Logo Image not found or type unknown

PROGRAMMA DEL CORSO DI PROGRAMMAZIONE PER LA STATISTICA E I BIG DATA

SETTORE SCIENTIFICO

ING-INF/05 (IINF-05/A)

CFU

9

SETTORE DISCIPLINARE

/**/
ING-INF/05

ANNO DI CORSO

/**/
I Anno

TIPOLOGIA DI ATTIVITÀ FORMATIVA

Base X

Caratterizzante q

Affine q

Altre attività q

NUMERO CREDITI

/**/
9 CFU

DOCENTE

/**/
Claudio Loconsole

OBIETTIVI FORMATIVI SPECIFICI

/**/

Python è un linguaggio intuitivo e un software open source, con una molteplicità di librerie disponibili per i diversi campi applicativi. L'obiettivo del corso è fornire allo studente gli strumenti basilari per poter programmare con tale linguaggio, fornendo in particolare una conoscenza approfondita delle librerie adatte nell'ambito della Data Science.

RISULTATI DI APPRENDIMENTO SPECIFICI

/**/

Conoscenza e capacità di comprensione

Il corso intende fornire le conoscenze utili per implementare programmi con Python, utili per l'analisi statistica dei Big data. Esse includono: conoscenza e comprensione di fondamenti del linguaggio, degli oggetti e funzioni usate, delle principali librerie statistiche e delle capacità grafiche e di visualizzazione di Python. L'attenzione è rivolta essenzialmente ai metodi applicati in caso di big data.

Capacità di applicare conoscenza e comprensione

Il corso permette allo studente di effettuare elaborazioni statistiche su dati standard e Big data, dando la capacità allo stesso di attingere alle opportune librerie/moduli offerte dal programma, per risolvere in modo più appropriato il problema e l'oggetto sotto studio.

Autonomia di giudizio

Autonome capacità di giudizio sulla qualità dei risultati applicativi ottenuti, nonché autonome capacità di giudizio, di valutazione comparativa e scelta di soluzioni, moduli e software, avendo una visione complessiva anche su altri software statistici (R, SAS, SPSS).

Abilità comunicative

Capacità di comunicare in maniera appropriata le elaborazioni svolte tramite Python, sia attraverso gli strumenti di visualizzazione e grafici, che in forma più approfondita, come output delle funzioni e applicazioni eseguite, con particolare attenzione ai casi di utilizzo di Big data. Lo sviluppo di abilità comunicative, sia orali che scritte, sarà anche stimolata attraverso la didattica interattiva (con la redazione di elaborati da parte dello studente) e i momenti di videoconferenza attivati, ivi compreso la prova finale di esame.

Capacità di apprendimento

Capacità di apprendere, sperimentare ed usare, in modo autonomo, linguaggi di programmazione anche di nuova definizione ed implementazione, rivolti all'uso di big data. Capacità di seguire l'evoluzione della programmazione Python in ogni suo aspetto attraverso la letteratura e la documentazione tecnica, rimanendo aggiornato sulle nuove librerie rese disponibili in forma open-source. Capacità di aggiornamento ed autoaggiornamento.

PROGRAMMA DIDATTICO

/**/

1 - IL LINGUAGGIO PYTHON

2 - FONDAMENTI DEL LINGUAGGIO PYTHON

3 - OPERAZIONI SUI DATI

4 - STRUTTURE DECISIONALI E ITERATIVE

5 - FUNZIONI

6 - CASO DI STUDIO: PENSARE CON LE FUNZIONI

7 - FILE

8 - ECCEZIONI

9 - CASO DI STUDIO: PROGRAMMARE UTILIZZANDO FILE E GESTENDO LE ECCEZIONI

10 - LISTE

11 - DIZIONARI E SET

12 - CLASSI E OGGETTI

13 - DATABASE RELAZIONALI

14 - SQL

15 - SQL AVANZATO

16 - CONFIGURAZIONE E USO DI UN DATABASE MYSQL CON PYTHON

17 - OPERARE SULLE STRINGHE

18 - ESERCITAZIONI DI PROGRAMMAZIONE STRUTTURATA

19 - ESPRESSIONI REGOLARI

20 - APPROFONDIMENTO SULLE ESPRESSIONI REGOLARI

21 - APPROFONDIMENTO SUI FILE

22 - OPERAZIONI SULLE SEQUENZE

23 - ANACONDA E LA LIBRERIA NUMPY

24 - ARRAY IN NUMPY

25 - OPERAZIONI SUI DATI

26 - RAPPRESENTAZIONE GRAFICA

27 - PROBABILITA' E STATISTICA

28 - PANDAS

29 - OPERAZIONI CON PANDAS

30 - OPERAZIONI AVANZATE CON PANDAS

31 - TECNOLOGIE PER I BIG DATA

32 - PYSPARK

33 - BIG DATA CON PYSPARK

34 - I GRAFI E LE RETI

35 - UTILITY PER IL RECUPERO E L'ANALISI DEI DATI DELLE RETI

36 - "NETWORKX"

37 - VISUALIZZAZIONE E ANALISI DI GRAFI

38 - ESEMPIO DI ANALISI DEI DATI DELLE RETI

39 - INTRODUZIONE ALLA GUI DI PYTHON CON TKINTER

40 - PROGETTAZIONE DEI FORM CON TKINTER

41 - WIDGET TKINTER E CLASSE LABELINPUT

42 - PROGRAMMAZIONE DEL FORM

43 - INTRODUZIONE A R

44 - FATTORI E MATRICI IN R

45 - DATAFRAME E LISTE IN R

46 - OPERAZIONI E RAPPRESENTAZIONE GRAFICA IN R

47 - SPSS

48 - SAS

49 - PROBLEMI DI DATA SCIENCE: DEFINIZIONE, VARIABILE TARGET E METRICHE

50 - GESTIONE DEI DATI MANCANTI E RIPRODUCIBILITÀ COMPUTAZIONALE

51 - IL FENOMENO BIG DATA

52 - BIG DATA, SOCIAL NETWORK E SISTEMA DELL'INFORMAZIONE

53 - AMAZON WEB SERVICES (AWS)

