



Analisi e modellazione dei dati e dei processi: metodi e modelli - AMDP

Master di I livello

UnitelmaSapienza.it



Indice

- 1 Finalità
- 2 Destinatari
- 3 Requisiti di accesso
- 4 Didattica
- 5 Durata e organizzazione del Master
- 6 Direzione e docenti
- 7 Assistenza e tutoraggio
- 8 Project work, prova finale e titolo
- 9 Quote e modalità di iscrizione
- 10 Programma



1. Finalità

Nell'odierna società della conoscenza i dati e la capacità di analizzarli per estrarre da essi informazione e conoscenza assumono una importanza strategica per il successo di una qualsiasi organizzazione, sia pubblica che privata. Questo Master intende fornire al discente le conoscenze e le competenze necessarie per modellare e analizzare i dati di interesse di una organizzazione e i processi che ne regolano il funzionamento e dai quali i dati si originano. La capacità di modellare e analizzare dati di varia natura e dimensione (mediante tecniche di text mining, social network analysis e machine learning) consentirà al discente, da un lato, di conoscere i dati in gioco nella realtà in esame, le relative modalità di archiviazione (con particolare riguardo ai database relazionali) e trattamento degli stessi (anche con riguardo agli aspetti di sicurezza e tutela della privacy), e dall'altro, di individuare correlazioni, pattern interpretativi e indici statistici utili all'estrazione di informazione e conoscenza utili per l'organizzazione. La capacità di modellare e analizzare i processi, oltre che a consentire al discente di comprendere e documentare le modalità correnti di funzionamento della realtà in esame, permetteranno di individuare eventuali criticità e opportunità di miglioramento nei processi stessi e definire le modalità di introduzione degli stessi. La figura professionale formata dal Master avrà capacità analitiche, e quindi una propensione al ragionamento matematico e statistico e all'analisi e soluzione di problemi. A queste capacità si aggiungono conoscenze di programmazione necessarie allo sviluppo di routine software e conoscenze relative all'uso di specifiche tecnologie informatiche.

2. Destinatari

Il Master si rivolge a dirigenti e funzionari pubblici; operatori e responsabili dei sistemi informativi e dei centri elaborazione dati; responsabili della comunicazione e del marketing; responsabili del controllo dei processi e delle procedure; laureati in discipline tecnico-scientifiche interessati a specializzarsi nell'analisi e modellazione dei dati e dei processi.

3. Requisiti di accesso

Possono accedere al master coloro che sono in possesso di: i. Laurea conseguita ai sensi del DM n. 270/2004; ii. Laurea conseguita ai sensi del DM n. 509/1999; iii. Laurea conseguita secondo gli ordinamenti previgenti; iv. Titolo rilasciato all'estero, riconosciuto idoneo dalla normativa vigente.



4. Didattica

Il master sarà erogato esclusivamente in rete, ai sensi del decreto interministeriale del 17 aprile 2003 e relativo allegato tecnico come modificato dal DM 15 aprile 2005, secondo il modello di "lezione digitale" adottato dall'Università UnitelmaSapienza, tramite la piattaforma e-learning dell'ateneo.

La didattica in e-learning prevede le lezioni audio-video dei docenti, metadate ed indicizzate, integrata dai relativi materiali didattici scaricabili in rete. Lo studente segue le lezioni quando vuole – h 24 – come vuole - PC, tablet, smartphone – nella modalità prescelta – video/audio video, audio-streaming/downloading MP3.

Il docente provvederà a monitorare il percorso formativo con prove in itinere on line: in particolare saranno effettuate delle prove scritte che possono essere strutturate come tesine o test a risposta aperta, chiusa o sintetica con la relativa votazione espressa in trentesimi, e che costituiscono un mezzo per la verifica e la misurazione delle conoscenze per il docente e di autovalutazione per lo studente.

La didattica tiene conto della necessità di sviluppare non solo conoscenze, ma anche di utilizzare strumenti e comportamenti professionali, con le seguenti caratteristiche:

- attività di progetto sul campo mirate a verificare la capacità di applicazione delle metodologie e delle tecniche in situazioni reali;
- testimonianze e casi di studio.

