

Application Research of Big Data Mining and Decision Analysis System in Colleges and Universities

Ying Chen

NanChang Institute of Science & Technology, Nanchang,330108

Keywords: Subject construction, Decision support, Data Warehouse, Data mining

Abstract. The subjects are the basic units to organize the teaching and the bonding point of cultivating talents and developing science, technology and culture in colleges and universities. It has become the developing tendency of the present world higher education to strengthen the subject construction. With the development of the computer and network techniques, colleges and universities have realized information management. A large number of management information systems used form a huge factory in which a large amount of effective profiling data is generated. It has become a problem to be solved for the subjects manager in colleges and universities and a hot research topic for the scholars at home and abroad how to use the informational data constantly accumulated to provide services for the subject construction and improve the objectivity and the scientific nature of the management decision of subject in colleges and universities. With the subject construction in colleges and universities as the application background, this paper effectively uses the data management information systems generate in colleges and universities to explore the phenomena and trends of the subject development, reveal the rules and characteristics of subject development in colleges and universities and inherit and develop the new applications of decision support system in the field of subject management in colleges and universities on the basis of the research on the data mining technology and decision support system.

Introduction

With the continual development of computer technology, network technology and management science in the information age, the information of business and management have been basically realized in nearly all the units and departments such as governments, enterprises, banks and universities with the help of computer and network technology. The application of management information system(MIS) not only greatly improves the management efficiency by freeing a large number of staff from the complex business but also makes the management convenient, unified and standardized. At the same time, the application of management information system can generate a large amount of effective data reflect the realistic status of business. The management decision of the senior leaders has imperceptibly become the most important factor for the success of enterprise because of the business information. The management information system provides convenience and fast for the business, but it also brings unprecedented pressure to the decision of the senior leaders. However, in the face of the modernization of enterprise management, the information of business, the leading of high level decision and the hugeness of business data, people urgently need the theories, methods and advanced tools which can help to make the scientific decision.

Decision Support System

Since the concept of decision support system (DSS) was first proposed by Michael S. Scott Morton in America in the early 1970s, it has attracted the attention of scholars in many fields such as management science, information science, mathematics and operational research. A large number of scholars make research on the theories and methods of decision support system, propose many excellent theories and greatly promote the development of the theory of decision support system. Soon R.H. Sprague proposed the three units structure (also known as two base structure model), shown as Fig. 1.

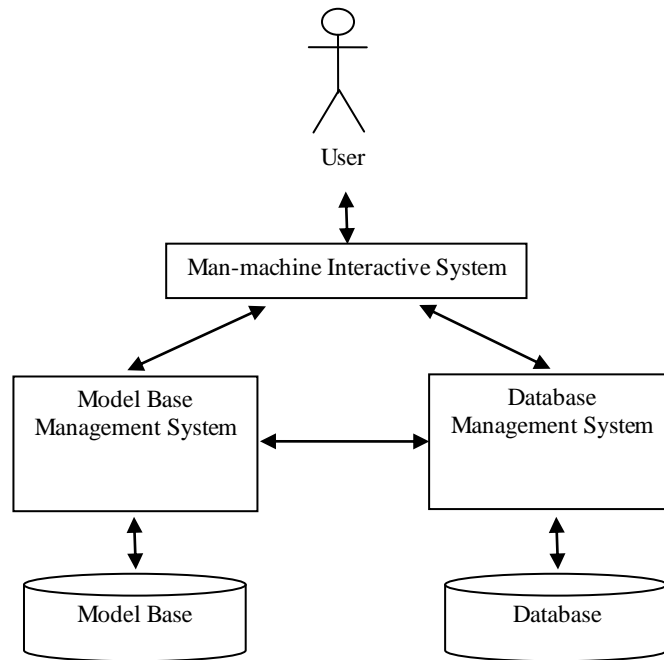


Figure 1. Structure Chart of Traditional Decision Support System

The structure includes man-machine interactive component, data components (database and database management system) and model components (model base and model base management system). Among the above components, (1) the man-machine interactive component is the interactive interface between the decision support system and the user. The user can control the operation of decision support system through the interactive component, and the system shows the result of decision support for the user through the interactive system; (2) The data components are used to store and manage data, and provide the language interface for the operational data; (3) the model components are used to build, store and manage the models, and provide the language interface for the operational models.

