

30/09/2009

## Processi di data mining

Prof. Stefano Lodi  
Facoltà di Scienze Statistiche – Sede di Rimini  
Dipartimento di Elettronica, Informatica e Sistemistica  
Università di Bologna  
[stefano.lodi@unibo.it](mailto:stefano.lodi@unibo.it)  
<http://www-db.deis.unibo.it/~slodi>

## Programma del corso

- **Sistemi per il supporto alle decisioni**
  - Processi decisionali
  - Classificazione delle decisioni
  - Definizione di sistema per il supporto alle decisioni, sue proprietà, struttura e fattori di criticità.

LMSIAF II anno 2009-2010      Processi di data mining      Introduzione e programma 2

## Programma del corso

- **Data warehousing**
  - Data warehousing e business intelligence
  - OLTP e OLAP
  - Architettura di un data warehouse
  - Schemi e operazioni in un data warehouse

LMSIAF II anno 2009-2010      Processi di data mining      Introduzione e programma 3

## Programma del corso

- **Knowledge Discovery**
  - Knowledge discovery e business intelligence. Il processo di knowledge discovery
  - Problemi e algoritmi di Data Mining
    - Regole associative
    - Clustering di dati vettoriali
      - algoritmi one-pass, algoritmo BIRCH; algoritmi di clustering basati sulla densità: algoritmi DBSCAN, DENCLUE
    - Clustering di dati categorici
    - Algoritmi di apprendimento di modelli di classificazione

LMSIAF II anno 2009-2010      Processi di data mining      Introduzione e programma 4

## Programma del corso

- Argomenti avanzati
  - Algoritmi distribuiti per le regole associative; algoritmi distribuiti per il clustering; algoritmi distribuiti per la classificazione
  - Algoritmi per stream di dati: calcolo di informazioni aggregate da stream di dati; clustering di stream di dati: time frame piramidali e algoritmo CluStream
- Esercitazioni di laboratorio
  - IBM Intelligent Miner
  - Microsoft SQL Server Business Intelligence Development Studio

LMSIAF II anno 2009-2010      Processi di data mining      Introduzione e programma 5

## Tesi di Laurea Magistrale

- Algoritmi di clustering distribuito di dati numerici in ambienti non centralizzati
  - Reti distribuite massive (peer-to-peer networks, sensor networks)
  - Streaming data (dati finanziari, dati di connessione web, dati ambientali automaticamente rilevamenti)
- Algoritmi di apprendimento di modelli di classificazione Support Vector Machine
  - Dati distribuiti
  - Streaming data

LMSIAF II anno 2009-2010      Processi di data mining      Introduzione e programma 6

### Testi di riferimento

- Han, J., & Kamber, M. (2000). Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers. (disponibile in biblioteca)
- Vercellis, C. (2006). Business intelligence. Modelli matematici e sistemi per le decisioni. Milano: McGraw-Hill.
- Articoli scientifici (a cura del docente)

LMSIAF Il anno 2009-2010

Processi di data mining

Introduzione e programma 7

### Modalità d'esame

- Prova pratica con strumenti di data warehousing e data mining
  - Importazione di dati in un sistema di data warehousing e data mining
  - Creazione di cubi in un data warehouse
  - Analisi di dati con strumenti di data mining
- Prova orale
  - Contenuto delle lezioni (diario del corso)

LMSIAF Il anno 2009-2010

Processi di data mining

Introduzione e programma 8