Data Mining: Future Trends and Applications

Annan Naidu Paidi

Asst.Prof of CSE Centurion University, Odisha, India.

Abstract: Knowledge has played a significant role on human activities since his development. Data mining is the process of knowledge discovery where knowledge is gained by analyzing the data store in very large repositories, which are analyzed from various perspectives and the result is summarized it into useful information. Due to the importance of extracting knowledge/information from the large data repositories, data mining has become a very important and guaranteed branch of engineering affecting human life in various spheres directly or indirectly. The purpose of this paper is to survey many of the future trends in the field of data mining, with a focus on those which are thought to have the most promise and applicability to future data mining applications.

Keywords: Current and Future of Data Mining, Data Mining, Data Mining Trends, Data mining Applications.

I.   Introduction

Originated from knowledge discovery from databases (KDD), also known as data Mining (DM),Data Mining (DM) is the extraction of new knowledge from large databases. Many techniques are currently used in this fast emerging field, including statistical analysis and machine learning based approaches. With the rapid development of the World Wide Web and the fast increase of unstructured databases, new technologies and applications are continuously coming forth it his field. The main challenges in data mining are:

- Data mining to deal with huge amounts of data located at different sites The amount of data can easily exceed the terabyte limit;
- Data mining is very computationally intensive process involving very large data sets. Usually, it is necessary to partition and distribute the data for parallel processing to achieve acceptable time and space performance;
- Input data change rapidly. In many application domain data to be mined either is produced with high rate or they actually come in streams. In those cases, knowledge has to be mined fast and efficiently in order to be usable and updated.

II.   The Scope of Data Mining

Data mining derives its name from the similarities between searching for valuable business information in a large database — for example, finding linked products in gigabytes of store scanner data — and mining a mountain for a vein of valuable ore. Both processes require either sifting through an immense amount of material, or intelligently probing it to find exactly where the value resides. Given databases of sufficient size and quality, data mining technology can generate new business opportunities by providing these capabilities:

- **Automated prediction of trends and behaviours**. Data mining automates the process of finding predictive information in large databases. Questions that traditionally required extensive hands-on analysis can now be answered directly from the data — quickly. A typical example of a predictive problem is targeted marketing. Data mining uses data on past promotional mailings to identify the targets most likely to maximize return on investment in future mailings. Other predictive problems include forecasting bankruptcy and other forms of default, and identifying segments of a population likely to respond similarly to given events.
- **Automated discovery of previously unknown patterns**. Data mining tools sweep through databases and identify previously hidden patterns in one step. An example of pattern discovery is the analysis of retail sales data to identify seemingly unrelated products that are often purchased together. Other pattern discovery problems include detecting fraudulent credit card transactions and identifying anomalous data that could represent data entry keying errors.
  The most commonly used techniques in data mining are:
- **Artificial neural networks**: Non-linear predictive models that learn through training and resemble biological neural networks in structure.
- **Decision trees**: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID).
- **Genetic algorithms**: Optimization techniques that use process such as genetic combination, mutation, and natural selection in a design based on the concepts of evolution.
- **Nearest neighbour method**: A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where $k \geq 1$). Sometimes called the k-nearest neighbour technique.
- **Rule induction**: The extraction of useful if-then rules from data based on statistical significance.

III.   ROOTS OF DATA MINING

A.   Statistics

The most important lines is statistics. Without statistics, there would be no data mining, as statistics are the foundation of most technologies on which data mining is built. Statistics embrace concepts such as regression analysis, standard distribution, standard deviation, standard variance, discriminate analysis, cluster analysis, and confidence intervals, all of which are used to study data and data relationships. These are the very building blocks with which more advanced statistical analyses are underpinned. Certainly, within the heart of today's data mining tools and techniques, classical statistical analysis plays a significant role.

B. Artiiciafl Intelligence & Machine Learning

Data mining's second longest family line is artificial intelligence and machine learning. AI is built upon heuristics as opposed to statistics, and attempts to apply human-thought like processing to statistical problems. Because this approach requires vast computer processing power, it was not practical until the early 1980s, when computers began to offer useful power at reasonable prices. AI found a few applications at thievery high end scientific/government markets, but the required supercomputers of the era priced AI out of the reach of virtually everyone else. Machine Learning could be considered as an evolution of AI, because it blends AI heuristics with advanced statistical methods. It let computer programs learn about the data they study and then apply learned knowledge to data.

