International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 18 Issue 09, SEP, 2018

EXPLORING THE POTENTIAL OF DATA MINING

[#1]**Mr.KANDUKURI CHANDRA SENA CHARY,** *Assistant Professor*

[#2]**Mr.BOLLI RAMESH,** *Assistant Professor*

Department of Computer Science and Engineering,

SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TS.

Abstract— The term "data mining" is defined here, along with some of its practical applications. As a new technology, data mining hasn't nearly realized its full potential yet. Nonetheless, it is now a common tool in many industries. Some examples of these establishments include shops, hospitals, banks, and insurance firms. Some of these companies are integrating data mining with statistical methods, pattern recognition software, and other cutting-edge tools. Data mining makes it possible to uncover previously hidden linkages and intricate patterns. This technology is favored by businesses because it allows for more in-depth insights into customer behavior, leading to more informed advertising decisions. Here is a rundown of common company issues and the data mining approaches that can aid remedy them. ***Keywords***—Data mining, Customer relationship management, Literature review, Classification

1. INTRODUCTION

There are various applications for data mining. There are currently a plethora of data mining tools available for purchase. However, there are many challenges in this area. In this tutorial, we'll explore the varied applications of Data Mining and examine the cutting-edge research being conducted in the field today. Data mining is the process of combing through massive volumes of data using various techniques to identify patterns that might be exploited to increase profits, decrease expenses, or do both. As more and more industries store and collect vast volumes of data, they are becoming more and more interested in detecting patterns in their databases, like association rules, correlations, clusters, and so on. Association rule mining is a powerful tool for analyzing large databases of consumer transactions in order to identify clusters of similar products. Transactions are what a customer buys while they are there. A wide range of tools and approaches are used in data mining to locate useful information. "Internet mining" refers to the practice of applying these techniques to the World Wide Web, either in their original form or after some adaptation to the online environment. The term "Internet mining" refers to the practice of searching for and consuming relevant content on the World Wide Web. The practice of "mining" the Internet can be categorized in three ways

- Content Mining
- Structure Mining
- Usage Mining

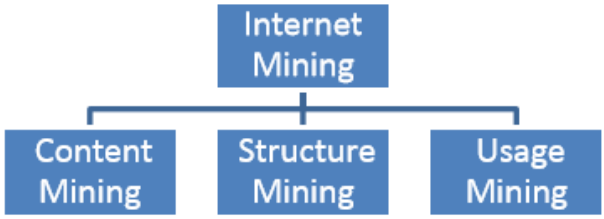

Fig 1. Types of Internet Mining

Material mining describes the method used to retrieve targeted content from the web. There are a wide variety of search engines available for content extraction, including Altavista, Lycos, WebCrawler, MetaCrawler, and many more.

B. Structure mining is the method of creating a summary of a website's or web page's structure by tracing the paths taken by its hyperlinks to other resources.

C. Usage Mining: This is the process of automatically mining data on website visitors. That which it contains,

2. DATA MINING APPLICATIONS

In this course, you'll learn about data mining's many applications, from healthcare and banking to logistics and advertising

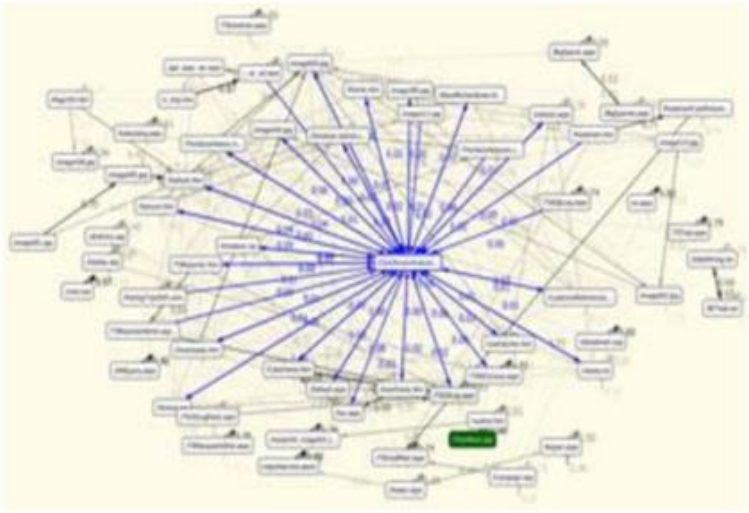

Fig 2. Data mining applications

Data mining is the practice of exploring large datasets for previously unknown relationships or patterns that can improve an organization's performance. Many businesses rely on data mining to help them expand and stay ahead of the competition. The applications of data mining are demonstrated across a variety of sectors in this course, from healthcare and insurance to sales and marketing to transportation.

3.DATA MINING APPLICATIONS IN SALES/MARKETING

Data mining allows companies to discover previously unseen regularities in customer spending. This facilitates the rapid and low-cost development of novel advertising approaches. Some applications of data mining in the business world are listed below.

