



# Dottorato di Ricerca in Ingegneria dell'Informazione

## Data Mining and Soft Computing

**Francisco Herrera**

**Research Group on Soft Computing and  
Information Intelligent Systems (SCI<sup>2</sup>S)**

**Dept. of Computer Science and A.I.  
University of Granada, Spain**

Email: [herrera@decsai.ugr.es](mailto:herrera@decsai.ugr.es)

<http://sci2s.ugr.es>

<http://decsai.ugr.es/~herrera>



# Data Mining and Soft Computing



## **In this course:**

We will introduce the Data Mining and Knowledge Discovery area: steps, task, challenges ..

We will introduce Soft Computing techniques:  
Fuzzy Logic, Genetic Algorithms, ...

... and we will present the use of Soft Computing techniques in Data Mining

# Data Mining and Soft Computing



## Material of this course at:

[http://sci2s.ugr.es/docencia/asignatura.php?id\\_asignatura=14](http://sci2s.ugr.es/docencia/asignatura.php?id_asignatura=14)

The screenshot shows a Mozilla Firefox browser window displaying the website for the course "Data Mining and Soft Computing". The browser's address bar shows the URL [http://sci2s.ugr.es/docencia/asignatura.php?id\\_asignatura=14](http://sci2s.ugr.es/docencia/asignatura.php?id_asignatura=14). The website has a dark blue background with a gold banner at the top that reads "Data Mining and Soft Computing". Below the banner, the text "Post-Graduate Program" and "Course: Data Mining and Soft Computing" is displayed, followed by "Dottorato di Ricerca in Ingegneria dell'Informazione". The instructor is listed as "Francisco Herrera (Dpto. de Ciencias de la Computación e I.A.)". A "Summary of Sessions" section lists nine sessions, each with a brief description and a link to a PDF file. The left sidebar contains a navigation menu with links to "Presentation", "Members", "Research Lines", "Advised PhD", "Publications", "Highly Cited Papers", "Review, Taxonomy & Position Papers", "Edited Books & Special Issues", "Editorial Boards", "Citation Reports & Indexes", "Conference Activities", and "Projects". Logos for DECSAI and ETSIT are visible in the top right corner of the website.

SCI2S (Soft Computing and Intelligent Information Systems) - Docencia - Mozilla Firefox

Archivo Editar Ver Historial Marcadores Herramientas Ayuda

[http://sci2s.ugr.es/docencia/asignatura.php?id\\_asignatura=14](http://sci2s.ugr.es/docencia/asignatura.php?id_asignatura=14)

Más visitados Comenzar a usar Firefox Últimas noticias IEEE TFS Soft Computing

<http://sci2s.ugr.es>

DECSAI Universidad de Granada

ETSIT Escuela Técnica Superior de Ingenieros Industriales de la Universidad de Sevilla

Data Mining and Soft Computing

Post-Graduate Program  
Course: Data Mining and Soft Computing  
Dottorato di Ricerca in Ingegneria dell'Informazione

Francisco Herrera (Dpto. de Ciencias de la Computación e I.A.)

Summary of Sessions

- Session 1: Introduction to Data Mining and Knowledge Discovery ([PDF: 1039 Kb](#))
- Session 2: Data Preparation.
- Session 3: Introduction to Prediction, Classification and Association.
- Session 4: Data Mining From the Top 10 Algorithms to the New Challenges.
- Session 5: Introduction to Soft Computing. Focusing our attention in Fuzzy Logic and Evolutionary Computation.
- Session 6: Soft Computing Techniques in Data Mining: Fuzzy Data Mining and Knowledge Extraction based on Evolutionary Learning.
- Session 7: Genetic Fuzzy Systems: State of the Art and New Trends.
- Session 8: Some Advanced Topics I: Classification with Imbalanced Data Sets
- Session 9: Some Advanced Topics II: Subgroup Discovery

