UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Social Media Analytics

2627-2-FDS02Q018

Obiettivi

Conoscenza e comprensione

L'insegnamento mira a fornire agli studenti i principali concetti che stanno alla base della gestione dei dati originati nei media sociali (accesso, pre-processing, modellazione) e della loro successiva analisi.

Capacità di applicare conoscenza e comprensione

Gli studenti saranno in grado, in particolare, di collezionare, processare e analizzare dati provenienti dai principali media sociali, utilizzando le tecnologie più adatte allo scopo. Sarà inoltre in grado di fornire una rappresentazione di strutture sociali complesse al fine di estrarne informazioni utili. Durante le lezioni e le attività di laboratorio vengono promosse e valutate le capacità degli studenti di applicare le conoscenze acquisite sugli argomenti trattati nell'insegnamento.

Autonomia di giudizio

L’insegnamento mira a sviluppare l’autonomia di giudizio e la capacità di analisi critica rispetto alle principali sfide legate alla raccolta, alla rappresentazione e all'analisi dei dati nelle piattaforme sociali. Tali competenze saranno stimolate anche 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 recupero, rappresentazione e analisi dei dati sociali.

Contenuti sintetici

1. Introduzione al Web sociale;
2. I dati nei media sociali e problematiche connesse (compreso il problema della privacy);
3. La rappresentazione dei dati sociali mediante strutture sociali complesse: teoria dei grafi e delle reti;
4. Analisi topologica delle reti sociali: le principali metriche;
5. Analisi topologica delle reti sociali: algoritmi per l'identificazione delle comunità nelle reti sociali;
6. Analisi dei contenuti nell reti sociali: soggettività, “sentiment” e ironia;
7. Analisi dei contenuti nell reti sociali: *Named-Entity Recognition* e *Linking*;
8. Visualizzazione di dati provenienti da social media e delle analisi su tali dati.

Programma esteso

1. Introduzione

- Introduzione al Web sociale e alla terminologia utilizzata: Web sociale, reti sociali, media sociali, contenuto generato dagli utenti;
- Gli “oggetti informativi” sul Web sociale: 1) testi: post, blog, microblog, 2) immagini, 3) audio, 4) video;
- *Social Media Analytics*: definizione e obiettivi, i concetti di auto-presentazione e auto-rivelazione.

2. I dati nei media sociali

- Le principali piattaforme, le tipologie di dato, le interfacce di programmazione, il processo di crawling.
- *Pre-processing* e memorizzazione di dati sociali.
- Cenni di problematiche legate alla raccolta dei dati, sia dal punto di vista legale (il GDPR) sia dal punto di vista tecnologico.

3. La rappresentazione dei dati sociali

- Strutture dati elementari e complesse;
- Rappresentazione di strutture sociali mediante grafi (teoria dei grafi, tipologie di reti).

4./5. Analisi delle reti sociali

- *Link analysis*, *Web link analysis*, principali metriche;
- *Network clustering*: algoritmi di *community detection*;
- Modelli di influenza e contagio nelle reti sociali.

6./7. Analisi dei contenuti nelle reti sociali

- Introduzione a concetti di *Natural Language Processing* nell'ambito delle reti sociali;
- Oggettività/soggettività, polarità, emozioni e ironia nelle reti sociali;
- Approcci lessicali e approcci semantici;
- *Named-entity Recognition* e *Linking*.

8. Visualizzazione di dati provenienti da social media: problemi aperti e tecniche

Prerequisiti

Conoscenza di base dei principi di algebra lineare, statistica, programmazione.

Modalità didattica

- Lezioni (28 ore).
- Esercitazioni in laboratorio (18 ore).
- L'insegnamento verrà tenuto in lingua inglese.
- 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.
- Alcune lezioni potranno essere svolte in modalità erogativa-interattiva da remoto.

Materiale didattico

- Rahman, Md Saidur. Basic graph theory. Springer, 2017.
- Reinhard Diestel. Graph Theory. Springer, 2017.
- Suliman Hawamdeh, Hsia-Ching Chang. Analytics and Knowledge Management. CRC Press. 2018.
- John Scott. Social Network Analysis. Sage, 2012.
- Bing Liu. Sentiment Analysis and Opinion Mining. Morgan & Claypool. 2016.
- Eneko Agirre, Philip Edmonds. Word Sense Disambiguation: Algorithms and Applications (Text, Speech and Language Technology). Springer. 2007.

Periodo di erogazione dell'insegnamento

Primo semestre.

Modalità di verifica del profitto e valutazione

Prova scritta con esercizi e domande aperte.
La prova scritta ha come obiettivo la valutazione estensiva ed intensiva delle competenze teoriche e teorico-pratiche acquisite durante l'insegnamento.

Progetto di gruppo (con presentazione orale).
Il progetto ha come obiettivo la valutazione della capacità degli studenti di applicare in ambiti reali le competenze acquisite durante l’insegnamento, attraverso lo sviluppo e l’impiego di soluzioni tecnologiche per l’analisi dei dati nei social media. Il progetto viene presentato di persona dagli studenti, per accertare tali competenze sia di natura tecnica sia critica e di giudizio, e al contempo per sviluppare le capacità comunicative.

- La prova scritta viene valutata su una scala da 0 a 24.

- Gli studenti devono ottenere una valutazione maggiore o uguale a 14,5 nella prova scritta affinchè l'esame scritto sia considerato superato.
- Il progetto, con relativa discussione orale, viene valutato su una scala da 0 a 8.

Il voto finale sarà dato dalla somma del voto ottenuto nell'esame scritto (solo se maggiore o uguale a 14,5) e dal voto relativo al progetto.

Non sono previste prove in itinere.

Orario di ricevimento

Su appuntamento: Marco Viviani (marco.viviani@unimib.it).

Sustainable Development Goals

SALUTE E BENESSERE | PARITÁ DI GENERE | RIDURRE LE DISUGUAGLIANZE