Logo Image not found or type unknown

PROGRAMMA DEL CORSO DI DATA MINING E BIG DATA

SETTORE SCIENTIFICO

SECS-S/01 (STAT-01/A)

CFU

9

SETTORE SCIENTIFICO DISCIPLINARE

STAT-01/A

ANNO DI CORSO

I Anno

TIPOLOGIA DI ATTIVITÀ FORMATIVA

/**/
Base q

Caratterizzante X

Affine q

Altre attività q

NUMERO DI CREDITI

9 CFU

DOCENTE

Raffaele Miele

MODALITÀ DI ISCRIZIONE E DI GESTIONE DEI RAPPORTI CON GLI STUDENTI

L’iscrizione ed i rapporti con gli studenti sono gestiti mediante la piattaforma informatica che permette l’iscrizione ai corsi, la fruizione delle lezioni, la partecipazione a forum e tutoraggi, il download del materiale didattico e la comunicazione con il docente. Un tutor assisterà gli studenti nello svolgimento di queste attività.

OBIETTIVI FORMATIVI SPECIFICI

/**/
L’obiettivo del corso è di far acquisire le necessarie conoscenze tecniche di analisi e trattamento dei dati, soprattutto con applicazioni pratiche su Big data, attraverso l’utilizzo di opportuni software per il Data Mining.

RISULTATI DI APPRENDIMENTO ATTESI

/**/
Conoscenza e capacità di comprensione

Il corso mira a far conoscere gli aspetti metodologici e applicativi del Data Mining, partendo dalla classica esplorazione dei dati, per arrivare a metodi più complessi, quali il machine learning e le reti neurali. La conoscenza teorica e applicativa dei differenti metodi viene presentata parallelamente, con particolare attenzione a casi studio con l’uso di Big data.

Capacità di applicare conoscenza e comprensione

Lo studente sarà in grado di applicare le conoscenze apprese durante il corso a casi pratici, valutando quali tecnica applicare e interpretando in maniera appropriata i risultati ottenuti.

Autonomia di giudizio

Il corso intende far acquisire allo studente la capacità di formulare in maniera autonoma la tecnica di analisi più appropriata alla tipologia di informazione disponibile e agli obiettivi di analisi.

Le conoscenze apprese nel corso permetteranno di prendere iniziative e decisioni in termini di estrazione di conoscenza dai dati, in un’ottica di risoluzione di problemi pratici a partire dai dati e dalle informazioni estratte da banche dati classiche oppure da fonti tipo Big data.

Abilità comunicative

Gli strumenti messi a disposizione del corso permettono di comunicare i risultati e le analisi nella maniera più efficace e appropriata. Gli argomenti teorici e le esercitazioni pratiche svolte nel corso supportano tale capacità.

Capacità di apprendimento

Il corso mira a far acquisisre allo studente la capacità di apprendimento e aggiornamento continuo, necessario per affrontare problemi di Data Mining per l’analisi dei dati nei diversi contesti applicativi.

PROGRAMMA DIDATTICO

1 - Data Mining, Apprendimento Automatico. Definizioni

2 - Cos'è un modello di Data Mining e come si usa

3 - L'importanza del Data Mining in azienda. Casi d'uso

4 - Database & Data Warehouse

5 - Big Data

6 - Data Mining. Buone pratiche in azienda. Il modello CRISP-DM

7 - Analisi esplorativa univariata

8 - Analisi esplorativa Bivariata

9 - Analisi esplorativa. Trasformare i dati

10 - Analisi esplorativa. Aspetti pratici.

11 - Introduzione all'SQL. Estrazione dei dati

12 - Introduzione all'SQL. Estrazione dei dati da più tabelle

13 - Introduzione al linguaggio R

14 - Le strutture dati e gli oggetti R più importanti

15 - Gli strumenti R per la manipolazione dei dati

16 - Gli strumenti R per la visualizzazione dei dati

17 - Analisi Esplorativa con R. Aspetti pratici

18 - Costruire il dataset di Training

19 - La costruzione delle Feature

20 - Tassonomia dei modelli di data mining

21 - Modelli di Classificazione

22 - Alberi di Classificazione

23 - Regressione

24 - Misurazione delle performance dei modelli supervisionati

25 - Classificazione: problema del Target Imbalance

26 - Overfitting e Iperparametri

27 - Ensemble Learning

28 - Esercitazione: Regressione. Demand Forecasting.

