

FACOLTA': Ingegneria

CORSO DI LAUREA: Ingegneria Informatica e dell'Automazione

INSEGNAMENTO: Sistemi Transazionali e Data Mining

NOME DOCENTE: Pietro Ducange

indirizzo e-mail: pietro.ducange@uniecampus.it

orario ricevimento via e-mail: il Mercoledì dalle 18 alle 20

Il ricevimento telefonico si effettua il Giovedì dalle 15 alle 17

(Il docente riceve comunque gli studenti frontalmente dopo il seminario previsto prima dell'esame)

OBIETTIVI DEL CORSO:

- Fornire allo studente gli strumenti basilari per l'analisi e l'estrazione delle conoscenze dai dati
- Studiare i principali modelli descrittivi e predittivi caratteristici del data mining e fornire allo studente le basi per diventare "Ingegnere della Conoscenza"
- Approfondire gli strumenti software a disposizione per applicare le tecniche di data mining presentate nel corso
- Presentare degli esempi pratici e reali di applicazione delle tecniche di data mining

CONTENUTI DEL CORSO:

Le lezioni saranno organizzate secondo i seguenti nuclei tematici:

1. Introduzione al corso e al data mining
2. Trattamento preliminare dei dati
3. Misure di distanza e similarità
4. Analisi descrittiva tramite clustering
5. Analisi descrittiva tramite regole associative
6. Analisi predittiva: classificazione e regressione
7. programma WEKA
8. USO delle API WEKA
9. Accenni alle tecniche di data mining per big data

MODALITA' DI SVOLGIMENTO ESAME:

L'esame sarà sostenuto dallo studente in due parti

1. Un test scritto con domande (principalmente aperte) teoriche e/o esercizi relativi agli argomenti del corso.
2. Un progettinio da presentare e discutere in sede di esame

Il test sarà subito corretto dal docente e seguirà discussione orale con lo studente.

Le specifiche del progetto devono essere concordate con il docente con adeguato anticipo. Il docente non valuterà progetti di cui non saranno inviate ed approvate le specifiche.

BIBLIOGRAFIA:

Susi Dulli, Sara Furini e Edmondo Peron "Data mining. Metodi e strategie", Springer, 2009
(caldamente consigliato, molte delle lezioni faranno riferimento a questo libro)

Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd ed., Morgan Kaufmann Publishers, 2006 (In Inglese, uno dei libri più completi sul data mining).

Ian H. Witten, Eibe Frank, Mark A. Hall, "Data Mining: Practical Machine Learning Tools and Techniques 3rd Edition", Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 2011 (In Inglese, fa riferimento all'uso del programma WEKA).

CONSIGLI DEL DOCENTE:

Risorse per lo studio:

- 1) **Slide fornite dal docente:** contengono tutte le informazioni necessarie per acquisire le conoscenze del corso. Contengono approfondimenti, note storiche, esercizi e riferimenti bibliografici.
- 2) **Alcuni libri di testo:** da consultare in base alle necessità di ciascuno studente. La molte delle informazioni contenute nelle slide fanno riferimento a specifici paragrafi o libri.
- 3) **Pagine Web verificate dal docente:** non tutto ciò che si trova su internet è corretto, anzi molte informazioni sono rumorose e contraddittorie. Il docente si preoccuperà di certificare le pagine web che riterrà utili ed interessanti per approfondire alcuni argomenti.

Prerequisiti Consigliati

1. Conoscenza del linguaggio di programmazione Java
2. Basi di Ingegneria del Software