AN INSIGHT ON DATA WAREHOUSING FOR PRODUCT MARKETING

DEEPTI TEKWANI[a1], SHUBHI AGRAWAL[b] AND RAJESH KESHAVRAO DESHMUKH[c]

[abc]Department of CSE, SSIPMT, Raipur, Chhattisgarh, India

ABSTRACT

In this information era, vast quantities of data are collected in organizations, yet the managers experience a great amount of difficulty in obtaining the information they need for decision-making. This problem has led to the emergence of the data warehouses. Data warehouses have been accepted as the heart of the modern decision support systems. It supports the decision-making process with variety of applications such as target marketing, analysing the trends, and many more. As the importance of extracting relevant information from the data cannot be denied, the data warehouse can be used to make information available to the decision makers. This paper examines the different tools and characteristics of the data warehouse, as well as its usage for the purpose of product marketing. With the integration of data from various operational data sources in the data warehouse, there is a large amount of data which can be used to express our past buying behaviours, predict our future behaviour and generate new ways for marketing the product to increase sales and further gain competitive advantage.

KEYWORDS: Data Warehouse, Analytics, Decision Making.

A Data Warehouse (DW) is defined as "a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management's decision-making process". [1]They are central repositories of integrated data from multiple data sources which store current and historical data. Data warehouse obtains the data from a number of operational database systems which are then converted into a suitable form. This process is known as Extraction, Transformation, and Loading (ETL). [2] Along with the target database, there is another database to store the metadata which is known as the metadata repository. The metadata repository contains the data about the data-description of the source data, the target data, and how the source data has been converted to the target data. Data warehouse can be used to analyse the data and observe patterns to gain competitive advantage. It can also be used to draw marketing insights from the customer-centric records which can help the organization to increase sales, revenue and develop lasting relationships with their customers.

TOOLS

There are different kinds of tools available which provide data warehousing and analysis services. These are:

Hadoop

Hadoop is an open-source software framework which allows to store and process big data across a cluster of computers. It provides massive storage and is designed to scale up from a single server to a thousand machines. Hadoop is free and uses commodity hardware to store large quantities of data. Hadoop's distributed computing model provides enormous processing power. It is an economical and reliable solution for data analysis.

OLAP:

OLAP (Online Analytical Processing) is based on the multidimensional data model which enables users to selectively extract and view data from different points of view. It allows decision-makers to get an insight of information through fast, consistent, and interactive access to information. Online Analytical processing can be used for the purpose of discovering previously undiscerned relationships between data items. It performs multidimensional analysis of data and provides the ability for complex calculations, trend analysis, and sophisticated data modelling.

Amazon Redshift

Amazon Redshift provides fast and fully managed petabyte scale data warehouse service. In this service, unlimited analytics can be used by unlimited users for just $1000 per terabyte per year. Amazon Redshift provides in-built security services and it can be resized according to the user's performance and capacity needs. It delivers input-output performance for virtually any size dataset by using columnar storage technology and monitoring and distributing queries across multiple nodes.

Teradata

The Teradata product is also known as a "data warehouse system" which is used to store and manage data. It can be used to help organizations to consolidate data from several sources and infer unique and important insights. [3] It is considered as one of the most popular data warehouse application as it allows companies to

[1]Corresponding author

analyse data in an easy manner. It consists of two divisions namely data analytics and marketing applications. The Teradata's product can be used for business analysis as the data warehouses can keep track of the company records such as sales, customer preferences, etc. One of the most important aspect of this application is that it partitions data into hot and cold, where hot data is frequently used.

CHARACTERISTICS

The different characteristics of the data warehouses are mainly divided in the following categories:

Subject – Oriented

The data in the data warehouse is arranged around specific subjects such as sales, customers and product. This arrangement is differs from the transactional systems where data is organized by business processes. The data warehouses store relevant data in accordance with the requirement of the organization's staff which will use the data. As the data warehouse is subject-oriented, it can easily be used to analyse a particular subject area.

Integrated

The data warehouse is used to integrate the data from multiple data sources. The integration of data is performed to define a unique representation of the data coming from different sources. This is an important feature of the data warehouse as the same data is defined in different ways in the different systems which could make the data analysis a time-consuming process. The problems such as naming conflicts and data inconsistencies can be resolved through this feature.

Time Variant

A data warehouse maintains historical data which spans over years unlike the transactional databases, where only the recent data such as for the day, week or month are stored. The data amassed in the data warehouse is identified with a specific time period. In order to observe and discover trends, analysts need large amounts of historical data. This historical data can be utilized to detect deviations, patterns and long-time relationships.

Non-Volatile

A data warehouse is non-volatile which means that the users cannot change or update the data. [4] The non-volatility feature of the data warehouse ensures that all the users are working with the same data. The warehouse is updated by the information technology department rather than by the users of the data warehouse. Only after the loading process from the operational systems is done, the users can access the data with the help of queries.

