



PERGAMON

Computers & Industrial Engineering 43 (2002) 657–659

---

---

**computers &  
industrial  
engineering**

---

---

[www.elsevier.com/locate/dsw](http://www.elsevier.com/locate/dsw)

## Preface

# A focused issue on data mining and knowledge discovery in industrial engineering

It has often been said that we live in the ‘information age’. This verdict is best manifested by the immense creation, availability, and use of enormous volumes of data. After expressing datasets in megabytes we are now expressing them in terms of gigabytes (it is interesting to note that in Greek, the term ‘mega’ stands for ‘large’ and the term ‘giga’ stands for giant). At the same time, we are increasingly expressing datasets in terms of terabytes. It should come as no surprise that ‘teras’ means monster in Greek! The proliferation of large masses of data has created many new opportunities for those working in science, engineering, and business, but also some daunting challenges. The opportunities are offered by the abundance and availability of data, and the challenges are posed by the problem of how to organize it, retrieve it, and extract knowledge from it.

The field of data mining (DM) and knowledge discovery from databases (KDD) has emerged as a new discipline in engineering and computer science to address these new opportunities and challenges. Some may claim that this is an old scientific field since people always wanted to be able to analyze vast amounts of data and extract useful information (new knowledge) from them. However, in the modern sense of DM and KDD the focus tends to be on extracting information that can be characterized as ‘knowledge’, from data can be very complex and in large quantities. Industrial Engineering (IE), with the diverse areas it encompasses, presents unique opportunities for the application of DM and KDD, and for the development of new concepts and techniques in this field.

This special focused issue aims at presenting some new theoretical results in DM and KDD and representative applications in the field of IE. The three guest editors believe that the papers presented here represent a selected anthology of some of the key developments in the interface of DM/KDD and IE.

This issue contains 13 papers. The first seven papers are more of a theoretical nature, while the remaining six are more application oriented. The first paper is written by L.-Y. Zhai, L.-P. Khoo, and S.-C. Fok. These authors study the fundamental problem of how to extract the features that are pertinent to a DM/KDD problem. The authors propose a novel approach, which is based on rough sets and genetic algorithms (GAs). The second paper by N. Ye and X. Li discusses a new classification approach, which is based on clustering. The potential of this approach is also studied on some benchmark datasets. A well-known approach for DM and KDD is to use neural

networks (NNs). The third paper, written by P. Wu, proposes a new search component for a NN approach that is based on fuzzy sets. The numerical results presented in this paper are very promising. The fourth paper, written by X. Huo, presents a rigorous statistical approach for the identification of embedded consecutive subsequences and their relation to some DM and spatial statistics problems.

The fifth paper, written by G. Chen, Q. Wei, D. Liu, and G. Wets presents a powerful approach for the power of this approach is based on its simplicity, which does not come at the sacrifice of its applicability and effectiveness. Most approaches that mine association rules from datasets suffer from high complexity times. The new approach, called SAR for simple association rules, offers an efficient and effective alternative to some of the current methods for solving some problems by means of association rules, which often suffer from extreme complexity. The sixth paper by Y.Y. Hu, R.-S. Chen, and G.-H. Tzeng also deals with the subject of association rules. Its focus is on the mining of fuzzy association rules. The seventh paper, written by T.-L. Sun and W.-L. Kuo, presents a rather intriguing approach to DM and KDD. This approach uses some visual representations of the data. In this way, the role of the human analyst becomes a critical one since computerized systems cannot fully comprehend these visual representations. All the aforementioned theoretical papers are accompanied with brief numerical explorations of the effectiveness of the proposed methods.

The second half of this special issue contains six papers. These papers are mostly application oriented although they describe some new theoretical developments as well. The eighth paper written by M. Kantardzic, B. Djulbegovic, and H. Hamdan, describes the application of DM and KDD to the medical problem of diagnosing polycythemia vera (PV). It also compares the new approach with the existing medical practice. The results are very interesting. The ninth paper presents an application of a new neural network (NN) model to weather forecasting. This paper is written by T.B. Trafalis, A. White, B. Santosa, and M.B. Richman. This application involves the processing of large and complex rainfall and weather datasets. The tenth paper describes an application of neural network methods in studying a macro-economic problem related to the assessment of the investment risks of a number of countries. The results of the NN approach are compared with results obtained by applying more traditional statistical approaches. The authors of this paper are I. Becerra-Fernandez, S.H. Zanakis, and S. Walczak. The eleventh paper, written by S.H. Ha, S.M. Bae, and S.C. Park presents the application of certain DM and KDD methods of discovering new marketing behaviors. This new knowledge is used next in designing new marketing strategies. The twelfth paper, written by D.E. Brown and J.R. Brence, presents the application of some advanced DM and KDD methods to diagnosing corrosion problems on airplanes. The last paper, written by S. Morris, Z. Wu, C. DeYong, S. Salman, and D. Yemenu, presents the development of DM and KDD system for clustering and analyzing text documents. Visual approaches play a central role on these developments.

The guest editors of this special issue are very thankful to all authors of these papers. Their patience and aim for excellence is highly appreciated. We hope that this special issue will become a forum for the dissemination of the current developments and also for initiating new partnerships among researchers and practitioners. This is why special issue has a dedicated webpage with URL: [http://cda4.imse.lsu.edu/books1/special\\_issue2/special2.htm](http://cda4.imse.lsu.edu/books1/special_issue2/special2.htm). Finally, the

guest editors wish to express their gratitude, for his patience and guidance to Dr. Mohamed I. Dessouky, the Chief Editor of the Computers and IE journal.

E. Triantaphyllou<sup>\*,a</sup>

T.W. Liao<sup>a</sup>

S.S. Iyengar<sup>b</sup>

<sup>a</sup>*Department of Industrial and Manufacturing Systems Engineering,  
College of Engineering, Louisiana State University,*

*3128 CEBA Building,*

*Baton Rouge, LA 70803, USA*

*E-mail address: trianta@lsu.edu*

*http://www.imse.lsu.edu/vangelis*

<sup>b</sup>*Department of Computer Science,  
Louisiana State University, Coates Hall,  
Baton Rouge, LA 70804, USA*

---

\* Corresponding author. Tel.: +1-255-578-5372; fax: +1-255-578-5990.