



UNIVERSITÀ DEGLI STUDI DI PALERMO

DIPARTIMENTO	Matematica e Informatica
ANNO ACCADEMICO OFFERTA	2023/2024
ANNO ACCADEMICO EROGAZIONE	2023/2024
CORSO DILAUREA MAGISTRALE	DATA, ALGORITHMS, AND MACHINE INTELLIGENCE
INSEGNAMENTO	INFORMATION THEORY AND DATA COMPRESSION
TIPO DI ATTIVITA'	B
AMBITO	50341-Discipline Informatiche
CODICE INSEGNAMENTO	22450
SETTORI SCIENTIFICO-DISCIPLINARI	INF/01
DOCENTE RESPONSABILE	SCIORTINO MARINELLA Professore Ordinario Univ. di PALERMO
ALTRI DOCENTI	
CFU	6
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	94
NUMERO DI ORE RISERVATE ALLA DIDATTICA ASSISTITA	56
PROPEDEUTICITA'	
MUTUAZIONI	
ANNO DI CORSO	1
PERIODO DELLE LEZIONI	2° semestre
MODALITA' DI FREQUENZA	Facoltativa
TIPO DI VALUTAZIONE	Voto in trentesimi
ORARIO DI RICEVIMENTO DEGLI STUDENTI	SCIORTINO MARINELLA Martedì 13:00 14:00 Ufficio n. 201, Il piano del DMI Venerdì 12:00 13:30 Ufficio n. 201, Il piano del DMI

PREREQUISITI	Capacità di usare un linguaggio di programmazione. Nozioni di base della teoria degli algoritmi e delle strutture dati.
RISULTATI DI APPRENDIMENTO ATTESI	Conoscenze e capacità di comprensione. Acquisizione dei concetti fondamentali della Teoria dell'Informazione e delle principali metodologie di Compressione Dati, con particolare riferimento alla Teoria Matematica della Comunicazione di C. Shannon, nel cui ambito svolgono un ruolo centrale le formalizzazioni delle nozioni di Informazione e di Codice. Conoscenze dei principali metodi di compressione dati. Acquisizione degli strumenti avanzati per leggere aspetti basilari della letteratura specialistica della disciplina. Capacità di utilizzare il linguaggio tecnico proprio della disciplina. Capacità di applicare conoscenze e comprensione. Capacità di riconoscere ed organizzare in autonomia argomenti base della teoria dell'informazione. Capacità di implementare e applicare i principali metodi di compressione dati. Capacità di utilizzare le conoscenze acquisite (in particolare, le metodologie di compressione dati) in campi applicativi specifici. Autonomia di giudizio. Essere in grado di valutare la rilevanza generale di argomenti della disciplina, e di collegare gli aspetti teorici della teoria dell'informazione con gli aspetti pratici della compressione dati. Gli studenti saranno guidati ad apprendere in maniera critica e responsabile gli argomenti che verranno loro proposti in aula e in laboratorio. Ciascuno studente avrà inoltre occasione di arricchire la propria autonomia di giudizio attraverso la realizzazione di un elaborato che sarà parte integrante della prova di valutazione. Abilità comunicative. Capacità di esporre in modo chiaro e rigoroso le tematiche generali della teoria dell'informazione anche ad un pubblico non esperto, mostrando come metodi e risultati teorici si rapportano ad ambiti applicativi specifici. Attraverso le attività di laboratorio previste, il corso tenderà a sviluppare negli studenti l'interazione e la capacità di saper lavorare in gruppo, di confrontarsi sulle problematiche al fine di individuare le soluzioni in base alle conoscenze acquisite durante il corso. L'acquisizione delle abilità comunicative sarà realizzata tramite la partecipazione attiva dello studente alle attività di laboratorio nonché l'esposizione dei risultati del lavoro individuale o di gruppo su argomenti o problematiche proposti dal docente. Capacità di apprendimento. Gli studenti svilupperanno la capacità di aggiornamento con la consultazione delle pubblicazioni scientifiche proprie del settore. Il materiale proposto in aula e in laboratorio svilupperà le capacità di apprendimento degli studenti che saranno in grado di "interrogare" in modo integrato le proprie conoscenze-competenze a fronte delle problematiche affrontate. Gli studenti saranno stimolati inoltre ad una conoscenza più approfondita e critica dei metodi di compressione a loro già noti.
VALUTAZIONE DELL'APPRENDIMENTO	La verifica finale mira a valutare se lo studente ha conoscenza e comprensione degli argomenti, se ha acquisito la capacità di applicare in pratica tale conoscenza e se possiede il linguaggio proprio della disciplina. La verifica finale consisterà di un progetto di studio nel quale gli studenti in gruppo dovranno mostrare di saper applicare gli argomenti studiati durante il corso e di una prova orale in cui descriveranno il lavoro svolto. Durante la prova orale gli studenti, oltre a descrivere il suo progetto di studio svolto, dovranno rispondere a tre domande su tutte le parti oggetto del programma. La valutazione finale, formulata sulla base delle seguenti condizioni, verrà espressa in trentesimi: a) non possiede una conoscenza accettabile dei contenuti degli argomenti trattati nell'insegnamento (insufficiente); b) sufficiente conoscenza di base degli argomenti principali dell'insegnamento e del linguaggio tecnico; sufficiente capacità di approfondimento e di elaborazione del progetto assegnato (18-20); c) discreta conoscenza degli argomenti principali dell'insegnamento, discreta proprietà di linguaggio, sufficiente capacità di applicare autonomamente le conoscenze acquisite, sufficiente capacità di approfondimento e di elaborazione del progetto assegnato (21-23); d) discreta conoscenza degli argomenti del corso, discreta proprietà di linguaggio, discreta capacità di applicare autonomamente le conoscenze alla soluzione dei problemi proposti, discreta capacità di approfondimento e di elaborazione del progetto assegnato (24-26); e) buona padronanza degli argomenti, piena proprietà di linguaggio, lo studente è in grado di applicare in modo adeguato le conoscenze per risolvere i problemi proposti; buona capacità di approfondimento e di elaborazione del progetto assegnato (27-29); f) ottima conoscenza degli argomenti, ottima proprietà di linguaggio, buona capacità analitica, lo studente è in grado di applicare in modo pienamente adeguato le conoscenze acquisite per risolvere i problemi proposti; ottima capacità di approfondimento e di elaborazione del progetto assegnato

