

Decision Support Systems and Data Mining - An Integrated Approach

Dr. Gagandeep Marken

Abstract: ‘A stitch in time saves nine’, is a well known proverb, which is true in ones’ individual life and as well as in the life of an organization. The survival and hence the success of an organization depends on the right decisions taken well in time. The organizations around the globe are starving hard to manage huge amount of data stored in their data repositories and hence are finding ways and means to explore data to enhance the quality of decisions through better decision support systems. OLAP based tools for used for decision making in the organization irrespective of their size and operations, but as the size of data bases is increasing at an alarming rate, these tools have become insufficient. The tools and techniques of data mining can prove to be boon for organizations to discover patterns and relationships hidden in data and hence they can improve the quality of data analysis resulting in better decisions. This paper throws light on the concept that how data mining methods can be integrated with the decision support systems to enhance the quality of decisions.

Keywords: Integrated Decision Support, Data Mining, Globalization, Clustering.

I. INTRODUCTION

OLAP based decision support systems are used in many organizations, these systems use a deductive approach of analyzing for interpreting data. These methods use the user’s ability to choose the right options by drilling down to find the most suitable information, trends or patterns required for the making decision in any particular context. So these systems are user centric where the user’s perception, ability and knowledge play a pivot role for making the decision. These systems suffer from the problem that these provide the answers to the quantitative queries, like what is the total sale in any particular region, or what is total output of a particular production unit? But the present business scenarios of globalization and cut throat competition require in-depth study and analysis of the core business processes where the study of market trends, consumer’s behavior, their taste and preferences and the prediction of their future actions and comprehensions are also very important. One wrong decision or the right decision taken at the wrong time may prove to be fatal for the organization. So the need is to find out a system other than OLAP which will try to overcome these deficiencies. Traditional isolated, stand-alone DSS has been recently facing new challenges. In order to improve the performance of DSS to meet the challenges, research has been actively carried out to develop integrated decision support systems (IDSS). The data mining techniques can be a boon in such situation; these techniques use an inductive approach of analyzing data. The use of data mining to facilitate decision support can lead to an improved performance of decision making and can enable the tackling of new types of problems that have not been addressed before. The integration of data mining and decision support can significantly improve current approaches and may create new approaches to problem solving, by enabling the fusion of knowledge from experts and knowledge extracted from data. The paper focuses on the concept that how the techniques and tools of data mining can be integrated with the decision support system that can be helpful for the organizations and individuals for taking right decision at the right time.

II. DATA MINING
"Now that we have gathered so much data, what do we do with it?"
This is the famous opening statement of the editorial by Usama Fayyad and Ramasamy Uthurusamy in
the Communications of the ACM, Special issue on Data Mining (Fayyad & Uthurusamy, 1996). With
the relentless growth of information available on the Internet, combined with the storage of gigantic
quantities of data on personal and corporate computers, individuals and businesses often find
themselves in situations where they feel overwhelmed by the sheer amount of data they must process to
make time constrained but informed decisions. Apart from this valuable information could be "hidden"
in personal/corporate/Web data, and data mining may allow you to uncover it. And so DM, and various
related methodologies, techniques and tools are now becoming more and more widely used in order to
uncover this implicit and potentially useful information or knowledge.

Data mining has been defined as “the nontrivial extraction of implicit, previously unknown, and
potentially useful information from data”. It is “the science of extracting useful information from large
databases”. Data mining is one of the tasks in the process of knowledge discovery from the database.
In the last decade, the digital revolution has provided relatively inexpensive and available means to
collect and store data. The increase in the data volume causes greater difficulties in extracting useful
information for decision support. The traditional manual data analysis has become insufficient, and
methods for efficient computer-based analysis have become indispensable. From this need, a new
interdisciplinary field of data mining has emerged. Data mining encompasses statistical, pattern
recognition, and machine learning tools to support the analysis of data and discovery of principles
that lie within the data. Data mining consists of five major elements; to extract, to transform, and to load
transaction data onto the data warehouse system, to store and manage the data in a multidimensional
database system, to provide data access to business analysts and information. Analyze the data by
application software, and finally to present the data in a useful format, such as a graph or table.

