

## SCHEDA DELL'INSEGNAMENTO (SI)

### "ALGORITMI, STRUTTURE DATI E MACHINE LEARNING"

SSD ING-INF/05\*

DENOMINAZIONE DEL CORSO DI STUDIO: CORSO DI LAUREA MAGISTRALE IN INGEGNERIA INFORMATICA

ANNO ACCADEMICO 2023-2024

#### INFORMAZIONI GENERALI - DOCENTE

DOCENTE: ROBERTO PIETRANTUONO, CARLO SANSONE

TELEFONO: 0817683880, 0817683640

EMAIL: ROBERTO.PIETRANTUONO@UNINA.IT, CARLO.SANSONE@UNINA.IT

#### INFORMAZIONI GENERALI - ATTIVITÀ

INSEGNAMENTO INTEGRATO (EVENTUALE):

MODULO (EVENTUALE): ALGORITMI E STRUTTURE DATI, MACHINE LEARNING

CANALE (EVENTUALE):

ANNO DI CORSO (I, II, III): I

SEMESTRE (I, II): ANNUALE

CFU: 12

## **INSEGNAMENTI PROPEDEUTICI (se previsti dall'Ordinamento del CdS)**

**Nessuno**

## **EVENTUALI PREREQUISITI**

**Nessuno**

## **OBIETTIVI FORMATIVI**

L'insegnamento si propone di fornire in primo luogo le nozioni necessarie per la progettazione e l'analisi di algoritmi e strutture dati nello sviluppo delle applicazioni informatiche. Tali nozioni includono i fondamenti teorici e le tecniche avanzate di progettazione ed analisi di algoritmi la cui applicazione spazia su tutti gli aspetti relativi ad un sistema di elaborazione, da quelli hardware a quelli software, dai sistemi operativi alle reti di elaboratori, dalle basi di dati ai sistemi informativi, dai linguaggi di programmazione all'ingegneria del software, dall'interazione uomo-macchina al riconoscimento dei segnali e delle immagini, all'elaborazione multimediale, all'ingegneria della conoscenza, alla robotica e all'intelligenza artificiale ed alla robotica. Con riferimento a quest'ultimo campo, l'insegnamento si propone di presentare nel dettaglio i principali algoritmi di Machine Learning per la soluzione di problemi di classificazione, predizione numerica e clustering, nonché le metodologie di gestione e sviluppo di un processo di Machine Learning, dalla preparazione dei dati alla valutazione dei risultati. L'insegnamento consentirà anche di sviluppare competenze pratiche nella soluzione di problemi reali di classificazione, predizioni numerica o clustering tramite algoritmi di Machine Learning, grazie ad esercitazioni svolte con tool open source e/o commerciali.

## **RISULTATI DI APPRENDIMENTO ATTESI (DESCRITTORI DI DUBLINO)**

### **Conoscenza e capacità di comprensione**

#### Modulo di Algoritmi e Strutture Dati

Al termine del corso, lo studente deve dimostrare di aver acquisito familiarità con una ampia varietà di strutture dati ed algoritmi noti che risolvono problemi di carattere fondamentale, di aver compreso le tecniche per la sintesi di nuovi algoritmi e di padroneggiare i metodi per analizzare la correttezza e la complessità asintotica degli algoritmi.

#### Modulo di Machine Learning

Al termine del corso, lo studente deve dimostrare di conoscere i principali algoritmi di Machine Learning e Deep Learning e di essere in grado di scegliere l'algoritmo di Machine Learning più adatto a risolvere uno specifico problemi di classificazione e/o predizione numerica e/o clustering, sulla base dei requisiti del problema medesimo. Lo studente deve inoltre dimostrare di essere in grado di scegliere le opportune tecniche di preparazione dei dati e deve conoscere le tecniche necessarie per la valutazione delle prestazioni di algoritmi di Machine Learning e Deep Learning.

### **Capacità di applicare conoscenza e comprensione**

#### Modulo di Algoritmi e Strutture Dati

Al termine del corso, lo studente deve dimostrare di sapere applicare e combinare le principali tecniche di progettazione, nonché strutture dati avanzate, per la sintesi di algoritmi corretti ed efficienti per la risoluzione di problemi nello sviluppo delle applicazioni informatiche, e di saperne analizzare formalmente la correttezza e la complessità asintotica. Il percorso formativo è orientato a fornire le capacità e gli strumenti necessarie a risolvere problemi nuovi o non familiari negli ampi contesti relativi ai sistemi di elaborazione dell'informazione.