29 - Calibrazione, Selezione delle variabili e Regolarizzazione

30 - Esercitazione: Rischio di credito

31 - Bias e Fairness dei Modelli. Model Explainability

32 - Esercitazione: Bias e Fairness

33 - Riduzione della dimensionalità

34 - Esercitazione: Data Preparation. ONG

35 - Esercitazione: Scoring. ONG.

36 - Cluster Analysis

37 - Anomaly Detection

38 - Association rules: algoritmo apriori

39 - Social Network Analysis: introduzione

40 - Social Network Analysis: concetti avanzati

41 - Esercitazione: Grafi. Member Get Member

42 - Esercitazione: Grafi. Investigazione AML

43 - Esercitazione: Grafi. Recommendation Engine

44 - Natural Language Processing. Introduzione

45 - Computer Vision. Reti Neurali. Cenni

TIPOLOGIE DI ATTIVITÀ DIDATTICHE PREVISTE E RELATIVE MODALITÀ DI SVOLGIMENTO

/**/
L’insegnamento è articolato in videolezioni di circa 30 minuti corredate da dispense, slide e questionario di autovalutazione.

Per ogni insegnamento è prevista1 videolezione di didattica erogativa in modalità sincrona a contenuto innovativo ed interattivo, secondo modalità definite dal docente di riferimento,vi è altresì la possibilità di redazionedi un elaborato per insegnamento, differenziato in termini di difficoltà rispetto all’ampiezza dei CFU assegnati.

Il modello didattico 2025-2026, in ottemperanza al D.M. 1835 del 6 dicembre 2024, prevede di norma, per ogni CFU, un totale di almeno 7 ore di didattica. La didattica erogativa è perciò effettuata dall’Anno Accademico 2025/2026 per l’80% in modalità asincrona, articolata in un numero di videolezioni coerente ai CFU complessivi del singolo insegnamento, corredate da materiale didattico adeguato allo studio individuale e, per almeno il 20%, in modalità

sincrona

La didattica erogativa asincrona prevede per ogni ora una videolezione registrata, una dispensa corredata da riferimenti bibliografici, note, tabelle, immagini, grafici ed un questionario di dieci domande di autovalutazione con quattro possibili risposte di cui solo una corretta e tre distrattori, oltre un file di riepilogo relativo agli obiettivi ed alla struttura in paragrafi della lezione, con l'aggiunta di alcune parole chiave. Nel dettaglio la videolezione corrisponde alla singola lezione teorica del docente.La didattica sincrona si compone di una web conferenze per CFU e di un elaborato per insegnamento, differenziato in termini di difficoltà rispetto all'ampiezza dei CFU assegnati.L'obiettivo della didattica erogativa in modalità sincrona è assicurare tutte quelle attività che tipicamente richiedono apprendimenti "in situazione" o rapporto "face to face", quali laboratori, seminari, esperienze sul campo, tirocini, ecc., tenendo conto anche delle metodologie a carattere innovativo e volte a favorire l'interazione docente-studenti e tra studenti

Sono previsti:

interventi didattici rivolti da parte del docente/tutor all'intera classe (o a un suo sottogruppo), tipicamente sotto forma di dimostrazioni o spiegazioni aggiuntive (ad esempio dimostrazione o suggerimenti operativi su come si risolve un problema, esercizio esimilari); gli interventi brevi effettuati dai corsisti (ad esempio in ambienti di discussione o di collaborazione); le e-tivity strutturate (individuali o collaborative), sotto forma tipicamente di report, esercizio, studio di caso, problem solving, web quest,progetto,produzionediartefatto(ovariantiassimilabili),effettuati dai corsisti, con relativo feed-back; le forme tipiche di valutazione formativa, con il carattere di questionari o test initinere; le esperienze di apprendimento in situazione realizzabili attraverso ambienti di simulazione, oppure attraverso la virtualizzazione di laboratori didattici.
Nelle suddette attività convergono molteplici strumenti didattici, che agiscono in modo sinergico sul percorso di formazione ed apprendimento dello studente. La partecipazione attiva alle suddette attività ha come obiettivo quello di stimolare gli studenti lungo tutto il percorso didattico e garantisce loro la possibilità di ottenere una valutazione aggiuntiva che si sommerà alla valutazione dell'esame finale.

Nel computo delle ore della didattica erogativa sono escluse le interazioni a carattere orientativo sui programmi, sul Corso di Studio, sull'uso della piattaforma e simili, che rientrano nei servizi di tutoraggio per l'orientamento. Sono altresì escluse le ore di tutorato didattico disciplinare, cioè la mera ripetizione di contenuti già proposti nella forma erogativa attraverso colloqui di recupero o approfondimento one-to-one.

