Faculty of Engineering and Information Technology
University of Technology, Sydney

Actionable Knowledge Discovery: Methodologies and Frameworks

A thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

by

Dan Luo

June 2009

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

Production Note:
Signature removed prior to publication.

Acknowledgments

I appreciate Professor Chengqi Zhang, my PhD supervisor, for offering me the chance to this study. Without his support, it would be very difficult for me to finish this work.

My second thanks goes to Associate Professor Longbing Cao for his technical suggestions and joint research in my study and to this thesis in particular.

My additional thanks goes to the following schoolfellows for their suggestions and help during my study: Dr. Yanchang Zhao, Dr. Huaifeng Zhang, Dr Jiarui Ni, Mr Yuming Ou, etc., all members in the Data Sciences and Knowledge Discovery Research Lab, as well as all relevant staff and students in the Faculty and the University Graduate School.

I am also grateful to Ms. Li Liu for her continuous support in various management issues.

Finally, I appreciate the support from the APA Scholarship for my research, which contributes to the delivery of this thesis.

Contents

List of Figures

List of Tables

Abstract

Most data mining algorithms and tools stop at the mining and delivery of patterns satisfying expected technical interestingness. There are often many patterns mined but business people either are not interested in them or do not know what follow-up actions to take to support their business decisions. This issue has seriously affected the widespread employment of advanced data mining techniques in greatly promoting enterprise operational quality and productivity.

In this thesis, a formal and systematic view of actionable knowledge discovery (AKD for short) has been proposed from the system and microeconomy perspectives. AKD is a closed-loop optimization problem-solving process from problem definition, framework/model design to actionable pattern discovery, and to deliver operationalizable business rules that can be seamlessly associated or integrated with business processes and systems. To support AKD, corresponding methodologies, frameworks and tools have been proposed with case studies in the real world to address critical challenges facing the traditional KDD and to cater for crucially important factors surrounding real-life AKD.

First, a comprehensive survey and retrospection on the existing data mining methodologies, issues and challenges in actionable knowledge discovery are reviewed.

Second, a practical data mining methodology: *domain driven data mining* is addressed.

Third, several frameworks have been proposed to support domain driven

actionable knowledge discovery.

Fourth, case studies of domain-driven actionable pattern mining in stock markets and social security data are presented to demonstrate the usefulness and potential of the proposed domain driven actionable knowledge discovery.

In summary, this thesis explores in detail how domain driven actionable knowledge discovery can be effectively and efficiently applied to the discovery and delivery of knowledge satisfying both technical and business concerns as well as to support smart decision-making in the real world. The issues and techniques addressed in this thesis have potential to promote the research on critical KDD challenges, and contribute to the paradigm shift from data-centered and technical significance-oriented hidden pattern mining to domain-driven and balanced actionable knowledge discovery. The proposed methodologies and frameworks are flexible, general and effective to be expanded and applied to mining real-life complex data for actionable knowledge.