Intelligent decision support system (IDDS) is developed from the two base structure model of the traditional decision support system combining with the expert system by bringing the knowledge base of the expert system to the decision support system. The structure of the intelligent decision support system is shown as Fig. 2.

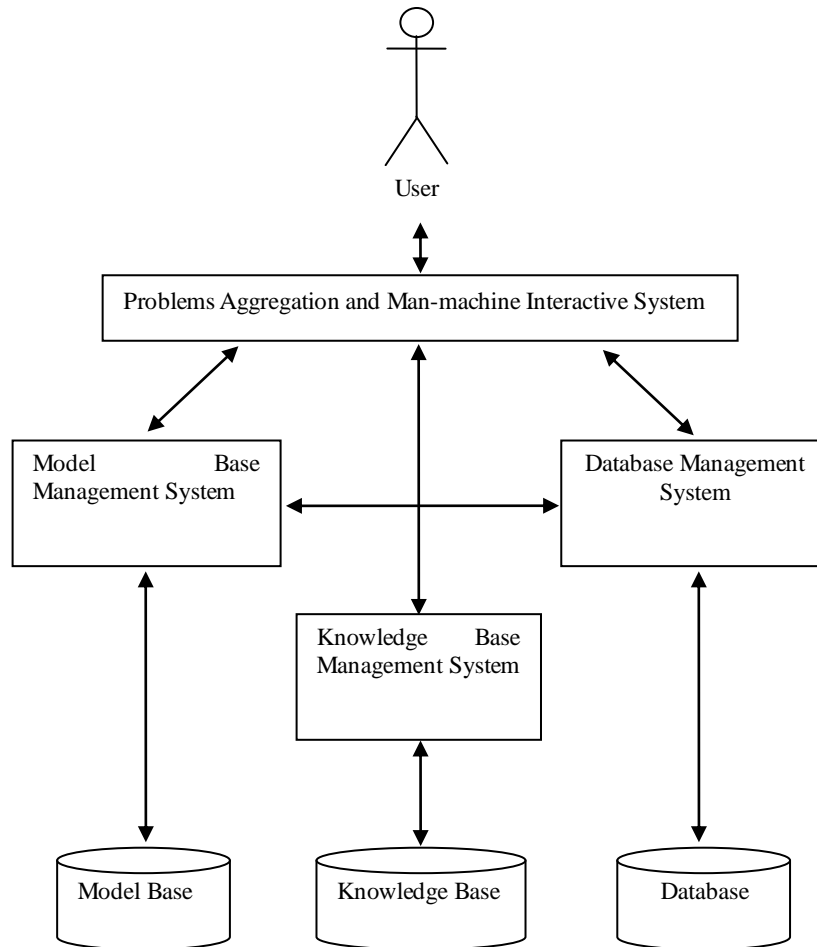


Figure 2. Structure Chart of Intelligent Decision Support System

As the combination of numerical analysis and knowledge processing, intelligent decision support system uses two analysis methods qualitative analysis and quantitative analysis which can solve the semi-structured and unstructured problems more effectively than the traditional decision support system. With the support of the expert system, IDSS can solve the partial qualitative problems in decision-making that not only can fully exert the characteristic of the expert system that the qualitative problems are solved with the form of knowledge reasoning but also can exert the characteristic of the traditional decision support system that the quantitative problems are solved with model calculation.

Design of Data Warehouse of Subject Construction in Colleges and Universities

The data warehouse is an organizational structure of different database systems which is developed on the basis of the original relational database. The basic data and the aggregate data it get from the original business database and other data sources are divided into four different layers the early historical data layer, the current basic data layer, the low-grade aggregate data layer and the high-grade aggregate data layer. The structure chart of a general data warehouse is shown as Figure 3.

