CRITICAL FACTORS FOR ACHIEVING DATA MINING SUCCESS

ANG EE LYNN

Research report in partial fulfillment of the requirements for the degree of Master of Business Administration

2009

DEDICATION

Specially dedicated to:

My beloved parents

for their sacrifices and unconditional love.

ACKNOWLEDGEMENT

There are many individuals whom I would like to thank for their support and guidance given throughout the course of my study.

First of all, my heartfelt appreciation and gratitude to my supervisor, Associate Professor DR. MD Harashid Harun for his valuable guidance and advice. Without his tireless support, I would not have learnt so much more. Truly, it is my honor to have him as my supervisor.

My deepest gratitude to my parents and brothers. The moral support and encouragement given throughout my study is indescribable.

To top up the list and not to be forgotten are the respondents of this research for their willingness and contribution towards the success of this research.

Last but not least, many thanks to my course-mates, colleagues and friends for their sharing and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRAK

Perlombongan data menukarkan data kepada bentuk pengetahuan yang bernilai dan berguna untuk meningkatkan proses membuat keputusan dan memperoleh kelebihan persaingan. Semua syarikat mahupun syarikat besar atau kecil sedang membangunkan dan menggunakan teknologi perlombongan data. Kajian ini bertujuan untuk menentukan faktor-faktor yang mempengaruhi kepuasan pengguna terhadap perlombongan data. Faktor-faktor yang dikaji adalah masa balasan sistem, kualiti data, kesenangan penggunaan, kebergunaan maklumat dan sokongan yang disediakan kepada pengguna. Data dikumpul daripada 102 responden di Semenanjung Malaysia. Keputusan kajian ini menunjukkan kepuasan penggunaan bagi perlombongan data adalan jelas dipengaruhi oleh masa balasan sistem, kualiti data, kesenangan penggunaan, kebergunaan maklumat dan sokongan yang disediakan kepada pengguna. Kajian ini juga membuktikan bahawa tahap kepuasan pengguna perlombongan data adalah lebih daripada purata.

ABSTRACT

Data mining is transforming data into valuable and actionable knowledge to enhance the decision making process and gain competitive advantage. Organizations of all sizes are developing and implementing data mining technologies. This research seeks to identify the critical factors influencing the user satisfaction on data mining. The factors studied are system throughput, data quality, ease of use, information utility and support provided to end-user. Data was collected from 102 respondents throughout Peninsular Malaysia. The result of this study indicates that user satisfaction with the data mining is significantly affected by system throughput, data quality, ease of use, information utility and support provided to end-user. This study also implies that the level of data mining user satisfaction among the respondents was above average.

Chapter 1

INTRODUCTION

1.1 Introduction

This chapter introduces the research outline of the study. The chapter illustrates an overview of the background, problem statement, research objectives, research question, definition of key terms and significance of study.

1.2 Background

Due to recent advances in information and database technologies, business organizations are able to collect huge quantities of both internal and external data that potentially can be processed to support decision-making (Vijayan and Ranjit, 1999). Enterprise decision makers are seeking out competitive advantages by eliminating inefficiencies, optimizing internal operations, and maximizing relationships with all organizational stakeholders (employees, customers, partners, and shareholders) in this intensely competitive global marketplace (Nemati & Barko, 2001).

According to Forcht and Cochran (1999), the effective use and analysis of customer data are now critically important to the success of organizations. In addition, a survey of North American CEOs by the Conference Board and Accenture claimed that the no. 1 management challenge for 1999 and 2000 was keeping focus on the customer (Hall, 2001). Furthermore, Adams (2001) suggested that the benefits of adopting a customer-focused strategy are many, including but not limited to customer loyalty and retention, personalization to aid purchase decisions, better marketing and sales coordination, increased organizational efficiency, and advantages gained through improved customer intelligence. Hence, organizations of all sizes are developing and deploying data mining technologies to leverage data-resources to enhance their decision-making capabilities (Eckerson & Watson, 2001).

In the mid-1980s, the Malaysian government made serious efforts to get the banks to merge due to the economic downturn and the problems besetting small banks. The Governor of Bank Negara (National Bank), Tan Sri Ali Abul Hassan Sulaiman, claimed that Malaysia's domestic banking institutions would be able to face the pressure and challenges arising from an increasingly competitive global environment (*Bank Negara explains rationale for bank merger*, 1999). Twenty-one domestic commercial banks were merged into ten anchor banks by 31 December 2001 as a result of the consolidation that took place in the Malaysian banking sector (*Bank Negara*, 2002).

Banks mergers have played an important role in the development of data warehousing and data mining. In fact, the banking industry has been the leader in the use of data warehouses (Gupta, 1997). For example, when a banking unit uses different operational systems in the different branches, the top management still needs to view the consolidated business and manage the associated risks accordingly. Meanwhile, data mining is the process of sorting out and analyzing in data warehouse (Teh and Zaitun, 2003).