#### Modulo di Machine Learning

Al termine del corso, lo studente deve dimostrare di essere in grado di risolvere problemi reali di classificazione, predizione numerica o clustering utilizzando algoritmi di Machine Learning o Deep Learning. Lo studente deve anche dimostrare di saper valutare in modo corretto le prestazioni dei sistemi che realizza.

## **PROGRAMMA-SYLLABUS**

## Modulo di Algoritmi e Strutture Dati (6 CFU)

- **CONCETTI INTRODUTTIVI (0.25 CFU)**  
Algoritmi e strutture dati, ricorsione, analisi e progettazione degli algoritmi.
- **TECNICHE DI ANALISI E STRUTTURE DATI ELEMENTARI (1 CFU)**  
Analisi di correttezza: invariante di ciclo, correttezza di algoritmi ricorsivi.  
Analisi di complessità: analisi asintotica, notazioni  $O$ ,  $\Omega$ ,  $\Theta$ ; analisi di algoritmi ricorsivi.  
Strutture dati elementari: dizionari; pile e code; code di priorità; liste; tabelle hash e stringhe; alberi binari di ricerca.
- **DIVIDE et IMPERA, ORDINAMENTO (1.25 CFU)**  
Problemi ed algoritmi comuni. Ordinamento: merge sort, heap sort, quick sort.  
Ordinamento in tempo lineare (counting sort, radix sort, bucket sort), mediane e statistiche d'ordine.
- **RICERCA COMBINATORIALE (0.25 CFU)**  
Ricerca esaustiva, ricerca combinatoriale; backtracking, pruning della ricerca
- **PROGRAMMAZIONE DINAMICA (1 CFU)**  
Introduzione alla programmazione dinamica. Ricerca esaustiva vs. ricerca greedy vs. programmazione dinamica.  
Applicazioni. Problemi ed algoritmi comuni. Problema di string matching, edit distance, longest increasing sequence. Fibonacci. Problema dello zaino. Ulteriori esempi.
- **STRUTTURE DATI E TECNICHE DI ANALISI AVANZATE (1 CFU)**  
Alberi RB, alberi auto-aggiustanti. Grafi. Rappresentazione, esplorazione in ampiezza e profondità, ordinamento topologico. Applicazioni. Tecniche di analisi di algoritmi avanzate: analisi ammortizzata.
- **PROBLEMI ED ALGORITMI COMUNI, ESEMPI APPLICATIVI (1 CFU)**  
Problemi di teoria dei numeri (es.: algoritmi DES ed RSA). Problemi su grafi: ricerca di cammini minimi. Algoritmi paralleli, multithreading dinamico. Esempi (Fibonacci, ordinamento). Analisi e confronto con algoritmi sequenziali.  
Traduttori ed interpreti: analisi lessicale, analisi sintattica, strutture dati usate nei traduttori.
- **PROBLEMI INTRATTABILI (0.25 CFU)**  
Introduzione a problemi NP ed NP-completi. Riducibilità. Esempi di problemi NP-completi.

Parte Esercitativa: Prevalentemente in C.

## Modulo di Machine Learning (6 CFU)

- **INTRODUZIONE AL MACHINE LEARNING (0.25 CFU)**
- **INPUT ED OUTPUT (0.5 CFU)**  
Concetti, istanze ed attributi. Rappresentazione della conoscenza.
- **METODI DI BASE (1.25 CFU)**  
Modelli probabilistici, alberi di decisione, regole di classificazione, modelli lineari, instance-based e multi-instance learning, clustering.
- **VALUTAZIONE DELLE PRESTAZIONI (0.5 CFU)**  
Training e Testing. CV, LOO, Cost-sensitive classification. ROC. Valutazione di algoritmi di predizione numerica.
- **METODI AVANZATI (2 CFU)**  
Alberi di decisione: C4.5. Regole di classificazione. Instance-based learning. Estensione dei modelli lineari: SVM. Epsilon-SVR. MLP. Predizione numerica con modelli lineari.
- **TRASFORMARE I DATI (0.25 CFU)**

Selezione degli attributi, PCA, Discretizzazione, Campionamento. One-class classification.