C. Databases

Third family is Databases. Huge amount of data needs to be stored in a repository, and that too needs to be managed. So, comes in light the databases. Earlier data was managed in records and fields, then in various models like hierarchical, network etc. Relational model served the needs of data storage for long while. Other advanced system that emerged is object relational databases. But in data mining, volume of data is too high, so we need specialized servers for it. We call the term as Data Warehousing. Data warehousing also supports OLAP operations to be applied on it, to support decision making.

D. Other Technologies

Apart from these, data mining inculcates various other areas e.g. pattern discovery, visualization, business intelligence etc. The table summarizes the evolution data mining on the grounds of development in databases.

Evolutionary Step	Business Question	Enabling Technologies	Product Providers	Characteristics
Data Collection (1960s)	"What was my total revenue in the last five years?"	Computers, tapes, disks	IBM, CDC	Retrospective, static data delivery
Data Access (1980s)	"What were unit sales in New England last March?"	Relational databases (RDBMS), Structured Query Language (SQL), ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Retrospective, dynamic data delivery at record level
Data Warehousing & Decision Support (1990s)	"What were unit sales in New England last March? Drill down to Boston."	On-line analytic processing (OLAP), multidimensional databases, data warehouses	Pilot, Comshare, Arbor, Cognos, Microstrategy	Retrospective, dynamic data delivery at multiple levels
Data Mining (Emerging Today)	"What's likely to happen to Boston unit sales next month? Why?"	Advanced algorithms, multiprocessor computers, massive databases	Pilot, Lockheed, IBM, SGI, numerous startups (nascent industry)	Prospective, proactive information delivery

Table 1. Steps in the Evolution of Data Mining.

IV. CURRENT TRENDS AND APPLICATIONS

Data mining is formally defined as the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [2]. The field of data mining has been growing rapidly due to its broad applicability, achievements and scientific progress, understanding. A number of data mining applications have been successfully implemented in various domains like fraud detection, retail, health care, finance, telecommunication, and risk analysis...etc. are few to name. The ever increasing complexities in various fields and improvements in technology have posed new challenges to data mining; the various challenges include different data formats, data from disparate locations, advances in computation and networking resources, research and scientific fields, ever growing business challenges etc.Advancements in data mining with various integrations and implications of methods and techniques have shaped the present data mining applications to handle the various challenges, the current trends of data mining applications are:

A. Fight against Terrorism

After 9-11 attacks, many countries imposed new laws against fighting terrorism. These laws allow intelligence agencies to effectively fight against terrorist organizations.USA launched Total Information Awareness program with the goal of creating a huge database of that consolidate all the information on population. Similar projects were also launched in European countries and rest of the world. This program faced several problems,
a. The heterogeneity of database, the target database had to deal with text, audio, image and multimedia data.
b. Second problem was scalability of algorithms. The execution time increases as size of data (which is huge). For example, 230 cameras were placed in London, to read number plates of vehicles. An estimated 40,000 vehicles pass camera every hour, in this way the camera must recognize 10 vehicles per second, which poses heavy loads on both hardware and software.

B. Bio-informatics and Cure for Diseases

The second most important application trend, deals with mining and interpretation of biological sequences and structures. Data mining tools are rapidly being used in finding genes regarding cure of diseases like Cancer and AIDS .

C. Web and Semantic Web

Web is the hottest trend now, but it is unstructured. Data mining is helping web to be organized, which is called Semantic web. The underlying technology is Resource Description Framework (RDF) which is used to describe resources. FOAF is also a supporting technology, heavily used in Face book and Orkut for tagging. But still there are issues like combining all RDF statements and dealing with erroneous RDF statements. Data mining technologies are serving a lot to make the web, a semantic web.

D. Business Trends

Today's business environment is more dynamic, so businesses must be able to react quicker, must be more profitable, and offer high quality services that ever before. Here, data mining serves as a fundamental technology in enabling customer's transactions more accurately, faster and meaningfully. Data mining techniques of classification, regression, and cluster analysis are used for in current business trends . Almost all of the current business data mining applications are based on the classification and prediction techniques for supporting business decisions, thus creating strong Business Intelligence (BI) system.