Sugumaran and Bose (1999) defined data mining is the process of discovering meaningful new correlations, patterns, and trends by sifting through large amounts of data stored in data warehouses and by using pattern recognition technologies as well as statistical and mathematical techniques.

Data mining consists of a wide array of technologies such as e-business intelligence, data analysis, digital dashboards, information portals, Customer Relationship Management (CRM) etc. According to Nemati and Barko (2003), data mining enables organizations to answer questions about the past (what has happened), the present (what is happening) and the future (what might happen). Armed with this capability, organizations are able to generate valuable knowledge form their data, which in turn enhances enterprise decisions and translates into a strategic competitive advantage.

1.3 Problem Statement

Many organizations are beginning to realize the importance of Data Mining in their strategic planning and the enormous organizational benefits attained by successfully implementing these systems (Lee and Siau, 2001). For the past 50 years, several issues related to project management, measurement criteria, and factors contributing to the performance of projects have been widely studied. The implementation of a data mining is both very expensive and high risky. Many projects exceed their budgets, miss deadlines, and fail to meet end-user needs or other objectives (Nemati & Barko, 2003). According to a recent study of 1,311 Organizational Data Mining (ODM) professionals by The Data Warehousing Institute, the average corporate budget for Customer Relationship Management (CRM) projects was over $4 million yet regrettably a large percentage of these projects were experiencing difficulties (Eckerson and Watson, 2001). Surveys (Copperdale, 1995; Willcocks and Griffiths, 1994) claimed that this scenario applies to more than 50 percent of IT projects. Consequently, Data Mining projects have not faired so well either in a similar fashion.

Meanwhile, academicians and practitioners have considered data mining to be a valuable decision support tool; however, little systematic research has been undertaken to measure its success in organizations (Nemati & Barko, 2003). Much of the literature to date discusses its popularity, rapid growth, and the critical issues, but without an accurate measure of success, organizations cannot evaluate the success of their costly data mining projects.

Therefore, the development of such a measurement should be of top priority for people interested in the data mining phenomenon.

Value of data has been appeared as a critical corporate asset. Information technology systems play an important role in supporting strategic decision making. However, a lot of projects have exceeded their budgets, miss deadlines, and fail to meet end-user needs or other objectives (Nemati & Barko, 2003). Based on literature review, we study on how we can enable data mining project success in an organization – “what are the factors that influence the user satisfaction on data mining?” This research aims to identify the critical factors affecting user satisfaction of data mining.

1.4 Objective of Study

In line with problems illustrated above, this study aims to accomplish two main objectives as follows:

(1) To examine the satisfaction level of end-users using data mining system.

(2) To determine the critical factors that influence user satisfaction in data mining.

1.5 Research Question

In order to achieve the above-mentioned objectives, the following research questions are formulated:

(1) What is the satisfaction level of end-users using data mining at workplace?

(2) What are the critical factors that affect the user satisfaction in data mining?

1.6 Research Scope

In this study, the target population is employees from insurance, finance and banking, telecommunication, information technology and manufacturing industry and they must have more than one year experience using data mining at their workplace. The population of study covers employees from private sector, profit-making organizations located in Peninsular Malaysia only. This research is to identify the critical factors influencing user satisfaction of data mining.

1.7 Significance of Study

As data mining gains attention all over the world, it is important to determine the level of data mining satisfaction especially in the context of developing country like Malaysia. This study will assess the level of data mining satisfaction among firms in Malaysia and its critical success factors. The critical success factors are conceptualized to have direct relationships with the level of data mining satisfaction.

This study identifies five factors that are hypothesized to be most critical in predicting the level of data mining satisfaction among firms in Malaysia: system throughput, data quality, ease of use, information utility and support provided to end-user. The identified critical factors were developed from the syntheses of relevant previous studies.

This study aims to help executives and managers that are currently planning to adopt data mining system to better understand the critical factors influencing user satisfaction on data mining system. In turn, it also aims to help the managers to emphasize on and prioritize the implementation of each of the critical factors accordingly by addressing those factors that they are trailing before making the data mining updating and enhancing investment decisions.

1.8 Definition of Key Terms

For sharing common understanding of the concepts and better understanding of further discussion, the following key terms' definition are referred specifically.

1) **Data Mining (DM):** the process of discovering meaningful new correlations, patterns, and trends by sifting through large amounts of data stored in data warehouses and by using pattern recognition technologies as well as statistical and mathematical techniques. (Sugumaran and Bose, 1999).

2) **System Throughput:** Turnaround time in retrieving the information sought by end-users and consequently closely associated to its processing speed (Sen and Jacob, 1998).