III. DECISION SUPPORT SYSTEM (DSS)
Decision support systems (DSS) are defined as interactive computer-based systems intended to help
decision makers to utilize data and models in order to identify problems, solve problems and make
decisions. They incorporate both data and models and these are designed to assist decision makers in
semi-structured and unstructured decision making processes. They provide support for decision
making; they do not replace the decision makers.

Simon’s model of decision making process consisted of three phases: intelligence, design and choice.
In the model, intelligence is concerned with the search for problems, design involves the development
of alternatives, and choice is about analyzing the alternatives and selecting one for implementation.
This classic problem-solving model of “intelligence-design-choice” has been widely accepted and
adopted. Later on a fourth monitoring phase was included. Diverse DSS were developed to support
decision makers at all levels in an organization including systems that could support problem
structuring, operations, financial management and strategic decision making, even extending to support
for optimization and simulation. Data mining methods also extend the possibilities of discovering
information, trends and patterns by using richer model representations (e.g. decision rules, trees, tables)
than the usual statistical methods, and are therefore well-suited for making the results more
comprehensible to the non-technically oriented business users.

DSS includes a body of knowledge that describes some aspects of the decision maker's world that
specify how to accomplish various tasks that indicates what conclusions are valid in different
The expected benefits of DSS that are discovered are higher decision quality, improved
communication, cost reduction, increased productivity, time savings, improved customer satisfaction
and improved employee satisfaction. DSS is a computer-based system consisting of three main
interacting components:

• A Language System: A mechanism to provide communication between the user and other components
of the DSS.
• A Knowledge System: A repository of problem domain knowledge embodied in DSS as either data or
procedures.
• A Problem Processing System: A link between the other two components, containing one or more of
the general problem manipulation capabilities required for decision-making.

IV. INTEGRATION OF DATA MINING AND DECISION SUPPORT

a. Historical Prospective:
Several authors discuss integration of data mining into decision support and they all confirm the value
of it. Chen argues that the use of data mining helps institutions to make critical decisions faster and
with a greater degree of confidence. The use of data mining lowers the uncertainty in the decision
process. Lavrac and Bohanec claim that the integration of data mining and decision support will help
the decision makers to better decisions even for the problems new to the organization or to the
individuals. They opine that this integration will redefine the present methods of decision making and
will open up the new approaches to problem solving, by enabling the fusion of knowledge from experts
and knowledge extracted from data. The ever growing competition among the enterprises has forced the
decision makers to switch from traditional one function or single user focused DSS to complex and
intelligent agent-based analysis, suggestion and judgement. The proposed integrated decision support
system proposes the integration of new technologies, processes and business environments into
decision support paradigm to enable improved performance. Integration achieved by establishing the
relationships between the different components in the form of models, services, tools, methods,
subsystems, systems, etc. It is not a property of a single element, but of its relationships with other
components of the environment. Earlier, DSS consists of three basic capability components: a data
management component, a model management component and a dialogue management component.
These three elements were often represented as a database management system (DBMS) to manage the
data, a model base management system (MBMS) that manages the models to formulate the problems
and solutions, and an interface that enables interactive queries, reporting, and graphing functions of
decision solutions. With the passage of time and by the development of technologies this traditional,
three-element focused, DSS have evolved into modern IDSS which has enriched the content of the
three basic components, and has improved the basic decision making process.

b. Various Types of Integration:

Data and Information Integration
In large enterprises, the data is generated, stored and processed by various users at the various sites,
these users process the data as per their requirements of the individual users. But ultimately in the end
when the final decision is to be taken the various factors and the data from the various sources is to be
brought under one roof, which requires the consistent view of data so that a decision beneficial for the
whole organization can be taken. At the first step, the data and information integration should reflect
the use of “same” data and information by system components even though the data and information is
represented differently or is deduced from various sources. Davenport argues that information
integration in companies should consider how it might strengthen or weaken competitive advantages, how it will affect company’s culture and how information management will be implemented. Further the knowledge management systems are to be integrated with decision support systems so that different types of knowledge available from the different sources can be combined together for making effective decisions.