Data mining is used in market basket analysis to discover the relationships between products, the times they were purchased, and the orders in which they were purchased. Using this data, businesses may better promote their highest-earning products. In addition, it encourages shoppers to acquire similar things that they might not have considered about before.

"Data mining" is a technique used by retailers to discover trends in consumers' purchasing habits. **4. DATA MINING APPLICATIONS IN BANKING / FINANCE**

Data mining methods, such as distributed data mining, have been applied to the investigation, development, and implementation of countermeasures against credit card fraud.

Data mining is used to determine consumer loyalty based on factors such as the frequency and value of purchases made throughout time, as well as the timing and duration of the most recent purchase. Each consumer's relative value is calculated using these parameters. Improved performance in the classroom correlates with a similar boost in consumer loyalty.

Banks may keep their credit card clients with the aid of data mining. Data mining examines historical information to make predictions about future events, such as a client switching credit card providers. The results of this research can be used by financial institutions to develop and publicize a variety of targeted promotions aimed at retaining these clients.

You can find out how much different types of customers spend on credit cards by employing data mining.

Financial data can be mined for hidden correlations via data mining.

Stock trading rules can be discovered through the mining of market data history.

What role may data mining play in the health care and insurance industries?

The ability of the insurance industry to mine data for insights about its clients, rivals, and markets is crucial to the sector's growth and prosperity. New data mining tools, when applied successfully in the insurance industry, give companies significant advantages over their rivals. A few applications of data mining in the insurance sector are as follows:

Claims analysis makes use of data mining to identify concurrent medical procedures.

Data mining facilitates forecasting consumer behavior when new plans are introduced.

Using data mining, insurance companies can learn about the typical actions of high-risk clients.

Data mining is a procedure that aids in the detection of fraud. Transportation data mining is a growing field.

Data mining helps determine the best times and methods for loading and transporting merchandise from storage facilities to retail outlets.

5.DATA MINING APPLICATIONS IN MEDICINE

Data mining can discover potential office visits by observing the patients' behaviors.

When used to medical data, data mining can reveal patterns that lead to life-saving diagnoses and therapies.

Data mining technologies are being used by many industries to unearth previously concealed information. Companies benefit from increased productivity and expansion as a result.

Fig 3. Data mining applications in S/W mgt.

6.FINANCIAL DATA ANALYSIS

Data mining and systematic analysis are only possible due to the consistent availability of high-quality financial data provided by the banking and finance industries. Some typical instances are listed below.

The process of constructing and populating a data warehouse tailored to the needs of multidimensional data processing and extraction. Examining credit policies and estimating loan costs. In order to conduct targeted marketing, buyers are segmented into subsets based on shared traits. There are instances of money laundering and other forms of illicit financial activity.

7.RETAIL INDUSTRY

Because of the vast amounts of data that can be collected on sales, customer purchase history, transportation, consumption, and services, data mining is widely used in the retail industry. The amount of information gathered will increase as the internet continues to develop in user-friendliness and accessibility.

Data mining facilitates the discovery of purchasing trends and patterns in the retail industry. The result is improved quality, increased client retention, and increased customer satisfaction. The following are some applications of data mining in retail:

Leveraging data mining's efficacy in making and utilizing data registries. Consider sales, consumers, products, timing, and location from a variety of perspectives. Assessing the efficacy of current sales strategies

8.CUSTOMER RETENTION.

recommendations for products and resources to which they may be linked. There is currently a period of rapid expansion in the telecommunications sector, which encompasses a wide variety of related but distinct businesses. As a result of rapid advancements in computing and communication technology, the telecoms sector is expanding at an unprecedented rate. Because of this, data mining has become a crucial tool for assisting and comprehending commercial processes.

Telecommunications data mining is vital for a number of reasons, including the detection of fraud, optimization of available resources, and enhancement of service quality. Some ways in which data mining enhances telecommunications services are:

9.MULTIDIMENSIONAL ANALYSIS OF TELECOMMUNICATION DATA.

- You can make phone calls from your mobile device.
- Analyzing telecommunications
- data with graphing tools.

10.BIOLOGICAL DATA ANALYSIS

Rapid progress is being made in the biological sciences. Genome research encompasses proteomics, functional genomics, and other areas of biomedical inquiry. Collecting and analyzing biological data is what bioinformatics is all about. Data mining has facilitated our comprehension of biological data by illuminating previously unseen connections and providing us with novel insights.

Synthesizing data from multiple genomic and proteomic sources. The procedure involves a number of steps, including alignment, indexing, comparison, and searching for commonalities between several nucleotide sequences.

Protein pathways, genetic networks, and structural patterns are identified and named. Connections and routes are studied carefully. Analysis of genetic data visualization tools.

11.OTHER SCIENTIFIC APPLICATIONS

As these applications typically include manageable amounts of consistent data, statistical methods shine. There is an abundance of information from the geosciences and astronomy, among other fields of study.

