

FACOLTA': Ingegneria

CORSO DI LAUREA: Ingegneria Informatica e dell'Automazione

INSEGNAMENTO: Data Mining

CFU: 9

NOME DOCENTE: Pietro Ducange

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

SI RICORDA AGLI STUDENTI CHE IN BASE ALLA COMUNICAZIONE DI SEGRETERIA DEL 12/12/2013, I DOCENTI SONO CONTATTABILI ATTRAVERSO IL SISTEMA DI MESSAGGISTICA DELLA PIATTAFORMA.

orario ricevimento ufficio virtuale: il Martedì dalle 17 alle 18

É molto apprezzato dal docente essere contattato in anticipo tramite il sistema di messaggistica per avvertirlo della necessità di ricevimento e delle tematiche da approfondire.

RISULTATI DI APPRENDIMENTO DELL'INSEGNAMENTO

Il corso ha l'obiettivo di far conseguire allo studente i seguenti risultati formativi

1. Con riferimento alla conoscenza e capacità di comprensione, il corso si propone di fornire ai partecipanti:
 - a. Gli strumenti basilari (modelli ed algoritmi) per l'analisi e l'estrazione delle conoscenza dai dati
 - b. Un insieme di competenze specialistiche per l'utilizzo di strumenti software per il data mining
 - c. Una panoramica di esempi pratici e reali di applicazione delle tecniche di data mining
2. Con riferimento alla conoscenza e capacità di comprensione applicate, al termine del corso lo studente sarà in grado di:
 - a. Risolvere semplici problemi di data mining
 - b. Formalizzare problemi di estrazione di conoscenza dai dati ed identificare gli opportuni modelli teorici (descritti e/o predittivi) e gli algoritmi per la loro soluzione
 - c. Progettare ed implementare applicazione software in grado di risolvere problemi reali di data mining
3. Con riferimento all'autonomia di giudizio
 - a. Saper impostare correttamente una progetto di estrazione delle conoscenza, integrandolo all'interno di una reale applicazione informatica e scegliendo in maniera opportuna le strategie di progetto e di implementazione.
4. Con riferimento alle abilità comunicativa

- a. Utilizzo del lessico e della terminologia appropriata per presentare, in forma scritta o verbale, le argomentazioni e i risultati del proprio studio
5. Con riferimento all'abilità ad apprendere
- a. Autonomia nell'utilizzo del materiale e negli strumenti didattici in piattaforma

CONTENUTI DELL'INSEGNAMENTO:

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
7. Il tool WEKA e uso delle sue API
8. Realizzazione del progetto

EVENTUALI PROPEDEUTICITA' CONSIGLIATE:

Si consiglia di aver almeno seguito il corso di Linguaggi di Programmazione e di essere a conoscenza dei fondamenti di Ingegneria del Software e Basi di Dati

MODALITA' DI SVOLGIMENTO ESAME:

Le modalità d'esame sono descritte sul Regolamento per lo svolgimento degli esami di profitto consultabile sul sito dell'ateneo.

[Regolamento per lo svolgimento degli esami di profitto](#)

Lo studente per essere ammesso alla prova **scritta (o alla prova orale ove previsto) deve realizzare preventivamente un progetto didattico** o una esercitazione alternativa. Sia il progetto sia l'esercitazione devono essere discussi (e approvati), almeno una settimana prima della prova scritta, con il docente nel suo ufficio virtuale. Vedere di seguito (didattica Interattiva) le **modalità di valutazione** del progetto didattico o dell'esercitazione alternative.

METODI DI ACCERTAMENTO DEI RISULTATI DI APPRENDIMENTO E MODALITA' DI VALUTAZIONE:

1. Con riferimento alla conoscenza e capacità di comprensione
 - Domande a risposta chiusa e aperta nella prova scritta

2. Con riferimento alla conoscenza e capacità di comprensione applicate

- Esercitazioni applicative durante il corso con consegna in ePortfolio
- Esercizi in sede di esame scritto
- Realizzazione del progetto didattico
- In sede di esame orale, allo studente sarà data la possibilità di discutere problematiche teorico/pratiche relative agli argomenti trattati nel corso

3. Con riferimento all'autonomia di giudizio:

-Lo studente dovrà essere in grado di impostare in maniera ottimale il progetto didattico. Ciò è valutato sulla base della documentazione fornita e della discussione del progetto in aula virtuale.

4. Con riferimento all'abilità comunicativa:

L'abilità viene valutata sulla base della proprietà di linguaggio tecnico utilizzato nelle risposte alle domande a risposta aperta e durante l'eventuale prova orale.

5. Con riferimento all'abilità ad apprendere:

L'abilità è valutata essenzialmente sulla base delle esercitazioni svolte in itinere, in modo da verificare che lo studente sia effettivamente in grado di utilizzare gli strumenti offerti in piattaforma di e-learning.

BIBLIOGRAFIA:

Susi Dulli, Sara Furini e Edmondo Peron "Data mining. Metodi e strategie", Springer, 2009 (Testo Obbligatorio da cui approfondire gli argomenti, molte delle lezioni fanno riferimento a questo libro)

J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd ed., Morgan Kaufmann Publishers, 2006 (Testo Consigliato, 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 (Testo Consigliato, In Inglese, fa riferimento all'uso del programma WEKA).

ATTIVITA' DIDATTICA EROGATIVA E INTERATTIVA:

Didattica Erogativa:

Durante il corso saranno erogate agli studenti un certo numero di lezioni audio/video. I contenuti di tali lezioni sono da ritenersi oggetto di esame e non devono essere assolutamente scartate dagli studenti. Si ricorda che anche tutto il materiale fornito sotto forma di slide e dispense deve essere studiato e sarà oggetto di esame.

