Chapter XV

Data Mining in Health Care Applications[1]

Fay Cobb Payton
North Carolina State University, USA

ABSTRACT

Recent attention has turned to the healthcare industry and its use of voluntary community health information network (CHIN) models for e-health and care delivery. This chapter suggests that competition, economic dimensions, political issues, and a group of enablers are the primary determinants of implementation success. Most critical to these implementations is the issue of data management and utilization. Thus, health care organizations are finding value as well as strategic applications to mining patient data, in general, and community data, in particular. While significant gains can be obtained and have been noted at the organizational level of analysis, much attention has been given to the individual, where the focal points have centered on privacy and security of patient data. While the privacy debate is a salient issue, data mining (DM) offers broader community-based gains that enable and improve healthcare forecasting, analyses, and visualization.

INTRODUCTION

In this chapter, I provide general insight into data mining with emphasis on the health care industry. My discussion focuses on earlier electronic commerce health care initiatives, namely community health information networks (CHINs) in three regions of the United States. From my exploration of the implementation issues that impacted the success of each case, I offer Figure 1 to illustrate a cluster of factors that capture the dynamics of these live scenarios.

While my prior work (Payton & Ginzberg, 2001) focused on the information technology implementation process, here I provide insight regarding the mining of health

data centered about new e-health models of care delivery. One such model for health care was the CHIN, which was a widely sought after model in the late 1990s, and continues to be widely debated by leading industry groups, such as The Healthy Cities Organization (http://www.healthycities.org) and the IEEE-USA Medical Technology and Policy Committee (http://www.ieeeusa.org/committees/MTPC).

While CHINs were said to enable multiple organizations to share health services data in order to meet common objective(s), ranging from profit maximization to improvement of public health conditions and wellness, health care organizations are now deploying e-health strategies to accomplish these goals while garnering data to address community and public health needs. Thus, health care organizations are finding value as well as strategic applications to mining patient data, in general, and community data, in particular. While significant gains can be obtained and have been noted at the organizational level of analysis, much attention has been given to the individual where the focal points have centered about privacy and security of patient data. While the privacy debate is a salient issue, data mining offers broader community-based gains that enable and improve health care forecasting, analyses, and visualization.

Thus, the objectives of this chapter are:

1)   To describe the uses of data-mining techniques in health care;
2)   To explore the intersection of data mining in the context of CHINs;
3)   To explain how the CHIN platform can be used for fraud detection, profitability analysis, patient profiling, and retention management;
4)   To discuss emerging and future trends; and
5)   To explore the technical and social issues that remain.

DATA MINING IN HEALTH CARE

Data mining has often been defined as a "process of extracting previously unknown, valid and actionable information from large databases and then using the information to make critical business decisions" (Cabena, Hadjinian, Stadler, Verhees, & Zanasi, 1998). This definition is based on the premise that an organization can link its myriad sources of data into a data warehouse (to potentially include data marts). Further, these data sources can evolve to a higher degree of analyses to include exploration using on-line analytical processing (OLAP), statistical analyses, and querying. One mechanism that can be used to migrate beyond exploration and permit organizations to engage in information discovery is the Sample, Explore, Modify, Model and Assess (SEMMA) Methodology, as developed by the SAS Corporation (http://www.sas.com).   In sum, SEMMA enables its users to: 1) test a segment of data (test data) to be mined for preliminary outcomes; 2) assess this sample data for outliers, trends, etc.; 3) revise the sample data set for model testing, which can include neutral networks, linear regression, and decision trees; 4) evaluate the model for accuracy; and 5) validate the model using the complete data set (Groth, 1998).

In general, the rationale for mining data is a function of organizational needs and type of firm's role in the industry. Moreover, researchers (Hirji, 2001; Keim, 2001) have established that data-mining applications, when implemented effectively, can result in strategic planning and competitive advantage. In an effort to minimize the collection and

14 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/chapter/data-mining-health-care-applications/7608

Related Content

Data Discovery Over Time Series From Star Schemas Based on Association, Correlation, and Causality
Wallace Anacleto Pinheiro, Geraldo Xexéo, Jano Moreira de Souzaand Ana Bárbara Sapienza Pinheiro (2020). *International Journal of Data Warehousing and Mining (pp. 95-111).*
www.irma-international.org/article/data-discovery-over-time-series-from-star-schemas-based-on-association-correlation-and-causality/265259

Information Retrieval in the Hidden Web
Shakeel Ahmed, Shubham Sharmaand Saneh Lata Yadav (2021). *New Opportunities for Sentiment Analysis and Information Processing (pp. 50-71).*
www.irma-international.org/chapter/information-retrieval-in-the-hidden-web/286904

Emotion-Drive Interpretable Fake News Detection
Xiaoyi Ge, Mingshu Zhang, Xu An Wang, Jia Liuand Bin Wei (2022). *International Journal of Data Warehousing and Mining (pp. 1-17).*
www.irma-international.org/article/emotion-drive-interpretable-fake-news-detection/314585

Computing Join Aggregates Over Private Tables
Rong She, Ke Want, Ada Waichee Fuand Xu Yabo (2008). *International Journal of Data Warehousing and Mining (pp. 22-41).*
www.irma-international.org/article/computing-join-aggregates-over-private/1816

Approaches and Applications for Sentiment Analysis: A Literature Review
M. Govindarajan (2022). *Data Mining Approaches for Big Data and Sentiment Analysis in Social Media (pp. 1-23).*
www.irma-international.org/chapter/approaches-and-applications-for-sentiment-analysis/293147