Foreword

Database systems have had great success during the past two decades. More and more data are being collected and saved in databases—a database with a pedabyte of data is no longer uncommon. Finding useful information in these databases has become an important focus of many enterprises; and more and more attention has turned to data mining as a key component to such information discovery. Data-mining algorithms and visualization tools are being used to find important patterns in data and to create useful forecasts. This technology is being applied in virtually all business sections including banking, telecommunication, manufacturing, marketing, and e-commerce.

Data-mining algorithms and visualization tools were introduced in SQL Server 2000. Since then, most relational database systems now include data mining features in their product. Data mining in SQL Server 2005 is the next big step in the integration of data-mining and database technologies—a culmination of five years of intensive collaboration between the SQL Server product team and Microsoft Research. Engineers and researchers from these two organizations have worked together to bring both classic and new, cutting-edge data mining tools to SQL Server. The authors, ZhaoHui Tang and Jamie MacLennan, have been the key drivers of this collaboration.

This book is an invaluable companion to SQL Server 2005 Data Mining. The authors explain the basic principles of each algorithm and visualization tool, and provide many hands-on examples. I am certain that many database developers, database administrators, IT professionals, and students of data mining will benefit from this book.

David Heckerman
Research Manager
Microsoft Research, Redmond