Inf Technol Manag (2012) 13:297–310
DOI 10.1007/s10799-012-0123-z

Business intelligence in enterprise computing environment

Li Zeng · Ling Li · Lian Duan

Published online: 9 May 2012

Abstract Business intelligence (BI) is the process of gathering correct information in the correct format at the correct time; and delivering the results for decision-making purposes, or have a positive impact on business operations, tactics, and strategy in the enterprises. This paper is intended as a brief review of BI in an enterprise computing environment, with an emphasis on the algorithms and methods. The review points out the challenges to the broad and deep deployment of business intelligence systems, and provide proposals to make business intelligence more effective.

Keywords Business intelligence · Intelligent computing · Data mining

1 Introduction

The most common types of information systems in the enterprise computing environment are transaction processing systems (TPS), management information systems (MIS), decision support systems (DSS), executive information systems, enterprise information systems (EIS) or called enterprise systems (ES) (Enterprise Resource Planning, ERP) [9]. Together, these systems help enterprises to accomplish both routine and special tasks–from recording sales, to processing payrolls, to supporting decisions in various departments, to providing alternatives for business operations. However, as businesses continue to use these systems for a growing number of functions in today's competitive environment, most enterprises are facing challenges processing and analyzing huge amounts of data and turning it into useful information. They have too much detailed operational data, yet they cannot get the satisfactory answers they need from large volumes of information to enable them to react quickly to changing circumstances because of the scattered nature of the data.

For delivering the correct information in the correct format to the correct people at the correct time for decision-making purposes, Business Intelligence (BI) is presented. This set of techniques, technologies, tools, and solutions is designed to enable users to efficiently extract useful business information from huge amounts of data. The concept of BI was first introduced in 1990s, and referred to tools and technologies including data warehouses. Now business intelligence is regarded as a powerful solution, an extremely valuable tool, and a key approach to increasing the value of the enterprise. More and more business enterprises are deploying advanced business intelligence systems to enhance their competitiveness.

2 Information systems in enterprises

No one can deny the fact that information systems are widely used in business operations in more organizations than ever before. For example, information systems are used in finance or accounting divisions to forecast

L. Zeng · L. Duan
Institute of Computing Technology, Chinese Academy
of Sciences, Beijing 100080, China

L. Li
Old Dominion University, Norfolk, VA 23529, USA

L. Duan (✉)
New Jersey Institute of Technology, Newark, NJ 07102, USA
e-mail: duanlian.cn@gmail.com

revenues [49], to monitor business activity, to determine the best uses of funds, to manage financial resources [102, 103], to analyze investments, and to perform audits in order to make sure that the financial reports and documents are accurate. In the sales or marketing departments, information systems are used to check the inventory, to seek the best approaches for sales, and to set appropriate product prices. In general, the types of information systems used within organizations can be classified into: [1] TPS; [2] MIS; [3] DSS, OLAP (Online Analytical Processing), and BI; [4] EIS/ES; and [5] other special-purpose systems.

As one of the most fundamental information systems in many enterprises, TPS is used to handle the large volume of business transactions that occur daily within an organization. Information systems that not only support business processes and operations, but also help competitive tactical decision-making are MIS. MIS uses the data from a TPS to generate useful information for management at the tactical level.

Decision support systems are a class of computer-based information systems or knowledge-based systems that support strategic decision-making activities [18, 19, 53, 81, 83, 89, 94, 97]. DSS is a collection of people, procedures, data, and models used to support specific business decision-making tasks. Distributed DSS, intelligent DSS, and web-based DSS appeared through integrating with networking technology, artificial intelligence, and the Internet. DSS differs from an MIS in the support given to users, the decision emphasis, the development and approach, the system components, and the outputs. DSS marked the beginning of information system specifically designed for decision support in complex environments.

Business intelligence focuses not only on real-time data, but on real-time analysis that can be performed and instantaneously changes parameters of business processes. BI does not provide the same functionality as the traditional information systems, but instead operates on data that is extracted from operational data sources, and provides an effective means to propagate actions back into business processes and operations.

Enterprise systems, Enterprise information systems or Enterprise Resource Planning is a set of integrated programs that is capable of managing a company's vital business operations for an entire enterprise [80, 95, 104]. ERP is a term originally derived from Manufacturing Resource Planning (MRP). MRP evolved into ERP when routings and company's capacity planning activity became major part of the standard software activity [46]. ERP systems typically handle the manufacturing, logistics, inventory, invoicing, accounting, and distribution for a company. ES or EIS software commonly aids in the control of many business activities, such as production, quality management, inventory, sales, marketing, delivery, and human resources management. As a component of ES, a workflow system is often times a rule-based management software that directs, coordinates, and monitors the execution of an interrelated set of tasks arranged to form a business process [99]. More specifically, workflow is the operational aspect of a work procedure, i.e., how tasks are structured, who performs then and how they are performed, what their relative order is, and how they are synchronized. Its primary purpose is to provide users with tracking, routing and other capabilities designed to improve business processes. Graph-based formalisms such as Petri nets are used to model and analyze workflow issues [13, 55].

Other special-purpose information systems include expert systems, knowledge-based systems, virtual systems, e-business systems, and other systems such as social networks that will not be discussed in this paper [8, 20–22, 44, 47, 52, 54, 74, 76, 88, 93, 100]. Electronic Business (e-business) is also referred to as E-Commerce [23, 45, 60, 78]. It mainly consists of the distributing, buying, selling, marketing, and servicing of products over electronic systems such as the Internet or other computer networks. E-commerce involves business transactions executed electronically between entities such as business-to-business, business-to-consumers, and others. E-commerce offers more opportunities to enterprises by enabling them to market and sell at a low cost worldwide, thus enabling those small enterprises to enter the global market right from start-up. However, many large companies and major retailers also offer their products online.

The BI market has been continusely growing and reached 10.7 billion dollars through 2011 [17]. Meanwhile we can see the significant boosting BI research in the area of information technology. Figure 1 and Table 1 show the number of BI related papers published in *Information Technology and Management* journal for the period of