KEEL Software

# Data Mining and Soft Computing

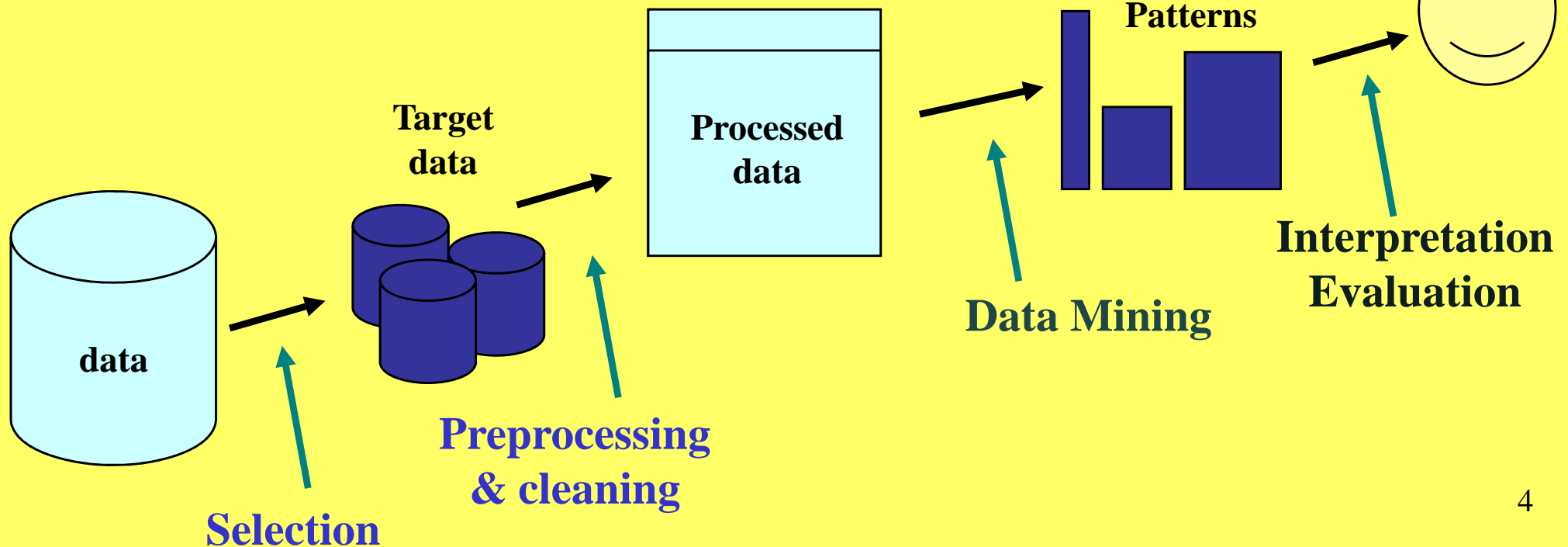


## ■ Data mining:

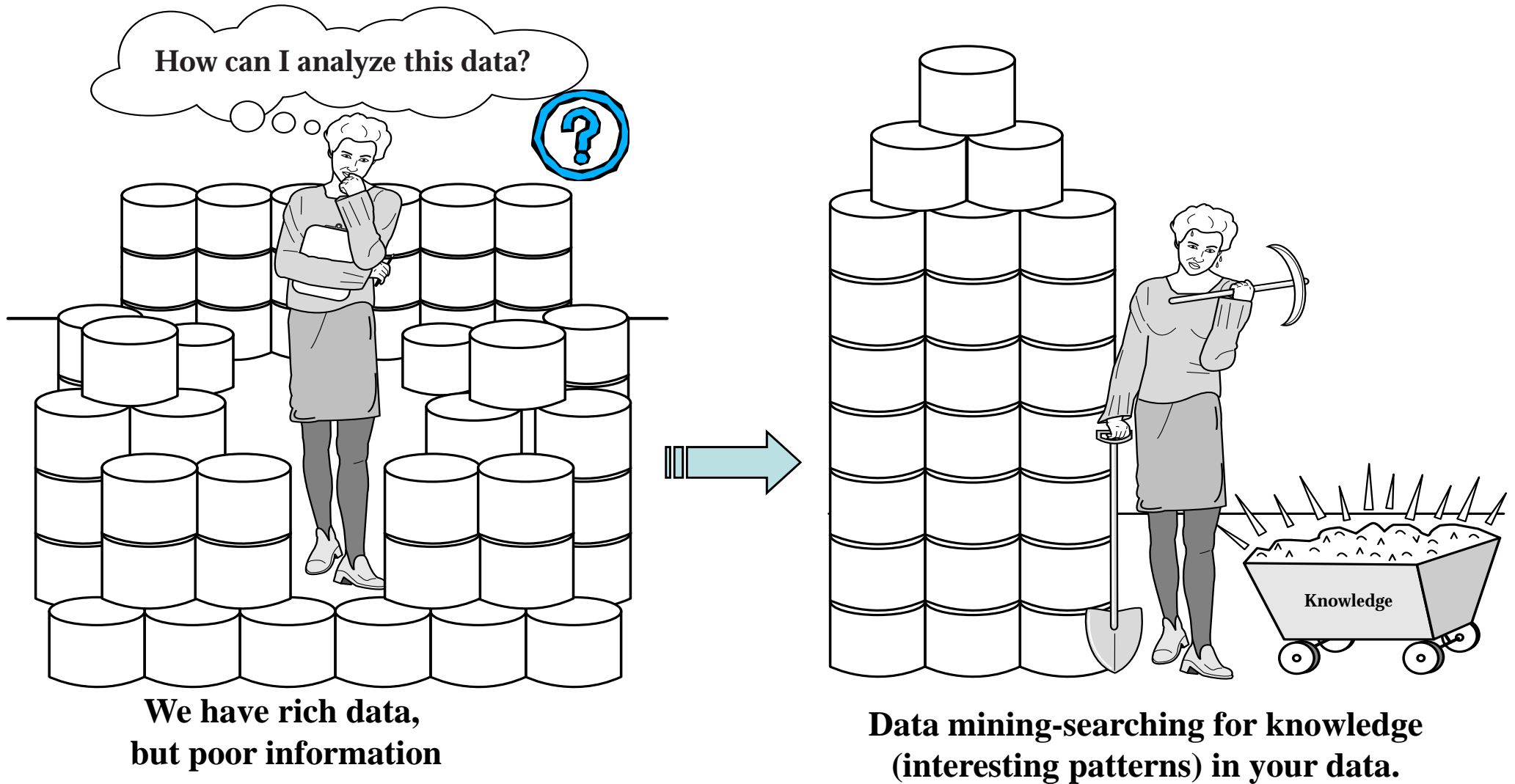
- Extraction of interesting (**non-trivial, implicit, previously unknown and potentially useful**) information or patterns from data in large databases



**Knowledge**



# Data Mining



# Soft Computing

Soft computing refers to a collection of computational techniques in computer science, machine learning and some engineering disciplines, which study, model, and analyze very complex phenomena: those for which more conventional methods have not yielded low cost, analytic, and complete solutions.

## Prof. Zadeh:

*"...in contrast to traditional hard computing, soft computing exploits the tolerance for **imprecision, uncertainty, and partial truth** to achieve tractability, robustness, low solution-cost, and better rapport with reality"*



Lotfi A. Zadeh

**Introduce  
“Fuzzy Logic” in 1965  
and “Soft Computing”  
in 1992.**

# Computational Intelligence



The Field of Interest of the Society shall be the theory, design, application, and development of biologically and linguistically motivated computational paradigms emphasizing **neural networks, connectionist systems, genetic algorithms, evolutionary programming, fuzzy systems, and hybrid intelligent systems** in which these paradigms are contained.

# Data Mining and Soft Computing



## ■ Contents:

### Part I. Principles of Data Mining

Introduction to Data Mining and Knowledge Discovery

Data Preparation

Introduction to Prediction, Classification, Clustering and Association

Data Mining - From the Top 10 Algorithms to the New Challenges



# Data Mining and Soft Computing



## ■ Contents:

### **Part II. Soft Computing Techniques in Data Mining**

- Introduction to Soft Computing. Focusing our attention in Fuzzy Logic and Evolutionary Computation
- Soft Computing Techniques in Data Mining: Fuzzy Data Mining and Knowledge Extraction based on Evolutionary Learning
- Genetic Fuzzy Systems: State of the Art and New Trends

