process reveals hidden patterns that can't be detected using traditional query and OLAP tools. So there is "Teradata Solution", unique approach to counselling and our optimal use of data warehouse technology.[4]

Keywords: OLAP (Online Analytical Processing), pattern identification, neural networks, data revelation, machine learning, mines interesting knowledge or patterns

1. INTRODUCTION

Data mining is not a one-time event. It's a process - an ongoing evolution of discovery and analysis. It's a process that uncovers new and meaningful patterns in collected data; patterns can use to address challenging business questions that require prediction and inference. And it's a process that demands a unique set of skills and resources. Data mining also plays a crucial role in raising the value of Teradata solution. The requirements from models for prediction, estimation, and other inferences involving uncertainty are built by analyzing the huge volumes of historical data which can deliver the knowledge to take more informed strategic business decisions and more effective interactions with individual customers.[3]
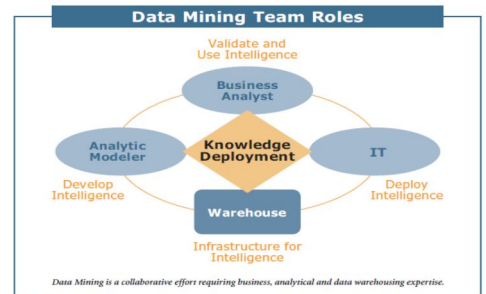
Figure 1.Data mining introduction

2. WHY DATA MINING

To transfer data into useful information and knowledge from the huge amounts of data, the concept comes in focus is "Data Mining".

It is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among "internal" factors such as price, product positioning, or staff skills, and "external" factors such as economic gauge, contest, and customer demographics. And, it enables them to determine the impact on consumer regarding its sales, customer fulfillment, and corporate profits. So company can drill down huge data into summary information to provide a detail view of transactional data.

A retailer can manage current records of customer, which will help them to promote their targets which are based on an individual's purchase history, by using data mining. Retailer can also mine demographic data which will help them to develop products & sponsorships to request for specific customer segments. For example, Blockbuster Entertainment mines its video rental history database to recommend rentals to individual customers. For example Book Library mines its rental history database to recommend rentals to individual customers.

3. DIFFICULTIES FACED BY DATA MINING

The volume and diversity of data being captured by companies today is overwhelming. To take smarter strategic decision, there is a need to turn this information into business processes and actionable plans by data analysis due to epidemic growth in collected data. The challenge or difficulty in Data Mining is to extract the exact data what you require from the enormous data collection. For e.g. the data provided by OLAP tool is three dimensional, but sometimes it happens that we don't require un-necessarily 3rd dimension of data.

4. APPROACHES TOWARDS THE TERADATA SOLUTION

Teradata product is a product of Teradata Corporation, an American computer company that sells analytic data platforms, applications and related services. The Teradata product is referred to as a "data warehouse system" and stores and manages data. The data warehouses use a "shared nothing" architecture which means that each server node has its own memory and processing power. Storage of data can be incremented by adding more servers and nodes. Servers spread their workload with the help of database software which is present at its top. To process different types of data teradata provides its applications and software's. In 2010, Teradata added text analytics to track unstructured data, such as word processor documents, and semi-structured data, such as spreadsheets. Teradata's product can be used for business analysis. Data warehouses can track company data, such as sales, customer preferences, product placement, etc. There is appointment for minority, women, veteran, or small business vendors due to supplier diversity program by teradata.

4.1 Teradata Master Data Manager

Master data is one of the critical parts from the database as it is going be used by business operations as well as reporting and analytics systems. There is requirement of proper master data management so that it will not put

business at risk in future. A good master data management solution mitigates that risk by managing data architecture, metadata, data quality, data hierarchies, master data workflow, and data governance. It also synchronizes master data so that changes are propagated across the entire enterprise. These all requirements can be meet by Teradata Master Data Manager and more also in a complete package which operates flawlessly with existing Teradata Solutions with lower cost than the other Master Data Management solutions from the market today.

To provide rapid analytics solution, Teradata helps us by providing data warehouse solution provider and KXEN. (A privately- held predictive analytics software vendor KXEN (pronounced as the initials K-X-E-N) is a relatively lesser - known competitor in the predictive analytics space, and has a distinct bent towards user - friendly and line – of - business predictive analysis empowerment). "KXEN's automated solutions were simple and easy to integrate, did not demand the high degree of skill and experience needed by traditional tools and team was up to speed on their use in less than two days.[5]

Rapid analytics is the practice of maximizing an existing data infrastructure and data mining technology to drive predictive intelligence into the business process. The ultimate goal is to improve the profitability of analytic projects by reducing cost and accelerating the data mining process.

