

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

An Implementation of Web Personalization using Web Mining Techniques

Kannan Subramanian^{*1}, Suresh Babu .G.N.K²

^{1*} Assistant Professor, Department of MCA, Bharath University, Chennai, India

²Head, Dept of MCA, GKM College of Engg, Vandular, Chennai, India

ABSTRACT: Web mining is the application of data mining techniques to extract knowledge from web. Web mining has been explored to a vast degree and different techniques have been proposed for a variety of applications that includes “web search”, classification and personalization etc. Most research on web mining has been from a ‘data-centric’ point of view. In its paper, we highlight the significance of studying the evolving nature of the web personalization. Web usage mining is used to discover interesting user navigation patterns and can be applied to many real-world problems, such as improving web sites/pages, making additional topic or product recommendations, user/customer behavior studies, etc. a web usage mining system performs five major tasks. i) Data Gathering, ii) Data Preparation, iii) Navigation Pattern Discovery, iv) Pattern Analysis and Visualization and v) Pattern Applications. Each task is explained in detail and its related technologies are introduced. The web mining research is a converging research area from several research communities, such as databases, information retrieval, and artificial intelligence. In this paper we implement how web mining techniques can be apply for the customization. i.e., Web personalization.

KEYWORDS: Usage Mining, Navigation Patterns, Pattern Analysis, Content Mining, Structure Mining.

I.INTRODUCTION

Data Mining has attracted a great deal of attention in the information industry and in society as a whole in recent years, due to the wide availability of huge amounts of data and the imminent need for truing such data into useful information and knowledge. This information and knowledge gained can be used for applications ranging from market analysis, fraud detection, and customer retention, to production control and science exploration.

Data mining can be viewed as a result of the natural evolution if information technology. The database system industry has witnessed as evolutionary path in the development of the following functionalities (Figure 1.1): Data Collection and database creation, data management (including data storage and retrieval and database transaction processing and advanced data analysis (involving data warehousing and data mining). For instance, the early development of data collection and database creation mechanisms, served as a prerequisite for later developments of effective mechanisms for data storage and retrieval, and query and transaction processing. With numerous database systems offering query and transaction processing, as common practice, advanced data analysis has naturally become the text target.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