A supporto dell'attività didattica rappresentata dalle videolezioni, verranno svolti dei web seminar, attività realizzata in modalità sincrona che prevede varie tipologie:

- webinar tematico, in cui saranno proposti approfondimenti dedicati ad alcuni temi di particolare interesse ed attualità, al fine di rendere più dinamica la didattica, promuovendo anche spazi di confronto interattivi tra docenti e studenti;
- webinar di discussione, che avranno per oggetto argomenti teorici tratti dal programma, su cui i partecipanti al corso saranno invitati a proporre il proprio punto di vista sulla base delle esperienze maturate nell'ambito delle rispettive organizzazioni di appartenenza, al fine di promuovere dinamiche di apprendimento proattive, in grado di rivelare le principali implicazioni e potenzialità applicative delle teorie studiate.

Per valorizzare in pieno le opportunità connesse alla didattica interattiva online, potranno essere utilizzati anche i seguenti strumenti:

- wiki, per costruire insieme agli studenti "alberi" ipertestuali di conoscenze condivise, dove i contributi teorici potranno essere integrati con apporti di natura empirica proposti da tutti i partecipanti al corso, i quali potranno, così, reciprocamente arricchire i propri percorsi di apprendimento;
- forum, nei quali potranno essere lanciati dal docente (ma anche dagli stessi studenti) provocazioni al dibattito su temi specifici o più generali trattati nell'ambito del Master, ma anche su iniziative tese a favorire un apprendimento critico e interattivo, come nel caso della



discussione di case study o di una rilettura problematica di eventi di attualità connessi agli argomenti teorici del corso;

5. Durata ed organizzazione del Master

La durata di un Master è di norma stabilita in 12 (dodici) mesi. Il corsista in regola con il superamento delle attività didattiche previste dal proprio piano di studi e che abbia assolto agli adempimenti didattici ed amministrativi previsti dall'Ateneo, potrà presentare apposita domanda di conseguimento titolo anticipato, comunque non prima di 6 (sei) mesi dalla data di prima iscrizione al Master. L'Ateneo, inoltre, può concedere una proroga per completare il Master entro il termine massimo di ulteriori 12 (dodici) mesi. La proroga comporta il rinnovo dell'iscrizione con pagamento contestuale di una tassa pari al 25% della tassa di iscrizione annuale originariamente corrisposta. Nel caso in cui lo studente non consegua il titolo entro tale termine di proroga, e voglia comunque proseguire il percorso, potrà procedere al rinnovo dell'iscrizione con pagamento della tassa di iscrizione annuale originariamente corrisposta.

6. Direzione e docenti

La direzione del Master è affidata al Prof. Damiano Distante, professore associato di Informatica (SSD INF/01) presso l'Università degli Studi di Roma UnitelmaSapienza, con esperienza più che decennale nelle materie del master e nella docenza universitaria.

Fanno parte della faculty del Master, oltre al Direttore, i seguenti docenti:

- Dott. Marco Cascio, Ricercatore di Informatica (SSD INF/01) presso l'Università degli Studi di Roma UnitelmaSapienza;
- Prof.ssa Rosella Castellano, Professore Ordinario di Metodi matematici dell'economia e delle scienze attuariali e finanziarie (SSD SECS-S/06) presso l'Università degli Studi di Roma UnitelmaSapienza;
- Prof. Carlo Cavicchia, ricercatore di Statistica presso l'Erasmus Rotterdam University e docente a contratto presso l'Università degli Studi di Roma Unitelma Sapienza;
- Prof.ssa Marta Cimitile, Professore Associato di Sistemi di elaborazione delle informazioni (SSD ING-INF/05) presso l'Università degli Studi di Roma UnitelmaSapienza;
- Prof. Pasquale Sarnacchiaro, Professore Associato di Statistica (SSD SECS-S/01) presso l'Università degli Studi Federico II di Napoli;
- Dott.ssa Fabiola Orlandi, Avvocato, esperta di Diritto civile e di Diritto amministrativo, cultore di diritto amministrativo e tutor presso l'Università degli Studi di Roma UnitelmaSapienza.
- Dott.ssa Giorgia Zaccaria, Assegnista di Ricerca (SSD SECS-S/01) presso l'Università degli Studi di Milano Bicocca.



7. Assistenza e tutoraggio

L'Ateneo garantisce in via continuativa un'assistenza ed un sostegno al processo di apprendimento degli allievi mediante la presenza di un tutor esperto e qualificato.

Il tutor, oltre a prestare assistenza agli studenti dell'Università, gestisce in modo proattivo il rapporto con la classe virtuale degli studenti, predisponendo ogni strumento didattico ed informativo ritenuto utile per supportare gli studenti, utilizzando anche modalità sincrone ed asincrone, quali chat, forum etc.