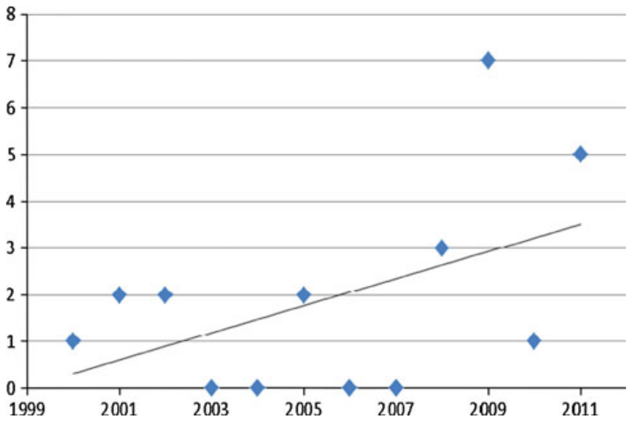


Fig. 1 The trend of BI related papers published in *IT&M*

Table 1 Papers related to BI published in *IT&M*

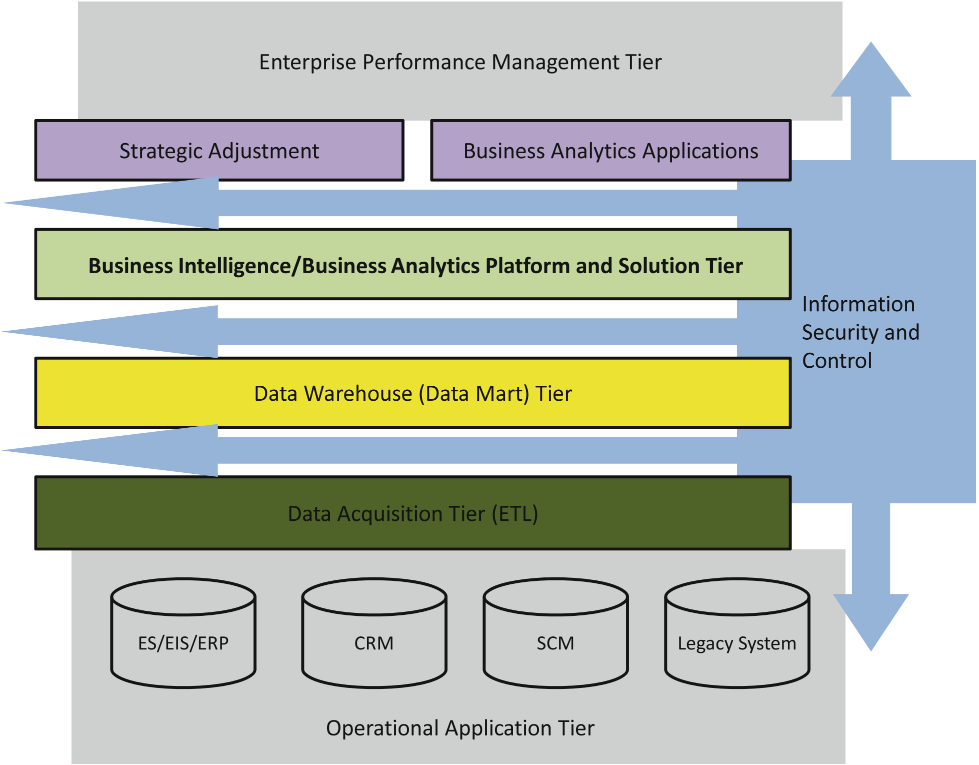
Boulicaut [5]
Agarwal et al. [3]
Nissen [58]
Datta and Thomas [12]
Rees and Koehler [63]
Piramuthu [61]
Sikora and Piramuthu [69]
Ko and Osei-Bryson [36]
Sugumaran et al. [70]
Zhou et al. [113]
Abbasi and Chen [1]
Jiang and Agarwal [31, 32]
Kim et al. [35]
Paruchuri et al. [59]
Raja and Tretter [62]
Takano et al. [72]
Tremblay et al. [77]
Boylu et al. [6]
Huang and Wan [30]
Liu et al. [51]
Shan et al. [64]
Wang et al. [83]
Wang et al. [86]

2000–2011 and the sources of these papers. This paper is intended as a brief review of BI with the emphasis on their relevancy to information technology and management.

3 Technical framework of BI in enterprise computing environment

It is normally accepted that the technology categories of a business intelligence system mainly encompass data warehouses, data marts, OLAP, and data mining. Figure 2 shows the architecture of BI systems in enterprise computing environment. Data warehouse is the fundamental infrastructure of business intelligence systems, and data mining is its core component—one that allows users to analyze data, identify data patterns, and detect trends. OLAP, on the other hand, is a set of front-end analyzing tools. Those who wish to construct benign enterprise intelligence computing environment should consider at least the following: the delivery of accurate, valid, integrated, and in-time data, and the means by which the data can be transformed into decision information. However, neither high quality data nor an effective means is easily acquired. An effective technical framework can be used to solve two issues above. The framework consists of an operational applications tier, a data acquisition tier, a data warehouse tier, a platforms-and-enterprise BI suites tier, and an extended corporate performance management tier. The operational applications tier includes, for example, systems such as legacy systems, CRM, ES, and SCM. Extraction, Transformation, and Loading belong in the data acquisition tier. Besides data warehousing, the data warehouse tier includes data marts and an operational data store. Data warehousing, OLAP, and data mining are three of the

Fig. 2 Architecture of BI systems in enterprise computing environment


Enterprise Performance Management Tier
Strategic Adjustment
Business Analytics Applications
Business Intelligence/Business Analytics Platform and Solution Tier
Information Security and Control
Data Warehouse (Data Mart) Tier
Data Acquisition Tier (ETL)
ES/EIS/ERP
CRM
SCM
Legacy System
Operational Application Tier

most significant technologies in the BI arena. A data warehouse can be defined as a large repository of historical data pertaining to an organization. OLAP refers to the technologies of performing complex analysis over the information stored in a data warehouse. The complexity of queries required to support OLAP applications makes it difficult to implement OLAP using standard relational database technology. Data mining is the process of identifying and interpreting patterns in data to solve a specific business problem.

A data warehouse can be viewed as a database that holds business information such as sales, products or other data about day-to-day operations, covering the aspects of the company's processes, production, and customers, or other data sources in the enterprises. This data is poured into a data warehouse on a regular schedule, and, after that, management can perform complex queries and analysis (normally data mining or OLAP) on the information without slowing down operational systems. The data warehouse provides users with a multi-dimensional view of the data they need to analyze business conditions. It is designed specifically to support managerial decision-making, rather than simply meeting the needs of transactions processing systems. A data warehouse typically starts out as a very large database, containing millions and even hundreds of millions of data records. To remain current and accurate, the data warehouse receives regular updates. The updating process must be efficient, automated or semi-automated, and as fast as possible owing to the colossal amount of data involved. It is common for a data warehouse to contain several years of current and historical data. Web warehousing [75] is the combination of data warehousing and the World Wide Web technology. The Internet has made it possible to apply web technology to traditional data warehousing, which has resulted in improved cost savings and productivity. According to Nemati et al. [57], the basic purpose of a data warehouse is to empower knowledge workers with information that allows them to make decisions based on a solid foundation of fact. However, only a fraction of information exists on computers, and the vast majority of a firm's intellectual assets such as institutional knowledge assets exist in the minds of people [100]. Therefore, the new generation of knowledge system is required to provide the capacity to capture, cleanse, store, organize, leverage, and disseminate not only data and information but also the knowledge of the firm. As an extension to the data warehouse model, the knowledge warehouse is likely to propose a new direction.

A data mart is a subset or a specialized version of a data warehouse. A data mart contains a subset of the data for a single aspect of company's business, instead of storing all of its enterprise data in one database, e.g., finance, inventory, or personnel. A data warehouse is used for summary data that can be accessed by an entire enterprise, whereas a data mart is helpful for small groups who want to access detailed data. Much like a data warehouse, the data marts typically can be deployed on less powerful hardware with small storage devices and helps business people to strategize based on analyses of past trends and experiences. However, the key difference between a data warehouse and a data mart is that the creation of a data mart is based on a specific, predefined need for a certain grouping and configuration of selected data. Since a data mart emphasizes easy access to relevant information, the star schema or multi-dimensional model is a fairly popular design choice, because it enables a relational database to emulate the structure and analytical functionality of a multi-dimensional database.

The notion of OLAP which was introduced in 1990s refers to the techniques of performing complex analysis of the information stored in a data warehouse. In general, OLAP applications are characterized by the rendering of enterprise data into multi-dimensional perspectives. This is achieved through complex queries that aggregate and consolidate data on a frequent basis, often using statistical formulae. For example, a supermarket may be interested in comparing its total sales for this year, or identifying sequences of 3 years or more when its sales have increased or decreased. It has been claimed that relational database technology is well suited to fulfilling the needs of OLAP. However, the major use of relational technology so far has been in traditional transaction management. Conversely, OLAP provides a quick approach to the answers to analytical queries that are dimensional in nature. OLAP can provide on-line analytical support, for which the relational model is ill-equipped. OLAP is part of the broader category of business intelligence, which also includes Extract Transform Load (ETL). As a matter of fact, readers can easily gauge the limitations of the relational model by trying to answer the queries in relational language such as SQL.