APPLICATIONS

As data mining matures, new and increasingly innovative applications for it emerge. Although a wide variety of data mining scenarios can be described. For the purpose of this paper the applications of data mining are divided in thefollowing categories:
- Healthcare
- Finance
- Retail industry
- Telecommunication
- Text Mining & Web Mining
- Higher Education

Healthcare

The past decade has seen an explosive growth in biomedical research, ranging from the development of new pharmaceuticals and in cancer therapies to the identification and study of human genome by discovering large scale sequencing patterns and gene functions. Recent research in DNA analysis has led to the discovery of genetic causes for many diseases and disabilities as well as approaches for disease diagnosis, prevention and treatment.

Finance

Most banks and financial institutions offer a wide variety of banking services (such as checking, saving, and business and individual customer transactions), credit (such as business, mortgage, and automobile loans), and investment services (such as mutual funds). Some also offer insurance services and stock services. Financial data collected in the

banking and financial industry is often relatively complete, reliable and high quality, which facilitates systematic data analysis and data mining. For example it can also help in fraud detection by detecting a group of people who stage accidents to collect on insurance money.

Retail Industry

Retail industry collects huge amount of data on sales, customer shopping history, goods transportation and consumption and service records and so on. The quantity of data collected continues to expand rapidly, especially due to the increasing ease, availability and popularity of the business conducted on web, or e-commerce. Retail industry provides a rich source for data mining. Retail data mining can help identify customer behaviour, discover customer shopping patterns and trends, improve the quality of customer service, achieve better customer retention and satisfaction, enhance goods consumption ratios design more effective goods transportation and distribution policies and reduce the cost of business.

Telecommunication

The telecommunication industry has quickly evolved from offering local and long distance telephone services to provide many other comprehensive communication services including voice, fax, pager, cellular phone, images, e-mail, computer and web data transmission and other data traffic. The integration of telecommunication, computer network, Internet and numerous other means of communication and computing are underway. Moreover, with the deregulation of the telecommunication industry in many countries and the development of new computer and communication technologies, the telecommunication market is rapidly expanding and highly competitive. This creates a great demand from data mining in order to help understand business involved, identify telecommunication patterns, catch fraudulent activities, make better use of resources ,and improve the quality of service.

Text Mining and Web Mining

Text mining is the process of searching large volumes of documents from certain keywords or key phrases. By searching literally thousands of documents various relationships between the documents can be established. Using text mining however, we can easily derive certain patterns in the comments that may help identify a commonest of customer perceptions not captured by the other survey questions. An extension of text mining is web mining. Web mining is an exciting new field that integrates data and text mining within a website. It enhances the web site with intelligent behaviour, such as suggesting related links or recommending new products to the consumer. Web mining is especially exciting because it enables tasks that were previously difficult to implement. They can be configured to monitor and gather data from a wide variety of locations and can analyze the data across one or multiple sites. For example the search engines work on the principle of data mining.

Higher Education

An important challenge that higher education faces today is predicting paths of students and alumni. Which student will enrol in particular course programs? Who will need additional assistance in order to graduate? Meanwhile additional issues, enrolment management and time-todegree, continue to exert pressure on colleges to search for new and faster solutions. Institutions can better address these students and alumni through the analysis and presentation of data. Data mining has quickly emerged as a highly desirable tool for using current reporting capabilities to uncover and understand hidden patterns in vast databases.

V. Future Trends and Applications

DISTRIBUTED/COLLECTIVE DATA MINING

One area of data mining which is attracting a good amount of attention is that of distributed and collective data mining. Much of the data mining which is being done currently focuses on a database or data warehouse of information which is physically located in one place. However, the situation arises where information may be located in different places, in different physical locations. This is known generally as distributed data mining (DDM). Therefore, the goal is to effectively mine distributed data which is located in heterogeneous sites. Examples of this include biological information located in different databases, data which comes from the databases of two different firms, or analysis of data from different branches of a corporation, the combining of which would be an expensive and time-consuming process.

Distributed data mining (DDM) is used to offer a different approach to traditional approaches analysis, by using a combination of localized data analysis, together with a "global data model. " In more specific terms, this is specified as:- performing local data analysis for generating partial data models, and-combining the local data models from different data sites in order to develop the global model. This global model combines the results of the separate analyses. Often the global model produced, especially if the data in different locations has different features or characteristics, may become incorrect or ambiguous. This problem is especially critical when the data in distributed sites is heterogeneous rather than homogeneous.

