

# DATA MINING IN THE INDUSTRY

Michal Kebísek



Universitätsverlag Ilmenau  
2012

## Impressum

### **Bibliographic information of the German National Library**

The German National Library lists this publication in the German national bibliography, with detailed bibliographic information on the Internet at <http://dnb.d-nb.de>.

Author's acknowledgement to Gabriela Chmelíková for translation.

This scientific monograph originated from the author's dissertation thesis defended at the Slovak University of Technology in Bratislava, Faculty of Materials Science and Technology in Trnava.

#### **Reviewers:**

Prof. Ing. Juraj Spalek, PhD.  
Prof. Dr. Peter Husar  
Doc. Ing. Anton Kachaňák, CSc.

#### **Author's contact address:**

Ing. Michal Kebísek, PhD.  
Slovak University of Technology in Bratislava  
Faculty of Materials Science and Technology in Trnava

Ilmenau Technical University / University Library

### **Universitätsverlag Ilmenau**

Postfach 10 05 65  
98684 Ilmenau  
[www.tu-ilmenau.de/universitaetsverlag](http://www.tu-ilmenau.de/universitaetsverlag)

### **Production and delivery**

Verlagshaus Monsenstein und Vannerdat OHG  
Am Hawerkamp 31  
48155 Münster  
[www.mv-verlag.de](http://www.mv-verlag.de)

**ISSN** 2193-6439 (Print)  
**ISBN** 978-3-86360-048-8 (Print)  
**URN** urn:nbn:de:gbv:ilm1-2012100209

---

Titelfoto: [photocase.com](http://photocase.com)

## CONTENTS

|   |            |
|---|------------|
| <b>Introduction .....</b>   | <b>8</b>   |
| <b>1. Overview of current state in the field.....</b>   | <b>10</b>  |
| 1.1 Knowledge Discovery in Databases.....   | 10         |
| 1.2 Process of Knowledge discovery in databases .....   | 13         |
| 1.3 Data mining.....  | 16         |
| 1.4 Basic methods and data mining techniques.....   | 19         |
| 1.5 Problems with the application of Knowledge discovery<br>in databases in practice .....            | 21         |
| 1.6 KDD process use in practical applications.....  | 25         |
| 1.7 KDD tools .....   | 27         |
| 1.7.1 Commercial tools.....   | 27         |
| 1.7.2 Non-commercial tools .....  | 32         |
| <b>2. Identification of production system problems resolvable via data<br/>    mining.....</b>        | <b>37</b>  |
| <b>3. Production system simulation model proposal and relational<br/>    database structure .....</b> | <b>40</b>  |
| 3.1 Production system simulation model proposal.....  | 40         |
| 3.2 Proposal of relational database structure.....  | 44         |
| 3.2.1 Development of UML information system model.....  | 44         |
| 3.2.2 Development of relational database data model .....   | 46         |
| <b>4. Analysis of acquired production system data.....</b>  | <b>48</b>  |
| 4.1 Definition of data mining objectives.....   | 48         |
| 4.2 Data selection from the database.....   | 49         |
| 4.3 Data modification and transformation.....   | 54         |
| 4.4 Data mining.....  | 56         |
| 4.5 Evaluation of discovered knowledge .....  | 67         |
| <b>5. Discovered knowledge application in the production system .....</b>                             | <b>69</b>  |
| <b>6. Methodology proposal of a data mining process in industry to<br/>    improve control.....</b>   | <b>77</b>  |
| <b>7. Conclusion .....</b>  | <b>89</b>  |
| 7.1 Monograph contributions.....  | 90         |
| 7.2 Future development prospects.....   | 91         |
| <b>References .....</b>   | <b>93</b>  |
| <b>List of publications .....</b>   | <b>100</b> |