Data mining is the process of identifying and interpreting patterns in data to solve a specific business problem [10, 14, 24, 87]. It is an information analysis tool that involves the automated discovery of patterns and relationships in a data source. Data mining makes use of advanced statistical techniques and machine learning to discover facts in data warehouses or data marts, including in databases on the Internet. Unlike query tools, which require users to formulate and test hypotheses, data mining uses analysis tools to automatically generate a hypothesis about the patterns found in the data and then to predict future behavior. The objective is to discover patterns, trends, and rules from data warehouses to evaluate business operations, tactics, or strategies, which in turn should improve the competitiveness and profitability of enterprises, and should optimize business processes. BI vendors such as Oracle, SAS, and others are all incorporating data mining functionality into their products. Data mining strategies for BI include classification, time series analysis, clustering, association analysis, decision tree induction, support vector machine [106], *k*-nearest neighbor, genetic algorithms [33, 37, 38, 85], rough set [108], fuzzy sets [96, 109], *k*-means, case-based reasoning [25, 26, 34, 71, 90–92], feature space theory [41, 42], Bayesian networks [84], particle swarm optimization [83].

4 Business intelligence algorithms

The algorithms of data mining are the major components for business intelligence systems. Data mining strategies include classification, clustering, association analysis and many others.

4.1 Association rule

Association rule learning is also known as market basket analysis. Market basket analysis is used to determine those items most likely to be purchased by a customer during a shopping experience. The questions like "Supposing a customer purchases product A. How likely is the customer to purchase product B?" or "What kinds of items is the customer likely to purchase together?" are answered by association-finding algorithms. The output of the market basket analysis is generally a set of associations about customers' purchasing behavior. These associations are given in the form of a special set of rules known as association rules. Association rules are used to help to determine appropriate product marketing strategies. Association rules are of the form wherein a set of *n* items appear in a group along with a set of *m* items in the same group. For example, association rules can help us determine that if saving and checking accounts are owned by a customer, the customer will own a certificate of deposit with a certain frequency. While association rules do not warrant inferences of causality, they may point to relationships among items or events that could be studied further using more appropriate analytical techniques to determine the structure and nature of causalities that may exist. Unlike traditional classification, association rule generators allow the consequence of a rule to contain one or several attribute values, whereas traditional classification rules usually limit the consequent of a rule to a single attribute. In addition, using an association rule generator, an attribute may appear as both precondition and consequent of different rules in traditional classification. However, when attributes are present after generating association rules, this process becomes unreasonable, owing to large number of possible conditions for the consequent of each rule.

Some candidate-generation-and-test algorithms such as the Apriori algorithm [2] have been developed to generate association rules efficiently. This influential algorithm is used to find or mine frequent item sets which have attribute value combinations that meet a specified coverage requirement. Those attribute value combinations that do not meet the user's requirements are discarded. By this, the rule generation process can be completed in a reasonable amount of time. Apriori association rule generation is often a two-step process. The first step is to generate item sets, so that the second step can use the generated item sets to create a set of association rules.

However, candidate-generation-and-test algorithms suffer from both the generation of huge numbers of candidates and the scanning of the database time and time again. Thus, another approach called pattern-growth has been proposed. FP-Tree algorithm is such a method, used to mine the complete set of frequent item sets but without candidate generation. It does not need to generate a huge number of candidate sets, but retains the item set association information for the sake of mining separately; it scans the data set only a few times. As a result, in most cases, algorithms based on the pattern-growth approach find frequent patterns faster than those based on the candidate-generation-and-test approach. Hirate proposed a new mining algorithm, called "TF2P-growth" [27], which does not require any thresholds. This algorithm mines patterns in the descending order of their support values without any thresholds and returns frequent patterns to users sequentially, with short response time. Contrast set learning [56] is another form of associative learning. Contrast set learners use rules that meaningfully differ in their distribution across subsets.

Association rules are particularly popular because of their ability to find relationships in large databases without having the restriction of having to choose a single dependent variable. However, it is still important to minimize the work required by an association rule algorithm since volumes of data are often stored for market basket analysis.

4.2 Classification and prediction

Classification algorithm is simply a model for predicting a categorical variable that assumes one of a predetermined set of values. These values can be either nominal or ordinal, though ordinal variables are typically treated the same way as nominal ones in these models. When a problem is easy to classify and its boundary function is more complicated than it needs to be, the boundary is likely over-fitting. Analogously, when a problem is hard and the classifier is not powerful enough, the boundary becomes under-fitting. Classification describes the assignment of data records into predefined categories and discovers the relationship between the other variables and the target category. When a new record is inputted, the classifier determines the category and the probability that the record belongs to. Examples of classification algorithms include: linear classifiers (e.g. Fisher's linear discriminant, logistic regression, naive Bayesian classifier), quadratic classifiers, *k*-nearest neighbor, boosting, decision trees, neural networks [11, 39, 40, 67, 82, 110–112, 114], Bayesian networks, support vector machines, hidden Markov models, and so on.

However, no classification method stands out over the others with regards to all data types and domains. Empirical comparisons of classification methods were introduced by Lim [50] and Shavlik [65]. Classification and prediction methods can be compared and evaluated according to the several kinds of criteria [24]: predictive accuracy, robustness, scalability, and interpretability. Predictive accuracy refers to the ability of the model to correctly predict the target category. Robustness indicates the ability of the model to make correct predictions given noisy data or data with missing values. Scalability refers to the ability to construct the model efficiently given large amounts of data. Interpretability denotes the ability of interpretation or visualization provided by the classified model.

The decision tree model is a flow-chart-like structure, in which leaves represent classifications, each inner node denotes a test on an attribute, and branches represent conjunctions of features that lead to those classifications. It is the ability of decision trees not only to predict the value of a categorical variable, but also to directly use categorical variables as input or predictor variables. This is perhaps the decision tree's single greatest advantage. Decision trees are by their very nature well-suited to deal with large numbers of input variables, to handle a mixture of data types, and to handle data that is not homogeneous, i.e., whose variables do not have the same interrelationships throughout the data space. They also provide insight into the structure of the data space and the meaning of a model, a result at times as important as the accuracy of a model. It should be noted that a variation of decision trees called regression trees can be used to build regression models rather than classification models, enjoying the same benefits just described. The basic algorithm for decision tree induction is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner [24]. This algorithm is known as a version of ID3 [7], a well-known decision tree induction algorithm. C4.5, a later version of the ID3 algorithm, uses the training samples to estimate the accuracy of each rule. Since this use results in an optimistic estimate of rule accuracy, C4.5 uses a pessimistic estimate to compensate for the bias. Alternatively, a set of test samples independent from the training set can be used to estimate rule accuracy.

Novel in the field of data mining, support vector machine (SVM) and kernel skill have been successfully applied to a variety of domains. SVM is a promising method for classification and regression analysis due to its solid mathematical foundations, which include two desirable properties: margin maximization and nonlinear classification using kernels. However, despite these two distinguishing properties, SVM is usually not chosen for large-scale data mining problems because its training complexity is highly dependent on the size of the data set. Unlike traditional pattern recognition and machine learning, real-world data mining applications often involve huge numbers of data records. Thus, it is too expensive to perform multiple scans on the entire data set, and it is also unfeasible to put all of the data set into memory.