	(30-30 cum laude).
OBIETTIVI FORMATIVI	Acquisizione dei concetti fondamentali della teoria dell'informazione e delle più comuni tecniche di codifica e di compressione lossless delle informazioni. Capacità di utilizzare il linguaggio tecnico proprio della disciplina. Capacità di utilizzare le conoscenze acquisite (in particolare, le metodologie lossless di compressione dati) in campi applicativi specifici. Capacità di aggiornamento con la consultazione delle pubblicazioni scientifiche proprie del settore.
ORGANIZZAZIONE DELLA DIDATTICA	La didattica è organizzata mediante lezioni frontali in aula e in laboratorio, in cui si prevede sia un trasferimento della conoscenza ma anche, con il supporto del mezzo informatico, la contestuale verifica dell'acquisizione delle competenze pratiche richieste. Per incrementare il coinvolgimento degli studenti si utilizzano anche metodologie di collaborative learning.
TESTI CONSIGLIATI	Testi di riferimento/Reference textbooks: Cover T.M. and Thomas J.A., Elements of Information Theory, Wiley-Interscience, 2006 (per gli argomenti di teoria dell'Informazione/for topics on Information Theory) ISBN 978-0471241959 Sayood K., Introduction to Data Compression, Morgan Kaufman, 2017 (Per argomenti su compressione dati lossless/for topics on lossless data compression) ISBN 978-0124157965 Salomon D., Motta G. Handbook of Data Compression, Fifth Edition, Springer 2010 (Per argomenti su compressione dati lossless/for topics on lossless data compression) ISBN 978-1-84882-902-2 Testi di consultazione/Consultation textbooks: Shannon, C.E. (July 1948). "A Mathematical Theory of Communication". Bell System Technical Journal 27: 379–423. DOI:10.1002/j.1538-7305.1948.tb01338 Nelson M. and Gailly, J.-L., The Data Compression Book, John Wiley & Sons, 1998. ISBN 978-1558514348

PROGRAMMA

ORE	Lezioni
4	Presentazione del corso. Introduzione alla Teoria dell'Informazione di Shannon. Entropia della sorgente come misure dell'informazione prodotta dalla sorgente nell'unità di tempo. Entropia e Compressione.
6	Codici. Codici univocamente decifrabili. Codici prefissi. Disuguaglianza di Kraft-McMillan. Algoritmo di Sardinas-Patterson.
6	Algoritmo per la codifica di Huffman. Ottimalità della codifica. Versione adattiva dell'algoritmo.
6	Codifica aritmetica. Tecniche per la codifica di interi.
5	Metodi di compressione basati su dizionari. Dizionari statici e il problema del Optimal Parsing. Dizionari Dinamici. Algoritmi di Lempel e Ziv: metodo di compressione LZ77.
2	Varianti dell'algoritmo LZ77. Algoritmo di compressione LZ78.
4	Metodi di compressione basati su contesto. La trasformata di Burrows e Wheeler (BWT). Invertibilità della BWT. Proprietà matematiche della BWT. BWT e Suffix array. Algoritmo per la costruzione di suffix array di un testo
5	Codifica Move-To-Front. Metodi di compressione e indicizzazione basati su BWT.
2	Complessità di Kolmogorov. Smallest grammar Problem.
ORE	Laboratori
4	Implementazione della codifica di Huffman, della codifica aritmetica e di tecniche per la codifica di interi.
5	Implementazione di algoritmi di compressione basati su dizionari.
4	Implementazione di tecniche di compressione e indicizzazione basate su BWT.
3	Compressione e motori di ricerca. Utilizzo di Compression Benchmarks.