UBIQUITOUS DATA MINING (UDM)

The advent of laptops, palmtops, cell phones, and wearable computers is making ubiquitous access to large quantity of data possible. Advanced analysis of data for extracting useful knowledge is the next natural step in the world of ubiquitous computing. Accessing and analyzing data from a ubiquitous computing device offer many challenges. For

example, UDM introduces additional cost due to communication, computation, security, and other factors. So one of the objectives of UDM is to mine data while minimizing the cost of ubiquitous presence.

Human-computer interaction is another challenging aspect of UDM. Visualizing patterns like classifiers, clusters, associations and others, in portable devices are usually difficult. The small display areas offer serious challenges to interactive data mining environments. Data management in a mobile environment is also a challenging issue. Moreover, the sociological and psychological aspects of the integration between data mining technology and our lifestyle are yet to be explored. The key issues to consider include theories of UDM, advanced algorithms for mobile and distributed applications, data management issues, mark-up languages, and other data representation techniques; integration with database applications for mobile environments, architectural issues: (architecture, control, security, and communication issues), specialized mobile devices for UDM, software agents and UDM (Agent based approaches in UDM, agent interaction--- cooperation, collaboration, negotiation, organizationalbehavior), applications of UDM (Application in business, science, engineering, medicine, and other disciplines), location management issues in UDM and technology for web-based applications of UDM.

HYPERTEXT AND HYPERMEDIA DATA MINING

Hypertext and hypermedia data mining can be characterized as mining data which includes text, hyperlinks, text mark-ups, and various other forms of hypermedia information. As such, it is closely related to both web mining, and multimedia mining, which are covered separately in this section, but in reality are quite close in terms of content and applications. While the World Wide Web is substantially composed of hypertext and hypermedia elements, there are other kinds of hypertext/hypermedia data sources which are not found on the web. Examples of these include the information found in online catalogues, digital libraries, online information databases, and the like.. Some of the important data mining techniques used for hypertext and hypermedia data mining include classification (supervised learning), clustering (unsupervised learning), semi-structured learning, and social network analysis.

In the case of classification, or supervised learning, the process starts off by reviewing training data in which items are marked as being part of a certain class or group. This data is the basis from which the algorithm is trained. One application of classification is in the area of web topic directories, which can group similar sounding or spelled terms into appropriate categories, so that searches will not bring up inappropriate sites and pages. The use of classification can also result in searches which are not only based on keywords, but also on category and classification attributes. Methods used for classification include naive Bayes classification, parameter smoothing, dependence modelling, and maximum entropy.
Unsupervised learning, or clustering, differs from classification in that classification involved the use of training data, clustering is concerned with the creation of hierarchies of documents based on similarity, and organize the documents based on that hierarchy. Intuitively, this would result in more similar documents being placed on the leaf levels of the hierarchy, with less similar sets of document areas being placed higher up, closer to the root of the tree. Techniques which have been used for unsupervised learning include k-means clustering, agglomerative clustering, random projections, and latent semantic indexing.

Semi-supervised learning and social network analysis are other methods which are important to hypermediabaseddata mining. Semi-supervised learning is the case where there are both labelled and unlabeled documents, and there is a need to learn from both types of documents. Social network analysis is also applicable because the web is considered a social network, which examines networks formed through collaborative association, whether it be between friends, academics doing research or service on committees, and between papers through references and citations. Graph distances and various aspects of connectivity come into play when working in the area of social networks.

MULTIMEDIA DATA MINING

Multimedia Data Mining is the mining and analysis of various types of data, including images, video, audio, and animation. The idea of mining data which contains different kinds of information is the main objective of multimedia data mining. As multimedia data mining incorporates the areas of text mining, as well as hypertext/hypermedia mining, these fields are closely related. Much of the information describing these other areas also applies to multimedia data mining. This field is also rather new, but holds much promise for the future. Multimedia information, because its nature as a large collection of multimedia objects, must be represented differently from conventional forms of data. One approach is to create a multimedia data cube which can be used to convert multimedia-type data into a form which is suited to analysis using one of the main data mining techniques, but taking into account the unique characteristics of the data. This may include the use of measures and dimensions for texture, shape, colour, and related attributes. In essence, it is possible to create a multidimensional spatial database. Among the types of analyses which can be conducted on multimedia databases include associations, clustering, classification, and similarity search. Another developing area in multimedia data mining is that of audio data mining (mining music). The idea is basically to use audio signals to indicate the patterns of data or to represent the features of data mining results. The basic advantage of audio data mining is that while using a technique such as visual data mining may disclose interesting patterns from observing graphical displays, it does require users to concentrate on watching patterns, which can become monotonous. But when representing it as a stream of audio, it is possible to transform patterns into sound and music and listen top itches, rhythms, tune, and melody in order to identify anything interesting or unusual.