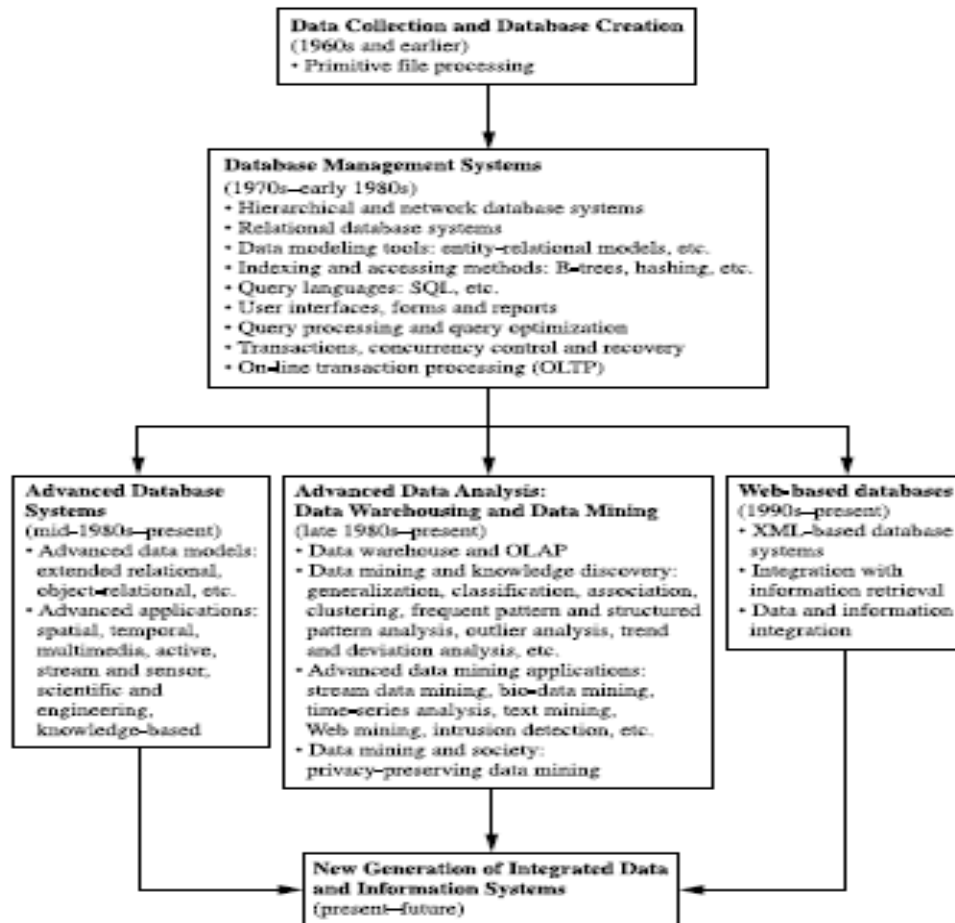


Figure 1.1 The evolution of database system technology.

II.MOTIVATED DATA MINING TECHNIQUES AND ANALYSIS

Since the 1960s, database and information technology has been evolving systematically from primitive file processing systems to sophisticated and powerful database systems [1]. The research and development in database systems since the 1970s has progressed from early hierarchical and network database systems to the development of relational database systems (Where data are stored in relational table and structures; see section 1.3.1), data modeling tools, and indexing and accessing methods [6]. In addition, users gained convenient and flexible data access through query languages, user interfaces, optimized query processing, and transaction management [2]. Efficient methods for online transactions processing (OLTP). Where a query is viewed as a read-only transaction, have contributed. Substantially, to the evolution

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

and wide acceptance of relational technology, as a major tool for efficient storage, retrieval, and management of large amounts of data [6].

III. DATA MINING USING KDD

Data mining refers to extracting or “mining” knowledge from large amounts of data. The term is actually a misnomer [4]. Remember that the mining of gold from rocks (or) sand is referred to as gold mining rather than rock or sand mining. Thus, data mining should have been more appropriately named “knowledge mining from data”, which is unfortunately, somewhat long. “Knowledge mining”, a shorter term, may not reflect the emphasis on mining from large amounts of data. Nevertheless, mining is a vivid term characterizing the process that finds a small set of precious nuggets from a great deal of raw material (Figure 1.3). Thus, such a misnomer that carries both “Data” and “Mining” became a popular choice [3]. Many other terms carry a similar or slightly different meaning to data mining, such as knowledge mining from data, knowledge extraction, data/pattern analysis, data archaeology, and data dredging. Many people treat data mining as a synonym, for another popularly used term, knowledge discovery from data (or) KDD. Alternatively, others view data mining as simply an essential step in the process of knowledge discovery [5]. Knowledge discovery as a process is depicted in Figure 1.4 and consists of an iterative sequence of the following steps.



Figure. 1.3. Data Mining – Searching for knowledge (interesting patterns) in your dat

1. **Data Cleaning:** (To remove noise and inconsistent data).
2. **Data integration** (where multiple data sources may be combined).
3. **Data selection** (where data relevant to the analysis task are retrieved from the database) [7].
4. **Data transformation:** (where data are transformed or consolidated into form appropriate for mining by performing summary or aggregation operations for instance).
5. **Data mining:** (an essential process where intelligent methods are applied in order to extract data patterns).
6. **Pattern evaluation** (to identify the truly interesting patterns representing knowledge based on some interestingness measures, section 1.5) [8].
7. **Knowledge presentation:** (where visualization and knowledge representation techniques are used to present the mined knowledge to the user).

