



UNIVERSITÀ DI PISA

MULTIMEDIA INFORMATION RETRIEVAL AND COMPUTER VISION

GIUSEPPE AMATO

Anno accademico

2021/22

CdS

ARTIFICIAL INTELLIGENCE AND
DATA ENGINEERING

Codice

886II

CFU

9

Moduli	Settore/i	Tipo	Ore	Docente/i
MULTIMEDIA INFORMATION RETRIEVAL AND COMPUTER VISION	ING-INF/05	LEZIONI	90	GIUSEPPE AMATO FABRIZIO FALCHI CLAUDIO GENNARO

Obiettivi di apprendimento

Conoscenze

Fornire agli studenti un solido background su Multimedia Information Retrieval, Multimedia Content Based Retrieval, Automated Multimedia Content Understanding, Multimedia Data Mining, Computer Vision.
Comprendere le sfide e le problematiche legate all'efficacia, all'efficienza e alla scalabilità quando si ha a che fare con set di dati multimediali molto grandi (su scala web).

Modalità di verifica delle conoscenze

La valutazione delle conoscenze sarà effettuata attraverso il progetto finale (opzionale) e la prova orale.

Capacità

Gli studenti impareranno a ragionare intorno ai problemi e le soluzioni per Text-Based Retrieval, Multimedia Content Analysis, Multimedia Content Description, Computer Vision, Multimedia, Content Based Indexing, Multimedia Content Based Retrieval, Multimedia Data Mining and Classification, Similarity Searching, Scalable Access Methods for Similarity Searching, Deep Learning techniques applied to multimedia

Modalità di verifica delle capacità

Durante le sessioni di laboratorio, agli studenti verranno proposti compiti mirati da risolvere utilizzando strumenti appresi nelle sessioni di teoria.

Comportamenti

Gli studenti impareranno metodologie di valutazione rigorose per valutare le prestazioni delle soluzioni per Multimedia Information Retrieval

Modalità di verifica dei comportamenti

I progetti finali richiederanno la valutazione delle prestazioni delle soluzioni sviluppate utilizzando misure rigorose di Multimedia Information Retrieval

Prerequisiti (conoscenze iniziali)

Nozioni di base sui sistemi di gestione dei database, programmazione Python

Programma (contenuti dell'insegnamento)

Text Indexing and Retrieval:

- Introduction to Information Retrieval. Information Retrieval Systems vs Database Systems. Boolean retrieval. Term-document incidence matrices. Inverted indexes. Indexing Inverted indexes. Merge Algorithms: AND, and, BUT. Issues about the boolean retrieval and introduction to the Ranked Retrieval. Scoring documents. Term frequency. Collection statistics Weighting schemes. Vector space scoring.
- Term Frequency and Inverse Document Frequency. Vector space representation. Cosine similarity, Computing cosine scores, Normalizations, Maximum tf normalization. Evaluation in



UNIVERSITÀ DI PISA

information retrieval, Evaluation in information retrieval: Precision, Recall, and Accuracy. F1 metric.

- Mean Average Precision. Text Classification. Evaluation of Text Classifiers. Confusion matrix. Binary Classification. Multiclass Classification. f-score. Macro- vs Micro-Averaging. Supervised Learning. k nearest neighbor classification. Query processing and phrase queries.
- Content-based and Text-based image retrieval. Global and Local features for Content-based image retrieval (CBIR). Metric Spaces. Triangle inequality and indexing. CBIR through permutation-based approach. Complemented permutations. Content-based image retrieval through standard text retrieval techniques.

Audio Indexing and Retrieval

- Audio retrieval and Fingerprinting. Notes on Fourier Transform, Discrete Fourier Transform, and Short Time Fourier Transform. Spectral Peak as audio fingerprint. Indexing Fingerprint with inverted lists.

Multimedia Content Indexing and Representation

- Multimedia Content Representation: color spaces, still images, video
- Global Features, Local Invariant Feature Detectors, Local Features
- Local Features Matching, Local Features Aggregation
- Deep Learning in Computer Vision, Convolutional Neural Networks
- Deep Learning Advanced Topics, Transfer Learning
- Detection and Segmentation: fully convolutional neural network, R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN
- Generative Models: PixelRNN, PixelCNN, Variational Autoencoder, Generative Adversarial Network

Multimedia Searching on a Large Scale:

- Foundation of Metric Space Searching, Distance Searching Problem, Metric Distance Measures, Similarity Queries, Basic Partitioning Principles, Principles of Similarity Query Execution, Policies for avoiding Distance Computations, Principles of Approximate Similarity Search
- Advanced issues: Statistics, Proximity, Performance Prediction, Tree quality Measures, Choosing reference points
- Exact Similarity Search: Vantage Point Trees, AESA/LAESA, The M-Tree Family, The M-Tree Bulk-Loading Algorithm, Multi-Way Insertion Algorithms, The Slim Tree, Slim-Down Algorithm, Pivoting M-Tree
- Hash Based Methods: D-Index. Approximate Similarity Search with M-Tree, Approximate Similarity Search: Relative Error Approximation, Good Fraction Approximation, Small Chance Improvement Approximation, Proximity-Based Approximation, PAC Nearest Neighbor Searching
- Recent Similarity Search Techniques: Permutation Based Methods, Permutation Spearman Rank, PP-Index, MI-File, Locality Sensitive Hashing (LSH), LSH based on p-stable distributions, LSH with Hamming Distance

Bibliografia e materiale didattico

Si consulti il sito del corso: <https://sites.google.com/site/unipimircv/>

Modalità d'esame

L'esame finale è costituito da un progetto pratico opzionale e da una prova orale.

Il progetto utilizzerà le nozioni apprese durante le sessioni teoriche e la sessione di laboratorio per sviluppare uno strumento software completamente funzionale per gestire uno scenario specifico relativo alla gestione e al recupero delle informazioni multimediali. Il progetto sarà assegnato a gruppi di studenti (tipicamente da 1 a 5).

Il test orale ha lo scopo di verificare che gli studenti abbiano acquisito le conoscenze e le competenze richieste.

Pagina web del corso

<https://sites.google.com/site/unipimircv/>

Ultimo aggiornamento 18/11/2021 22:36