

Preface

The intent of this book is to describe some recent data mining tools that have proven effective in dealing with data sets which often involve uncertain description or other complexities that cause difficulty for the conventional approaches of logistic regression, neural network models, and decision trees. Among these traditional algorithms, neural network models often have a relative advantage when data is complex. We will discuss methods with simple examples, review applications, and evaluate relative advantages of several contemporary methods.

Book Concept

Our intent is to cover the fundamental concepts of data mining, to demonstrate the potential of gathering large sets of data, and analyzing these data sets to gain useful business understanding. We have organized the material into three parts. Part I introduces concepts. Part II contains chapters on a number of different techniques often used in data mining. Part III focuses on business applications of data mining. Not all of these chapters need to be covered, and their sequence could be varied at instructor design.

The book will include short vignettes of how specific concepts have been applied in real practice. A series of representative data sets will be generated to demonstrate specific methods and concepts. References to data mining software and sites such as www.kdnuggets.com will be provided.

Part I: Introduction

Chapter 1 gives an overview of data mining, and provides a description of the data mining process. An overview of useful business applications is provided.

Chapter 2 presents the data mining process in more detail. It demonstrates this process with a typical set of data. Visualization of data through data mining software is addressed.

Part II: Data Mining Methods as Tools

Chapter 3 presents memory-based reasoning methods of data mining. Major real applications are described. Algorithms are demonstrated with prototypical data based on real applications.

Chapter 4 discusses association rule methods. Application in the form of market basket analysis is discussed. A real data set is described, and a simplified version used to demonstrate association rule methods.

Chapter 5 presents fuzzy data mining approaches. Fuzzy decision tree approaches are described, as well as fuzzy association rule applications. Real data mining applications are described and demonstrated

Chapter 6 presents Rough Sets, a recently popularized data mining method.

Chapter 7 describes support vector machines and the types of data sets in which they seem to have relative advantage.

Chapter 8 discusses the use of genetic algorithms to supplement various data mining operations.

Chapter 9 describes methods to evaluate models in the process of data mining.

Part III: Applications

Chapter 10 presents a spectrum of successful applications of the data mining techniques, focusing on the value of these analyses to business decision making.

University of Nebraska-Lincoln

David L. Olson

Oklahoma State University

Dursun Delen