UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Machine Learning

2526-1-F1802Q101

Obiettivi

Il corso si propone di introdurre gli elementi di base e di fornire i concetti e gli strumenti fondamentali, basati su metodi computazionali, per rappresentare l'apprendimento, la conoscenza e il ragionamento in condizioni di incertezza.

Lo studente sarà in grado di progettare e sviluppare strumenti software basati su tecniche di apprendimento automatico (machine learning) per risolvere semplici problemi di analisi di dati, knowledge discovery e supporto alle decisioni in presenza di informazione incerta o incompleta.

Gli obiettivi specifici rispetto ai Descrittori di Dublino (DdD) sono i seguenti:
1. Conoscenza e capacità di comprensione

- le basi teoriche dell'apprendimento automatico;
- la comprensione dei principali modelli (alberi di decisione, reti neurali, SVM, modelli bayesiani, k-means, ecc.);

2. Conoscenza e capacità di comprensione applicate

- progettare e sviluppare strumenti software basati su modelli di machine learning;
- applicare metodi di classificazione supervisionata e non supervisionata a problemi reali;

3. Autonomia di giudizio

- selezionare i metodi più appropriati in relazione alla natura del problema e dei dati;
- valutare la qualità delle soluzioni proposte in base ai risultati ottenuti dai modelli rispetto a metriche di performance;
- analizzare criticamente le assunzioni dei modelli (es. separabilità lineare, indipendenza delle variabili) e i risultati ottenuti rispetto a tali assunzioni.

4. Abilità comunicative

- comunicare le scelte progettuali e metodologiche in modo argomentato
- presentare e discutere in modo chiaro i risultati di un progetto in sede di prova orale;
- lavorare in gruppo, condividendo conoscenze.

5. Capacità di apprendere

- sviluppare capacità autonome di studio attraverso il completamento di assignment facoltativi derivanti da attività di laboratorio focalizzate all'uso di modelli di machine learning

Contenuti sintetici

Il corso fornirà una approfondita presentazione della vasta collezione di metodi e strumenti che si sono resi disponibili negli anni per lo sviluppo di sistemi autonomi di apprendimento e di supporto alla analisi di dati multivariati. Questi strumenti comprendono gli alberi di decisione, le reti neurali, le macchine a vettori di supporto, così come algoritmi di classificazione non supervisionata e combinazioni sempre più sofisticate di queste architetture. Si presenteranno esempi di recenti applicazioni.

Programma esteso

Introduzione e terminologia

Concept learning
· Alberi di decisione
· Algoritmo ID3

Reti neurali
· Il percettrone; separabilità lineare
· Algoritmo di apprendimento del percettrone
· Delta rule. Concetto di discesa del gradiente
· Il percettrone multistrato
· Algoritmo di retropropagazione dell'errore

Support vector machines
· Separazione ottima
· Hard and Soft SVM
· Kernels

Apprendimento bayesiano
· Teorema di Bayes
· Classificatore bayesiano ingenuo

Apprendimento non supervisionato
· L'algoritmo k-means

Valutazione dei risultati dell'apprendimento
· valutazione per i metodi supervisionati
· valutazione del clustering

Reti neurali profonde (Deep Learning)
· Introduzione
· Tecniche di riferimento

Prerequisiti

Conoscenze di base su algoritmi e strutture di dati.

Elementi di calcolo delle probabilità e statistica

Modalità didattica

- 8 lezioni da 2 ore svolte in modalità erogativa in presenza;
- 10 esercitazioni da 2 ore svolte in presenza in modalità 50% erogativa 50% interattiva;
- 12 attività di laboratorio da 2 ore svolte in modalità interattiva in presenza;

Materiale didattico

Materiali riferiti sulla pagina elearning dell'insegnamento

Periodo di erogazione dell'insegnamento

Primo semestre

Modalità di verifica del profitto e valutazione

L'esame consiste di una prova scritta e una prova orale, ciascuna valida per il 50% del voto, e si svolge secondo le seguenti modalità:

1. **Prova scritta** con domande teoriche ed esercizi. Per accedere all'orale è necessario ottenere un punteggio >= 6 nella prova scritta. La prova scritta avrà una valutazione da 0 a 30.
2. **Prova orale** centrata sulla discussione di un progetto svolto tipicamente in gruppo. La valutazione del progetto si basa sulla qualità complessiva dell'elaborato e su domande di contenuto specifico sul programma dell'insegnamento. La prova orale avrà una valutazione da 0 a 30.

Il **voto finale** sarà determinato come media tra la prova scritta e la prova orale.

E' data la possibilità agli studenti di sostenere due prove intermedie come sistitutive della prova scritta. Ogni prova intermedia sarà costituita da domande teoriche ed esercizi ed avrà una valutazione da 0 a 30.

Orario di ricevimento

Su appuntamento

Sustainable Development Goals