Steps 1 to 4 are different of data preprocessing, where the data are prepared for mining [9]. The data mining step may interact with the user or a knowledge base. The interesting patterns are presented to the user and may be stored as new knowledge in the knowledge base. Note that according this view, data mining is only one step in the entire process,

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

albeit an essential one because it uncovers hidden patterns for evaluation. We agree that data mining is a step in the knowledge discovery process [10]. However, in industry in media, and in the database research milieu, the term data mining is becoming more popular than the longer term of knowledge discovery from data. Therefore, in this book, we choose to use the term data mining [13]. We adopt a broad view of data mining functionality: data mining is the process of discovering interesting knowledge from large amounts of data stored in databases, data warehouses, or other information repositories [12]. Based on this view, the architecture of a typical data mining system may have the following components (Figure 1.5):

- Database, data warehouse, world wide web, or other information repository: this is one or a set of databases, data warehouses, spreadsheets, or other kinds of information repositories. Data cleaning and data integration techniques may be performed on the data [11].
- Database or data warehouse server: the database or data warehouse server is responsible for fetching the relevant data, based on the user's data mining request.
- Knowledge Base: This is the domain knowledge that is used to guide the search or evaluate the interestingness of resulting patterns [14]. Such knowledge can include concept hierarchies, used to organize attributes or attribute values into different levels of abstraction. Knowledge such as user beliefs, which can be used to assess a pattern's interestingness based on its unexpectedness, may also be included [15]. Other examples of domain knowledge are additional interestingness constraints or thresholds, and metadata. (e.g., describing data from multiples heterogeneous sources).
- Data mining engine: this is essential to the data mining system and ideally consists of a set of functional modules for tasks such as characterization, outlier analysis and evolution analysis [16].
- Pattern evaluation module: this component typically employs interestingness measures (section 1.5) and interacts with the data mining modules so as to focus the search toward interesting patterns. It may use interestingness thresholds to filter out discovered patterns. Alternatively, the pattern evaluation module may be integrated mining method used. For efficient data mining, it is highly recommended to push the evaluation of pattern interestingness deep as possible into the mining process so as to confine the search to only the interesting patterns [17].
- User interface: this module communicates between users and the data mining system, allowing the user to interact with the system by specifying data mining query or task, providing information to help focus the search, and performing exploratory data mining based on the intermediate data mining results. In addition, this component allows the user to browse database and data warehouse schemas or data structures, evaluate mined patterns, and visualize the patterns in different forms.

IV. DATA MINING - ON WHAT KIND OF DATA

In this section, we examine a number of different data repositories on which mining can be performed [18]. In principle, data mining should be applicable to any kind of data repository, as well as to transient data, such as data streams. Thus the scope of our examination of data repositories will include relational databases, data warehouses, and transactional databases, advanced database systems, flat files, data streams, and the World Wide Web. Advanced database systems include object-relational databases and specific application-oriented databases, such as spatial databases, time-series databases, text databases, and multimedia databases. The challenges and techniques of mining may differ for each of the repository systems.

Although this book assumes that readers have basic knowledge of information systems, we provide a brief introduction to each of the major data repository systems listed above. In this section, we also introduce the fictitious *AllElectronics* store, which will be used to illustrate concepts throughout the text [19].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

Relational Databases:-

A database system, also called a database management system (DBMS), consists of a collection of interrelated data, known as a database, and a set of software programs to manage and access the data. The software programs involve mechanisms for the definition of database structures; for data storage; for concurrent, shared, or distributed data access; and for ensuring the consistency and security of the information stored, despite system crashes or attempts at unauthorized access [20].

A relational database is a collection of tables, each of which is assigned a unique name. Each table consists of a set of attributes (*columns* or *fields*) and usually stores a large set of tuples (*records* or *rows*). Each tuple in a relational table represents an object identified by a unique *key* and described by a set of attribute values. A semantic data model, such as an entity-relationship (ER) data model, is often constructed for relational databases [21].

Data Warehouses:-

Suppose that *AllElectronics* is a successful international company, with branches around the world. Each branch has its own set of databases. The president of *AllElectronics* has asked you to provide an analysis of the company's sales per item type per branch for the third quarter. This is a difficult task, particularly since the relevant data are spread out over several databases, physically located at numerous sites.