8. Project work, prova finale e titolo

Per ciascun partecipante è previsto, durante lo svolgimento del Master, lo svolgimento di prove Intermedie (una per ciascun modulo didattico) e di un project work finale. Il project work (PW) è una attività complessa e strutturata che comprende la progettazione e la realizzazione di una ricerca sul campo, finalizzata ad applicare, su casi concreti, le conoscenze acquisite, ovvero attività di ricerca, studio, interazione con il docente e il tutor, e la stesura e discussione di un elaborato finale di PW.

La durata del PW, che prevede l'attribuzione di 20 CFU, può essere indicativamente valutata in 500 ore, considerando l'impegno complessivo dello studente in attività di ricerca, studio, interazione con il docente e il tutor, stesura e discussione dell'elaborato finale.

La scelta dell'argomento su cui il PW dovrà vertere dovrà essere effettuata, con riferimento ai moduli del master, in coerenza con le modalità definite dal Direttore del Master.

Il PW potrà essere svolto singolarmente o in gruppo in relazione alla complessità dell'argomento e alla provenienza degli iscritti, ma la valutazione sarà comunque individuale. L'elaborato finale sarà presentato e sottoposto alla valutazione della Commissione per il conseguimento del titolo di Master. Il voto verrà espresso in centodecimi.

A seguito della valutazione positiva dell'elaborato finale da parte della Commissione di Master, l'Università rilascerà il diploma di Master Universitario di I livello in "Analisi e modellazione dei dati e dei processi: metodi e modelli - AMDP" con l'attribuzione di 60 CFU.

9. Quote e modalità di Iscrizione

L'iscrizione al Master è possibile in qualsiasi momento dell'anno, on line, sul sito www.unitelmasapienza.it, attraverso l'apposita sezione "Iscriviti".

La quota integrale di iscrizione al Master è stabilita in € 1.600,00 frazionabili in 3 rate di cui la prima di € 600,00 all'atto dell'iscrizione e le successive, di € 500,00 cadauna, a 30 gg e 60 gg dalla data di iscrizione.

Al termine del percorso è previsto il pagamento della tassa conseguimento titolo pari a € 200,00.

Si informa che il bollo per l'iscrizione (€ 16,00) e i due bolli per il conseguimento titolo (tot. € 32,00) sono previsti per normativa vigente dell'Agenzia dell'Entrate. La proroga annuale (come da art. 5 del presente Bando) comporta il rinnovo dell'iscrizione con pagamento contestuale di una tassa pari al 25% della tassa di iscrizione annuale originariamente corrisposta.



10. Programma

Il master si articola nei seguenti 11 moduli formativi:

Modulo I. Analisi e modellazione di processi di business, SSD ING-INF/05 - 4 CFU

Docente: Prof.ssa Marta Cimitile

Obiettivi formativi: Il modulo intende fornire al discente la base di conoscenze e metodi necessari per la modellazione, la simulazione e l'analisi di modelli di business. Le lezioni contengono nozioni teoriche e pratiche (alcuni dei principali tool attualmente disponibili per la modellazione e l'analisi dei processi sono presentati durante il corso). Inoltre, sono mostrati esempi di modellazione e analisi di processi reali.

Argomenti trattati: Modellazione di Processi di Business; Data Flow Diagram; BPMN; UML per la modellazione dei processi; Process Mining.

Software utilizzati nelle esercitazioni: Disco

Modulo II. Progettazione, implementazione e gestione di basi di dati, SSD INF/01 - 4 CFU

Docente: Prof. Damiano Distante

Obiettivi formativi: Il modulo fornisce al discente conoscenze sulle metodologie e i modelli per la progettazione di basi di dati relazionali, sul linguaggio SQL sia per la definizione che per la manipolazione dei dati, e sui DBMS per la realizzazione e gestione di basi di dati. Il modulo fornisce inoltre nozioni di base e una panoramica sui database NoSQL.

Argomenti trattati: Sistemi informativi, Basi di dati e DBMS; Il processo di progettazione di una base di dati; Il modello Entità-Relazione; La progettazione concettuale; Il modello relazionale; La progettazione logica; Il linguaggio SQL per la definizione dei dati; Il linguaggio SQL per la manipolazione dei dati; I database NoSQL.