3) **Data Quality:** Refers to the currency, accuracy, consistency and level of detail of data. (Bongsik, 2003).

4) **Ease of Use:** The degree to which a person believes that using a particular system would be free of effort (Davis, 1989).

5) **Information Utility:** System capability to support end-users to satisfy their information requirements (DeLone and McLean,1992).

6) **Support Provided to End-user:** Ability of Information System Department to deliver adequate training and support to the end-user (Chen, Soliman, Mao, and Frolick, 2000).

7) **User Satisfaction:** An affective state that is the emotional reaction to a product or service experience is influenced by a consumer's satisfaction with the product itself and with the information used in choosing the product (Spreng, MacKenzie, & Olshavsky, 1996).

1.9 Organization of the Remaining Chapters

This report is organized into five chapters. Chapter 1 provides an overview to the topic of interest, research problem, research objectives, research scope and significance of the study.

Chapter 2 presents the comprehensive literature reviews on previous studies; and ends with the development of theoretical framework and formulation of hypotheses. Chapter 3 focuses on describing the methodology deployed in conducting the research questionnaires development, measures, sampling design, data collection and analysis. Chapter 4 presents the statistical analyses and hypotheses testing. Finally, Chapter 5 concludes with a discussion of the findings, implication, limitation of the study, and provides some suggestion for future research.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

In order to better understanding of the present study, a comprehensive search of previous literature is carried out. To start with, this chapter introduces the concept of Data Mining. Then, it focuses on discussing related underlying theories, and the expansion of the theories to derive at the current theoretical framework and subsequently, the hypotheses development.

2.2 Data Mining (DM)

Data Mining is an important area of practice and research. Project management dimensions and factors affecting the implementation of information technology projects have been widely studied, yet few studies assessed its practices in general and critical success factors in particular (Nemati & Barko, 2003). Thus, the nature of the variables influencing the performance of Data Mining (DM) technologies is not well understood as well. Based on literature review, we study on how we can enable data mining project success in an organization – "what are the factors that influence the user satisfaction on data mining?" This study will help us to determine the user satisfaction factors and better understand the success factors in data mining.

In the intensely competitive global marketplace, the need for removing inefficiencies becomes even more critical to success. Knowledge management due to an inefficient information system was found as a prevalent obstacle that hindered the organizations' decision-making process in a recent survey of 41 top executives who explored the role information technology systems play in supporting strategic decision making (Hedelin and

Allwood, 2002). The formidable challenge facing organizations involves the collection, management and presentation of its data to enable management to make well-informed and timely decisions.

In recent years, organizations have realized the value of data as a critical corporate asset and the essential role it perform (Kumar and Palvia, 2001). Hence, organizations of all sizes are developing and deploying data mining technologies to leverage data-resources to enhance their decision-making capabilities (Eckerson, 2001). Multiple studies have revealed that executives rank their organizational data as a top priority (Nemati & Barko, 2003).

2.2.1 Introduction of DM

Data mining is a concept that has been establishing itself since the late 1980's. Data mining is the process of extracting valid, previously unknown, comprehensible, and actionable information from large databases and using it to make crucial business decisions (Sugumaran and Bose, 1999). Generally, data mining (i.e. data or knowledge discovery) can be viewed as a process of analyzing data from different perspectives and summarizing it into useful information. According to Berry and Linoff (1997), the obtained information can be then used to increase revenue, cuts costs or both. The importance of collecting data that reflect our business or scientific activities to achieve competitive advantage is widely recognized now. However, the bottleneck of turning this data into our success is the difficulty of extracting knowledge from the collected data (Bose & Mahapatra, 2001).

Data might be one of the most valuable assets for corporation. However, the important thing is that we must know how to reveal valuable knowledge hidden in raw data. Data mining allows us to extract diamonds of knowledge from our historical data and predict outcomes of future situations. It will help us optimize our business decisions, increase the value of each customer and communication, and improve satisfaction of customer with the necessary services. Data that require analysis differ for companies in different industries. That is, data mining can help us reveal knowledge hidden in data and turn this knowledge into a crucial competitive advantage. (Chen, Han & Yu, 1996).

2.2.2 Concept and Techniques of DM

Data Mining is a powerful new technology with great potential to help companies focus on the most important information in their data warehouses by extracting the hidden predictive information from large databases. Fayyad (1996) use the term Knowledge Discovery in Databases (KDD) to refer to the overall process of discovering useful knowledge from large datasets. KDD is defined as "the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data". Data mining is classified as the particular step in this process dealing with the application of specific algorithms for extracting patterns (models) from data. Data mining utilizes techniques such as neural networks, rule-based systems, case based reasoning, machine learning and statistical programs, either alone or in combination, to analyze and extract patterns from the data. KDD process comprises many steps, which involve data selection, data preprocessing, data transformation, DM (search for patterns), and interpretation and evaluation of patterns. The steps depicted start with the raw data and finish with the extracted knowledge, which was acquired as a result of the KDD process. The set of DM tasks used to extract and verify patterns in data is the core of the process.