If *AllElectronics* had a data warehouse, this task would be easy. A data warehouse is a repository of information collected from multiple sources, stored under a unified schema, and that usually resides at a single site. Data warehouses are constructed via a process of data cleaning, data integration, data transformation, data loading, and periodic data refreshing. This process is discussed in Chapters 2 and 3.

Figure 1.7 shows the typical framework for construction and use of a data warehouse for *AllElectronics*.

Transactional Databases:-

In general, a transactional database consists of a file where each record represents a transaction [22]. A transaction typically includes a unique transaction identity number (*trans ID*) and a list of the items making up the transaction (such as items purchased in a store).

Advanced Data and Information Systems and Advanced Applications:-

Relational database systems have been widely used in business applications. With the progress of database technology, various kinds of advanced data and information systems have emerged and are undergoing development to address the requirements of new applications [23].

The new database applications include handling spatial data (such as maps), engineering design data (such as the design of buildings, system components, or integrated circuits), hypertext and multimedia data (including text, image, video, and audio data), time-related data (such as historical records or stock exchange data), stream data (such as video surveillance and sensor data, where data flow in and out like streams), and the WorldWideWeb (a huge, widely distributed information repository made available by the Internet). These applications require efficient data structures and scalable methods for handling complex object structures; variable-length records; semistructured or unstructured data; text, spatiotemporal, and multimedia data; and database schemas with complex structures and dynamic changes [24].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

In response to these needs, advanced databases stems and specific application-oriented database systems have been developed. These include object-relational database systems, temporal and time-series database systems, spatial and spatiotemporal database systems, text and multimedia database systems, heterogeneous and legacy database systems, data stream management systems, and Web-based global information systems.

Object Relational Databases:-

Object-relational databases are constructed based on an object-relational data model. This model extends the relational model by providing a rich data type for handling complex objects and object orientation. Because most sophisticated database applications need to handle complex objects and structures, object-relational databases are becoming increasingly popular in industry and applications.

Conceptually, the object-relational data model inherits the essential concepts of object-oriented databases, where, in general terms, each entity is considered as an object. Following the *AllElectronics* example, objects can be individual employees, customers, or items. Data and code relating to an object are *encapsulated* into a single unit. Each object has associated with it the following:

- a. A set of variables that describe the objects. These correspond to attributes in the entity-relationship and relational models.
- b. A set of messages that the object can use to communicate with other objects, or with the rest of the database system.

A set of methods, where each method holds the code to implement a message. Upon receiving a message, the method returns a value in response. For instance, the method for the message *get photo(employee)* will retrieve and return a photo of the given employee object.

Temporal Databases, sequence databases, and time-series databases:-

A temporal database typically stores relational data that include time-related attributes. These attributes may involve several timestamps, each having different semantics [20]. A sequence database stores sequences of ordered events, with or without a concrete notion of time. Examples include customer shopping sequences, Web click streams, and biological sequences. A time-series database stores sequences of values or events obtained over repeated measurements of time (e.g., hourly, daily, weekly). Examples include data collected from the stock exchange, inventory control, and the observation of natural phenomena (like temperature and wind). Data mining techniques can be used to find the characteristics of object evolution, or the trend of changes for objects in the database [19]. Such information can be useful in decision making and strategy planning. For instance, the mining of banking data may aid in the scheduling of bank tellers according to the volume of customer traffic. Stock exchange data can be mined to uncover trends that could help you plan investment strategies (e.g., when is the best time to purchase *AllElectronics* stock?). Such analyses typically require defining multiple granularity of time. For example, time may be decomposed according to fiscal years, academic years, or calendar years. Years may be further decomposed into quarters or months.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

Spatial Databases and spatial temporal databases:-

Spatial databases contain spatial-related information. Examples include geographic (map) databases, very large-scale integration (VLSI) or computer-aided design databases, and medical and satellite image databases. Spatial data may be represented in raster format, consisting of n -dimensional bit maps or pixel maps [19]. For example, a 2-D satellite image may be represented as raster data, where each pixel registers the rainfall in a given area. Maps can be represented in vector format, where roads, bridges, buildings, and lakes are represented as unions or overlays of basic geometric constructs, such as points, lines, polygons, and the partitions and networks formed by these components.