Linguaggi e DBMS di riferimento: SQL, MySQL, Oracle

Modulo III. Open Data, SSD ING-INF/05 - 4 CFU

Docente: Prof.ssa Marta Cimitile

Obiettivi formativi: Il modulo intende fornire al discente la base di conoscenze e metodi necessari per la raccolta, la gestione e l'analisi degli Open Data. Le lezioni sono di tipo teorico e pratico. I contenuti del Corso sono inoltre approfonditi con casi di studio e esempi di applicazioni di Open Data in contesti reali.

Argomenti trattati: Introduzione agli Open Data; I portali Open Data; La sostenibilità dei portali Open Data; La scala di Berners Lee; Open Source vs Open Data; Linked Open Data; Semantic Web.



Modulo IV. Analisi statistica dei dati - Inferenza statistica, SSD SECS-S/01 - 4 CFU

Docente: Prof. Pasquale Sarnacchiaro

Obiettivi formativi: Il modulo si propone di fornire le basi concettuali e operative del ragionamento statistico. In particolare, esso prevede un'introduzione all'analisi statistica dei dati, l'esposizione delle nozioni del calcolo delle probabilità e dell'inferenza statistica evidenziando l'utilità delle tecniche statistiche per l'analisi empirica

Argomenti trattati: Probabilità concetti di base; Principali teoremi sulle probabilità; Variabili casuali discrete; Variabili casuali continue; Distribuzioni e campionamento; Principali metodi di campionamento; Teoria della stima; Intervalli di confidenza; Determinazione della numerosità campionaria; Test d'ipotesi per medie; Test d'ipotesi per varianze; Test d'ipotesi per proporzioni

Linguaggi e software utilizzati nelle esercitazioni: R, RStudio

Modulo V. Analisi statistica dei dati - Analisi statistica multivariata, SSD SECS-S/01 - 4 CFU

Docente: Prof. Carlo Cavicchia

Obiettivi formativi: Lo scopo del modulo è quello di sviluppare nello studente le abilità necessarie per l'analisi di sistemi complessi di dati e per la successiva interpretazione dei risultati ottenuti. Al termine del corso, lo studente avrà acquisito nozioni e tecniche necessarie all'analisi di dati di grandi dimensioni (big data) tramite tecniche di riduzioni dimensionale, modelli a variabili latenti e analisi dei gruppi. Il corso prevede l'utilizzo del software R (e la sua interfaccia Rstudio) per l'applicazione delle tecniche e dei modelli trattati.

Argomenti trattati: Richiami di statistica e correlazione, Analisi delle componenti principali, determinazione delle componenti principali, rappresentazioni grafiche dell'analisi in componenti principali, modelli fattoriali, differenze tra l'analisi delle componenti principali e l'analisi fattoriale, cluster analysis: metodi gerarchici, metodi non gerarchici, uso congiunto di analisi delle componenti principali e cluster analysis.

Linguaggi e software utilizzati nelle esercitazioni: R, RStudio

Modulo VI. Analisi statistica dei dati - Modelli statistici di regressione, SSD SECS-S/01 - 4 CFU

Docente: Dott.ssa Giorgia Zaccaria

Obiettivi formativi: Il modulo mira ad approfondire teoria e applicazioni dei modelli di regressione con particolare riferimento ai modelli lineari. Sono trattati modelli per dati continui e binari.

Argomenti trattati: Regressione lineare semplice: introduzione e stima dei parametri, intervalli di confidenza, test d'ipotesi e bontà di adattamento del modello; Regressione lineare multipla: introduzione e stima dei parametri, proprietà degli stimatori e bontà di



adattamento del modello, test d'ipotesi, predittori qualitativi e analisi dei residui; Regressione logistica: introduzione e stima dei parametri.

Linguaggi e software utilizzati nelle esercitazioni: R, RStudio

Modulo VII. Analisi di dati testuali, SSD INF/01 - 4 CFU

Docente: Dott. Marco Cascio

Obiettivi formativi: Il modulo fornisce al discente le conoscenze di base sull'elaborazione automatica dei testi. In questo modulo sono introdotte le principali applicazioni e sono forniti gli strumenti teorici e metodologici di base per la comprensione e la realizzazione sia di modelli linguistici sia di espressioni regolari. Infine, sono trattate applicazioni avanzate quale l'analisi del sentimento per aspetti e dove verranno mostrate le molteplici criticità dell'intero flusso di elaborazione.

