Preface

Managing Data Mining Technologies in Organizations: Techniques and Applications reflects a potpourri of chapters that demonstrate diverse use of techniques and their applications for data mining. The chapters illustrate applications of data mining and visualization for credit screening, forecasting, medical diagnosis and many others. The chapters discuss research results and applications of data mining to deliver rich decision-making information.

In chapter 1, titled "Bayesian Networks as a Decision Support Tool in Credit Scoring Domain," Abramowicz et al. discuss the applicability of Bayesian belief networks (BBN) for credit scoring in commercial banks. The authors show how to use a domain expert and the Bayesian probability theory to build BBNs. According to the authors, a domain expert may be very valuable in designing a causal model for the BBN. After a causal model is obtained, Bayesian probability theory and historical data can be used to compute conditional probabilities at various nodes. The authors provide an example towards the end of the chapter.

In chapter 2, titled "Frontier Versus Ordinary Regression Models for Data Mining," Troutt et al. discuss the differences between frontier and ordinary least-square regression models. The authors mention that frontier regression models have several desirable properties and may be useful in certain data mining applications. At the end of the chapter, the authors provide several examples ranging from supply chain management to marketing, where frontier models may be useful. Several suggested guidelines for the appropriate selection of frontier models are also provided.

In chapter 3, titled "An Evolutionary Misclassification Cost Minimization Approach for Medical Diagnosis," Pendharkar et al. discuss how to incorporate asymmetric misclassification cost in a genetic-algorithm-based classification system. The authors develop a linear genetic-algorithm-based classification system that incorporates asymmetric misclassification cost and compare the performance of their approach with the traditional statistical and genetic algorithm approaches that don't incorporate misclassification costs. Using simulated and real-life data on medical diagnosis, the authors show that their approach performs better than the competing approaches.

In chapter 4, titled "Guiding Knowledge Discovery Through Interactive Data Mining," Ceglar et al. highlight that present data mining algorithms constrain the generation of rules that may be of interest to the users. They provide an alternative interactive data mining approach that involves the user in the data mining process. The chapter provides a state-of-the-art overview of interactive data mining research.

In chapter 5, titled "A Proposed Process for Performing Data Mining Projects," Hirji discusses some of the data mining project management issues. Specifically, Hirji proposes a data mining process to jump-start a data mining project and to mitigate failure risks associated with performing data mining projects. The author uses a case study approach to illustrate a process to perform data mining projects.

In chapter 6, titled "Data Mining for Optimal Combination Demand Forecasts," Chan et al. discuss data mining issues related to the optimal combination of demand forecasts. The authors use an inventory management example of a major bank in Hong Kong and propose a quadratic programming model to estimate weights for combining the forecasts. Using different weighting methods and performance measures, the authors show that combined forecast performs better than any of the individual forecasts.

In chapter 7, titled "The Myth of Enterprise Database Redesign," Paper et al. discuss some of the business process reengineering issues related to data warehousing and data mining. Using a case study, Paper et al. develop an emergent theory about the factors related to enterprise database integration success.

In chapter 8, titled "New Information Technologies and Other Pertinent Issues Impacting the Strategic Dimension of CRM for Business Excellence," Kuppuraju and Subramanian describe the application of data mining for customer relationship management (CRM). The authors investigate some relevant issues related to the development and deployment of CRM applications. Additionally, the authors identify future research areas for CRM.

In chapter 9, titled "Utilization of Data Mining Techniques to Detect and Predict Accounting Fraud: A Comparison of Neural Networks and Discriminant Analysis," Rodger describes the use of data mining for predicting accounting fraud. The chapter highlights proactive information collection and analysis to detect accounting fraud.

In chapter 10, titled "A Multidimensional Data Warehouse Development Methodology," Cavero et al. discuss a methodology to develop a multidimensional data warehouse. According to the authors there is no generally accepted methodology for development of a data warehouse. The authors propose a data warehouse development methodology, which is integrated within a traditional software development methodology.

In chapter 11, titled "A Telecommunications Model for Managing Complexity of Voice and Data Networks and Services," Ghahramani describes how data mining algorithms reduce information overload, increase data integrity and accuracy, and aid effective management of global telecommunication networks.

In chapter 12, titles "Combination Forcasts Based on Markov Chain Monte Carlo Estimation of the Mode," Pang et al. propose use concepts from optimization and statistics to develop an approach for the combination of forecasts.

In chapter 13, titled "Web Mining: Creating Structure Out of Chaos," Lee proposes an overview of web mining. He illustrates how link topology can be used to analyze web site usage.

In chapter 14, titled "Connectionist and Evolutionary Models for Learning, Discovering and Forecasting Software Effort," Pendharkar and Subramanian illustrate how connectionist and evolutionary models can be used to learn nonlinear forecasting functions for predicting software effort.

I hope that the audience of this book finds the chapters interesting.

Parag C. Pendharkar
July 2002