Master Degree in Computer Science
Master Degree in Data Science and Economics

Information Retrieval

Course Introduction

Prof. Alfio Ferrara

Department of Computer Science, Università degli Studi di Milano
Room 7012 via Celoria 18, 20133 Milano, Italia alfio.ferrara@unimi.it

sed noli modo

Prof. Alfio Ferrara
Course Chair
alfio.ferrara@unimi.it

Francesco Periti
Instructor
francesco.periti@unimi.it

Sergio Picascia
Instructor
sergio.picascia@unimi.it

Davide Riva
Instructor
davide.riva1@unimi.it

Elisabetta Rocchetti
Instructor
elisabetta.rocchetti@unimi.it

Darya Shlyk
Instructor
darya.shlyk@unimi.it

1. **Introduction to Information Retrieval** $\sim 6\ hrs$
2. **Evaluation in information retrieval and query expansion** $\sim 6\ hrs$
3. **Hidden Markov Models, Conditional Random Fields and Expectation Maximization** $\sim 3\ hrs$
4. **Latent Semantic Indexing and Latent Dirichlet Allocation** $\sim 3\ hrs$
5. **Clustering and Classification for Information Retrieval** $\sim 3\ hrs$
6. **Introduction to Neural Networks** $\sim 3\ hrs$
7. **Introduction to PyTorch functional** $\sim 3\ hrs$
8. **Word embeddings and Word2Vec** $\sim 3\ hrs$
9. **Sequence to sequence modeling** $\sim 9\ hrs$
10. **Attention, Transformers and BERT** $\sim 3\ hrs$
11. **Text2Images** $\sim 3\ hrs$

UNIVERSITÀ DEGLI STUDI DI MILANO
LA STATALE

Notice board, General info, Materials

Official Website
https://aferrarair.ariel.ctu.unimi.it/

Code, Use cases

GitHub
https://github.com/afflint/inforet

Schedule

Monday, 15:30-18:30
Aula Alpha, via Celoria 18

Thursday, 15:30-18:30
Room V4, via Venezian

Final assignment

The final project consists in the preparation of a short study on one of the topics of the course, identifying a **precise research question** and **measurable objectives**. The project will propose a methodology for solving the research question and provide an experimental verification of the results obtained according to results evaluation metrics. **The emphasis is not on obtaining high performance but rather on the critical discussion of the results obtained** in order to understand the potential effectiveness of the proposed methodology.

The results must be documented in a **short article of not less than 4 pages and no more than 8**, composed according to the guidelines provided by the lecturers. Students have also to **provide access to a GitHub repository containing the code and reproducible experimental results**. Finally, the project will be discussed by presenting a **10 minute presentation with slides**.

Exam dates are just for the registration of the final grade

The project discussion will be set by appointment

PROCEDURE

1. Subscribe to any available date
2. Contact Prof. Ferrara as soon as
    1. The project is finished and ready to be discussed
    2. After the date of your subscription
3. Setup an appointment and discuss your work

If you are interested in doing your final master thesis on these topics, the final project may be a preliminary work in view of the thesis. In this case, discuss the contents with Prof. Ferrara.

Coding is a fundamental part of the course

We will present **several use cases** and **programming examples**

Students are required to **complement the theoretical study with a hands on approach** to the proposed case studies

Python 3.x is the reference programming language

If needed, just refresh your Python knowledge with the **official tutorial** (https://docs.python.org/3/tutorial)