2.2.3 Current Trend on DM

Data Mining has a wide array of technologies, including e-business intelligence, data analysis, SQL, CRM, e-CRM, EIS, digital dashboards, information portals, etc. The adoption of Data Mining technologies is increasing since they find their way into more mainstream applications. There are a number of reasons for this growing trend. Main reason is data mining enabling organizations to answer questions about the past (what has happened), the present (what is happening), and the future (what might happen). Moreover, data mining technology has provided most organizations with the ability to manage and analyze vast amounts of data.

Besides that, larger, faster, and cheaper storage technologies that house mountains of data such as Web click streams, telephone calls and credit card transactions have been a key factor in the growth of Data Mining over the past few years (Hedberg, 1999). As an example, a small skin-care products vendor developed a predictive model to raise the efficiency of its mail-order business by reducing catalogs mailed and increasing response rates. Furthermore, the retailer reduced catalog circulation by 50 percent in 2001 whereas increased revenue per catalog by almost 20 percent by intelligently targeting consumers who were prefer to buy (Whiting, 2002). Lastly, the recent economic slump is forcing organizations to squeeze more return out of their telemarketing, advertising, sales and other marketing efforts.

2.3 Model Delone and Mclean (2003)

The theoretical notion of information systems success has formed the basis of this research. Numerous researches have been conducted to investigate the concept from different angles including its measurement and representation (Pitt., Watson & Kavan, 1995; Raghunathan & Raghunathan, 1994; Segars & Grover, 1998), and the technological, end-user, and organizational variables associated with it (Li 1997; Yap, Soh, & Raman 1992; Yoon, Guimares & Oneal 1995). These studies suggested that information systems success is a multidimensional concept that includes system factors (i.e. ease of use and throughput), data factors (i.e. data quality), organizational factors (i.e. policy, training and support), and user factors (i.e. user satisfaction, individual impact and utility) (Jain 1997; Li 1997; Williams & Ramaprasad 1996).

To explore potential success factors of the data mining, a part of DeLone and McLean's (2003) IS success model was applied in this research since data mining is a special type of Information System. DeLone and McLean (1992) made an important step towards consolidation of prior research. They introduced a model of information systems success based on a study of more than 180 published papers, which addressed the issue of information systems success. DeLone and McLean identified six major dimensions of success which were system quality, information quality, use, user satisfaction, individual impact and organizational impact. The Delone and McLean model of information system success has attracted much attention amongst researchers. From 1993 to mid 2002, nearly 300 articles had referenced and cited the DeLone and McLean model in refereed journals and researches from past literature on information system success. However, as DeLone and McLean argued 'this success model clearly needs further development and validation before it could serve as a basis for the selection of appropriate IS measures.'

Due to the original IS success model needed further validation, DeLone and McLean proposed an updated model in 2003. They added Service Quality (e.g. IS support) as one important dimension. In addition, they added Intention to Use as an alternative measure because an attitude is worthwhile to measure in some context. Finally, DeLone and McLean (2003) had extended and streamlined the original model called 'Net Benefits' to broaden the impacts of IS also to groups, industries and nations, depending on the context as depicted in figure 2.2.