But SVM is good at supervised learning that tries to maximize the generalization by maximizing the margin while supporting nonlinear separation using advanced kernels, by which SVM tries to avoid over-fitting and under-fitting. The margin in SVM denotes the distance from the boundary to the closest data in the feature space. In SVM, the problem of computing a margin maximized boundary function is specified by the following quadratic programming (QP) problems [79]:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j k(x_i \cdot x_j) - \sum_{j=1}^{l}\alpha_j \\ s.t. \quad & \sum_{i=1}^{l} y_i \alpha_i = 0 \\ & \forall i:\ 0 \le \alpha_i \le C, i = 1, \ldots l \end{aligned}$$

where l denotes the number of training data, and α denotes a vector of l variables, and each α_i corresponds to training data (x_i, y_i). C is the soft margin parameter, controlling the influence of the outliers (or noise) in training data. The kernel k (x_i, y_i) for linear boundary function is (x_i, y_j), a scalar product of two data points. The nonlinear transformation of the feature space is performed by replacing k (x_i, y_i) with an advanced kernel, such as the polynomial kernel $(x^T x_i + 1)^p$or the RBF kernel $\exp\left(-\frac{|x-x_i|^2}{\sigma^2}\right)$. The use of an advanced kernel is an attractive computational short-cut. An advanced kernel is a function that operates on the input data but has the effect of computing the scalar product of their images in what is usually a much higher dimensional feature space, or even an infinite dimensional space, which allows one to work implicitly with hyper-planes in highly complex spaces.

However, as mentioned before, most of the existing support vector machines are not feasible to run very large data sets due to their high complexity on the data size or to the frequent accesses on such large data sets causing expensive I/O operations. Yu [105] presents a novel method, called Clustering-Based SVM (CB-SVM), which maximizes the SVM performance for very large data sets given a limited amount of resource, e.g., memory. CB-SVM applies a hierarchical micro-clustering algorithm that scans the entire data set only once to provide an SVM with high quality samples. These samples carry statistical summaries of the data and maximize the benefit of learning. The analyses show that the training complexity of CB-SVM is quadratically dependent on the number of support vectors, which is usually much fewer than that of the entire data set. The experiments on synthetic and real-world data sets show that CB-SVM is highly scalable for very large data sets and very accurate in terms of classification. However, CB-SVM is currently limited to the usage of linear kernels since the hierarchical micro-clusters would not be isomorphic to a new high-dimensional feature space once the space is transformed by a nonlinear kernel. That is, the statistical summaries of data such as radius and distances computed in the input space will not be preserved in the transformed feature space. Constructing effective indexing structures for a non-linear kernel is an interesting direction for future work since it has high practical value, especially for classifying large business data sets. Hong and Weiss [28] reviewed the key theoretical developments in PAC and statistical learning theory that have led to the development of support vector machines and to the use of multiple models for increased predictive accuracy. Training support vector machines involves a huge optimization problem. Boley and Cao [4] proposed an algorithm called Cluster SVM that accelerates the training process by exploiting the distributional properties of the training data.

In order to improve the performances of traditional SVM on a dataset with unbalanced class distribution, an improved SVM was presented. Genetic algorithm-SVM (GA-SVM) was constructed by combining the genetic algorithm and the simple support vector machine. The parameters of SVM were coded into chromosomes with gray coding strategy. The results by Huang et al. [29] indicate that GA-SVM can gain higher classification accurately and with a faster learning speed, and that it works well with a faster learning speed on a perfectly constructed dataset.

The key differences between prediction and classification are that we use prediction to predict a continuous value, rather than a categorical label. The prediction of continuous values can be modeled by statistical techniques of regression. For example, we might like to develop a model to predict the salary of college graduates with 10 years of work experience, or the potential sales of a new product given its price. Many problems can be solved by linear regression, and even more can be tackled by applying transformations to the variables so that a nonlinear problem can be converted to a linear one. Regression models include linear and multiple regressions, and nonlinear regression. Other regression models include generalized linear models and log-linear models. Generalized linear models represent the theoretical foundation on which linear regression can be applied to the modeling of categorical response variables. Logistic regression and Poisson regression are two examples of generalized linear models. Log-linear models approximate discrete multidimensional probability distributions. They may be used to estimate the probability values associated with data cube cells. Some commercial BI software packages are devoted to solving regression problems.

4.3 Clustering analysis

Clustering analysis is a common technique used in many fields, including machine learning, data mining, pattern recognition, image analysis, bioinformatics, and market research [15, 16, 43]. Clustering is a typical form of unsupervised learning which classifies similar objects into different groups, or more precisely, partition a data set into clusters, so that the data in each subset ideally share some common trait. In other words, clustering analysis is “the process of grouping the data into classes or clusters so that objects within a cluster have high similarity in comparison to one another, but are very dissimilar to objects in other cluster” [24]. Cluster analysis is a statistical process used to identify homogeneous groups of data objects. By clustering, one can identify dense and sparse regions and therefore, can discover overall distribution patterns and interesting correlations among data attributes. Due to the massive sizes of enterprise data today, implementation of any clustering algorithms must be scalable to complete analysis within a reasonable amount of time. Analogously, most clustering statistical algorithms do not work well with large databases due to memory limitations and to the execution times required, as well as to classification algorithms.

In business applications, clustering helps marketers discover distinct groups and characterize customer groups based on purchasing patterns. As a data mining function, cluster analysis can be used as a stand-alone tool to gain insight into the distribution of data, to observe the characteristics of each cluster, and to focus on a set of clusters for further analysis. Alternatively, it may serve as a preprocessing step for other algorithms, such as characterization and classification, which then operate on the detected clusters. As a branch of statistics, cluster analysis has been studied extensively for many years, focusing mainly on distance-based cluster analysis. Cluster analysis tools based on *k*-means, and several other methods have also been built into many statistical analysis software packages or systems. In machine learning, clustering is an example of unsupervised learning. As opposed to classification, clustering and unsupervised learning do not rely on predefined classes and class-labeled training examples. For this reason, clustering is a form of learning by observation, rather than learning by examples. In conceptual clustering, a group of objects forms a class only if it is describable by a concept. This differs from conventional clustering; which measures similarity based on geometric distance. In general, major clustering analysis methods can be classified into several categories [24] such as partitioning methods, hierarchical methods, density-based methods, grid-based methods, and model-based methods.

The *k*-means algorithm takes the input parameter *k*, and partitions a set of *n* objects into *k* clusters so that intra-cluster similarity is high but the inter-cluster similarity is low. Cluster similarity is measured in regard to the mean value of the objects in a cluster, which can be viewed as the cluster's center of gravity. The *k*-means algorithm, however, can be applied only when the mean of a cluster is defined, but this may not always be the case. Similar to the *k*-means algorithm, the *k*-medoid algorithm represents each cluster by the center of the objects of the cluster located near the center. Ng and Han proposed the CLARANS algorithm [66], which is an improved *k*-modoid method.

Wavelet-based Clustering (or Wave Cluster) [66] can be efficiently applied to detect clusters of arbitrary shape. A good clustering analysis approach should be insensitive to the noise, the outliers, and the input order of data. What is more, it should be efficiently used by both low dimensional and high dimensional large datasets. Wave Cluster is a grid-based and density-based algorithm, which uses the multi-resolution property of wavelet transform. It can handle large datasets efficiently and identify arbitrary shaped clusters at varying degrees of detail; furthermore, it can efficiently perform on very large databases. This approach meets most of the desirable properties of a good clustering technique as mentioned above. Here is an example of arbitrary shape data distribution. Figure 3 presents the clustering result produced by Wave Cluster. From this, it is evident that WaveCluster is powerful in handling any type of sophisticated patterns and removing noise.

4.4 Time-related analysis and mining

Time-series databases are popular in many applications, such as studying daily fluctuations of the stock market, traces of a dynamic production processes, and the like. Time-related analysis and mining comprises of mining techniques that are applied to the analysis of time-ordered data records. These data mining techniques attempt to detect similar sequences or subsequences in the ordered data.

Time-series databases and sequence databases include two typical time-related data. A time-series database consists of sequences of values or events changing with time. The values are typically measured at equal time intervals. A time-series database is also a sequence database. However, a sequence database is any database that consists of sequences of ordered events, with or without concrete notions of time. Trend analysis, similarity search, and the mining of sequential patterns and periodic patterns are several important aspects of time-related analysis and mining.