SPATIAL AND GEOGRAPHIC DATA MINING

The data types which come to mind when the term data mining is mentioned involves data as we know it—statistical, generally numerical data of varying kinds. However, it is also important to consider information which is of an entirely different kind—spatial and geographic data which could contain information about astronomical data, natural resources, or even orbiting satellites and spacecraft which transmit images of earth from out in space. Much of this data is image-oriented, and can represent a great deal of information if properly analyzed and mined.

A definition of spatial data mining is as follows: "the extraction of implicit knowledge, spatial relationships, or other patterns not explicitly stored in spatial databases." Some of the components of spatial data which differentiate it from other kinds include distance and topological information, which can be indexed using multidimensional structures, and required special spatial data access methods, together with spatial knowledge representation and data access methods, along with the ability to handle geometric calculations. Analyzing spatial and geographic data include such tasks as understanding and browsing spatial data, uncovering relationships between spatial data items (and also between non-spatial and spatial items), and also analysis using spatial databases and spatial knowledge bases. The applications of these would be useful in such fields as remote sensing, medical imaging, navigation, and related uses. Some of the techniques and data structures which are used when analyzing spatial and related types of data include the use of spatial warehouses, spatial data cubesandspatialOLAP.Spatialdatawarehousescanbedefinedasthosewhicharesubjectoriented, integrated, nonvolatile, and time-variant . Some of the challenges in constructing a spatial data warehouse include the difficulties of integration of data from heterogeneous sources, and also applying the use of on-line analytical processing which is not only relatively fast, but also offers some forms of flexibility. In general, spatial data cubes, which are components of spatial data warehouses, are designed with three types of dimensions and two types of measures. The three types of dimensions include the noncapital dimension (data which is noncapital in nature), the spatial to noncapital dimension (primitive level is spatial but higher level-generalization is noncapital), and the spatial-to-spatial dimension (both primitive and higher levels are all spatial). In terms of measures, there are both numerical (numbers only), and spatial (pointers to spatial object) measured used in spatial data cubes. A side from the implementation of data warehouses for spatial data, there is also the issue of analyses which can be done on the data. Some of the analyses which can be done include association analysis, clustering methods, and the mining of raster databases There have been number of studies conducted on spatial data mining.

TIME SERIES/SEQUENCE DATA MINING

Another important area in data mining centres on the mining of time series and sequence-based data. Simply put, this involves the mining of a sequence of data, which can either be referenced by time (time-series, such as stock market and production process data), or is simply a sequence of data which is ordered in a sequence. In general, one aspect of mining time series data focuses on the goal of identifying movements or components which exist within the data (trend analysis).These can include long-term or trend movements, seasonal variations, cyclical variations, and random movements (Han and Kamber, 2001).Other techniques which can be used on these kinds of data include similarity search, sequential pattern mining, and periodicity analysis. *Similarity search* is concerned with the identification of a pattern sequence which is close or similar to a given pattern, and this form of analysis can be broken down into two subtypes: whole sequence matching and subsequence matching. Whole sequence matching attempts to find all sequences which bear a likeness to each other, while subsequence matching attempts to find those patterns which are similar to a specified, given sequence.

Sequential pattern mining has as its focus the identification of sequences which occur frequently in a time series or sequence of data. This is particularly useful in the analysis of customers, where certain buying patterns could be identified, such as what might be the likely follow-up purchase to purchasing a certain electronics item or computer, for example.

Periodicity analysis attempts to analyze the data from the perspective of identifying patterns which repeat or recur in a time series. This form of data mining analysis can be categorized as being full periodic, partial periodic or cyclic periodic. In general, full periodic is the situation where all of the data points in time contribute to the behaviour of the series. This is in contrast to partial periodicity, where only certain points in time contribute to series behaviour.

CONSTRAINT- BASED DATA MINING

Many of the data mining techniques which currently exist are very useful but lack the benefit of any guidance or user control. One method of implementing some form of human involvement into data mining is in the form of constraint-based data mining. This form of data mining incorporates the use of constraints which guides the process. Frequently this is combined with the benefits of multidimensional mining to add greater power to the process. There are several categories of constraints which can be used, each of which has its own characteristics and purpose. These are:

Knowledge-type constraints. This type of constraint specifies the "type of knowledge" which is to be mined, and is typically specified at the beginning of any data mining query. Some of the types of constraints which can be used include clustering, association, and classification.