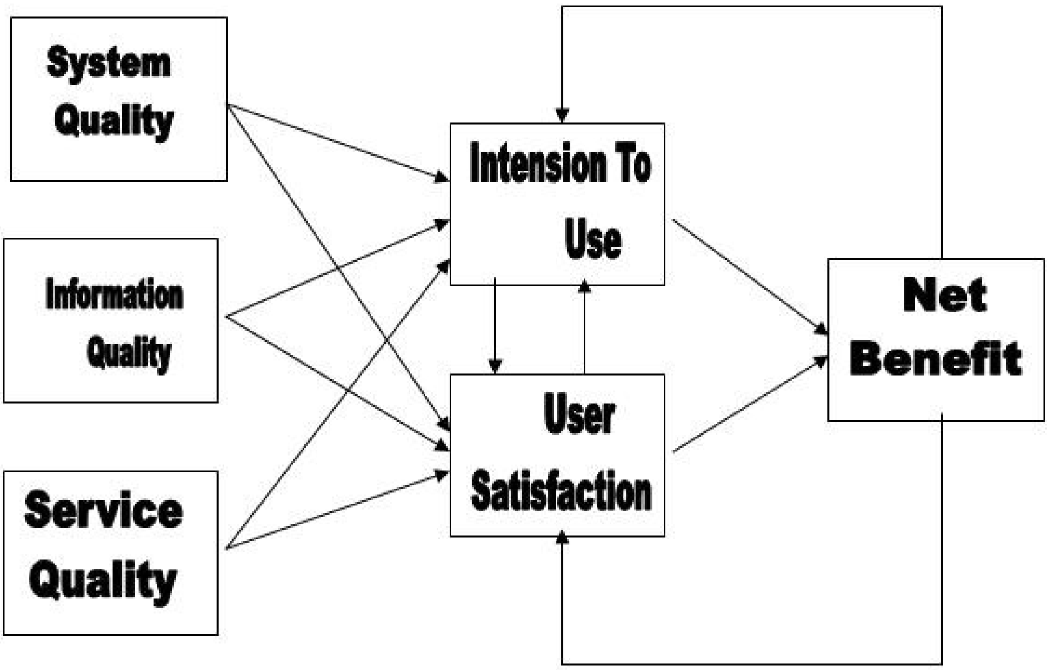

Figure 2.1 **Updated DeLone and MeLean IS success model**

(DeLone & McLean, 2003)

Hence, the dimensions used in this research are System Quality (System Throughput, Data Quality and Ease of Use), Information Quality (Information Utility), Service Quality (End-User support) and User Satisfaction.

2.4 User satisfaction

As studying such a wide range of issues for an emerging technology is difficult, my research focuses only on user satisfaction. User satisfaction has been extensively acknowledged as a salient dependent variable to determine an information system's success (Delone & McLean, 1992). Thus user satisfaction is an important requirement in that it comprises one of the main considerations upon which managers can take corrective measures to raise system acceptance by end-users. Powers and In addition, Dickson (1973) studied the factors affecting the success of management information systems. They identified user satisfaction as one of the key factors affecting management information systems success.

DeLone & McLean (2003) identified three reasons why user satisfaction has been widely used as the single measure of IS success: (1) High degree of face validity; (2) Development of reliable tools for measure, and (3) conceptual weakness and unavailability of other measures. Hence this variable was applied here to investigate the relative impact of the selected independent variables on the user satisfaction in data mining.

Satisfaction is a subjective measure, usually self-reported by users of their impressions and feelings about the outcome (solution satisfaction) and process (process satisfaction) used to reach the course outcome (Ocker & Yaverbaum, 1990). According to Bond and Fink (2003), user satisfaction dimension is recognized in the management and marketing literature as one of the indicative measure closely tied to product and service quality. In addition, Kotler and Keller (2006) defined satisfaction is a person's feelings of pleasure or disappointment result from comparing a product's perceived performance or outcome in relation to his or her expectations. The user satisfaction gauges opinions of the users about Data Mining System based on their previous or current experience with the system. This is rated on perceptions of satisfaction, enjoyment, success and recommends ability.

Few studies from previous researchers have identified contributors that influence user satisfaction in information system. These are quality of the information content and feedback (Swan, 2001; Bolliger & Martindale, 2001). In addition, a minimum of technical problems with the delivery is another factor (Serban, 2000; Bolliger & Martindale, 2001; Arbaugh, 2000). This is correct and logical as it would reduce the anxiety and frustration with the Data Mining System which is one of source of dissatisfaction.

From the many past research above focusing on user satisfaction, it is unquestionable that user satisfaction is an important measure in Information System as well as data mining environment.

2.5 System success factors in data mining

Next, we look at what are the factors influencing user satisfaction on data mining. There are various factors that determine the user satisfaction on data mining.

2.5.1 System Throughput

For system quality, it is "concern with whether or not there are bugs in the systems, the consistency of user interface, ease of use, response rate in interactive system, documentation and sometimes, quality and maintainability of the program code" (Seddon & Kiew, 1994 cited in Almutairi & Subramanian, 2005). Among many potential variables of system quality, this survey included system throughput, ease of use and data quality, which were regarded as crucial for the success of the data mining. System throughput of the data mining is the turnaround time in retrieving the information sought by end-users and consequently closely associated to its processing speed (Sen & Jacob, 1998).

Some researchers have found that lengthy system response times may cause lower satisfaction and poor productivity among users (Shneiderman, 1998). According to John and Chris (2000), if the system responds within acceptable limits, dissatisfaction will be minimal (John & Chris, 2000). The results clearly show a drop in user satisfaction as the response time increases (john & Chris, 2000).

According to Bonsik (2003), queries submitted to the data mining, in general, require extensive computing because of their decision support nature if compared to those of transactional databases. On the other hand, too much processing may force users to abandon the data mining because it may not be the only information source and its usage may not be as mandatory since operational systems are serving daily transactions. DeLone and McLean (2003) acknowledge that the degree of system use may constitute a good indicator for system success.