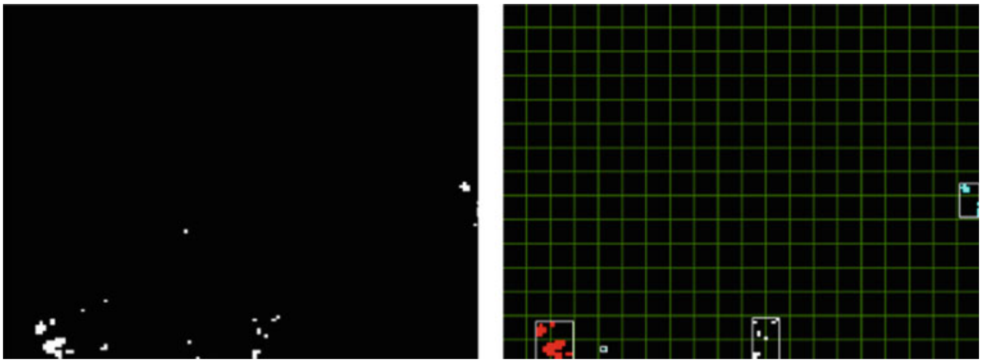
There are four major components or movements that are used to characterize time-series data [24]: long-term or trend movements, cyclic movements or cyclic variations, seasonal movements or seasonal variations, and irregular or random movements. Similarity searches in time-series analysis are typically helpful for the analysis of financial markets (like stock data analysis). Sequential pattern mining is the discovery of frequent patterns related to time or other sequences. Since many business transactions, telecommunications records, and production process are time sequenced data, sequential pattern mining is useful in the analysis of such data for understanding marketing, developing customer retention strategies, and so on.

5 Analysis process of business intelligence

Business intelligence generally follows a continuous cycle that begins with a precise description of the business issue. With the data and the mining techniques selected, data miners will conduct mining and evaluate the results. It is likely that further iterations of the data selection and an application of different mining techniques may be necessary to provide a satisfied solution. If the mining effort effectively addresses the original business problem, it becomes necessary to deploy the results so that the work leads to concrete actions taken.

As users use standard statistical techniques or reporting tools to explore the data in databases, they are making a hypothesis about a business issue that they were addressing and then attempting to prove or disprove their hypothesis by looking for data to support or contradict their hypothesis.

Data mining uses an alternative approach that begins with the premise that we do not know what patterns of data exist. Many business and research fields have been proven to be excellent candidates for data mining; for example, banking, insurance, retail, telecommunications, manufacturing, pharmaceuticals, biotechnology and the like, where significant benefits have also been derived. Well-known



(a) Original space

(b) Wave Cluster results

Fig. 3 **a** Original space. **b** WaveCluster results

applications are customer profiling in retail, loan delinquency and fraud detection in banking and finance, customer retention in telecoms, and patient profiling in health care. Data mining is about the discovery of patterns and relationships in data. All of the different applications are using the same data mining concepts and applying them in different ways. That is not to say that data mining is magic and omnipotent. We still have to understand the overall business process. The process starts with defining the business problem that we want to solve. Then a mining expert can concentrate on the right solution. This involves gathering relevant data and discovering hidden patterns using mining algorithms. Once the analysis is complete, the new knowledge extracted from the data can be put into action. The process is depicted in Fig. 4.

5.1 Step1: Create a precise description of the business issue

The first step is to identify the business issue that we want to address and then determine how the business issue can be translated into a question, or set of questions, that data mining can tackle. As we are formulating the business issue, we need to also think about whether we have access to the right data. It is important to recognize that the data we hold may not contain the information that is required to answer the question.

5.2 Step2: Map the business issue to model

When the data is being used routinely to support a specific business application, the data and meta data together form what we call data model that supports the application. It is a complex task to define data models for any application. The challenge is that very often we are not sure at the outset which variables are important and therefore exactly what is required. Mapping the business issue to a data model can therefore become a time-consuming activity. The alternative is to use common data models that have been developed to solve business issues similar to the ones we are trying to address. While these types of models may not initially provide us with all of the information we require, they are usually designed to be extendable to include additional variables.

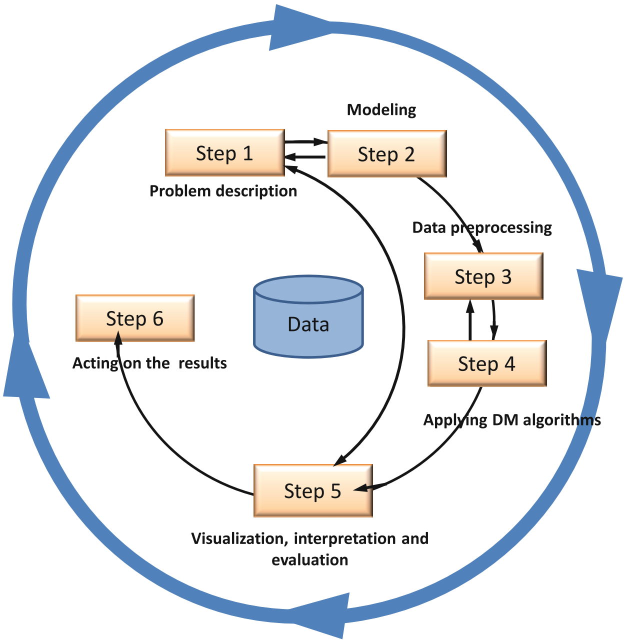


Fig. 4 Analysis process in applying business intelligence

The main advantage of using a common data model is that it provides us with a way of quickly seeing how data mining can be used.

5.3 Step3: Data preprocessing

Most data preprocessing comes in the form of data cleaning, which involves dealing with missing information. Ideally, the majority of data preprocessing takes place before data is permanently stored in a structure such as a data warehouse. Common concerns with noisy data, often represented by random error, include incorrect attribute values, duplicate records, and data smoothing. In very large datasets, noise can come in many shapes and forms. Though some automated graphical tools may assist with data cleaning, the responsibility for data transition still lies in the hands of the data warehouse specialist. For example, a numeric value of -1 for an attribute such as weight or blood pressure is an obvious error. Such errors often occur when data is missing and when default values are assigned to fill in for missing items. If the dataset is large and only a few incorrect values exist, finding such errors can be difficult. Some data analysis tools allow the user to input a valid range of values for numerical data. Data smoothing is both a data cleaning and a data transformation process. Several data smoothing techniques attempt to reduce the number of values for a numeric attribute. Some classifiers, such as neural networks, use functions that perform data smoothing during the classification process. Another data smoothing process which takes place prior to classification is external data smoothing. Rounding and computing mean values are two simple externals. Mean value smoothing is appropriate when we wish to use a classifier that does not support numerical data. In this case, all numerical attribute values are replaced by a corresponding class mean. Another common data smoothing technique attempts to find the possibility of removing atypical instances from the dataset. In most cases, missing attribute values indicate lost information.

For example, a missing value for the attribute “age” certainly indicates a data item that exists but is unaccounted for. However, a missing value for “salary” may be taken as an un-entered data item, but it could also indicate an individual who is unemployed. Some data mining techniques are able to deal directly with missing values. However, many classifiers require all attributes to contain a value. Possible options for dealing with missing data before the data is presented to a data mining algorithm include: discarding records with missing values, replacing missing values with the class mean for real-valued data, and replacing missing attribute values with the values found within other highly similar instances.

5.4 Step4: Apply matched data mining algorithms

When it comes to solving a particular problem, we have several techniques to choose from. The question becomes, how do we know which data mining technique to use?