Argomenti trattati: Testi digitali; Analisi dei testi; Elementi di calcolo delle probabilità; Occorrenze, Co-occorrenze, Collocazioni; Modelli linguistici; Smoothing; Esercizi su Modelli Linguistici; Espressioni Regolari; Analisi del Sentimento; Analisi del Sentimento per aspetti.

Modulo VIII. Analisi di reti sociali, SSD INF/01 - 4 CFU

Docente: Dott. Marco Cascio

Obiettivi formativi: Il modulo intende fornire al discente le conoscenze di base sull'analisi delle reti sociali. Saranno introdotte quindi le principali applicazioni e verranno forniti gli strumenti teorici e metodologici di base per la comprensione e la realizzazione di esperimenti che anche grazie al supporto di tool allo stato dell'arte permetteranno al discente di sviluppare una piena consapevolezza sui limiti e le potenzialità dettati dalla disciplina. Il modulo prevede oltre alle videolezioni delle e-tivity rivolte alla sperimentazione diretta delle metodologie.

Argomenti trattati: Introduzione alle reti sociali; Definizioni; Misure sulle reti sociali; Il software Pajek; Reti sociali con Pajek; Graph Database.

Software utilizzati nelle esercitazioni: Pajek

Modulo IX. Machine Learning e Big Data, SSD SECS-S/06 - 4 CFU

Docente: Prof.ssa Rosella Castellano

Obiettivi formativi: Il modulo ha come obiettivo quello di discutere i principali modelli quantitativi utilizzati per l'analisi e la classificazione dei Big Data. In particolare, tali modelli non saranno discussi solo per via teorica, ma anche attraverso applicazioni pratiche di Machine Learning con il software open source WEKA.

Argomenti trattati: Introduzione. Concetti Essenziali: cosa si intende per advanced analytics; focus on predictive; dizionario e terminologia. I Modelli Predittivi: regressione lineare e logistica; reti neurali; alberi decisionali; ensemble modeling; algoritmi di clustering. Metodi per la valutazione delle performance dei modelli. Il processo di costruzione del modello.



Software utilizzati nelle esercitazioni: Weka.

Modulo IX bis. Laboratorio di Machine Learning con Python (SSD INF/01)

Lezioni a cura di: Dott. Lorenzo Madeddu

Obiettivi formativi: A fine laboratorio lo studente sarà in grado di applicare, utilizzando il linguaggio e le librerie Python, i modelli di Machine Learning più popolari, comprendendone la costruzione, valutazione e interpretazione. Inoltre, i numerosi esempi pratici offriranno gli strumenti per comprendere ed affrontare problematiche dei dataset reali come, ad esempio, casistiche di dati sbilanciati o sparsi.

Argomenti trattati: Richiami sui concetti base di Python. Uso di librerie di analisi dati: Numpy e Pandas. Esempi pratici e interattivi di Data Visualization con la libreria Plotly. Uso pratico di Scikit-Learn per sviluppare modelli predittivi supervisionati e non, quali: regressione lineare, regressione logistica, naive bayes, support vector machine, alberi di decisione, random forest, boosting, reti neurali e k-means. Tecniche di valutazione di un modello predittivo come, ad esempio, Precision, Recall, AUROC, AUPRC e K-Cross Fold-Validation. Introduzione all'interpretazione dei modelli "white box" per individuare, ad esempio, le feature più importanti per un determinato modello.

Linguaggi e software utilizzati: Python, Anaconda, Jupyter, librerie Python per il machine learning.

Modulo X. I dati personali: Profili normativi di tutela e sicurezza, SSD IUS/01 - 4 CFU

Docente: Dott.ssa Fabiola Orlandi

Obiettivi formativi: In questo modulo formativi si discutono i profili normativi di tutela e sicurezza dei dati personali, prendendo in esame la più recente normativa in materia.

Argomenti trattati: Banche dati e normativa europea; La nozione di Open Data; I diritti fondamentali; La protezione dei dati personali; Il d.lgs. 30 giugno 2003, n. 196; La responsabilità per illecito trattamento dei dati personali; Il Garante per la protezione dei dati personali; Il Regolamento UE n. 679/2016; Il decreto legislativo n. 101 del 2018; Cyber security e protezione dei dati personali.

INFORMAZIONI

Università degli Studi di Roma
UnitelmaSapienza
www.unitelmasapienza.it
info@unitelmasapienza.it