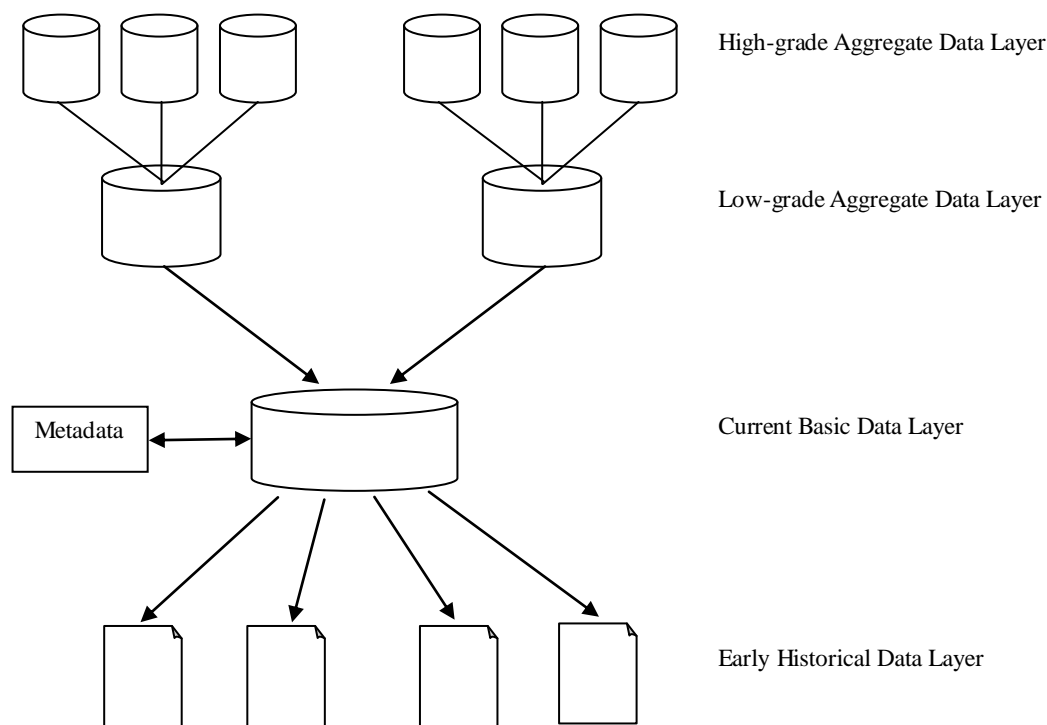


Figure 3. Structure Chart of Data Warehouse

The metadata goes into the current basic data layer after integration. As the most interested part of the system for the user of data warehouse, the current basic data layer with the most data stores the current business data. The low-grade aggregate data is extracted from the current basic data, and the high-grade aggregate data is a higher aggregation. The current basic data is saved to the early historical data over time. The whole organizational structure of the data warehouse is organized with the metadata.

Overall Framework of Decision Support System of Subject Construction in Colleges and Universities

The structure of decision support system of subject construction in colleges and universities designed in this paper is built on the basis of fully combining with the characteristics of the subject management in colleges and universities, the applicability of decision support system and the user-friendly. Firstly, the subject management in colleges and universities is an exchange process of data and information in fact. All aspects involved in the subject management are generally quantified by certain data. In design of this system, data analysis and data mining take the more important position, and the whole process of making decisions is seen as a process of data mining or a process of data analysis. In addition, this system provides the interface for data mining algorithms. Aiming at the domain problems need to be solved, the reasonable data mining algorithms should be designed or chosen to get the decision-making model or the decision-making knowledge with the historical data and practice, and the decision-making knowledge can give a reference to the scientific decision of the decision maker.

The system is mainly divided into the following modules data integration module, application module, business database module, Web server and Web client. The data integration module is mainly used to collect data from the business database and other historical databases, and then transfer and integrate the data, lastly save the data to the data warehouse by ETL server; Application module is mainly used to carry out data processing, data analysis and data mining to get the decision information and knowledge; Business database module is mainly used to store and manage the informational data from the application module; Web server is mainly used to accept the user's request, get information from the business database or the application server and return

information to the user; Web client is mainly used to send request and receive reply as the interface of user and system.