Didattica Erogativa: Webinar

Saranno proposti agli studenti dei webinar di approfondimento a partecipazione facoltativa, il cui calendario sarà a disposizione dello studente sulla propria piattaforma

informatica e sulla pagina personale del docente. In generale gli argomenti dei webinar (soggetti a cambiamento, previo avviso agli studenti) saranno:

- Introduzione al corso (Livello Base)
- Dati e preprocessing (Livello Base)
- Algoritmi di classificazione (Livello Intermedio)
- Algoritmi di clustering (Livello Intermedio)
- Algoritmi di generazione di regole di associazione (Livello Intermedio)
- Simulazione prova di esame (Livello Base)
-

Attenzione: la partecipazione alle aule virtuali non è obbligatoria e non dà diritto a nessun vantaggio né a punteggi aggiuntivi in sede di esame.

Didattica Interattiva: Esercizi ed esempi di domande di esame (fino alla lezione 50)

- Prestare attenzione alle sessioni di approfondimento con **esercizi** e alle sessioni di **esempi di domande di esame**.
- In alcune di queste sessioni verrà richiesto allo studente di caricare sul proprio ePortfolio (sessioni di studio contrassegnate dall'azione COMPILA) la soluzione dello specifico esercizio o le risposte alle domande di esame.
- La soluzione e/o le risposte caricate devono essere farina del sacco dello studente. È inutile caricare la soluzione fornita dal docente nelle slide o copiare pezzi di slide e incollarle come risposta alle domande.
- Sarà premura del docente fornire un feedback allo studente sul materiale caricato.

*Attenzione: lo svolgimento delle attività di cui sopra, previste della didattica interattiva, **sono obbligatorie**, ma non danno diritto a nessun vantaggio né a punteggi aggiuntivi in sede di esame. Esse servono come metodo di verifica della preparazione in itinere dello studente.*

Didattica Interattiva: Progetto Didattico o Esercitazione Alternativa (dalla lezione 51 alla lezione 72)

Durante il corso, viene proposta la realizzazione di un progettino didattico. 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.

È **obbligatorio** caricare nell'ePortfolio dello studente un certo numero di artefatti relativi al progetto didattico (a partire dalla lezione 51). Tali artefatti sono relativi a tutte le fasi di realizzazione del progetto e **saranno valutati ai fini dell'esame finale**.

Il progetto didattico può essere sostituito da una esercitazione assegnata dal docente. Anche in questo caso è obbligatorio caricare artefatti sull'ePortfolio e discutere in aula virtuale del docente l'esercitazione. **Anche tali artefatti saranno valutati ai fini dell'accesso all'esame finale.**

Didattica Interattiva: Webinar di Supporto allo svolgimento di esercitazioni

- Saranno calendarizzati degli incontri in aula virtuale in cui gli studenti potranno chiedere chiarimenti sullo svolgimento delle esercitazioni di cui sopra e per analizzare lo stato di avanzamento dei lavori relativi al progetto didattico.

Discussione dei progetti e delle esercitazioni

É obbligatorio discutere il progetto o l'esercitazione alternativa, fissando un appuntamento in aula virtuale del docente, entro la settimana che precede l'appello d'esame. In caso di mancata discussione, lo ***studente non sarà ammesso alla prova di esame.***

Maggiori informazioni relative al progetto didattico e alla esercitazione alternativa si trovano nella prima lezione, sulle FAQ predisposte dal docente e nei documenti pubblicati sulla pagina personale del docente.

Valutazione del progetto didattico o dell'esercitazione sul voto finale

- In base al regolamento di Ateneo, il test scritto conterrà 21 domande a risposta multipla e 3 domande a risposta aperta (ciascuna sarà valutata al massimo 3 punti).
- Le 21 domande a risposta multipla ed una domanda a risposta aperta serviranno per valutare le ***conoscenze e le capacità di comprensione delle studente.***
- ***Due*** delle domande a risposta aperta ***saranno relative al progetto*** e consentiranno di valutare conoscenze e le capacità ***di comprensione applicate*** delle studente e ***il livello di autonomia di giudizio*** raggiunto dallo studente.
- ***Gli studenti che decideranno di non svolgere il progetto,*** non potranno dimostrare completamente di aver acquisito capacità ***di comprensione applicate.*** Inoltre, il livello di ***autonomia di giudizio*** non sarà completamente valutabile. Per questo motivo, se avranno svolto correttamente la loro esercitazione alternativa (da concordare con il docente almeno 10 giorni prima dell'esame), ***potranno raggiungere al massimo la valutazione di 24/30*** (rispondendo correttamente alle 21 domande a crocette e alla domanda aperta).
- Gli studenti che avranno svolto il progetto, potranno ambire al massimo dei voti.

In tutti i test si ritroveranno sempre le seguenti domande:

- Quale è stata la soluzione proposta per soddisfare i requisiti del progetto?
- Quali sono stati i risultati raggiunti con la soluzione proposta per soddisfare i requisiti del progetto?

Sarà premura del ***docente valutare*** le risposte fornite, facendo riferimento a quanto svolto nel progetto.

Ovviamente, ad un ***progetto mediocre*** seguirà ***una valutazione mediocre*** delle risposte fornite (anche se verranno scritte delle cose apparentemente bellissime). Ad un ***progetto di ottimo*** livello verrà assegnata la ***massima valutazione.***

CONSIGLI DEL DOCENTE: (facoltativo)

Si invitano gli studenti a leggere attentamente la prima lezione del corso, le FAQ e tutti i documenti predisposti dal docente sulla sua pagina.