Geographic databases have numerous applications, ranging from forestry and ecology planning to providing public service information regarding the location of telephone and electric cables, pipes, and sewage systems. In addition, geographic databases are commonly used in vehicle navigation and dispatching systems. An example of such a system for taxis would store a city map with information regarding one-way streets, suggested routes for moving from region A to region B during rush hour, and the location of restaurants and hospitals, as well as the current location of each driver.

Text Databases and Multimedia Databases:-

Text databases are databases that contain word descriptions for objects. These word descriptions are usually not simple keywords but rather long sentences or paragraphs, such as product specifications, error or bug reports, warning messages, summary reports, notes, or other documents. Text databases may be highly unstructured (such as some Web pages on the WorldWideWeb). Some text databases may be somewhat structured, that is, *semi structured* (such as e-mail messages and many HTML/XML Web pages), whereas others are relatively well structured (such as library catalogue databases). Text databases with highly regular structures typically can be implemented using relational database systems.

Heterogeneous Databases and Legacy Databases: -

A heterogeneous database consists of a set of interconnected, autonomous component databases. The components communicate in order to exchange information and answer queries. Objects in one component database may differ greatly from objects in other component databases, making it difficult to assimilate their semantics into the overall heterogeneous database. Many enterprises acquire legacy databases as a result of the long history of information technology development (including the application of different hardware and operating systems). A legacy database is a group of *heterogeneous databases* that combines different kinds of data systems, such as relational or object-oriented databases, hierarchical databases, network databases, spreadsheets, multimedia databases, or file systems. The heterogeneous databases in a legacy database may be connected by intra or inter-computer networks [16].

Data Streams:-

Many applications involve the generation and analysis of a newkind of data, called stream data, where data flow in and out of an observation platform (or window) dynamically. Such data streams have the following unique features: *huge or possibly infinite volume, dynamically changing, flowing in and out in a fixed order, allowing only one or a small number of scans, and demanding fast (often real-time) response time*. Typical examples of data streams include various kinds of scientific and engineering data, time-series data, and data produced in other dynamic environments, such as power supply, network traffic, stock exchange, telecommunications, Web click streams, video surveillance, and weather or environment monitoring [17].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

The World Wide Web:-

The World Wide Web and its associated distributed information services, such as Yahoo!, Google, America Online, and AltaVista, provide rich, worldwide, on-line information services, where data objects are linked together to facilitate interactive access. Users seeking information of interest traverse from one object via links to another. Such systems provide ample opportunities and challenges for data mining [8]. For example, understanding user access patterns will not only help improve system design (by providing efficient access between highly correlated objects), but also leads to better marketing decisions (e.g., by placing advertisements in frequently visited documents, or by providing better customer/user classification and behavior analysis). Capturing user access patterns in such distributed information environments is called Web usage mining (or Weblog mining).

Data Mining Functionalities what kinds of patterns can be mined :-

Data mining functionalities are used to specify the kind of patterns to be found in data mining tasks. In general, data mining tasks can be classified into two categories: descriptive and predictive. Descriptive mining tasks characterize the general properties of the data in the database. Predictive mining tasks perform inference on the current data in order to make predictions.

Concept/Class Description: Characterization and Discrimination:-

Data can be associated with classes or concepts. For example, in the *AllElectronics* store, classes of items for sale include *computers* and *printers*, and concepts of customers include *bigSpenders* and *budgetSpenders*. It can be useful to describe individual classes and concepts in summarized, concise, and yet precise terms. Such descriptions of a class or a concept are called class/concept descriptions. These descriptions can be derived via.

