UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Information Retrieval

2526-2-F1801Q110

Obiettivi

Conoscenza e comprensione
L'obiettivo del corso è fornire un'introduzione ai concetti fondamentali, ai modelli formali, e alle tecniche per la realizzazione di sistemi per il reperimento automatico di documenti testuali in forma digitale (sistemi di "Information Retrieval", detti anche Motori di Ricerca, o Motori di Ricerca su Web quando i documenti da reperire sono costituiti da pagine Web). Dopo un'introduzione all'analisi automatica di testi e alla loro rappresentazione (tecniche base di NLP), si definirà il problema del reperimento automatico di informazioni, che presuppone la stima della rilevanza dei documenti rispetto alle necessità informative dell'utente espresse in una query. Al termine del corso lo studente sarà in grado di comprendere tecniche relative all'analisi e alla rappresentazione di testi in liguaggio naturale, nonchè il funzionamento di un motore di ricerca. Lo studente sarà inoltre in grado di utilizzare software "open source" per la definizione di applicazioni di Information Retrieval. Il laboratorio sarà finalizzato alla realizzazione di un'applicazione.
Capacità di applicare conoscenza e comprensione
Durante le lezioni e le attività di laboratorio verranno promosse e valutate le capacità degli studenti nell applicare le conoscenze acquisite sugli argomenti trattati nell'insegnamento.
Autonomia di giudizio
L’insegnamento ha l'obiettivo specifico di sviluppare autonomia di giudizio e capacità di analisi critica degli studenti rispetto alle principali tematiche del corso. Ciò sarà anche stimolato attraverso discussioni in aula e attività di laboratorio.
Abilità comunicative
Sviluppo della capacità di comunicare in modo chiaro, consapevole e privo di ambiguità contenuti tecnici, idee, problemi e relative soluzioni a interlocutori diversi. Tali abilità saranno promosse durante l’insegnamento e valutate in sede d’esame.
Capacità di apprendimento
L'insegnamento è concepito per fornire sia conoscenze teoriche sia competenze pratiche, costituendo un solido punto di partenza anche per eventuali approfondimenti individuali sui principi di rappresentazione, analisi e ricerca testuale.

Contenuti sintetici

Il corso introdurrà dapprima il problema della rappresentazione di testi per un loro trattamento automatico, e le tecniche utilizzate in molte applicazioni di NLP per il pre-processing di testi e per una loro rappresentazione formale (incluse tecniche neurali). Si presenteranno poi le tecniche pricipali per la progettazione e la realizzazione di motori di ricerca, con una accenno ai sistemi per la rappresentazione di informazioni.
Verranno introdotti i principali modelli per la determinazione della stima di rilevanza di un documento rispetto alle necessità informative dell'utente, dal modello vettoriale ai più recenti modelli neurali. Si presenteranno poi le tecniche alla base del funzionamento di motori di ricerca su Web.
Verranno inoltre presentate alcune tecniche per la personalizzazione della ricerca e si introdurrà la tecnica di Retrieval Augmented Generation.
Il corso introdurrà anche il tema della valutazione dell'efficacia dei motori di ricerca.

Programma esteso

1. Intoduzione al text processing e al Natural Language Processing (NLP)

2. Introduzione al Text Mining e ad alcune applicazioni correlate al NLP.

3. Pre-Processing, indicizzazione e rappresentazione formale di testi (bag of words, word embeddings, statistical language models, neural language models - large language models)

4. Modelli di sistemi di Information Retrieval: i modelli base (Booleano, Vettoriale, modelli Probabilistici). Modelli avanzati (modelli neurali). Accenni a motori di ricerca per documenti multimediali.

5. I motori di ricerca su Web: crawling, link analysis e altri fattori per la stima della rilevanza di pagine Web.

6. La valutazione dei motori di ricerca.

7. Argomenti avanzati

8. Introduzione a software open source per la definizione di motori di ricerca

Prerequisiti

Nozioni di base di Statistica e di Algebra Lineare.

Modalità didattica

Il corso verrà tenuto in lingua inglese e prevede sia lezioni (40 ore) sia esercitazioni in laboratorio (12 ore). Le lezioni sono da 2 ore e vengono svolte sia in modalità erogativa (specie nella parte iniziale della lezione) sia in modalità interattiva (durante la lezione) per il coinvolgimento attivo degli studenti. Durante le attività di laboratorio verrà spiegato e sperimentato l'utilizzo di software “open source”. Si organizzeranno seminari tenuti da esperti a livello internazionale. Alcune lezioni potranno essere svolte in modalità erogativa-interattiva da remoto.

Materiale didattico

Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press, 2008.

Periodo di erogazione dell'insegnamento

Primo semestre

Modalità di verifica del profitto e valutazione

Prova scritta individuale costituita da esercizi e domande aperte sui contenuti del corso. Realizzazione di un progetto di laboratorio che è possibile svolgere in gruppo (sino a tre studenti).
La prova scritta ha come obiettivo la valutazione del livello di comprensione degli aspetti teorici e tecnici di base dell'insegnamento erogato.
L'obiettivo del progetto di gruppo, attraverso l'utilizzo di sofrware open source, è lo sviluppo di soluzioni tecnologiche a problemi affrontati a lezione. In particolare, si considerano ambiti applicativi reali che necessitano della definizione di sistemi i cui fondamenti sono stati presentati a lezione.

Orario di ricevimento

Previo appuntamento con la docente.

Sustainable Development Goals