**Model integration**
A model is representation of reality that allow the decision maker to have the holistic view of the problem under consideration, by having an idea about the possible inputs, processes and the outputs generated and the impact of each and every process and the decision undertaken. The model creation differentiates the decision support system from the traditional information processing systems. It has been proved in practice that a DSS should create multiple models rather than a single larger model for the benefit of DSS implementation, maintenance and flexibility. By the process of model integration the individually developed sub-models are logically combined to create a large unified model, this integration will provide the support for making rational decision making. For model integration, two operations were defined and implemented: projection and join. Projection operation extracts a desired sub-model from a larger model for use in other applications. Join operation merges two models that have certain compatibility.

**Service integration**
Service integration is essential to support flexible function combinations in an integrated decision support environment. Ideally, all the functions offered by all the components should be accessible as per the need and use to all other components. This can be seen from two different views: first view is service integration with respect to provision, i.e. components offer services that other components in the environment require and use. Second view is service integration with respect to use, i.e. components appropriately use the services offered by other components in the environment. The components must be able to communicate the operations to be performed in order to share functionality, and operations will require data and information, the components must also communicate data and information or their references. In this sense, service integration complements data and information integration. Service integration addresses control-transfer and service-sharing issues.

**Process Integration**
Thomas defined three dimensions for understanding process integration: process unit, process event and process constraint. A process unit is a unit of work that yields a result. A process event is a condition that arises during a process step that may result in the execution of an associated action. A process constraint imposes constraints on some aspects of the process. Two components inside an IDSS are considered well integrated with respect to process if their assumptions about the process are consistent..

**Presentation Integration**
The users often interact with the different components of a system or subsystem, the interaction with these components should be consistent as far as possible, the presentation integration strives to achieve this by providing a consistent view throughout the interaction over the various parts of the system. Presentation integration can be achieved through appearance and behavior integration, as well as by interaction-paradigm integration. Appearance and behaviour integration deals with the concept that how easy it is for users to interact with the different components. This can be achieved by having the
similarities of component screen appearance and interaction behaviour. Components are considered well integrated with respect to appearance and behaviour integration if a user has a consistent view of all components visited and being interacted.

c. Role of Data Mining in DSS

Independently, data mining and DSS are well-developed areas. As a part of data mining collective tools, data warehouse and OLAP have been often used in IDSS to achieve data and information integration. Data warehouse was promoted as a solution for integrating data from diverse operational databases to support management of decision making. OLAP is multi-dimensional analysis which allows decision makers (analysts, managers and executives) to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information. This information has often been transformed from raw data (through integration of enterprise aggregate data across many dimensions such as product, time and location, etc.) to reflect the real dimensionality of the enterprise as understood by the users. The organizations have started using the concept of data mining for integrated decision support systems, but still this concept is at its initial level. Mladenic et al. has presented a conceptual framework for integrating data mining into IDSS, with methods and tools, and applying them to business problems in a collaborative setting. Most IDSS integrated with data mining technology are found in medical problem domains such as for automated cardiac diagnosis, quality assessment of haemodialysis services and predicting survival time for kidney dialysis patients. Integrating data mining technology into IDSS for dynamic manufacturing process was first explored by Lee. Later on, Lee and Park designed a customized sampling DSS (CSDSS) employing data mining to carry out knowledge mining task for customer-centric electronic commerce. Recently, Shi et al. has explored and developed a generic data mining platform for integrated decision support called MSMiner. The major strength of the MSMiner is that it provide an integrated, extensible DSS by employing data warehousing and data mining technologies, in which the integration of an entire decision making process (from data collection, through data modeling, data pre-processing and data mining to data visualization) is emphasized. In the MSMiner, a subsystem called extraction, transformation and loading (ETL) is designed to clean and transform data from multiple source databases before integrating them into a data warehouse. The data mining process can be seen as a knowledge creation phase in the whole data mining enhanced decision making process, in which the data sets undergo cleaning and pre-processing for removing discrepancies and inconsistencies to improve its quality. The selected data set is then analyzed to identify patterns that represent relationships among data.