- (1) *data characterization*, by summarizing the data of the class under study (often
- (2) called the target class) in general terms, or *data discrimination*, by comparison of the target class with one or a set of comparative classes (often called the contrasting classes), or
- (3) both data characterization and discrimination.

Data characterization is a summarization of the general characteristics or features of a target class of data. The data corresponding to the user-specified class are typically collected by a database query. For example, to study the characteristics of software products whose sales increased by 10% in the last year, the data related to such products can be collected by executing an SQL query.

Mining Frequent Patterns, Associations, and Correlations:-

Frequent patterns, as the name suggests, are patterns that occur frequently in data. There are many kinds of frequent patterns, including itemsets, subsequences, and substructures. A *frequent itemset* typically refers to a set of items that frequently appear together in a transactional data set, such as milk and bread. A frequently occurring subsequence, such as the pattern that customers tend to purchase first a PC, followed by a digital camera, and then a memory card, is a (*frequent*) *sequential pattern*. A substructure can refer to different structural forms, such as graphs, trees, or lattices, which may be combined with itemsets or subsequences. If a substructure occurs frequently, it is called a (*frequent*) *structured pattern*. Mining frequent patterns leads to the discovery of interesting associations and correlations within data.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

Classification and prediction:-

Classification is the process of finding a model (or function) that describes and distinguishes data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown. The derived model is based on the analysis of a set of training data (i.e., data objects whose class label is known).

Cluster Analysis:-

“What is cluster analysis?” Unlike classification and prediction, which analyze class-labeled data objects, clustering analyzes data objects without consulting a known class label [4].

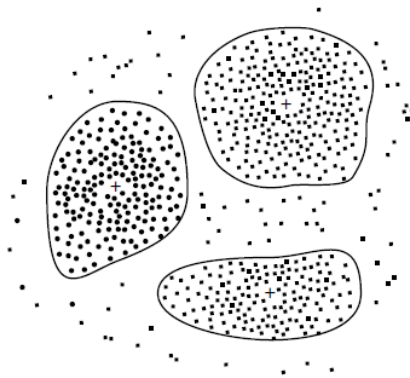


Figure: 1.11 A 2-D plot of customer data with respect to customer locations in a city, showing three data cluster

In general, the class labels are not present in the training data simply because they are not known to begin with. Clustering can be used to generate such labels. The objects are clustered or grouped based on the principle of *maximizing the intraclass similarity and minimizing the interclass similarity*. That is, clusters of objects are formed so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other clusters. Each cluster that is formed can be viewed as a class of objects, from which rules can be derived. Clustering can also facilitate taxonomy formation, that is, the organization of observations into a hierarchy of classes that group similar events together [7].

Outlier Analysis:-

A database may contain data objects that do not comply with the general behavior or model of the data. These data objects are outliers. Most data mining methods discard outliers as noise or exceptions. However, in some applications such as fraud detection, the rare events can be more interesting than the more regularly occurring ones. The analysis of outlier data is referred to as outlier mining.

Evolution Analysis:-

Data evolution analysis describes and models regularities or trends for objects whose behavior changes over time. Although this may include characterization, discrimination, association and correlation analysis, classification, prediction, or clustering of *timereLATED* data, distinct features of such an analysis include time-series data analysis, sequence or periodicity pattern matching, and similarity-based data analysis [15].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

Classification of Data Mining Systems:-

Data mining is an interdisciplinary field, the confluence of a set of disciplines, including database systems, statistics, machine learning, visualization, and information science (Figure 1.12). Moreover, depending on the data mining approach used, techniques from other disciplines may be applied, such as neural networks, fuzzy and/or rough set theory, knowledge representation, inductive logic programming, or high-performance computing. Depending on the kinds of data to be mined or on the given data mining application, the data mining system may also integrate techniques from spatial data analysis, information retrieval, pattern recognition, image analysis, signal processing, computer graphics, Web technology, economics, business, bioinformatics, or psychology. Because of the diversity of disciplines contributing to data mining, data mining research is expected to generate a large variety of data mining systems. Therefore, it is necessary to provide a clear classification of data mining systems, which may help potential users distinguish between such systems and identify those that best match their needs. Data mining systems can be categorized according to various criteria, as follows:

