

Preface

We are in the midst of a drastic transformation in the realm of commerce from that which prevailed over the past 50 years. Innovations in telecommunications, computer processing, and software technology have helped create the Information Economy. This term generally refers to the increased utilization of various forms of information technology (IT) to capture, store, extract, manipulate, analyze and communicate data and information of all forms by firms across industry sectors. As a result, organizations around the globe have greater accessibility to increased amounts of information than any time in the past. Because of the complementary nature of the IT spectrum mentioned above, firms can better transform vast amounts of data into a more vital asset, information, that ultimately enhances the knowledge level of individuals across functional areas of an organization.

As the information economy has evolved, the noteworthy progression of which began in the mid-1990s, economists, analysts and business leaders have devoted time and effort to identifying how the implementation of IT can increase the efficiency or productivity of a given enterprise. Many have referred to such innovations that have replaced factors of production in a direct sense, (e.g., labor displacing technology such as ATMs) as a primary driver of productivity growth. However, another source of corporate productivity comes in the form of “reducing the uncertainty of the business process.” This idea refers to the process of accessing value added, firm-relevant information in a timely manner. The availability of accurate information enables decision makers across functional areas to better understand the important factors that impact the bottom line of their activities. A clearer picture of those factors enhances their ability to devise and implement policies that more accurately address the problems of a given process or augment successful processes to new levels.

The proper utilization of information technology therefore increases the overall “business intelligence” of a given organization. Enhanced business intelligence helps reduce the uncertainty of those issues that really affect day to day operations at the firm level. Keep in mind, however, that one of the pitfalls of the evolution of the information economy has been the proliferation of a variety of buzzwords and phrases that depict nothing more than a rehashing of commonly accepted practices. Does business intelligence fall into this category? The answer is no, for all one needs to do is analyze the growth, innovation and implementation of the spectrum of technologies that comprise this space to see the dynamic and tangible value added it provides to corresponding organizations. Firms of all sizes and

industry types are utilizing these technologies to help augment their operations to compete, survive and thrive in this new dynamic economy.

Data Mining and Business Intelligence: A Guide to Productivity helps describe the process by which firms can increase their efficiency by implementing state-of-the-art IT. More specifically, it focuses on the high-end analytical software technologies, referred to as data mining, and how this technology, along with other applications such as On Line Analytical Processing (OLAP), can help decision makers extract information and knowledge from the vast amounts of data they collect on a day by day and minute by minute basis. This work is not written in a technical style but rather addresses the applied methodology behind properly implementing data mining techniques in the corporate environment. It provides an introduction to where the technology evolved (its theoretical base), an overview of the dominant methodologies that comprise the data mining spectrum and every day business applications where it can produce a value added. By doing so it bridges the gap between the important theoretical academic world and that of the applied side of the business environment. As was mentioned previously, we are undergoing a transformation in the world of commerce, which involves the evolution of e-commerce. This work has not ignored this growing phenomenon and addresses the issue of data mining in an e-commerce environment as well, connecting the more traditional “brick and mortar” firm structure to the growing “click and mortar” enterprise.

Chapter 1 provides the reader with an overall background of what the information economy and productivity at the firm level entails, which avoids a common problem that other works many times experience. Many works loosely refer to complex terminology such as the information age and productivity without providing an adequate description to ensure the audience’s understanding of such principles. *Data Mining and Business Intelligence: A Guide to Productivity* not only connects IT to business but relates it to the theoretical underpinnings of where it originated. After providing the audience with this information, it then delves into a description of business intelligence and the technology that comprises this spectrum. The focus turns to a particular sector of the business intelligence spectrum, namely data mining, and offers an introduction to the various methodologies included in this “high-end” analytical area. Finally, it provides an understanding of how the various forms of complementary information technologies work together and with the addition of business strategy, can help increase productivity for a given business entity.

Once the reader has a general idea of what data mining, business intelligence and productivity are about, this work then provides a more in-depth description of the theoretical base behind some of the methodologies along with a more technical description of the main mining approaches. In order to illustrate the applicability of the various mining methodologies in a business environment some common examples are offered. At this point, the reader should have a good understanding of the concept of data mining, but to drive the issue home, we provide a definition of the technology and classify approaches according to corresponding functionality. At the end of Chapter 2, the reader should be better equipped to undertake the process of conducting a “high-end” analytical application. However, before one can take that step there is one more major requirement to consider.