MODALITÀ E CRITERI DI VALUTAZIONE DELL'APPRENDIMENTO

/**/
La partecipazione alla didattica erogativa ha la finalità, tra le altre, di valutare lo studente durante l'apprendimento in itinere.

L'esame finale può essere sostenuto in forma scritta o in forma orale; lo studente può individuare, in autonomia, la modalità di svolgimento della prova, sempre rispettando la calendarizzazione predisposta dall'Ateneo.

L'esame orale consiste in un colloquio nel corso del quale il docente formula almeno tre domande.

L'esame scritto consiste nello svolgimento di un test a risposta multipla con 31 domande. Per ogni domanda lo studente deve scegliere una delle 4 possibili risposte. Solo una risposta è corretta.

Sia la verifica in forma orale che i quesiti in forma scritta sono formulati per valutare il grado di comprensione delle nozioni teoriche e la capacità di sviluppare il ragionamento utilizzando le nozioni acquisite per verificare la capacità di apprendimento ovvero il livello di apprendimento raggiunto dallo studente. I quesiti che richiedono l'elaborazione di un ragionamento consentiranno di valutare il livello di competenza e l'autonomia di giudizio maturati dallo studente.

Le abilità di comunicazione e la capacità di apprendimento saranno valutate attraverso le interazioni dirette tra docente e studenti che avranno luogo durante la fruizione del corso proposte dal docente o dal tutor.

CRITERI DI MISURAZIONE DELL'APPRENDIMENTO E ATTRIBUZIONE DEL VOTO FINALE

/**/
La didattica sincrona garantisce una premialità massima di 2 punti che si somma al voto dell'esame finale, suddivisa in 1 punto per la didattica erogativa sincrona (Webconference) ed 1 punto didattica erogativa sincrona (Elaborato). La premialità massima per le Webconference è di un punto sul voto di esame. Ogni studente può partecipare a tutte le Webconference erogate. Per ciascuna di esse, il superamento del test finale di apprendimento –che richiede almeno quattro risposte corrette su cinque domande relative al tema trattato – consente di ottenere un punteggio pari a 0,5. Una volta raggiunto un punteggio totale di 1, allo studente viene riconosciuta la premialità. La redazione dell'elaborato consente una premialità pari ad 1 punto sul voto dell'esame, se considerato sufficiente. Saranno rese disponibili due tracce di elaborati.

È data facoltà allo studente di partecipare alla didattica erogativa sincrona.

La valutazione finale ha lo scopo di misurare il grado di comprensione delle nozioni teoriche e la capacità di sviluppare il ragionamento utilizzando le nozioni acquisite per verificare la capacità di apprendimento ovvero il livello di apprendimento raggiunto dallo studente. Il giudizio riguarda l'intero percorso formativo del singolo insegnamento ed è di tipo sommativo.
Il voto finale dell'esame di profitto tiene conto del punteggio ottenuto nella verifica di profitto al quale si sommano le premialità che lo studente può aver ottenuto partecipando alla didattica erogativa sincrona e deriva, quindi, dalla somma delle due valutazioni. Il voto derivante dalla didattica sincrona verrà sommato al voto dell'esame se quest'ultimo sarà pari o superiore a diciotto trentesimi.
Il voto finale è espresso in trentesimi. Il voto minimo utile al superamento della prova è di diciotto trentesimi.

Ciascun test dovrà essere composto da 31 domande, così da garantire la possibilità di conseguire la lode, in ottemperanza alle norme Europee sul Diploma Supplement. L'attribuzione della lode è concessa esclusivamente allo studente che ha risposto positivamente alle prime 30 domande ed anche all'ultima domanda.

ATTIVITÀ DI DIDATTICA EROGATIVA ASINCRONA

Di norma massimo l'80% delle lezioni è svolto in modalità asincrona.

ATTIVITÀ DI DIDATTICA EROGATIVA SINCRONA CON RELATIVO FEED-BACK AL SINGOLO STUDENTE DA PARTE DEL DOCENTE O DEL TUTOR

Almeno il 20% delle lezioni è svolto in modalità sincrona e possono prevedere:

è Partecipazione web conference

è Redazione di un elaborato

è Svolgimento delle prove in itinere con feedback

è Svolgimento della simulazione del test finale

MATERIALE DIDATTICO UTILIZZATO

è Videolezioni

è Dispense predisposte dal docente e/o slide del docente

è Questionario di autovalutazione

è Materiali predisposti per le lezioni sincrone

è Testo di riferimento suggerito dal docente (facoltativo)

Il materiale didattico è sempre disponibile in piattaforma e consultabile dallo studente nei tempi e nelle modalità ad egli più affini.