On the other hand, in the study conducted by Chen et al. (2000), the result indicated that there are no significant relationships between system throughput and user satisfaction in the context of information system success. According to Chen et al. (2000), the nature of data mining is supporting strategic decision making and users often assume that it will take some time before getting the answer of their enquiries. Thus, the requirement on the response time is not as critical as quality of the result.

In summary, system throughput is crucial in data mining environment. Employees will be getting frustrated and their productivity will be low if the system response time is slow. Therefore, quicker system response time is believed to have higher satisfaction on the data mining system.

2.5.2 Data Quality

Data quality is the quality of data. Data are considered as high quality if they accurately represent the real-world construct to which they refer.

Data quality is measured as the level of accuracy in the data. Data quality can be affected by missing values, erroneous metadata and incorrect values (Nemati & Barko, 2001). Examples of bad data are customer records with negative ages, males designated as females, incorrect address information, etc. Nemati and Barko have suggested that the level of data quality and data integration is expected to have a positive influence on data mining project outcomes.

Since there is a considerable amount of data quality research involves investigating various categories of desirable attributes (or dimensions) of data, four data quality variables (currency, accuracy, consistency, and level of detail) are included in the survey. Data currency, accuracy, and consistency are considered some of the most important attributes of data quality (Fox, Levitin, & Redman, 1994; Wang & Strong, 1996). Fox et al. (1994) suggested that level of details (or granularity) reflects another quality feature of data representation.

Wand, Y. and Wang, R. claimed that inconsistent data is a problem with data quality as well and it is not just only arise from incorrect data. In order to ensure data consistency, company can take inititiatives to eliminate data shadow systems and centralize data in data mining. Morever, the market is improving data quality assurance such as making tools for analyzing and repairing poor quality data, cleaning the data on a contract basis and advising on fixing processes or systems to avoid data quality problems in the first place.

In reality, data quality is a concern for professionals involved with wide range of information system which is from data warehousing and business intelligence (data mining) to customer relationship management and supply chain management. Eckerson (2002) has revealed that the estimation of the total cost to the US economy of data quality problems at over US $ 600 billion per annum. Consequently, the problem is so serious that companies are beginning to build a data governance team to be responsible for data quality in the corporation. In fact, in some organizations, this data governance team has been established as part of a larger Regulatory Compliance function which is recognition of the importance of Data/Information Quality to organizations.

In addition, companies with an emphasis on marketing often focus their quality efforts on name and address information. However, data quality is recognized as an important property of all types of data. Principles of data quality can be applied to supply chain data, transactional data, and nearly every other category of data found in the enterprise

Pikkarainen, Pikkarainen, Karjaluoto. and Pahnila (2008) argued that data quality is intuitively important as a factor in analyzing user satisfaction. The result of their study found that the higher quality of the data, the higher satisfaction among the users in the information system environment.

DeLone and McLean (1992) and Atkinson (1999) each claimed data quality and its use are linked to project success. Other studies have revealed that data are an important part of a decision support system since they form the basis of the information that is delivered to decision makers (Watson, Rainer & Koh, 1991).

Chen et al. (2000) investigated the factors influencing the user satisfaction on data warehouse which is one type of Information System. However, the results showed that there is no significant relationship between data quality and user satisfaction.

In summary, data quality plays an important role in the data mining environment. System which provides accurate, current, consistent, and detail data will support employees in making better decision. As a result, it will increase user satisfaction among employees.

2.5.3 Ease of Use

Perceived ease of use has also been frequently identified as an important indicator for information systems acceptance by end-users (Adams, Nelson, & Todd, 1992; Davis 1989). Davis (1989) introduced technology acceptance model (TAM) to predict user's acceptance of IT. He also proposed two important constructs - perceived usefulness and perceived ease of use.

According to Davis (1989), perceived ease of use is "the degree to which a person believes that using a particular system would be free of effort”. In the TAM Model, complexity is referred to as perceived ease of use. The perception of a system’s ease of use may significantly affect the level of its adoption by prospective users due to the complexity of data processing for decision support. System which is less complicated and easy to learn will increase adoption by users and the user satisfaction level. John and Chris, (2000) also suggested that how “easy” an application is to use may be a factor that can determine the user’s satisfaction level.

Information system theory stated that system quality is the user’s judgment of dealing the system familiarly and easily (Rai, Lang, & Welker, 2002). Some researchers have acknowledged that system quality in information systems success model is substituted for perceived ease of use (Rai, Lang, & Welker, 2002; Seddon & Kiew, 1997). Perceived ease of use can be justified as the perception of system usage effort since it is an important variable in information system attitudes (Davis, 1989).