Many fields, including chemical engineering, fluid dynamics, climate and ecosystem modeling, and others, generate massive amounts of data through fast numerical simulations. Some instances of data mining's scientific uses are as follows

12.GRAPH-BASED MINING

Visualization and domain specific knowledge.

Any behavior that compromises the confidentiality, integrity, or availability of a network's resources is considered an incursion. Since everything is now interconnected, security has become paramount. As the amount of people using the Internet and the ease with which hackers can get into and attack networks has increased, intrusion detection has become an integral aspect of network administration. Some ways in which data mining techniques can be used to spot intrusions are listed below:

The development of an intrusion detection system by employing data mining methods.

Aggregation, in conjunction with association and correlation research, can aid in the identification and improvement of salient features.

13.DATA MINING SYSTEM PRODUCTS

Products for data mining systems and data mining apps tailored to certain fields abound today. More applications and tools for data mining are being included into the existing systems at this time. Additionally, work is being done to standardize data extraction languages.

Choosing a Data Mining Platform

Consider the following criteria while making your final decision on a data mining system:

Types of Data - The data mining system can handle a variety of data formats, including records, relational databases, and structured text. Data can also be kept in a relational database or a data warehouse in addition to the more common ASCII text format. Therefore, it is crucial to make sure the data mining system can handle the correct format.

14.SYSTEM ISSUES

An examination of the Data Mining system's interoperability with multiple operating systems is necessary. A data mining system has the capability to work on one or more operating systems. In addition, there are data mining platforms that let users submit information via the web and XML.

Data Sources are the various collections of information from which the data mining process will draw. Some data mining tools can only read in ASCII text files, while others can process data from a wide variety of relational databases. A data mining system should also enable ODBC or OLE DB connections.

Classification is just one data mining function offered by various systems; others provide concept description, discovery-driven OLAP analysis, association mining, linkage analysis, statistical analysis, prediction, clustering, outlier analysis, similarity search, and many more.

Integrating databases or data warehouses is essential for data mining. Together, these parts provide a unified system for processing data. The many different types of links are detailed here. No coupling

- Loose Coupling
- Semi tight Coupling
- Tight Coupling

15. SCALABILITY - THERE ARE TWO SCALABILITY ISSUES IN DATA MINING AS FOLLOWS:

In a row (database storage) When talking about data mining systems, scalability is the ability to process more data without drastically increasing query execution times. To be more precise, a system is considered row scalable if it can handle a tenfold increase in the number of rows without resulting in a tenfold increase

in the execution time of queries.

Column scalability refers to the property of a data mining system where the time it takes to execute mining queries increases in a linear manner as the number of columns increases.

Information visualization instruments - The field of data mining visualization can be broken down into the following groups:

Mine Operations Visualization The term "data visualization" is the practice of presenting information visually, for as through charts, graphs, or maps, in order to improve readability, comprehension, and analysis. Mining Data Visualization Using mining techniques to remove minerals or other valuable resources from the earth. Data presented graphical form Data mining that is guided by the user requires a user-friendly graphical user interface and a query language developed for this purpose. Unlike relational database systems, data mining tools don't have a standard query language. New Directions in Data Mining

Methods for Data Mining That Are Both Dynamic and Highly Scalable

Database, data warehouse, and web-based storage solutions with built-in data mining capabilities.

Queries in data mining should be standardized. Data mining in visual contexts Extraction of information from complex data structures using novel methods Data mining in the life sciences

16.CONCLUSION

In this research, we took a deep dive into the whys and hows of several data mining approaches. There are essentially only a few things in the world: Structure mining, user mining, and data mining applications can all be found in fields as varied as banking, marketing, retail, medical, finance, customer retention, graphical database mining, and basic scientific inquiry.

REFERENCE

1. Kantardzic, Mehmed (2003). Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons. ISBN 0-471- 22852-4. OCLC 50055336.
2. "Microsoft Academic Search: Top conferences in data mining". Microsoft Academic Search.
3. "Google Scholar: Top publications - Data Mining & Analysis". Google Scholar.
4. Proceedings, International Conferences on Knowledge Discovery and Data Mining, ACM, New York.
5. SIGKDD Explorations, ACM, New York.
6. Gregory Piatetsky-Shapiro (2002) KDnuggets Methodology Poll
7. Gregory Piatetsky-Shapiro (2004) KDnuggets Methodology Poll
8. Gregory Piatetsky-Shapiro (2007) KDnuggets Methodology Poll
9. Óscar Marbán, Gonzalo Mariscal and Javier Segovia (2009); A Data Mining & Knowledge Discovery Process Model. In Data Mining and Knowledge Discovery in Real Life Applications, Book edited by: Julio Ponce and Adem Karahoca, ISBN 978-3-902613-53-0, pp. 438–453, February 2009, I-Tech, Vienna, Austria.
10. Lukasz Kurgan and Petr Musilek (2006); A survey of Knowledge Discovery and Data Mining process models. The Knowledge Engineering Review. Volume 21 Issue 1, March 2006, pp 1–24, Cambridge University Press, New York, NY,