

Principali informazioni sull'insegnamento		
Denominazione dell'insegnamento	Elementi di Data Mining	
Corso di studio	Informatica e Tecnologie della Produzione del Software	
Anno Accademico	2025-2026	
Crediti formativi universitari (CFU) / European Credit Transfer and Accumulation System (ECTS)		
Settore Scientifico Disciplinare	ING-INF/05 - Sistemi di Elaborazione delle Informazioni	
Lingua di erogazione	Italiano	
Anno di corso	3	
Periodo di erogazione	2 ^a semestre, le date esatte sono riportate nel manifesto/regolamento	
Obbligo di frequenza	No	
Sito web del corso di studio	https://www.uniba.it/it/corsi/informatica-tecnologie-informazione-taranto/informatica-e-tecnologie-dellinformazione	

Docente/i	
Nome e cognome	Giuseppina Andresini
Indirizzo mail	giuseppina.andresini@uniba.it
Telefono	0805542407
Sede	Dipartimento di Informatica, Via Orabona 4, 70125, Bari. Stanza n. ???, V piano
Sede virtuale	Piattaforma ADA - https://elearning.di.uniba.it/
Sito web del docente	https://persone.ict.uniba.it/rubrica/giuseppina.andresini
Ricevimento	Su appuntamento

Syllabus	
Obiettivi formativi	L'insegnamento si propone di introdurre: • una metodologia standard per la scoperta di nuova conoscenza dai dati; • gli algoritmi di data mining più noti in letteratura; • la scelta di algoritmi di data mining per specifici compiti; • l'interpretazione dei risultati di un algoritmo di data mining Lo studente acquisirà una metodologia per la scoperta di nuova conoscenza dai dati, sarà in grado di condurre le varie fasi del processo di scoperta di nuova conoscenza dai dati
Prerequisiti	Le seguenti conoscenze preliminari facilitano ed accelerano la comprensione degli argomenti dell'insegnamento. Dall'insegnamento di Progettazione di Basi di Dati: la progettazione concettuale e logica; il linguaggio SQL. Dall'insegnamento di Matematica discreta: il principio di induzione, il concetto di vettori e matrici. Dall'insegnamento di Analisi Matematica: il concetto di funzione, variabile dipendente e indipendente Dall'insegnamento di Statistica per l'Ingegneria del Software: Probabilità elementare. Variabili aleatorie continue e discrete. Media, varianza e covarianza
Contenuti di insegnamento (Programma)	1. Scoperta di conoscenza nelle basi di dati: il processo. La scoperta di conoscenza nelle basi di dati: definizione. Il processo della scoperta di conoscenza nelle basi di dati. Il processo CRISP-DM: business understanding, data understanding, data preparation, modelling, evaluation, deployment. 2. Analisi di associazione. regola di associazione, itemset frequenti, principali esempi di pattern di associazione

	3. Classificazione: algoritmi ad albero, basati su regole, probabilistici e basati sulla distanza. Valutazione dei modelli di classificazione 4. Regressione: algoritmi lineari e ad albero. Valutazione dei modelli di regressione 5. Analisi di raggruppamento (clustering). 6. Introduzione ad algoritmi basati su reti neurali Laboratorio: Python scikit-learn.
Testi di riferimento	- Charu C. Aggarwal, Data Mining, Springer 2015 - Richard J. Roiger, Michael W. Geatz, Introduzione al Data Mining. McGraw-Hill, 2003 - Dmitry Zinoviev Data Science con Python. Dalle stringhe al machine learning, le tecniche essenziali per lavorare sui dati, Apogeo, 2017 Gli studenti che lo desiderano possono ottenere i testi in prestito dalla Biblioteca. Può convenire verificarne la disponibilità mediante il Sistema Bibliotecario di Ateneo https://opac.uniba.it/easyweb/w8018/index.php? e contattare la biblioteca per concordare il prestito
Note ai testi di riferimento	I testi di riferimento sono supportati da articoli scientifici e dispense forniti dal docente durante lo svolgimento del corso.

Organizzazione della didattica			
Ore			
Totali	Didattica frontale	Laboratorio ed esercitazioni	Studio individuale
	32 ore	30 ore	88 ore
CFU/ETCS			
	4	2	

Metodi didattici	
	Le lezioni frontali saranno dedicate all'apprendimento dei modelli teorici e dei concetti di base, coadiuvati da alcuni esempi. Le ore di esercitazione saranno dedicate all'esecuzione di esercizi in classe coinvolgendo direttamente gli studenti nella risoluzione degli stessi.

Risultati di apprendimento previsti	
Conoscenza e capacità di comprensione	- Acquisizione di conoscenze relative agli algoritmi di data mining più noti in letteratura. - Comprensione delle scelte di algoritmi di data mining per specifici compiti. Capacità di interpretazione dei risultati di un algoritmo di data mining.
Conoscenza e capacità di comprensione applicate	- Capacità di realizzazione di un semplice progetto di scoperta di conoscenza in una collezione di dati mediante: - Utilizzo di strumenti per la selezione, pre-elaborazione e trasformazione dei dati, e per la validazione dei pattern estratti. - Utilizzo di strumenti di data mining per l'estrazione di conoscenza finalizzata a scopi predittivi e descrittivi in diversi contesti applicativi (aziendali e scientifici).
Competenze trasversali	<i>Autonomia di giudizio</i> - Gli studenti sono in grado di apprezzare l'uso di algoritmi di data mining in processi di scoperta della conoscenza.

	○ L'autonomia di giudizio viene acquisita attraverso lo studio e l'interpretazione critica dei testi. ○ Il raggiungimento dell'adeguata autonomia è verificato attraverso le esercitazioni, che si tengono durante il corso, e con l'esame finale di profitto. • <i>Abilità comunicative</i> ○ Gli studenti sono in grado di esporre le tematiche incluse nel programma del corso mediante il lessico specifico della disciplina. • <i>Capacità di apprendere in modo autonomo</i> ○ Gli studenti sono in grado di approfondire in autonomia le tematiche incluse nel programma del corso anche ricorrendo a risorse non direttamente coinvolte nella erogazione delle ore di lezione.
--	--

Valutazione	
Modalità di verifica dell'apprendimento	Prova scritta riguardante i contenuti introdotti durante le lezioni. Consegna di un caso di studio come prova pratica, con applicazione della metodologia CRISP-DM a un insieme di dati per un particolare dominio applicativo
Criteri di valutazione	• <i>Conoscenza e capacità di comprensione:</i> ○ esposizione critica dei concetti appresi relativi al processo di scoperta della conoscenza e capacità di affrontare semplici esercizi di data mining. • <i>Conoscenza e capacità di comprensione applicate:</i> ○ analisi di dataset con applicazione di algoritmi di data mining e comprensione dei risultati ottenuti, in un ciclo finalizzato al miglioramento delle prestazioni • <i>Autonomia di giudizio:</i> ○ Capacità di svolgere semplici esercizi assegnati durante il corso delle lezioni • <i>Abilità comunicative:</i> ○ Uso del lessico specifico della disciplina informatica • <i>Capacità di apprendere:</i> ○ Sviluppo di argomenti su data mining non direttamente trattati nel corso ma assegnati dal docente
Criteri di misurazione dell'apprendimento e di attribuzione del voto finale	Il voto finale sarà attribuito sulla base della valutazione comparativa di tutti i criteri di valutazione sopra citati
Altro	

Data, 14/01/2026