Ease of use of data mining system is very crucial in workplace. The feature of system such as easy operating the system, easy understanding instruction for employees to get the information they need easily and quickly etc will definitely enhance their satisfaction in using the data mining system. Hence, this study extended existing work by investigating the implication of this variable in the data mining context.

The results from the study by Pikkarainen et al. (2008) indicated that ease of use was positively related to user satisfaction in the information system environment. In addition, Liu., Chen, & Zhou (2006) also found that there was significant relationship between perceived easy of use and user satisfaction. The results showed that users’ overall satisfaction, specially explained by perceived easy of use, most significantly affect their intention to use information system service.

On the contrary, there were also mixed results from previous studies in examining the influence of ease of use in the computing environment. In the study by Chen et al. (2000) to examine the factors influencing data warehouse user satisfaction, ease of use was found to be not related to user satisfaction. According to Chen et al. (2000), a possible explanation could be that users of data warehouses are mainly business managers with a reasonable amount of business training and a higher educational background than the majority of end-users; thus, the ease f use of IS seemed to carry less weight.

In summary, ease of use is equally important in the data mining environment as a factor in influencing the user satisfaction of the data mining system. System with easy navigation, clear instruction and online help menu will help employees to access the information they need easily and quickly. Thus, this will increase their productivity which in turn enhance their satisfaction when using data mining system.

2.5.4 Information Utility

In the DeLone and McLean (2003) information systems success (ISS) model, information quality is one of the six dimensions identified. "Information quality" is a measure of the value which the information provides to the user of that information. Information quality is described as a quality of system outputs of the product. Information quality dimension is "concern with such issues as timeliness, accuracy, relevance and format of information generated by information system" (Seddon & Kiew, 1994 as cited in Almutairi & Subramanian, 2005). To encourage repeat visits, visitors need to be provided with appropriate, complete and clear information (DeLone & McLean, 2003).

Information quality dimension consists of utility of information. Information utility of a system means its capability to support end-users to satisfy their information requirements (DeLone & McLean,1992). The information utility evaluates the Data Mining System whether it has the current, comprehensive and correctness in content.

Content is considered to be most important element of system (Turban & Gehrke, 2000) and is seen to be directly related to system success (Liu & Arnett, 2000). Content that has the expected of quality by users will lead to user's satisfaction. Seddon and Kiew (1994) as cited in Delone and Mclean (2003) have found that there was a significant relationship between Information Utility with User Satisfaction through a survey conducted on 104 users of university accounting system.

In regards to Information Quality Dimension, the information utility evaluates the data mining system whether it is useful, critical to the task and enhancing the employees' productivity with the information provided. As a result, this variable is closely associated with the degree to which using a particular system could enhance the user's job performance (Davis, 1989).

Pikkarainen et al. (2008) found that information utility can be utilized in analyzing user satisfaction with information system. Another study conducted by Liu et al. (2006) indicated that user satisfaction is affected through information utility (perceived usefulness) which in turn most significantly affects their intention to use IT service.

In summary, information utility is essential in the data mining environment. Information which is useful and critical to the employees' task will enhance their job performance and productivity. Hence, it will make employees more satisfied when using the data mining system.

2.5.5 Support provided to end-user

Many researches emphasized the importance of service quality for Information Systems success (Barquin and Edelstein 1997). Support provided to end-user, as a representative service quality variable, has been repeatedly investigated as an effective way to attract potential users, enhance their understanding on the subject system, and increase user satisfaction (Chen et al., 2000).

Due to the complexity of a data mining system and its data structure, end user training could be very crucial for its successful adoption and company-wide diffusion. Considering the fact that data mining is fairly new to many organizations, Information System (IS) Department should deal with the data mining project when the system is developed in the initialization stage. Thus, the end-users' performance and satisfaction is highly dependent on the training and support provided by the IS Department (Chen et al., 2000).

Chen et al. (2000) also revealed that success of the information system project depends on both the IS department and the end-users. User satisfaction depends on IS Department's ability to deliver adequate training and support to the end-users. On the other hand, the end-users have to understand the meaning of the data included in the data mining (King, 1996). Accordingly, it is the IS Departments' responsibility to increase the awareness among the end-users regarding the data in the data mining. ICs must understand the differences among the end-users: their ability, functions, and needs. Ford (1996) also found that demographic factors, prior computer training, and experience had significant impacts on the participants' use of computers.

According to Ettlie (1998), both formal training and regular review sessions are necessary to ensure the managers and employees are always updated with ongoing system and process changes. Definitely, lack of user training and failure to properly understand the changes of business process in enterprise application often seem to be responsible for failures and problem during the implementation (Wilder & Davis, 1998 and Crowley, 1999). A research done by Bingi, Sharma, and Godla (1999) reported that about 30% to 40% of front-line workers will not be able to handle the demands of a new information system that due to lack training.