Teradata provides a unique new program, Teradata® Rapid Insight Service, that's exactly what its name implies: a quick and easy low-cost way for retailers to uncover answers to some of their most pressing business questions. It's a unique combination of technology, people, and processes that can

help us to become more flexible, agile, and sophisticated in our approach to solving business problems.

4.2 A Low-Risk Approach
There is a specialized team of Teradata experts that understand industry's challenges and best practices, a team that can uncover low-risk options for attacking high impact issues and identify areas in which further investment is warranted. Teradata consultants can:[6]

- Identify high impact questions, and build the business case for further analysis.
- Provide analytic results that deliver new business insight.
- Suggest actions (based on industry best practices) implied from results.
- Provide the analytical framework, models, and methodology for you to extend your own analytic expertise and resources.
- Recommend additional ways to extend analytics for further insights with data you already have.

4.3 A Low Cost Approach
Teradata Accelerate for Finance - Today's economic pressures increase the demand for quick visibility to detailed revenue and cost driver data. To enable such insights, Teradata offers a financial analytic package that leverages pre-built intelligence to turn enterprise resource planning (ERP) data into actionable information. Teradata Accelerate for Finance quickly integrates ERP data from source to end report significantly reducing the time, risk and cost of financial analytics projects. The package is designed to eliminate the usual challenges, such as limited data transparency, time consuming and costly maintenance, and disconnected financial and

operational data. This package helps companies more easily reveal economic drivers and opportunities for cost and expense reduction, for revenue growth and for improved cash flow management.[6]

5. WHY TERADATA?

5.1 Data Mining and Analytics

The volume and diversity of data being captured by companies today is staggering.[1] This exponential growth in collected data increases the demand for data analysis and the need to turn this information into business processes and actionable plans to make smarter strategic decisions.

5.2 Teradata Warehouse Miner

Teradata Warehouse Miner will help you discover meaningful new data patterns and trends. It contains an array of data profiling and mining functions ranging from data exploration and transformation to analytic model development and deployment that are performed directly in the Teradata Database. While most data mining solutions require analysts to extract data samples to build and run analytic models, Teradata Warehouse Miner allows you to analyze detailed data without data movement, streamlining the data mining process.

5.3 Features & Benefits

Teradata Warehouse Miner makes the most of your data warehousing capabilities with analytic technology that:[2]

- Reduces the analytic modeling development cycle, allowing faster iterations to refine your model and increase analytic intelligence
- Simplifies data profiling and creation of an analytic data set with built-in intelligence, complementing any data mining tool

- Allows you to quickly and easily integrate models into business applications
- Accelerates model scoring, allowing you to analyze and score data in your warehouse efficiently
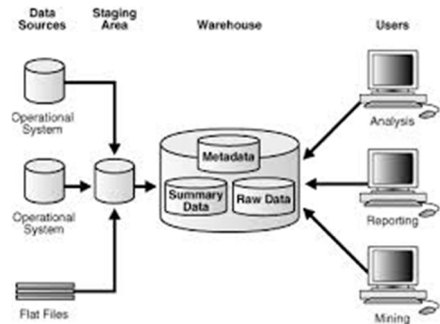

Figure 2. Working of data mining with teradata

6. CONCLUSION

By using Teradata Data Mining Services, we can mine data as well as integrate the process into our existing business procedures and technologies. The use of Data Mining makes it difficult to manage, decide and analyze the data as per the use of business. Teradata Solution focused on to explore how - and if - data mining can help our business. It's a low-risk, low-cost introduction to data mining services. We can use hardware, software and expertise - and data - to help to define a business problem, cleanse the data, develop a corresponding model and then deploy that model against our data. As a result, we get a better understanding of data mining and a clearer view of how it can benefit our business with the help of teradata solution.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] http://www.teradata.com/business-needs/data-mining-and-analytics/

[2] http://www.teradata.com/products-and-services/teradata-warehouse-miner/?ICID=Ptwm

[3] https://www.teradata.com/.../Teradata-Data-Mining-Services-eb1719/

[4] http://www.teradatatech.com/?p=103

[5] http://decisionfirst.files.wordpress.com/2013/09/sap-acquires-kxen.pdf

[6] www.teradata.com/brochures/Teradata-Rapid-Insight-Service-eb6161/