- **METODI PROBABILISTICI (0.25 CFU)**  
Reti bayesiane. Probability Density Estimation e Clustering. Modelli sequenziali e temporali (cenni).
- **DEEP LEARNING (0.5 CFU)**  
Addestramento e valutazione delle prestazioni di reti deep, Convolutional Neural Networks, Autoencoders. Reti Neurali ricorrenti e GAN (cenni).
- **OLTRE L'APPRENDIMENTO SUPERVISIONATO E NON-SUPERVISIONATO (0.25 CFU)**  
Semi-supervised learning. Multi-instance learning.
- **ENSEMBLE LEARNING (0.25 CFU)**  
Bagging, Randomization, Boosting, Stacking, ECOC.

Parte Esercitativa: Utilizzo di tool open source e/o commerciali (e.g., Knime, Weka).

## **MATERIALE DIDATTICO**

### **Modulo di Algoritmi e Strutture Dati**

#### LIBRO DI TESTO ADOTTATO

- 1) Thomas Cormen, Charles Leiserson, Ronald Rivest, and Clifford Stein. **Introduction to Algorithms**. 3rd ed. MIT Press, 2009. ISBN: 9780262033848.

TRASPARENZE DALLE LEZIONI ED ESERCITAZIONI disponibili sul sito web docenti di Ateneo sulla piattaforma Microsoft Teams.

#### LIBRO CONSIGLIATO

- 2) Steven Skiena. **The Algorithm Design Manual**, 3rd ed, Springer, 2020. ISBN-13: 978-3030542559, ISBN-10: 3030542556.

### **Modulo di Machine Learning**

#### LIBRO DI TESTO ADOTTATO

- 1) Ian H. Witten, Eibe Frank, Mark A. Hall and Christopher J. Pal. **Data mining: practical machine learning tools and techniques**. 4th ed. The Morgan Kaufmann, 2017. ISBN: 9780128042915.

## **MODALITÀ DI SVOLGIMENTO DELL'INSEGNAMENTO**

### **Modulo di Algoritmi e Strutture Dati**

Il docente utilizzerà: a) lezioni frontali per circa l'80% delle ore di lezione totali, b) esercitazioni per circa il 20% delle ore di lezione totali. È prevista l'assegnazione di esercizi da svolgere autonomamente e consegnare al docente.

### **Modulo di Machine Learning**

Il docente utilizzerà: a) lezioni frontali per circa il 60% delle ore totali, b) esercitazioni per approfondire aspetti teorici per circa il 35% delle ore totali, c) seminari per circa il 5% delle ore totali.

## **VERIFICA DI APPRENDIMENTO E CRITERI DI VALUTAZIONE**

**a) Modalità d'esame:**

L'esame si articola in prova	
scritta e orale	X
solo scritta	
solo orale	
discussione di elaborato progettuale	X
altro	X

In caso di prova scritta i quesiti sono	A risposta multipla	
	A risposta libera	X
	Esercizi numerici	

Per il modulo di Algoritmi e strutture Dati, l'esame si articola in n. 2 prove scritte ed in una prova orale che include una discussione sugli esercizi assegnati durante il corso. La parte scritta si articola in n.1 prova intercorso più n.1 prova finale. Le prove richiedono, per uno o più problemi, di implementare algoritmi per la loro risoluzione ed analizzarne la complessità asintotica.

La prova orale consta di un colloquio sostenuto dopo l'ultima prova scritta. Esso include anche la discussione sulla risoluzione degli esercizi assegnati durante il corso.

Per il modulo di Machine Learning, ciascuno studente dovrà sviluppare durante il corso 3 progetti (c.d. mini-contest), che, in caso di valutazione positiva, lo esonerano dallo sviluppo dell'elaborato progettuale finale.

La prova orale consta di un colloquio, che include anche la discussione di uno dei 3 progetti sviluppati durante il corso o dell'elaborato progettuale finale.

**b) Modalità di valutazione:**

Per il modulo di Algoritmi e Strutture Dati, le n.2 prove scritte pesano ciascuna il 30% sul giudizio finale (per un totale del 60%). La prova orale pesa per il restante 40% sul giudizio finale (con un peso del 30% per la valutazione e discussione degli esercizi e del 10% per l'interrogazione sulla parte del programma non coperta dalle prove e dagli esercizi).

Per il modulo di Machine Learning, la discussione dell'elaborato progettuale finale (o del progetto sviluppato durante il corso) pesa per il 10% sul giudizio finale, mentre la prova orale per il 90%.

Il voto finale dell'insegnamento sarà ponderato in base ai CFU di ciascun modulo e quindi così composto:

- Modulo di Algoritmi e Strutture Dati, 6 CFU, 50%
- Modulo di Machine Learning, 6 CFU, 50%