

# Benefits and Barriers in Mining the Healthcare Industry Data

*John Wang, Department of Information & Operations Management, Montclair State University, Montclair, NJ, USA*

*Bin Zhou, College of Business, University of Houston-Downtown, Houston, TX, USA*

*Ruiliang Yan, Department of Management & Marketing, Texas A&M University-Commerce, Commerce, TX, USA*

---

## ABSTRACT

*The authors' paper addresses the applications of data mining within the healthcare industry. Healthcare data are seen as one of the more rewarding and most difficult of all data to analyze. Proper data mining techniques provide the methodology and technology to transform the voluminous amounts of data into useful information for decision making. Data mining can be utilized to help find cures for existing diseases, uncovering patterns for genetic diseases and the causes of new diseases across the globe. By implementing data mining techniques the industry is finally gaining control over the inadequacy of readily available records. Data mining has been used in patient care, healthcare plans, and administration. By utilizing these methods, hospitals and healthcare insurance providers alike are able to save millions of dollars, administration headaches, and most importantly, countless lives.*

*Keywords: Clinical Care, Data Mining, Disease Management, Fraud Detection, Healthcare, Patient Profiling, Treatment Effectiveness*

---

## INTRODUCTION

Data mining is the science of managing and analyzing large datasets and discovering novel patterns (Ngai, Hu, Wong, Chen, & Sun, 2011; Wang, 2008). The progress of data mining has been around for decades. Although data mining is an evolution of a field that can be traced back many years ago, the terminology itself has been only been introduced in the 1990's. By using pattern recognition technologies and statistical and mathematical techniques to sift through

warehoused information, data mining helps analysts recognize significant facts, relationships, trends, patterns, exceptions and anomalies. The use of data mining can advance a company's position by creating a sustainable competitive advantage. Day to day operations of businesses in the modern healthcare industry create massive flows of valuable information in regards to their patients. Without a well-designed method of compiling data collected by physicians during their practices, it would serve little or no purpose in aiding fellow practitioners. Traditionally, decision making in health care is based on ground information, lessons learned in the past,

DOI: 10.4018/jsds.2012100103

resources and funds constraints. However, data mining techniques can be applied to create a knowledge rich healthcare environment (Kaur & Wasan, 2006; Vallejo et al., 2012). Over the past few decades the healthcare industry has been faced with the problems that are directly related to the inefficient storage of and access to data among their facilities. Huge amounts of data are generated daily in the processing of healthcare transactions from all healthcare providers, such as hospitals, clinics, physicians, patients and even insurance providers. Data mining is becoming more prevalent in the health care industry because of the vast quantities of data stored in a multitude of medical systems, more specifically systems of health care providers, hospitals, and other medical institutions (Veletos, 2003; Harris et al., 2012).

Healthcare has a diagnosis for its ailments: inefficiency. But the cure is not so straightforward. While organizations are becoming information-rich, they remain knowledge-poor, without the tools to better integrate information across their clinical, operations, research and financial systems. But rather than focus on what is wrong with healthcare, it would be more productive to imagine how we might make a smarter healthcare system. Connecting doctors, patients and insurers to share information seamlessly and securely should be the ultimate goal. That means that providing a smarter healthcare system is optimized around the patient to increase efficiency, reduce errors, to achieve better quality of clinical care and ultimately save more lives (Christensen & Oldenburg, 2009; Sharkey, Hsu, Batra, & Rigamonti, 2011).

Recognizing patterns of data in order to discover valuable information, new facts, and relationships among variables are important in making business decisions that would best minimize costs, maximize returns, and create operating efficiency without compromising the quality of patient care. Data mining has revolutionized the way the healthcare industry manages its information. It is helping to decrease the nearly 71,000 people who die in U.S. hospitals as a result of medical error. By implementing data mining in health insurance

plans, insurers have been able to identify more false claims than ever before, which allow them to save millions of dollars. In addition, data mining systems have eased the pains and efforts of operating a hospitals' health care administration (Manchur, 1998; Lee et al., 2011).

The workflow of healthcare organizations involves the generation and collection of various kinds of data relating to clinical practices, clinical trials, patient information, resource administration, policies and research. Traditionally, statistical techniques are used to derive some operational information from the data. Data mining provides the opportunity to derive, in an exploratory and interactive manner, valuable healthcare knowledge in terms of associations, sequential patterns, classifications, predictions and symbolic rules. Such inductively derived healthcare knowledge can provide strategic insights into the practice delivery of healthcare (Kumar, 2012).

As the medical field expands, it is the duty of each physician to evaluate and protect each patient from diseases, side effects and medical mishaps. Armed with a scalpel, stethoscope and other accrements, physicians are now armed with data mining as a tool for expanding their knowledge base. Data mining is available to every aspect within the healthcare industry. It is multifaceted and used in areas like insurance to detect fraud, the pharmaceutical industry to evaluate side effects of drugs and even detection of certain diseases based on genetics (Negev et al., 2012).

Data mining applications can benefit the healthcare industry with endless possibilities of overall healthcare improvement, aiding in better diagnosis and improved clinical care. However, the data applications are not without substantial limitations. Healthcare data mining is dependent on accumulating data information from different settings and systems, often on non-standardized data. The quality of data is also impacted by missing information and information recorded in different formats (Koh & Tan, 2005). The lack of standard clinical vocabulary and the need for greater data warehousing in the

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/benefits-barriers-mining-healthcare-industry/74355](http://www.igi-global.com/article/benefits-barriers-mining-healthcare-industry/74355)

## Related Content

---

### Top Management's Role in Promoting Decision Support Systems Efficiency: An Exploratory Study in Government Sector in Saudi Arabia

Abdullah Ibrahim Alkrajji (2021). *Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering* (pp. 1409-1429).

[www.irma-international.org/chapter/top-managements-role-in-promoting-decision-support-systems-efficiency/282648](http://www.irma-international.org/chapter/top-managements-role-in-promoting-decision-support-systems-efficiency/282648)

### Towards an Assessment of Public Library Value: Statistics on the Policy Makers' Agenda

Kristine Paberza (2010). *International Journal of Decision Support System Technology* (pp. 42-51).

[www.irma-international.org/article/towards-assessment-public-library-value/43910](http://www.irma-international.org/article/towards-assessment-public-library-value/43910)

### IDSSE-M: A Software System Engineering Methodology for Developing Intelligent Decision-Making Support Systems

Manuel Mora, Fen Wang, Ovsei Gelman and Miroljub Kljajic (2013). *Engineering Effective Decision Support Technologies: New Models and Applications* (pp. 43-71).

[www.irma-international.org/chapter/idsse-software-system-engineering-methodology/75687](http://www.irma-international.org/chapter/idsse-software-system-engineering-methodology/75687)

### Decision Support in Participatory Contexts: The Organisational Design Dimension

Mazri Chabane, Alexis Tsoukias and Katherine A. Daniell (2019). *International Journal of Decision Support System Technology* (pp. 47-80).

[www.irma-international.org/article/decision-support-in-participatory-contexts/230316](http://www.irma-international.org/article/decision-support-in-participatory-contexts/230316)

### Deploying Decision Support Systems Using Semantic Web Technologies

Lars Ludwig and David O'Sullivan (2010). *International Journal of Decision Support System Technology* (pp. 49-59).

[www.irma-international.org/article/deploying-decision-support-systems-using/40918](http://www.irma-international.org/article/deploying-decision-support-systems-using/40918)