We can determine an appropriate data mining technique given a set of data containing attributes and values mined alongside information about the nature of the data and the problem to be solved. For a given business issue, the step of selecting mining techniques or algorithms not only includes defining the appropriate technique or mix of techniques to use, but also the way in which the techniques must be applied. Data mining techniques can be generally divided into these two broad categories: discovery mining and predictive mining. Discovery mining applies to a range of techniques whose primary objective is to find patterns inside the business data without any prior knowledge of what patterns exist. Clustering and sequence analysis are typical examples of discovery mining techniques. Predictive data mining is applied to a range of techniques that find relationships between a specific variable, called the target variable, and the other variables in the data. Classification and regression are examples of predictive mining techniques.

5.5 Step5: Visualization, interpretation, and evaluation

Visualization, interpretation, and evaluation are to determine whether a learning model is acceptable and robust enough to be applied to problems outside the test environment [48]. If acceptable results are achieved, it is at the stage where acquired knowledge is translated into terms understandable by users. It should be noted that, vendors and products like IBM Intelligent Miner Visualization, NCR Teradata, and MSMiner [68, 98] are extremely good for this.

Performing any type of data mining can provide a wealth of information that can be difficult to interpret. This interpretation step often requires assistance from a business expert who can translate the mining results back into the business context, since it is unlikely that the business analyst will be a mining expert. Therefore, it is important that the results be presented in such a way that they are relatively easy to interpret. Users need a range of tools that enable them to visualize the results in order to provide the necessary statistical information necessary for facilitating the interpretation.

5.6 Step6: Act on the analysis results and reach goals

We create mathematical representations of the data and call them models. They contain the rules and patterns found by the mining algorithm. These models provide us with a deeper insight into our business; and can be deployed or used by other business processes. A number of possible actions may result from successful application of the knowledge discovery process. To apply what has been learned or mined is the ultimate goal of business intelligence. The deployment of the results of data analysis and mining is possibly the most important of all.

6 Making BI solutions more effective

As mentioned above, business intelligence might give users the ability to gain insights into business or organization by helping them understand the company's information assets. These data assets can include customer data; supply chain data; manufacturing, sales, and marketing data as well as any other sources of data critical to operation. It also allows users to integrate disparate data sources into a single coherent framework for real-time reporting and detailed analysis. Here are some trends which are dramatically driving the market need for better business intelligence tools: daily rising data volumes, geographically dispersed users, and the existing tools that are difficult to use. The existing business intelligence systems still lack the maturity and breadth of deployment which is need to meet business demands. Broader deployment of business intelligence systems throughout the enterprises will only occur if users can learn an application, deploy it, and manage it effectively. These are some of the reasons for the difficulties in the broad delivery of business intelligence systems in every enterprise.

Very often, business intelligence systems take a long time to install, build, and deploy. The average implementation time for some larger BI solutions even reaches about 6 months. Requirements and budgets often change after a long installation and implementation cycle. Many business intelligence applications are still difficult to use. A majority of BI projects are focused on the implementation, and adequate user training is often overlooked. As a result, a lack of end user acceptance is an critical factor that hampers business intelligence systems. In most cases, business intelligence systems have actually increased the workload, even though they were originally conceived as a means to relieve workload through intuitive reporting and analysis. This can eventually limit the wider deployment of BI throughout the enterprise. Furthermore, the cost and benefit can also be questioned. Often, after the completion of a rather lengthy and costly implementation, demands have changed. Once the applications cannot demonstrate a return on investment in time, or once few benefits are realized, the end users are likely to be disenchanted with business intelligence.

True "enterprise-wide" intelligence would definitely ensure that users who need access and analysis of information to support a business process or a decision will have a powerful yet intuitive solution or platform at their fingertips. Thus, business intelligence solutions should be broadly distributed to all users who need access to information. The more people using a technology, the more valuable the technology becomes. Once the business intelligence system is easier to use and to understand, and allows the user to evaluate alternatives, to draw conclusions, and to make decisions, it will be more broadly implemented. The tools and techniques used to access and analyze the information must be powerful, yet easy both to learn and to use. This can only happen if the information is easy to understand, is timely, and is relevant to the user. In addition, the user should have access to the resources necessary to perform his or her job. For true insight and effectiveness, understanding of the data across boundaries will help the user make business more productive. Analysis tools of business intelligence should be powerful, yet simple for users to learn, to deploy, and to maintain. These solutions need to be more flexible and adaptable to changes in the on-demand competitive business environment.

7 Summary

Although we are convinced that business intelligence is the proper road for business enterprises to follow, it will still be quite a journey. The work described in this paper is only a part of related research. Business intelligence systems should provide not only the capabilities to analyze what has occurred, but more importantly to tell the users what is going to happen in their enterprises, by using intelligent information processing techniques [107]. As ubiquitous computing technology increases the demand for enterprise computing, it is expected that this trend will apply to BI as well [34]. BI will continue to embrace cutting-edge technology and techniques, and will open new applications that will impact industrial sectors [73, 101].

Acknowledgments This work was partially supported by the Chinese Academy of Sciences under the Grant 20040402 (Distinguished Overseas Scholar Grant), Changjiang Scholar Program of the Ministry of Education of China, National Natural Science Foundation of China under the Grant 71132008, and US National Science Foundation under the Grant 1044845.

References

1. Abbasi A, Chen H (2009) A comparison of fraud cues and classification methods for fake escrow website detection. Inf Technol Manage 10(2–3):83–101
2. Aflori C, Craus M (2007) Grid implementation of the Apriori algorithm. Adv Eng Softw 38(5):295–300
3. Agarwal A, Davis J, Ward T (2001) Supporting ordinal four-state classification decisions using neural networks. Inf Technol Manage 2(1):5–26
4. Boley D, Cao D (2004) Training support vector machine using adaptive clustering. In: Proceddings of the SIAM international conference on data mining, Lake Buena Vista, FL, April 2004, pp 126–137
5. Boulicaut J (2000) A KDD framework to support database audit. Inf Technol Manage 1(3):195–207
6. Boylu F, Aytug H, Koehler G (2010) Data mining with agent gaming. Inf Technol Manage 11(1):1–6
7. Brown D, Corruble V, Pittard C (1993) A comparison of decision tree classifiers with back-propagation neural networks for multimodal classification problems. Pattern Recogn 26(6):953–961
8. Cao X, Yang F (2011) Measuring the performance of Internet companies using a two-stage data envelopment analysis model. Enterp Inf Syst 5(2):207–217

