UNIVERSITÀ DI PISA

FUNDAMENTALS OF DATA MINING AND MACHINE LEARNING

FRANCESCO MARCELLONI

Academic year	2023/24
Course	BANCA, FINANZA AZIENDALE E MERCATI FINANZIARI
Code	1093I
Credits	6

Modules	Area	Type	Hours	Teacher(s)
FUNDAMENTALS OF DATA MINING AND MACHINE LEARNING	ING-INF/05	LEZIONI	0	FRANCESCO MARCELLONI

Obiettivi di apprendimento

Conoscenze
Gli studenti che completeranno con successo l'insegnamento avranno una solida conoscenza delle principali tecniche per pre-processare i dati, frequent pattern mining, classificazione, predizione, clustering, e outlier detection. Questa conoscenza permetterà loro di affrontare diversi possibili problemi inerenti il data mining e di identificare la tecnica più adatta per risolverli.

Modalità di verifica delle conoscenze
Durante la verifica delle conoscenze, gli studenti devono dimostrare di aver appreso le diverse tecniche insegnate durante lo svolgimento del corso e devono essere capaci di identificare la soluzione più adatta per problemi di data mining specifici.
I metodi sono:

- esame orale
- Prova pratica

Capacità
Al termine del corso,

- lo studente saprà affrontare i più comuni problemi di data mining, trovando le soluzioni più idonee per risolverli
- lo studente saprà valutare e confrontare più soluzioni e scegliere la più adatta

Modalità di verifica delle capacità
Lo studente dovrà risolvere un problema di data mining durante una prova pratica.

Comportamenti
Lo studente potrà acquisire un metodo per affrontare problemi di data mining e per selezionare le migliori soluzioni da adottare

Modalità di verifica dei comportamenti
Durante le sessioni di laboratorio saranno valutati il grado di accuratezza e precisione delle attività svolte dallo studente

Prerequisiti (conoscenze iniziali)
Conoscenze di base di matematica
Conoscenze di programmazione

Indicazioni metodologiche
Le lezioni verranno svolte frontalmente con l'ausilio di slide
Le esercitazioni verranno svolte in laboratorio con l'ausilio di slide e esempi di programmazione
L'intero corso è tenuto in Inglese

UNIVERSITÀ DI PISA

Programma (contenuti dell'insegnamento)

Data Preprocessing: data cleaning, integration, reduction, transformation and discretization.
Frequent pattern mining: basic concepts, A-priori algorithm, Pattern-Growth approach, vertical data format, pattern evaluation methods, constraint-based frequent pattern mining, colossal pattern.
Classification: basic concepts, decision tree induction, Bayes classification methods, rule-based classification, lazy learners, techniques for improving accuracy, model evaluation and selection.
Clustering: basic concepts, partitioning methods, hierarchical methods, density-based methods, grid-based methods, model evaluation and selection, clustering with constraints.
Outlier detection: statistical, proximity-based, clustering-based and classification-based approaches.

Bibliografia e materiale didattico

Slides
Libro: J. Han and M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd ed., 2011 Papers on the different algorithms described during the course Slides of the lectures
Articoli forniti dal docente

Modalità d'esame

L'esame è composto da una prova pratica e una prova orale.
La prova pratica sarà sviluppata dando un problema allo studente che dovrà risolverlo utilizzando gli strumenti presentati a lezione e durante i laboratori
La prova orale consiste in un colloquio tra il candidato e il docente su alcune domande che possono essere anche assegnate in forma scritta al candidato.
La prova orale è superata se il candidato mostra padronanza degli argomenti trattati, si esprime in modo chiaro e con terminologia corretta, mostra capacità di analisi e sintesi.

Ultimo aggiornamento 09/09/2023 12:56