

Principali informazioni sull'insegnamento	A.A. 2020/2021
Titolo insegnamento	Data Mining
Corso di studio	LT Informatica e Tecnologie per la Produzione del Software
Crediti formativi	6
Denominazione inglese	Data Mining
Obbligo di frequenza	No
Lingua di erogazione	Italiana

Docente responsabile	Nome Cognome	Indirizzo Mail
	Corrado Loglisci	corrado.loglisci@uniba.it
Luogo ed Orario di Ricevimento	Dip. Informatica 5° Piano – Stanza 525. Venerdì 15:00- 17:00	

Dettaglio credi formativi	Ambito disciplinare	SSD	Crediti
	Tecnologie dell'Informatica	ING-INF/05	62

Modalità di erogazione	
Periodo di erogazione	Secondo Semestre
Anno di corso	III
Modalità di erogazione	Lezioni frontali + Didattica online Esercitazioni in aula e Laboratorio

Organizzazione della didattica	
Ore totali	150
Ore di corso	62
Ore di studio individuale	88

Calendario	
Inizio attività didattiche	1/3/2021
Fine attività didattiche	4/6/2021

Syllabus	
Prerequisiti	Si assume che il discente abbia conoscenze di base di: Strutture dati fondamentali e basi di dati. Programmazione. Analisi matematica. Spazi di ricerca e algoritmi di ricerca

<p>Risultati di apprendimento previsti (declinare rispetto ai Descrittori di Dublino) (si raccomanda che siano coerenti con i risultati di apprendimento del CdS, riportati nei quadri A4a, A4b e A4c della SUA, compreso i risultati di apprendimento trasversali)</p>	<p><i>Conoscenza e capacità di comprensione:</i></p> <ul style="list-style-type: none"> -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. <p><i>Conoscenza e capacità di comprensione applicate:</i></p> <ul style="list-style-type: none"> - Capacità di realizzazione di un semplice progetto di scoperta di conoscenza in una collezione di dati mediante: - Utilizzo e/o sviluppo di strumenti per la selezione, preelaborazione e trasformazione dei dati, e per la validazione di modelli e pattern. - Utilizzo di strumenti di data mining per l'estrazione di conoscenza finalizzata a scopi descrittivi e predittivi in diversi contesti applicativi (aziendali e scientifici). <p><i>Autonomia di giudizio</i></p> <ul style="list-style-type: none"> - 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. <p><i>Abilità comunicative:</i></p> <ul style="list-style-type: none"> - Gli studenti sono in grado di esporre le tematiche incluse nel programma del corso mediante il lessico specifico della disciplina. <p><i>Capacità di apprendere:</i></p> <ul style="list-style-type: none"> - 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.
<p>Contenuti di insegnamento</p>	<ol style="list-style-type: none"> 1. Knowledge discovery: The process of knowledge discovery in databases. The process CRISP-DM. 2. Data preparation: Data pre-processing, Feature selection. Feature extraction 3. Classification: Tree-based, rule-based, probabilistic and distance-based algorithms. Evaluation of the classification models 4. Regression: Linear and tree-models algorithms. Evaluation of the regression models. 5. Clustering: partition-based and hierarchical algorithms. Evaluation of the clusters.

	<p>6. Pattern discovery: Itemset mining. Association rules discovery.</p> <p>7. Advanced algorithms: Graph mining. Spatio-temporal data mining.</p> <p>Laboratory: Weka data mining toolkit. Knowledge flow. Explorer.</p>
--	--

Programma	
Testi di riferimento	<p>Richard J. Roiger, Michael W. Geatz. Introduzione al Data Mining. McGraw-Hill, 2003 Capitoli: 1-2-3-4-5-6-8-9</p> <p>A. Azzalini, B. Scarpa Analisi dei dati e data mining Springer, 2004 Capitoli: 1-4-5-6</p>
Note ai testi di riferimento	I testi di riferimento saranno supportati da materiale di lezione e articoli scientifici forniti dal docente durante lo svolgimento del corso
Metodi didattici	Lezioni frontali, esercitazioni guidate
Metodi di valutazione (indicare almeno la tipologia scritto, orale, altro)	<p>Prova orale riguardante:</p> <ul style="list-style-type: none"> -progetto caso di studio reale di un un processo KDD -concetti presentati a lezione
<p>Criteria di valutazione (per ogni risultato di apprendimento atteso su indicato, descrivere cosa ci si aspetta lo studente conosca o sia in grado di fare e a quale livello al fine di dimostrare che un risultato di apprendimento è stato raggiunto e a quale livello)</p>	<p>Lo studente sarà valutato sulla base delle capacità di analisi dei dati, di applicazione di algoritmi di data mining e di comprensione dei risultati ottenuti, in un ciclo finalizzato al miglioramento delle prestazioni. Gli strumenti che si possono all'uopo utilizzare sono quelli illustrati durante le ore di esercitazione/laboratorio.</p>
Altro	