9. Capozucca A, Guelfi N (2010) Modelling dependable collaborative time-constrained business processes. Enterp Inf Syst 4(2): 153–214
10. Chiang D, Lin C, Chen M (2011) The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres. Enterp Inf Syst 5(2):219–234
11. Curram S, Mingers J (1994) Neural networks, decision tree induction and discriminant analysis: an empirical comparison. J Oper Res Soc 45(4):440–450
12. Datta A, Thomas H (2002) Querying compressed data in data warehouses. Inf Technol Manage 3(4):353–386
13. Du Y, Qi L, Zhou M (2011) A vector matching method for applying logic Petri nets. Enterp Inf Syst 5(4):449–468
14. Duan L, Street W, Xu E (2011) Healthcare information systems: data mining methods in the creation of a clinical recommender system. Enterp Inf Syst 5(2):169–181
15. Duan L, Xu L, Guo F, Lee J, Yan B (2007) A local-density based spatial clustering algorithm with noise. Inf Syst 32(7): 978–986
16. Duan L, Xu L, Liu Y, Lee J (2009) Cluster-based outlier detection. Ann Oper Res 168(1):151–168
17. Duan L, Xu L (2012) Business intelligence for enterprise systems: a survey. IEEE Trans Industr Inform, published online. doi:10.1109/TII.2012.2188804
18. Feng S, Xu L (1999) Decision support for fuzzy comprehensive evaluation of urban development. Fuzzy Sets Syst 105(1):1–12
19. Feng S, Xu L (1999) An intelligent decision support system for fuzzy comprehensive evaluation of urban development. Expert Syst Appl 16(1):21–32
20. Fu C, Zhang G, Yang J, Liu X (2011) Study on the contract characteristics of internet architecture. Enterp Inf Syst 5(4): 495–513
21. Gong Z, Muyeba M, Guo J (2010) Business information query expansion through semantic network. Enterp Inf Syst 4(1): 1–22
22. Governatori G, Iannella R (2011) A modelling and reasoning framework for social network policies. Enterp Inf Syst 5(1): 145–167
23. Guo J, Xu L, Gong Z, Che C, Chaudhry S (2012) Semantic inference on heterogeneous e-marketplace activities. IEEE Trans SMC Part A Syst Hum 42(2):316–330
24. Han J, Kamber M (2006) Data ming: concepts and techniques. Morgan Kaufmann, Los Altos, CA
25. He W, Wang F, Means T, Xu L (2009) Insight into interface design of web-based case-based reasoning retrieval systems. Expert Syst Appl 36(3):7280–7287
26. He W, Xu L, Means T, Wang P (2009) Integrating web 2.0 with the case-based reasoning cycle: a systems approach. Syst Res Behav Sci 26(6):717–728
27. Hirate Y, Iwahashi E, Yamana H (2004) TF2P-growth: an efficient algorithm for mining frequent patterns without any thresholds. In: Proceedings of IEEE ICDM 2004 workshop on alternative techniques for data mining and knowledge discovery, November 1st 2004, Brighton, UK
28. Hong S, Weiss S (2001) Advances in predictive models for data mining. Pattern Recogn Lett 22:55–61
29. Huang J, Ma L, Qian J (2004) Improved support vector machine for multi-class classification problems. J Zhejiang Univ (Eng Sci) 38(12):1633–1636
30. Huang K, Wan S (2011) Application of enhanced cluster validity index function to automatic stock portfolio selection system. Inf Technol Manage 12(3):213–228
31. Jiang W, Agarwal A (2009) Special issue devoted to papers presented at the second INFORMS workshop on artificial intelligence and data mining, Seattle, November 3rd 2007. Inf Technol Manag 10(1):39
32. Jiang W, Agarwal A (2009) Special issue devoted to papers presented at the second INFORMS workshop on artificial intelligence and data mining, Seattle, November 3rd 2007. Inf Technol Manag 10(4):221
33. Jiang Y, Xu L, Wang H, Wang H (2009) Influencing factors for predicting financial performance based on genetic algorithms. Syst Res Behav Sci 26(6):661–673
34. Kakousis K, Paspallis N, Papadopoulos G (2010) A survey of software adpatation on mobile and ubiquitous computing. Enterp Inf Syst 4(4):355–389
35. Kim S, Huo X, Tsui K (2009) A finite-sample simulation study of cross validation in tree-based models. Inf Technol Manage 10(4):223–233
36. Ko M, Osei-Bryson K (2008) Reexamining the impact of information technology investment on productivity using regression tree and multivariate adaptive regression splines (MARS). Inf Technol Manage 9(4):285–299
37. Li F, Xu L, Jin C, Wang H (2011) Intelligent bionic genetic algorithm (IB-GA) and its convergence. Expert Syst Appl 38(7):8804–8811
38. Li F, Xu L, Jin C, Wang H (2011) Structure of multi-stage composite genetic algorithm (MSC-GA) and its performance. Expert Syst Appl 38(7):8929–8937
39. Li H, Li L (1999) Representing diverse mathematical problems using neural networks in hybrid intelligent systems. Expert Syst 16(4):262–272
40. Li H, Xu L (2000) A neural network representation of linear programming. Eur J Oper Res 124(2):224–234
41. Li H, Xu L (2001) Feature space theory—a mathematical foundation for data mining. Knowl-Based Syst 14(5–6):253–257
42. Li H, Xu L, Wang J, Mo Z (2003) Feature space theory in data mining: transformations between extensions and intensions in knowledge representation. Expert Syst 20(2):60–71
43. Li J, Wang K, Xu L (2009) Chameleon based on clustering feature tree and its application in customer segmentation. Ann Oper Res 168(1):225–245
44. Li L (1999) Knowledge-based problem solving: an approach to health assessment. Expert Syst Appl 16(1):33–42
45. Li L (2011) Introduction: advances in e-business engineering. Inf Technol Manage 12(2):49–50
46. Li L, Chaudhry S, Chaudhry P, Wang Y (2001) An evaluation of acquiring and implementing a manufacturing resource planning system. Prod Invent Manag J 42(3–4):1–8
47. Li L, Warfield J, Guo S, Guo W, Qi J (2007) Advances in intelligent information processing. Inf Syst 32(7):941–943
48. Li T, Feng S, Li L (2001) Information visualization for intelligent decision support systems. Knowl-Based Syst 14(5–6): 259–262
49. Liang L (2008) Earnings forecasts in enterprise information systems environment. Enterp Inf Syst 2(1):1–19
50. Lim T, Loh W, Shih Y (2000) A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. Mach Learn 40(3):203–228
51. Liu B, Cao S, He W (2011) Distributed data mining for e-business. Inf Technol Manage 12(2):67–79
52. Liu D, Deters R, Zhang W (2010) Architectural design for resilience. Enterp Inf Syst 4(2):137–152
53. Luo J, Xu L, Jamont J, Zeng L, Shi Z (2007) A flood decision support system on agent grid: method and implementation. Enterp Inf Syst 1(1):49–68
54. Lykourentzou I, Dagka F, Papadaki K, Lepouras G, Vassilakis C (2012) Wikis in enterprise settings: a survey. Enterp Inf Syst 6(1):1–53
55. Ma J, Wang K, Xu L (2011) Modelling and analysis of workflow for lean supply chains. Enterp Inf Syst 5(4):423–447