Data mining reveals implicit relationships behind data itself, which can assist decision makers obtaining overall perceptions of decisions in an organization. Information flowing through an IDSS without data mining process can be disjoint with discrepancy and inconsistency, which leads to poor decision support. Therefore, the performance of IDSS can be greatly improved through data mining to provide better decision support—knowledge with quality, business relevance and understood by users from multi-dimensions.

V. CHALLENGES IN INTEGRATION PROCESS:

The present changes in the technology from database to data warehouse and on-line analysis processing (OLAP), from mainframe to client/server architecture, and from single user model to World Wide Web access; and the growing interconnection with more dynamic business environment and further increasing complexity of the decision situations which puts enormous cognitive workload on decision
makers has totally changed the basic way of taking decisions. One common key issue behind the above challenges is that the traditional problem solving characterization of DSS has to be expanded and integrated to be compatible with new technologies, business environments and intelligence, to allow more transparent interaction between decision makers and systems, not only for improvement of the efficiency and effectiveness of the decisions, but also for collaborative support and virtual team working. Performing analysis through data mining follows an inductive approach of analyzing data where machine learning algorithms are applied to extract non-obvious knowledge from data. The major challenges in this context may be enumerated as:

1. The first important challenge for the integration is the data and information integration, as data is the basic format for decision variables, constraints and objective functions. New data management technologies such as data warehousing, OLAP and data mining have been widely employed in IDSS to handle data and information integration from multiple sources, to provide multiple presentation and business intelligence/knowledge to users. The biggest challenge related to data is to maintain the consistency between the data, information, and knowledge when they are transformed to different formats, and how to maintain efficiency when there are huge quantities of data being processed.

2. Another major challenge is the integration of qualitative modeling methods in IDSS. Qualitative modeling methods are becoming increasingly important for IDSS to explore symbolic qualitative aspects of the decision process: experience, intuition, judgment, and expertise. The integrated decision support systems should try to integrate qualitative and quantitative methods. Another major issue is the applicability of the various models as per the need of the organizations.

3. A big challenge for process integration is the creation of a common decision making process across different organization levels that top manager, middle managers and technical personnel.

4. The next important challenge is to provide the presentation integration for technical users and public users taking into account user’s domain knowledge, experience and behaviour, so that the users at the various levels should have a unified and consistent view of system.

Oracle has offered Oracle Data Mining (ODM) option integrated in the database enabling the use of data mining methods on data in Oracle database. Apart from ODM there is also Java based API available which enables the development of J2EE applications which use data mining. The motive behind the development of DMDSS was based on the fact that the traditional use of data mining through data mining software tools does not bring data mining closer to business users because of complexity of data mining tools.

VI. CONCLUSION

The primary purpose of a stand-alone DSS is to improve the performance of individual decision maker by improving the quality of his or her decision by enabling them to make effective and efficient decisions, in such as situation, an integrated decision systems have demonstrated its advantages by providing consistent, co-ordinated, active and global support for multiple users on varied decisions in an organization. The integration within an IDSS can be addressed not only from data, model, and presentation or user interface perspectives, but also from service and process viewpoints. The above integration can be enhanced through the use of latest technologies like data mining, intelligent agents and the Web technology.
REFERENCES:

2. Shaofeng Liu, Alex H. B. Duffy[2007], Integration of decision support systems to improve decision support performance, Robert Ian Whitfield - Iain M. Boyle, Knowledge Information System, DOI 10.1007/s10115-009-0192-4,
3. Rok Rupnik, Matjaž Kukar[2007], Data Mining Based Decision Support System to Support Association Rules, Electrotechnical Review: Ljubljana, Slovenija