DATA STORAGE

The storage of data in the data warehouse is done with the help of the ETL(Extract, Transform and Load) process. ETL are an important component in the data warehouse environment to ensure dataset in the data warehouse are cleansed from several OLTP(Online Transaction Processing) systems. The ETL process begins with the extraction of data from various operational databases. The extracted data is then transformed into a consistent format for the data warehouse. The cleansing of data and improvement of the data's quality are also performed in this step. Ultimately, in the next step, the data is uploaded and saved in the target data warehouse. This step is known as loading. Thus, the data is cleaned, transformed, and catalogued before being made available for data mining and analytical functions. [5]During the uploading of new data, the data warehouse is not available for query session. [6] The ETL process are an essential component which ensures the data integrity in the warehouse.

ANALYTICS AND DECISION MAKING

Nowadays, people just don't want to collect data, rather they want to understand the meaning underlying the data and use it to aid in the decision-making process. As the quantity of data generated from business systems and other operational devices is increased, the data warehouses are regularly updated so that analysis can be performed on the current statistics of the business to make business decisions. The decision makers are required to read quickly to mission critical needs due to the rapidly changing volatile and competitive markets. [7] Analytics have had a significant impact on research and technologies, since decision makers have become more and more interested in learning from previous data, thus gaining competitive advantage. [8] This can allow them to make more informed decisions, and market to different segments based on their preferences along with the recognition of sales and marketing opportunities. [9] Data analytics is the process of applying algorithms in order to analyse sets of data and extract useful and unknown patterns, relationships, and information. [10] Big data analytics can also enable the construction of predictive models for customer behaviour and purchase patterns, which eventually raises overall profitability.

PRODUCT MARKETING

Product marketing is about understanding the requirements of a potential buyer so that you can make your product great. The data stored in the data warehouse can be analysed to reveal valuable insights, allowing for decision-makers to capitalize upon the resulting opportunities from the wealth of historic and real-time data. By examining the customer purchasing patterns and observing the psychographics and demographics of the customers, we can easily market products to the right customers. For example, if we work in a company which makes trousers, we can analyse the data in the data warehouse and discover patterns of the potential buyers by observing various attributes such as age, gender, occupation, income levels, location and other preferences. Suppose we find a pattern in the data which suggests that the young adults residing in Mumbai who belong from upper class families can become the potential buyers, we can easily target them with the right kind of promotion and advertising. We can also segment our database in accordance with our products. This can help us to customize our products and promotions which can satisfy the needs of the audience in a specific way. It can also help us understand who our competitors are in each of the segments and how we can gain competitive advantage by utilizing our data warehouse in an efficient manner. By segmenting the customer data, we can improve the conversion rates as we focus our promotions and advertising in a tight, highly-interested market.

CONCLUSION

There is a rapid growth of data which is being produced on a daily basis, and within them lay the intrinsic details and sequences of hidden knowledge which should be extracted and utilized. For this purpose, data warehouse has become an essential part of the decision-making support systems. In this paper, we have examined the various tools and characteristics of the data warehouse, which has recently gained lots of interest due to its unprecedented opportunities and benefits. We have also examined its utilization for the purpose of marketing of the products. The analysis of the customer data can give us insights into their shopping habits which can reveal trends and associations which can benefit your business. It can also lead to promotion and programs which will result in higher revenues and better customer loyalty.

REFERENCES

Inmon, W.H., Hackathorn, and R.D, Using the data warehouse. Wiley-QED Publishing, Somerset, NJ, USA.

Manjunath T. N., Ravindra S. Hegadi, Umesh I. M., and Ravikumar G. K., "Realistic Analysis of Data Warehousing and Data Mining Application in Education Domain", International Journal of Machine Learning and Computing, vol. 2, 2012.

Top 10 Popular Data Warehouse Tools and Technologies, https://www.educba.com/10-popular-data-warehouse-tools/, 2016.

Features of Data Warehouse, http://vibhorraman.blogspot.in/2011/07/what-is-datawarehouse-and-features-of.html, 2011.

Bakshi, K., Considerations for Big Data: Architecture and Approaches, IEEE Aerospace Conference, pp. 1–7, 2012.

Uzair Sultan, Literature Review Real Time Data Warehousing, pp. 1-4, 2016.

Aditya Kumar Gupta, Chakit Gupta, Analyzing Customer Behaviour Using Data Mining Techniques: Optimizing Relationships with Customer, vol. 6, 2010.

Z. Song, A. Kusiak, Optimizing Product Configurations with a Data Mining Approach, International Journal of Production Research, vol. 47(7), pp. 1733–1751,2009.

P. Russom, Big Data Analytics, TDWI Best Practice Report, pp. 1–25, 2011.

M.N Adams, Perspectives on Data Mining, International Journal of Market Research, vol. 52(1), pp. 11–19, 2010.