56. Menzies T, Hu Y (2003) Data mining for very busy people. IEEE Comput 36(11):22–29
57. Nemati H, Steiger D, Iyer L, Herschel R (2002) Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. Decis Support Syst 33(2):143–161
58. Nissen M (2001) Agent-based supply chain integration. Inf Technol Manage 2(3):289–312
59. Paruchuri P, Pearce J, Marecki J, Tambe M (2009) Coordinating randomized policies for increasing security of agent systems. Inf Technol Manage 10(1):67–79
60. Piao C, Han X, Wu H (2010) Research on e-commerce transaction networks using multi-agent modelling and open application programming interface. Enterp Inf Syst 4(3):329–353
61. Piramuthu S (2005) Feature selection for reduction of tabular knowledge-based systems. Inf Technol Manage 6(4):351–362
62. Raja U, Tretter M (2009) Antecedents of open source software defects: a data mining approach to model formulation, validation and testing. Inf Technol Manage 10(4):235–251
63. Rees J, Koehler G (2002) Evolution in groups: a genetic algorithm approach to group decision support systems. Inf Technol Manage 3(3):213–227
64. Shan S, Wang L, Wang J, Hao Y, Hua F (2011) Research on e-Government evaluation model based on the principal component analysis. Inf Technol Manage 12(2):173–185
65. Shavlik J, Mooney R, Towell G (1991) Symbolic and neural learning algorithms: an experimental comparison. Mach Learn 6(2):111–143
66. Sheikholeslami G, Chatterjee S, Zhang A (1998) WaveCluster : a multi-resolution clustering approach for very large spatial databases. In: Proceedings of the international conference on very large data bases 1998, vol M, Issue 24. Publisher: Citeseer, pp 428–439
67. Shi S, Xu L, Liu B (1999) Improving the accuracy of nonlinear combined forecasting using neural networks. Expert Syst Appl 16(1):49–54
68. Shi Z, Huang Y, He Q, Xu L, Liu S, Qin L, Jia Z, Li J, Huang H, Zhao L (2007) MSMiner-a developing platform for OLAP. Decis Support Syst 42(4):2016–2028
69. Sikora R, Piramuthu S (2005) Efficient genetic algorithm based data mining using feature selection with hausdorff distance. Inf Technol Manage 6(4):315–331
70. Sugumaran V, Tanniru M, Storey V (2008) A knowledge-based framework for extracting components in agile systems development. Inf Technol Manage 9(1):37–53
71. Sun B, Xu L, Pei X, Li H (2003) Scenario-based knowledge representation in case-based reasoning systems. Expert Syst 20(2):92–99
72. Takano K, Chen X, Masuda K (2009) A framework for a feedback process to analyze and personalize a document vector space in a feature extraction model. Inf Technol Manage 10(2–3):151–176
73. Tan W, Shen W, Xu L, Zhou B, Li L (2008) A business process intelligence system for enterprise process performance management. IEEE Trans SMC Part C 38(6):745–756
74. Tan W, Xu Y, Xu W, Xu L, Zhao X, Wang L, Fu L (2010) A methodology toward manufacturing grid-based virtual enterprise operation platform. Enterp Inf Syst 4(3):283–309
75. Tan X, Yen D, Fang X (2003) Web warehousing: web technology meets data warehousing. Technol Soc 25(1):131–148
76. Tang C, Xu L, Feng S (2001) An agent-based geographical information system. Knowl-Based Syst 14(5–6):233–242
77. Tremblay M, Berndt D, Luther S, Foulis P, French D (2009) Identifying fall-related injuries: text mining the electronic medical record. Inf Technol Manage 10(4):253–265
78. van Sinderen M, Almeida J (2011) Enpowering enterprises through next-generation enterprise computing. Enterp Inf Syst 5(1):1–8
79. Vapnik V (1998) Statistical learning theory. Wiley, London
80. Viriyasitavat W, Xu L, Martin A (2012) SWSpec, service workflow requirements specification language: the formal requirements specification in service workflow environments. IEEE Trans Industr Inform, in press. doi:10.1109/TII.2011.2182519
81. Wang L, Xu L, Wang X, You W, Tan W (2009) Knowledge portal construction and resources integration for a large scale hydropower dam. Syst Res Behav Sci 26(3):357–366
82. Wang L, Xu L, Liu R, Wang H (2010) An approach for moving object recognition based on BPR and CI. Inf Syst Front 12(2):141–148
83. Wang L, Zeng J, Xu L (2011) A decision support system for substage-zoning filling design of rock-fill dams based on particle swarm optimization. Inf Technol Manage 12(2):111–119
84. Wang P, Xu L, Zhou S, Fan Z, Li Y, Feng S (2010) Novel Bayesian learning method for information aggregation in modular neural networks. Expert Syst Appl 37(2):1071–1074
85. Wang P, Zhang J, Xu L, Wang H, Feng S, Zhu H (2011) How to measure adaptation complexity in evolvable systems-a new synthetic approach of constructing fitness functions. Expert Syst Appl 38(8):10414–10419
86. Wang X, Wang H, Zhang L, Cao X (2011) Constructing a decision support system for management of employee turnover risk. Inf Technol Manage 12(2):187–196
87. Wetzstein B, Leitner P, Rosenberg F, Dustdar S, Leymann F (2011) Identifying influential factors of business process performance using dependency analysis. Enterp Inf Syst 5(1):79–98
88. Wu S, Xu L, He W (2009) Industry-oriented enterprise resource planning. Enterp Inf Syst 3(4):409–424
89. Xu E, Wermus M, Bauman B (2011) Development of an integrated medical supply information system. Enterp Inf Syst 5(3):385–399
90. Xu L (1995) Case-based reasoning for AIDS initial assessment. Knowl-Based Syst 8(1):32–38
91. Xu L (1995) Case-based reasoning—a major paradigm of artificial intelligence. IEEE Potentials 13(5):10–13
92. Xu L (1996) An integrated rule- and case-based approach to AIDS initial assessment. Int J Biomed Comput 40(3):197–207
93. Xu L (1999) Editorial. Expert Syst Appl 16(1):1–2
94. Xu L (2006) Advances in intelligent information processing. Expert Syst 23(5):249–250
95. Xu L (2011) Information architecture for supply chain quality management. Int J Prod Res 49(1):183–198
96. Xu L, Li Z, Li S, Tang F (2005) A polychromatic sets approach to the conceptual design of machine tools. Int J Prod Res 43(12):2397–2422
97. Xu L, Li Z, Li S, Tang F (2007) A decision support system for product design in concurrent engineering. Decis Support Syst 42(4):2029–2042
98. Xu L, Liang N, Gao Q (2008) An integrated approach for agricultural ecosystem management. IEEE Trans SMC Part C 38(4):590–599
99. Xu L, Liu H, Wang S, Wang K (2009) Modeling and analysis techniques for cross-organizational workflow systems. Syst Res Behav Sci 26(3):367–389
100. Xu L, Wang C, Luo X, Shi Z (2006) Integrating knowledge management and ERP in enterprise information systems. Syst Res Behav Sci 23(2):147–156
101. Xu S, Xu L (2011) Management: a scientific discipline for humanity. Inf Technol Manage 12(2):51–54
102. Yang B, Li L (2001) Development of a KBS for managing bank loan risk. Knowl-Based Syst 14(5–6):299–302

103. Yang B, Li L, Xu J (2001) An early warning system for loan risk assessment using artificial neural networks. Knowl-Based Syst 14(5–6):303–306

104. Yin Y, Fan Y, Xu L (2012) EMG & EPP-integrated human-machine interface between the paralyzed and rehabilitation exoskeleton. IEEE Trans Inf Technol Biomed, in press. doi:10.1109/TITB.2011.2178034

105. Yu H, Yang J, Han J (2003) Classifying large data sets using SVMs with hierarchical clusters. In: Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining, Washington, D.C., 24–27 Aug 2003

106. Yuan R, Li Z, Guan X, Xu L (2010) An SVM-based machine learning method for accurate internet traffic classification. Inf Syst Front 12(2):149–156

107. Zhang H, Wang D, Wu W, Hu H (2012) Term frequency-function of document frequency: a new term weighting scheme for enterprise information retrieval. Enterp Inf Syst, published online. doi:10.1080/17517575.2012.665945

108. Zhang M, Xu L, Zhang W, Li H (2003) A rough set approach to knowledge reduction based on inclusion degree and evidence reasoning theory. Expert Syst 20(5):298–304

109. Zhou S, Gan J, Xu L, John R (2009) Fuzziness index driven fuzzy relaxation algorithm and applications to image processing. Ann Oper Res 168(1):119–131

110. Zhou S, Li H, Xu L (2003) A variational approach to intensity approximation for remote sensing images using dynamic neural networks. Expert Syst 20(4):163–170

111. Zhou S, Xu L (1999) Dynamic recurrent neural networks for a hybrid intelligent decision support system for the metallurgical industry. Expert Syst 16(4):240–247

112. Zhou S, Xu L (2001) A new type of recurrent fuzzy neural network for modeling dynamic systems. Knowl-Based Syst 14(5–6):243–251

113. Zhou Y, Huang F, Chen H (2008) Combining probability models and web mining models: a framework for proper name transliteration. Inf Technol Manage 9(2):91–103

114. Zhu X, Wang X, Xu L, Li H (2008) Predicting stock index increments by neural networks: the role of trading volume under different horizons. Expert Syst Appl 34(4):3043–3054