Normally training starts with the education of the project team in system, line, project management and ends with system's users (Welti, 1999). User participation in the development process was found to be crucial for user satisfaction (Powers and Dickson, 1973). Nemati and Barko (2001) also claimed that there is a positive influence from end user expertise and involvement on data mining project implementation. Liu et al. (2006) suggested that increasing the relevant knowledge and skill of the users effectively enhance their satisfaction and the actual usage as well.

In summary, support provided to end-user is crucial in data mining environment. Information System Department plays an important role in providing adequate support and training to end-user. Employees will be more confident and satisfied when using data mining system if they are provided proper training and preferred support.

2.6 Theoretical Framework

The purpose of this study is to determine the critical factors that influence the satisfaction level of employees using data mining system in the organizations. The Updated Delone and Mclean's (2003) information system success model was used for the development of the theoretical framework for this research. Based on the comprehensive study of literature above, a framework to guide this study is developed, as shown in Figure 2.3. There are five factors (independent variables) which are believed to be most significant to influence the user satisfaction (dependant variable) in using data mining system: System Throughput, Data Quality, Ease of Use, Information Utility and Support Provided to End-user.

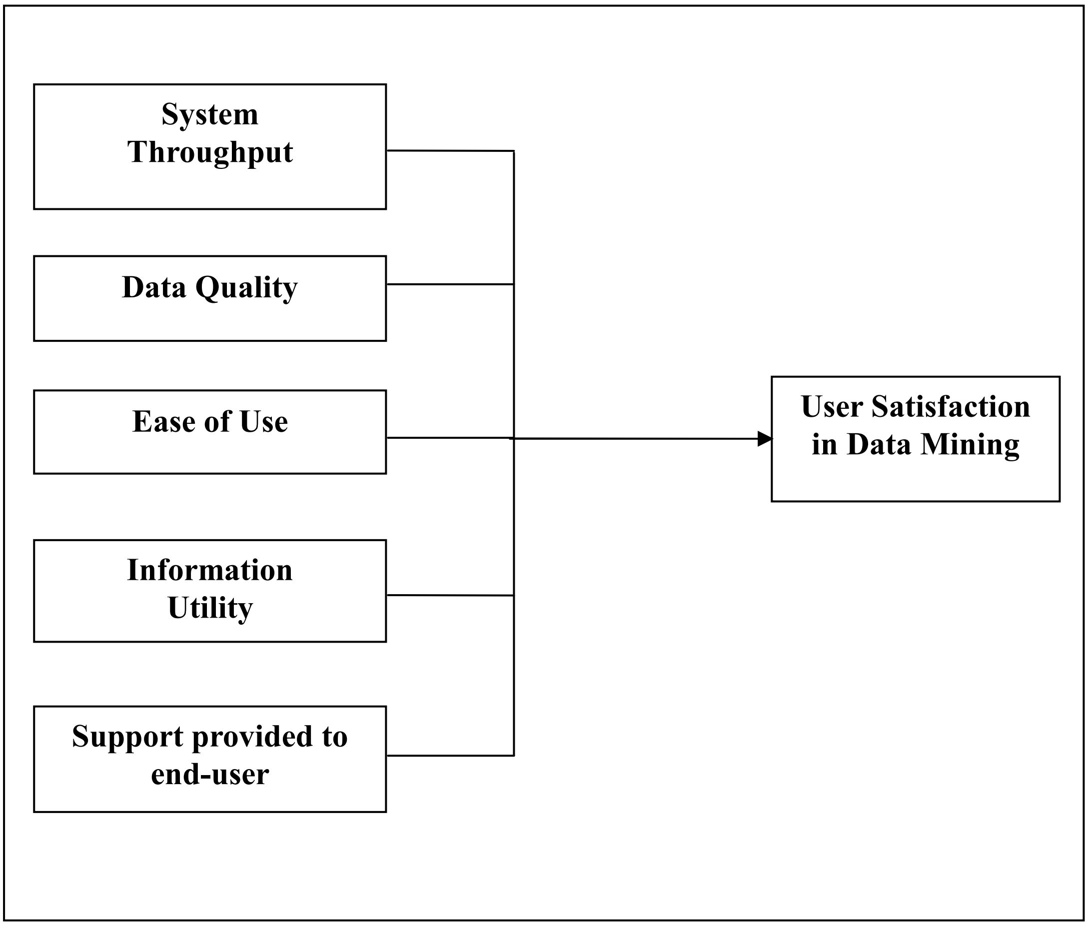

Figure 2.2 Theoretical Framework

2.7 Development of Research Hypothesis

Based on the theoretical framework shown above, five hypotheses will be tested in this study.

2.7.1 System Throughput and User Satisfaction

Previous research has found that lengthy system response times may cause lower satisfaction