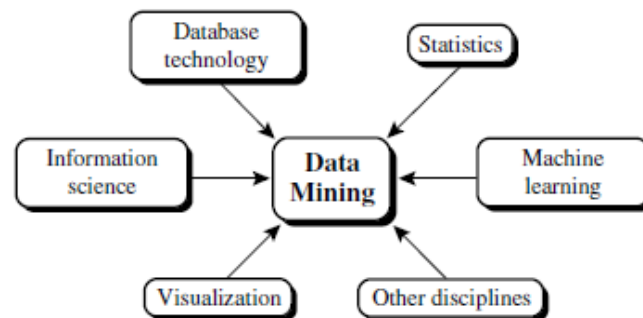


Figure 1.12. Data Mining as a confluence of multiple disciplines.

V.CONCLUSION

Database technology has evolved from primitive file processing to the development of database management systems with query and transaction processing. Further progress has led to the increasing demand for efficient and effective advanced data analysis tools. This need is a result of the explosive growth in data collected from applications, including business and management, government administration, science and engineering, and environmental control. Data mining is the task of discovering interesting patterns from large amounts of data, where the data can be stored in databases, data warehouses, or other information repositories. It is a young interdisciplinary field, drawing from areas such as database systems, data warehousing, statistics, machine learning, data visualization, information retrieval, and high-performance computing. Other contributing areas include neural networks, pattern recognition, spatial data analysis, image databases, signal processing, and many application fields, such as business, economics, and bioinformatics. A knowledge discovery process includes data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge presentation. The architecture of a typical data mining system includes a database and/or data warehouse and their appropriate servers, a data mining engine and pattern evaluation module (both of which interact with a knowledge base), and a graphical user interface. Integration of the data mining components, as a whole, with a database or data warehouse system can involve either no coupling, loose coupling, semi tight coupling, or tight coupling. A well-designed data mining system should offer tight or semi tight coupling with a database and/or data warehouse system. Data patterns can be mined

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2015

from many different kinds of databases, such as relational databases, data warehouses, and transactional, and object-relational databases. Interesting data patterns can also be extracted from other kinds of information repositories, including spatial, time-series, sequence, text, multimedia, and legacy databases, data streams, and the Worldwide Web. A data warehouse is a repository for long-term storage of data from multiple sources, organized so as to facilitate management decision making. The data are stored under a unified schema and are typically summarized. Data warehouse systems provide some data analysis capabilities, collectively referred to as OLAP (on-line analytical processing).

