



# Course Introduction

Data Mining and Text Mining (UIC 583 @ Politecnico di Milano)

- Prof. Pier Luca Lanzi  
Dipartimento di Elettronica e Informazione  
pierluca.lanzi@polimi.it  
voce: 02 23993472  
<http://www.pierlucalanzi.net>
- Office Hours  
Wednesday, from 15:00 until 17:00
- Teaching Assistant  
Dr. Daniele Loiacono  
Dipartimento di Elettronica e Informazione  
loiacono@elet.polimi.it

- Basics (24 hours)

Introduction to Data Mining and Text Mining

- Advanced Techniques and Applications (16 hours)

Advanced Data Mining techniques and applications

- Final Project (optional)

An application to real-world data

# Tecniche di Apprendimento Automatico per Applicazioni di Data Mining

- This shorter course is covered by the first 20 hours
- The topics are more or less the ones covered by the previous editions
- Lectures are in English
- Final test is in Italian (or in English upon request)

- What is Data Mining?
- Data and knowledge representation
  
- Data Mining tasks
  - Associations
  - Clustering
  - Classification
- Aggregate methods
  
- Preprocessing
- Advanced techniques and applications
  - Text Mining
  - Graph Mining
  - Data Streams
  
- Evaluation
  - One written exam
  - Course project

- Lecture slides available at <http://www.pierlucalanzi.net>
- Syllabus
  - Jiawei Han, Micheline Kamber. “Data Mining: Concepts and Techniques” Second Edition. Morgan Kauffman, 2006.
  - Ian H. Witten, Eibe Frank. “Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations” 2nd Edition.
  - Tom Mitchell. “Machine Learning”, McGraw Hill 1997
- Software
  - Rapid Miner <http://rapid-i.com/>

- May 2011 First Midterm (15 points)
- June 2011 Second Midterm (18 points)
- July 2011 Full exam for those who failed midterms