As is the case with any new technology, individuals must learn the important steps to success for its proper implementation in order to achieve maximum benefit. Data mining technology is grounded in mathematics and statistics and generally requires a particular skill base in order to achieve appropriate results. Chapter 3 provides the steps to success and pitfalls to avoid when conducting a mining application. It then offers an introduction to a widely accepted productivity enhancing technique referred to as Six Sigma and illustrates parallels to data mining applications and the reduction of business uncertainty and operational variances.

In chapter 4 we begin to focus on using various mining techniques to help solve specific problems in the business environment. This section introduces such applications as time series forecasting and cross sectional analysis, methods which examine trends over time and over a “snap shot” in time. It provides some basic, non-mining approaches to time series forecasts and steps up to the more sophisticated methods, and offers common examples to fully illustrate the applicability of the technology. A more in-depth view of cross-sectional methods are addressed, which include examples involving B2B and CRM applications.

The following chapter focuses on econometric based mining which includes regression and neural network analysis. Since previous chapters of this book provided the background behind these techniques, Chapter 5 focuses more on some essential business applications which are common to organizations across industry sectors. These include in-depth illustrations of how the high-end data mining or information mining techniques can be used to measure advertising, marketing and promotional impacts on a firm’s bottom line and provides important insights on pricing strategies. Finally, to illustrate the diverse applicability of these mining methods, we offer a number of common business scenarios in which these approaches can increase the users understanding of those factors that drive their process and help augment productivity.

Up to Chapter 6, this book concentrates more on the traditional “brick and mortar” structure of the firm. Chapter 6 bridges the gap to the evolving e-commerce “click and mortar” environment. The material in this chapter was contributed by an e-business organization (Engage Inc.), which provided some key insights on how to effectively utilize the Internet to help improve the performance of an organization. This chapter follows a logical progression from the previous chapter, which focused on advertising and marketing applications, as it highlights key issues on how to utilize data mining to more effectively devise marketing and advertising programs and better understand consumer behavior.

Chapter 7 then goes a step further by describing how such mining methodologies as Market Basket Analysis and related techniques, can be used to identify consumer preferences, which helps organizations focus on more accurately providing goods and services to individuals on a timely basis. The information in this chapter was contributed by another IT organization, Macromedia. At this stage, the reader has a better understanding of data mining and business intelligence; what it means and how it is used both in an e-commerce and non e-commerce environment. Chapter 8 then brings it all together by describing how organizations can utilize the entire business intelligence technological spectrum to reduce the uncertainty in particular business processes, which helps promote more efficient alloca-

tion of available resources.

In order to implement the BI spectrum effectively, one must not only know how to use the component technologies individually but must be able to integrate them properly. This can result in a number of obstacles that are difficult to overcome. Chapter 8 highlights some of the problems that exist in the implementation of data mining on an organizational level, one of which refers to a common issue involving the knowledge gap between “high-end” modelers and the non-mining community. It also stresses the importance of maintaining the BI strategic cycle. This refers to the procedure of identifying problems or areas of potential improvements, implementing appropriate policies to address these, and then monitoring the results of corresponding strategies to identify their success or failure, which then starts the cycle once again. The BI cycle incorporates all the components of the spectrum (data extraction and reporting, OLAP, mining and complementary Internet-related technologies) which help transform data into usable information that describes the underpinnings of corresponding processes and enlightens decision makers to devise and implement more accurate strategies to improve process performance.

To sum up this work, the final chapter takes a look at what the future holds for the data mining spectrum. This is never any easy task, especially when considering the dynamic nature of the information economy. However, we have concluded that the primary source of innovation in data mining will not revolve around the development of new algorithms, but rather the wider application of existing technology and greater integration with such complementary technologies as the Internet and reporting tools.

Data Mining and Business Intelligence: A Guide to Productivity, seeks to provide a greater understanding of what various forms of information technology offer to the world of business in the evolving information economy. By connecting the technological functionality to prevailing underlying business applications, which incorporate traditional business and economic theory, we hope to illustrate the full potential of data mining and business intelligence in achieving increased efficiency for the firm.