54 - DOMINI APPLICATIVI DEI BIG DATA E INTEGRAZIONE CON L'AI

TIPOLOGIE DI ATTIVITÀ DIDATTICHE PREVISTE E RELATIVE MODALITÀ DI SVOLGIMENTO

/**/

Ogni Macro-argomento è articolato in 15-17 videolezioni da 30 min. corredate da dispense, slide e test di apprendimento.

Per ogni insegnamento sono previste sino a 6 videolezioni (n.1 CFU) di didattica innovativa secondo modalità definite dal docente di riferimento.

Le videolezioni sono progettate in modo da fornire allo studente una solida base di competenze culturali, logiche e metodologiche atte a far acquisire capacità critiche necessarie ad esercitare il ragionamento matematico, anche in una prospettiva interdisciplinare, a vantaggio di una visione del diritto non meramente statica e razionale, bensì quale espressione della società e della sua incessante evoluzione.

Il modello didattico adottato prevede sia didattica erogativa (DE) sia didattica interattiva (DI):

La didattica erogativa (DE) prevede l'erogazione in modalità asincrona delle videolezioni, delle dispense, dei test di autovalutazioni predisposti dai docenti titolari dell'insegnamento; la metodologia di insegnamento avviene in teledidattica. La didattica interattiva (DI) comprende il complesso degli interventi didattici interattivi, predisposti dal docente o dal tutor in piattaforma, utili a sviluppare l'apprendimento online con modalità attive e partecipative ed è basata sull'interazione dei discenti con i docenti, attraverso la partecipazione ad attività didattiche online.
Sono previsti interventi brevi effettuati dai corsisti (ad esempio in ambienti di discussione o di collaborazione, in forum, blog, wiki), e-tivity strutturate (individuali o collaborative), sotto forma tipicamente di produzioni di elaborati o esercitazioni online e la partecipazione a web conference interattive.

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.

Per le attività di autoapprendimento sono previste 162 ore di studio individuale.

L'Ateneo prevede 7 h per ogni CFU articolate in 6 h di didattica erogativa (DE) e 1 h di didattica interattiva (DI).

Nel computo delle ore della DI sono escluse le interazioni a carattere orientativo sui programmi, sul cds, sull'uso della piattaforma e simili, che rientrano un semplice tutoraggio di 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 interattiva (DI) 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 i quesiti 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. 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 studente che avranno luogo durante la fruizione del corso (videoconferenze, e-tivity report, studio di casi elaborati) proposti dal docente o dal tutor.

CRITERI DI MISURAZIONE DELL'APPRENDIMENTO E ATTRIBUZIONE DEL VOTO FINALE

/**/

Sia lo svolgimento dell'elaborato, sia la presenza attiva durante le web conference prevedono un giudizio, da parte del docente, fino a un massimo di 2 punti. Lo studente può prendere parte ad entrambe le attività ma la votazione massima raggiungibile è sempre di 2 punti.

La valutazione proveniente dallo sviluppo dell'elaborato può essere pari a 0, 1 o 2 punti.

La valutazione derivante dalle web conference è strutturata tramite lo svolgimento, al termine della stessa, di un test finale a risposta multipla che può garantire da 0 a 1 punto.

È data facoltà allo studente di partecipare o meno alla didattica interattiva.

La valutazione finale ha lo scopo di misurare il raggiungimento degli obiettivi di apprendimento definiti alla base dell'insegnamento. 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 che lo studente può aver ottenuto partecipando correttamente alla didattica interattiva e deriva, quindi, dalla somma delle due valutazioni. Il voto derivante dalla didattica interattiva 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.

ATTIVITÀ DI DIDATTICA EROGATIVA (DE)

/**/

- è 54 Videolezioni + 54 test di autovalutazione

Impegno totale stimato: 54 ore

ATTIVITÀ DI DIDATTICA INTERATTIVA (DI) ED E-TIVITY CON RELATIVO FEED-BACK AL SINGOLO STUDENTE DA PARTE DEL DOCENTE O DEL TUTOR

- è Applicazione di Python per risolvere un problema pratico
- è Partecipazione a una web conference
- è Svolgimento delle prove in itinere con feedback
- è Forum sui diversi software statistici

Totale 9 ore

MATERIALE DIDATTICO UTILIZZATO

Videolezioni

Dispense predisposte dal docente e/o slide del docente

Testo di riferimento suggerito dal docente (facoltativo):

Gaddis, T. (2016). Introduzione a Python. Ediz. mylab. Con espansione online. Pearson. Zinoviev, D. (2017). Data science con Python. Apogeo. Maggi, G. (2020). Data science con Python. Edizioni LSWR. Marin, I., Shukla, A. Sarang, VK. (2019). L’analisi dei Big Data con Python. Tecniche nuove. Moore, A. D. (2018). Python GUI Programming with Tkinter. Packt. Botto, M., Raganelli, V. (2021). Programmazione e analisi statistica con R. in riga edizioni.Di Franco, G. (2009). L'analisi dei dati con SPSS. FrancoAngeli

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