Data constraints. This constraint identifies the data which is to be used in the specific data mining query. Since constraint-based mining is ideally conducted within the framework of an ad-hoc, query driven system, data constraints can be specified in a form similar to that of a SQL query.

Dimension/level constraints. Because much of the information being mined is in the form of a database or multidimensional data warehouse, it is possible to specify constraints which specify the levels or dimensions to be included in the current query.

Interestingness constraints. It would also be useful to determine what ranges of a particular variable or measures are considered to be particularly interesting and should be included in the query.

Rule constraints. It is also important to specify the specific rules which should be applied and used for a particular data mining query or application. One application of the constraint-based approach is in the Online Analytical Mining Architecture (OLAM) developed by Han, Lakshamanan, and Ng, 1999, and is designed to support the multidimensional and constraint based mining of databases and data warehouses.

PHENOMENAL DATA MINING

Phenomenal data mining is not a term for a data mining project that went extremely well. Instead, it focuses on the relationships between data and the phenomena which are inferred from the data . One example of this is that using receipts from cash supermarket purchases, it is possible to identify various aspects of the customers who are making these purchases. Some of these phenomena could include age, income, ethnicity, and purchasing habits. One aspect of phenomenal data mining, and in particular the goal to infer phenomena from data, is the need to have access to some facts about the relations between these data and their related phenomena. These could be included the program which examines data for phenomena, or also could be placed in a kind of knowledge base or database which can be drawn upon when doing the data mining. Part of the challenge in creating such a knowledge base involves the coding of common sense into a database, which has proved to be a difficult problem so far.

VI. Conclusions

In this paper I briefly reviewed the various data mining trends and applications from its inception to the future.This review puts focus on the hot and promising areas of data mining. Though very few areas are named here in this paper, yet they are those which are commonly forgotten. This paper provides a new perspective of a researcher regarding applications of data mining in social welfare.

References

[1]   Heikki, Mannila, "Data mining: machine learning, statistics, and databases", Statistics and Scientific Data Management, pp. 2-9. 1996.
[2]   Knowledge Discovery in Databases, AAAI Press / the MIT Press, Massachusetts Institute of Technology. ISBN 0–26256097–6. MIT1996.
[3]   Chakrabarti, van den Berg, and Dom. "Distributed Hypertext Resource Discovery through   Examples, "Proceedings of the 25thVLDB (International Conference on Very Large Data Bases), Edinburgh Scotland, 1999.
[4]   Han, J. and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2001.
[5]   Han, J., M. Kamber, and A. K. H. Tung, "Spatial Clustering Methods in Data Mining: A Survey", H. Miller and J. Han (eds.), Geographic Data Mining and Knowledge Discovery, Taylor and Francis, 2001.
[6]   Miller and J. Han (eds.), Geographic Data Mining and Knowledge Discovery, Taylor and Francis, 2001.
[7]   Kotsiantis, S., Kanellopoulos, D., Pintelas, P., "Multimedia mining", SEAS Transactions on   Systems, No 3, s. 3263-3268, 2005.
[8]   Huysmans, Baesens, Martens, Denys and Vanthienen, "New Trends in Data Mining", Tijdschrift voor Economie en Management, vol. L, 4, 2005.
[9]   Olfa Nasraoui and Maha Soliman, "Market-Based Profile Infrastructure: Giving Back to the User", Next Generation of Data Mining, Taylor and Francis, 2008.
[10]   Salmin, Sultana et al., "Ubiquitous Secretary: A Ubiquitous Computing Application Based on Web Services Architecture", InternationalJournal of Multimedia and Ubiquitous Engineering Vol. 4,   No. 4, October, 2009.
[11]   Jing He, "Advances in Data Mining: History and Future", Third international Symposium on Information Technology Application, 978-0-7695-3859-4/09 IEEE 2009 DOI 10.1109/IITA.2009.204.
[12]   M.S. Chen, J. Han, and P.S. Yu. "Data mining: An overview from database perspective", IEEE Transactions on Knowledge and Data Eng., 8(6):866-883, December 1999
[13]   Venkatadari M., Dr. Lokanataha C. Reddy, "A Review on Data Mining From Past to Future", International Journal of Computer Applications, pp.19-22, vol. 15, No. 7, Feb 2011.