

Business Intelligence per i Big Data



Tania Cerquitelli

AA. 2018-2019

1

Docenti

Docente: Tania Cerquitelli

Dip. Automatica e Informatica

Tel: 011 090-7178

e-mail: tania.cerquitelli@polito.it

Esercitori:

Evelina Di Corso

Dip. Automatica e Informatica

e-mail: evelina.dicorso@polito.it

Stefano Proto

Dip. Automatica e Informatica

e-mail: stefano.proto@polito.it

Ricevimento: su appuntamento



2

2

Organizzazione del corso

- Inizio lezioni: 04/03/2019

- Fine lezioni: 13/06/2019

- Orario delle lezioni ed esercitazioni

| Giorno | Orario | Aula |
|---------|-------------|---------|
| Lunedì | 11:30-14:30 | Aula 11 |
| Giovedì | 11:30-14:30 | LEP |



3

3

Obiettivi

- Business Intelligence: attività di analisi dei dati per il supporto delle decisioni aziendali

- Studio delle basi di dati destinate all'analisi dei dati (*data warehouse*)

- Architettura dei sistemi di data warehouse e delle metodologie di progettazione e sviluppo

- Introduzione alle basi di dati non relazionali (no-relational databases) a supporto dei Big Data

- Tecniche di analisi

- OLAP (On Line Analytical Processing) dei dati

- Tecniche ed algoritmi di data mining

- Sviluppo di casi applicativi mediante strumenti open Source



4

4

Lezioni

- Data warehouse

- architetture
- progettazione e sviluppo
- analisi OLAP (On Line Analytical Processing) dei dati
- estensioni del linguaggio SQL per interrogare i data warehouse

- Basi di dati NO-SQL

- Tipi di basi di dati NO-SQL
- Concetti di base
- Esempi di interrogazioni



5

5

Lezioni

- Data mining

- Tecniche di preparazione dei dati
- Algoritmi di classificazione e tecniche di validazione
- Algoritmi di clustering ed indici di qualità
- Algoritmi di estrazione di regole di associazione ed indici di qualità

- Studio di casi applicativi

- progettazione di data warehouse
- interrogazioni di sistemi no relazionali
- analisi di dati raccolti



6

6



Esercitazioni in aula

- Progettazione concettuale, logica e fisica di un data warehouse
- Interrogazione di un data warehouse mediante il linguaggio SQL esteso
- Interrogazione di un sistema NO-SQL mediante paradigma Map-Reduce

7



Esercitazioni in laboratorio

- Interrogazione di una base di dati relazionale mediante il linguaggio SQL e le sue estensioni
 - DBMS: Oracle
- Progettazione di un data warehouse e analisi OLAP
 - Tool open source (i.e., Pentaho) che si interfaccia con il DBMS Oracle
- Progettazione di un sistema NO-SQL e analisi
 - Tool open source: Mongo DB
- Applicazioni di tecniche di data mining
 - Tool open source (i.e., RapidMiner)

8



Esame

- Prova scritta (durata: 2 ore)
 - Progettazione concettuale, logica e fisica di un data warehouse
 - Interrogazione del data warehouse progettato mediante il linguaggio SQL
- Progetto
 - Utilizzo di una tecnica di data mining per analizzare un dataset reale mediante RapidMiner e/o SAS Enterprise Miner
- Consegna di 2 esercitazioni di laboratorio
 - Uso di algoritmi di classificazione in un caso di studio
 - Esercitazione con un sistema di basi di dati no relazionale

9



Materiale

- Sito web del corso
 - <http://dbdmg.polito.it/wordpress/teaching/business-intelligence/>
 - Copia dei lucidi utilizzati a lezione
 - Testi e risoluzioni di esercizi proposti in aula
 - Testi e materiale utile per le esercitazioni in laboratorio
 - Materiale utile per la risoluzione dei progetti
 - Materiale integrativo
 - Modello relazionale e linguaggio SQL

10



Testi

- Data warehousing
 - Golfarelli, Rizzi. Data warehouse: teoria e pratica della progettazione, McGraw-Hill 2006
 - Kimball e altri, numerosi testi su metodologia e casi di studio, Wiley
- Data mining
 - Han, Kamber. Data mining: concepts and techniques, Morgan Kaufmann 2006
 - Tan, Steinbach, Kumar. Introduction to data mining, Pearson 2006
- Basi di dati no relazionali
 - K. Chodorow and M. Dirolf. MongoDB: the definitive guide. O'Reilly Media, 2010

11



Link utili

- Data warehouse
 - <http://www.dwinfocenter.org>
 - <http://www.dwreview.com>
 - <http://kimballuniversity.com>
- Data mining
 - <http://www.kdnuggets.com>
- Tool open source
 - Pentaho: <http://www.pentaho.com/>
 - RapidMiner: <http://rapid-i.com/>
 - Mongo-DB: <https://www.mongodb.com/it>

12