It is the biggest characteristic of the system architecture to effectively integrate the data warehouse system, database system, application system (mathematical model calculation, data mining and data analysis) and Web system. The effective integration of these systems extends the traditional architecture on the basis of following the traditional architecture of decision support system. The system architecture carries out the coupling of training knowledge and using knowledge by using a relational database of business as the decision information base (knowledge base). The operation of knowledge mining is time consuming which need a great deal of reasoning and calculating. If there must be a long calculation process when the system is used, it is hard to ensure the real-time performance of system. The knowledge mined is stored through the business database, the user can directly get the existing knowledge from the knowledge base when he need make a decision for the decision problem to improve the real-time performance of the system. The architecture is similar with the decision process of human. We save the experience and knowledge transferred from much usual practice to our brains, our brains will give a quick decision based on the past experience when we need solve some problems.

Conclusion

Based on the research on the relative technologies, theories and applications of decision support system and the domain object of subject research, this paper considers that the building of decision support system of subject construction in colleges and universities should put emphasis on the research on the data-driven decision support system combining data warehouse technology and data mining technology. In other words, it also is the inevitable choice to carry out the decision support system of subject construction in colleges and universities. However, this paper completes and realizes the prototype system of decision support system of subject construction in colleges and universities based on data mining by applying the decision support system based on data warehouse to the subject construction in colleges and universities and combining data mining technology and Web technology. The system mainly concludes: (1) ETL server: to clean and integrate the raw data; (2) data warehouse server: to store and manage multi-dimensional data; (3) OLAP and data mining tools: to mine the data to get the internal mode and decision knowledge; (4) decision knowledge base: to store and manage the analyzed knowledge and the decision knowledge; (5) Web server: to provide business services; (6) Web browser: composed of man-machine interface and other modules.

Acknowledgements

Project Fund: science and technology planning project of the Education Department in Jiangxi Province (project number: GJJ151233).

References

- [1]. Pani S K. Decision tree analysis on j48 and random forest algorithm for data mining using breast cancer microarray dataset”,[C]// Sustech-15, Isbn. 2015.
- [2]. He M. A Data Analytics Framework for Smart Grids: Spatio-temporal Wind Power Analysis and Synchrophasor Data Mining[J]. Arizona State University, 2013.
- [3]. D’Oca S, Hong T. Occupancy schedules learning process through a data mining framework[J]. Energy & Buildings, 2015, 88:395-408.
- [4]. Hu S, Li Y, Zhang D. Fault-tolerant Mining Algorithm of Sampling Data from Dynamic System[C]// the 25th chinese control and decision conference. 2013:4927-4931.
- [5]. Dai T, Hu H, Wan Y. Design of health decision support system with data analysis and mining functions[C]// International Conference on Fuzzy Systems and Knowledge Discovery. IEEE, 2016:2266-2270.

- [6]. Chen C. Research on Unstructured Big Data Analysis and Decision Making System for Library Based on Hadoop[J]. *Information Science*, 2017.
- [7]. Liu Q J, Huang Z G, Wei L I. Best Dispatching Smart Information Analysis and Construction of Integrated Decision-making System Based on Data Mining[J]. *East China Electric Power*, 2013.
- [8]. Kaushal A, Shukla M. Comparative Analysis to Highlight Pros and Cons of Data Mining Techniques-Clustering, Neural Network and Decision Tree[J]. *International Journal of Computer Science & Information Technolo*, 2014.
- [9]. Lin J D, Huang W H, Hung C T, et al. Using Decision Tree for Data Mining of Pavement Maintenance and Management[J]. *Applied Mechanics & Materials*, 2013, 330(9):1015-1019.
- [10]. Tang X, Xu K, Xue X, et al. Study on Decision Analysis System of Gas Cylinder Basing on Data Mining Technology[J]. *China Special Equipment Safety*, 2015.