REFERENCES

- [1] M. S. Chen, J. Han, and P. S. Yu. Data mining: An overview from a database perspective. IEEE Trans. Knowledge and Data Engineering, 8:866-883, 1996.
- [2] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996.
- [3] W. J. Frawley, G. Piatetsky-Shapiro and C. J. Matheus, Knowledge Discovery in Databases: An Overview. In G. Piatetsky-Shapiro et al. (eds.), Knowledge Discovery in Databases. AAAI/MIT Press, 1991.
- [4] Udayakumar R., Khanaa V., Saravanan T., "Analysis of polarization mode dispersion in fibers and its mitigation using an optical compensation technique", Indian Journal of Science and Technology, ISSN : 0974-6846, 6(S6) (2013) pp. 4767-4771.
- [5] J. Han and M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.
- [6] T. Imielinski and H. Mannila. A database perspective on knowledge discovery. Communications of ACM, 39:58-64, 1996.
- [7] G. Piatetsky-Shapiro, U. M. Fayyad, and P. Smyth. From data mining to knowledge discovery: An overview. In U.M. Fayyad, et al. (eds.), Advances in Knowledge Discovery and Data Mining, 1-35. AAAI/MIT Press, 1996.
- [8] Udayakumar R., Khanaa V., Saravanan T., "Chromatic dispersion compensation in optical fiber communication system and its simulation", Indian Journal of Science and Technology, ISSN : 0974-6846, 6(S6) (2013) pp. 4762-4766.
- [9] G. Piatetsky-Shapiro and W. J. Frawley. Knowledge Discovery in Databases. AAAI/MIT Press, 1991.
- [10] The Past and Future of 1997: Database Systems: A Textbook Case of Research Paying Off. By: J.N. Gray, Microsoft 1997.
- [11] The Future of 1996: Database Research: Achievements and Opportunities Into the 21st Century. By: Silberschatz, M. Stonebraker, J. Ullman. Eds. SIGMOD Record, Vol. 25, No. pp. 52-63 March 1996.
- [12] Uma Mageswaran, S., Guna Sekhar, N.O., "Reactive power contribution of multiple STATCOM using particle swarm optimization", International Journal of Engineering and Technology, ISSN : 1793-8236, 5(1) (2013) pp. 122-126.
- [13] Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy, "Advances in Knowledge Discovery and Data Mining", AAAI Press/ The MIT Press, 1996.
- [14] J. Ross Quinlan, "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers, 1993.
- [15] Michael Berry and Gordon Linoff, "Data Mining Techniques (For Marketing, Sales, and Customer Support), John Wiley & Sons, 1997.
- [16] Udayakumar, R., Khanaa, V., Saravanan, T., "Synthesis and structural characterization of thin films of SnO₂ prepared by spray pyrolysis technique", Indian Journal of Science and Technology, ISSN : 0974-6846, 6(S6) (2013) pp.4754-4757.
- [17] Sholom M. Weiss and Nitin Indurkha, "Predictive Data Mining: A Practical Guide", Morgan Kaufmann Publishers, 1998.
- [18] Alex Freitas and Simon Lavington, "Mining Very Large Databases with Parallel Processing", Kluwer Academic Publishers, 1998.
- [19] A. K. Jain and R. C. Dubes, "Algorithms for Clustering Data", Prentice Hall, 1988.
- [20] Vidyakshmi K., Kamalakannan P., Viswanathan S., Ramaswamy S., "Antinociceptive effect of certain dihydroxy flavones in mice", Pharmacology Biochemistry and Behavior, ISSN : 0091-3057, 96(1) (2010) pp. 1-6.
- [21] V. Cherkassky and F. Mulier, "Learning From Data", John Wiley & Sons, 1998.
- [22] Usama Fayyad, "Mining Databases: Towards Algorithms for Knowledge Discovery", Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, vol. 21, no. 1, March 1998.
- [23] Christopher Matheus, Philip Chan, and Gregory Piatetsky-Shapiro, "Systems for Knowledge Discovery in Databases", IEEE Transactions on Knowledge and Data Engineering, 5(6):903-913, December 1993.
- [24] Rakesh Agrawal and Tomasz Imielinski, "Database Mining: A Performance Perspective", IEEE Transactions on Knowledge and Data Engineering, 5(6):914-925, December 1993.
- [25] R.Karthikeyan, Dr.S.R.Suresh, An Approach for Real Time Testing Reliability & Usability Testing Process, International Journal of Innovative Research in Computer and Communication Engineering, ISSN (Online): 2320 - 9801, pp 644-651, Vol. 1, Issue 3, May 2011
- [26] P.Ramya, dr.nalini, effective navigation queryresults based on biomedicaldatabase, International Journal of Innovative Research in Computer and Communication Engineering, ISSN: 2249-0183, pp 36-39, volume 1 Issue 3 No4- Dec 2011
- [27] P.Kavitha, Authentic Learning Activities with Pedagogical Stylistics - Enhancing InE-Learning Websites, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 1211-1215, Volume 1, Issue 6, August 2013
- [28] N.Priya, VANET Based Adaptive Traffic Signal Control, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 1201-1208, Vol. 1, Issue 6, August 2014
- [29] N.Sakthi Priya, Cervical Cancer Screening and Classification Using Acoustic Shadowing, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 1676-1679, Volume 1, Issue 8, October 2013