# Data Mining and Soft Computing

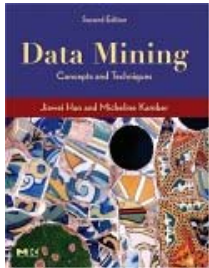
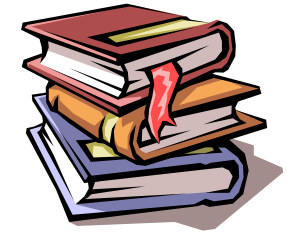


## ■ Contents:

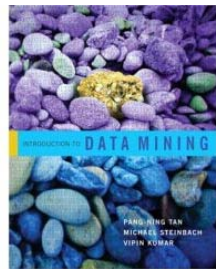
### **Part III. Data Mining: Some Advanced Topics**

- Some Advanced Topics I: Classification with Imbalanced Data Sets**
- Some Advanced Topics II: Subgroup Discovery**
- Some advanced Topics III: Data Complexity**
- Final talk: How must I Do my Experimental Study?  
Design of Experiments in Data Mining/  
Computational Intelligence. Using Non-parametric  
Tests. Some Cases of Study.**

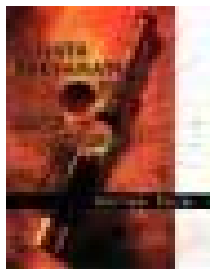
# Bibliography



**J. Han, M. Kamber.**  
**Data Mining. Concepts and Techniques**  
**Morgan Kaufmann, 2006 (Second Edition)**  
<http://www.cs.sfu.ca/~han/dmbook>

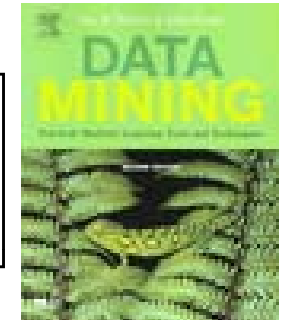


**Pang-Ning Tan, Michael Steinbach, and Vipin Kumar**  
**Introduction to Data Mining (First Edition)**  
**Addison Wesley, (May 2, 2005)**  
<http://www-users.cs.umn.edu/~kumar/dmbook/index.php>

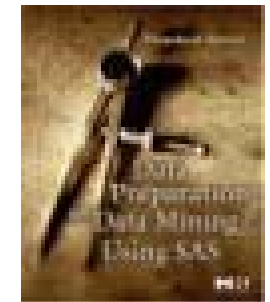


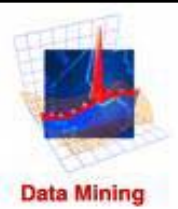
**Dorian Pyle**  
**Data Preparation for Data Mining**  
**Morgan Kaufmann, Mar 15, 1999**

**Mamdouh Refaat**  
**Data Preparation for Data Mining Using SAS**  
**Morgan Kaufmann, Sep. 29, 2006)**



**I.H. Witten, E. Frank.**  
**Data Mining: Practical Machine Learning Tools and Techniques,**  
**Second Edition, Morgan Kaufmann, 2005.**  
<http://www.cs.waikato.ac.nz/~ml/weka/book.html>





# Data Mining and Soft Computing

## Summary

1. Introduction to Data Mining and Knowledge Discovery
2. Data Preparation
3. Introduction to Prediction, Classification, Clustering and Association
4. Data Mining - From the Top 10 Algorithms to the New Challenges
5. Introduction to Soft Computing. Focusing our attention in Fuzzy Logic and Evolutionary Computation
6. Soft Computing Techniques in Data Mining: Fuzzy Data Mining and Knowledge Extraction based on Evolutionary Learning
7. Genetic Fuzzy Systems: State of the Art and New Trends
8. Some Advanced Topics I: Classification with Imbalanced Data Sets
9. Some Advanced Topics II: Subgroup Discovery
10. Some advanced Topics III: Data Complexity
11. Final talk: How must I Do my Experimental Study? Design of Experiments in Data Mining/Computational Intelligence. Using Non-parametric Tests. Some Cases of Study.