# Syllabus

## *Descrizione corso*

| | |
|---|---|
| **Titolo insegnamento** | Natural Language Processing and Recommender Systems |
| **Codice insegnamento** | 73071 |
| **Titolo aggiuntivo** | |
| **Settore Scientifico-Disciplinare** | INFO-01/A |
| **Lingua** | Inglese |
| **Corso di Studio** | Corso di laurea magistrale in Informatica per la Data Science |
| **Altri Corsi di Studio (mutuati)** | |
| **Docenti** | prof. Raffaella Bernardi,<br>Raffaella.Bernardi@unibz.it<br>https://www.unibz.it/en/faculties/engineering/academic-staff/person/2311 |
| **Assistente** | |
| **Semestre** | Secondo semestre |
| **Anno/i di corso** | 1 |
| **CFU** | 6 |
| **Ore didattica frontale** | 40 |
| **Ore di laboratorio** | 20 |
| **Ore di studio individuale** | 90 |
| **Ore di ricevimento previste** | |
| **Sintesi contenuti** | - Symbolic approaches to NLP<br>- Reccommender Systems<br>- N-gram Language Models and neural networks approaches<br>- Large Language Models and Post-training methods<br>- Conversational Systems<br>- NLP application to RS |
| **Argomenti dell'insegnamento** | The course will provide an overview of Natural Language Processing (NLP) and give some insights on its application to Recommender Systems (RS). It will start by introducing the NLP |

| | |
|---|---|
| | field, its long-standing goals and motivations. Then it will consists of three main parts:<br>I) it introduces:<br>Language Models (n-grams);<br>the move from LM to Neural LM by first diving into Distributional Semantics.<br>the raise of Large Language Models.<br><br>II) it discusses weakness and strengths of LLM with respect to the NLP long-standing challenges. In particular:<br>Syntax,<br>Semantics and Reasoning,<br>Pragmatics and Dialogues,<br>Visually Grounded Conversational systems<br><br>III) it gives an overview of Recommender Systems (RS) and some insights on the application of NLP to RS.<br><br>Parts I and II will complement the frontal classes with labs in which students will have hands-on experience on the topics introduced through the frontal classes. |
| **Parole chiave** | Natural Language Processing, Symbolic and Statistical approaches to Syntax, Semantics, Language Models and Large Language Models, Recommender Systems |
| **Prerequisiti** | Some knowledge of Linear algebra, probability, basic machine learning concepts is required; some basic knowledge of logic will be useful but it is not mandatory. |
| **Insegnamenti propedeutici** | |
| **Modalità di insegnamento** | Frontal lectures, exercises, labs, projects, seminars. |
| **Obbligo di frequenza** | Attending lectures is not compulsory but highly recommended. Final marks will be awarded on the basis of the outcome of the final exam and assignments (project and seminar). Students who are not attending the lectures must contact the lecturer to receive their mandatory assignments. |
| **Obiettivi formativi specifici e risultati di apprendimento attesi** | The course belongs to the type "caratterizzanti – discipline informatiche" in the curricula “Data Analytics”.<br>The main objective of the course is to a) reach an overview of Natural Language Processing (NLP) and its application to |

| | |
|---|---|
| | Recommender Systems (RS); and to familiarize with b) a variety of approaches and c) evaluation methods.<br>Hence, the student will familiarize with today's NLP models and their application to RS situating them in the broader and long-standing context of the field.<br><br>Applying knowledge and understanding:<br>- D1.2 - Understanding of the skills, tools and techniques required for an effective use of data science<br>- D1.7 - Knowledge of artificial intelligence techniques and methods for the implementation of intelligent systems<br>Applying knowledge and understanding:<br>- D2.1 - Practical application and evaluation of tools and techniques in the field of data science<br>- D2.2 - Ability to address and solve a problem using scientific methods<br>- D2.6 - Ability to apply innovative techniques of data mining and machine learning to extract knowledge from complex and heterogeneous data<br>Making judgments<br>- D3.2 - Ability to autonomously select the documentation (in the form of books, web, magazines, etc.) needed to keep up to date in a given sector<br>Communication skills<br>- D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology<br>- D4.3 - Ability to structure and draft scientific and technical documentation<br>Learning skills<br>- D5.2 - Ability to autonomously keep oneself up to date with the developments of the most important areas of data science<br>- D5.3 - Ability to deal with problems in a systematic and creative way and to appropriate problem solving techniques |
| **Obiettivi formativi specifici e risultati di apprendimento attesi (ulteriori info.)** | |
| **Modalità di esame** | - Mid-term assignments (e.g. Quiz, code assignment, Seminars), 20%<br>- Project in a small team (2 students), 40% |

| | |
|---|---|
| | - Final exam, written, 40 % of mark<br><br>The mid-term assignments will help students monitor their understanding of the topics introduced during the course. To this end, we will use quiz, and presentations of assignments such as coding and/or understanding scientific paper, chosen from the list of papers given by the lecturer. The presentation is aimed at assessing to what extent the student has achieved the above-mentioned learning outcomes related to: applying knowledge and understanding, making judgments, communication and learning skills.<br><br>The project will consist on the proposal for the evaluation of an NLP model against a specific task. The project domain, the attacked problem, the techniques and the obtained results, if any, must be described in a report (~4 pages of ACL latex style). The project is aimed at assessing to what extent the student has achieved the above-mentioned learning outcomes related to: applying knowledge and understanding, making judgments, and communication skills.<br><br>The written exam is based on verification questions and exercises, and it will assess to what extent the student has achieved the above-mentioned learning outcomes related to: knowledge and understanding, applying knowledge and understanding, and learning skills. |
| Criteri di valutazione | The mid-term assignment will be evaluated at the end of the semester and it is a prerequisite for attending the written exam. The project report must be submitted one week before the written exam. The student must pass each single component of the exam (mid-term assignment, project and written exam), by obtaining at least 50% of the mark.<br>Presentation evaluation: clarity of the presentation and capability to reply to clarification questions of the teacher.<br>Project evaluation criteria: clarity of the presentation, novelty of the problem and the solution, appropriateness of the methods, significance of the solution and capability to refer to the methods and techniques discussed in the course.<br>Written exam: correctness and clarity of the replies to the posed questions. |

| | |
|---|---|
| | Students who do not attend classes have to contact the lecturer at least one month before the written exam. |
| **Bibliografia obbligatoria** | Speech and Language Processing (3rd ed. 2025), Dan Jurafsky and James H. Martin<br><br>All the required reading material will be provided during the course and will be available in electronic format. Copy of the slides will be available as well.<br><br>Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it |
| **Bibliografia facoltativa** | Scientific papers and additional books will be specified on each course topic during lectures. |
| **Altre informazioni** | Software used: Python |
| **Obiettivi di Sviluppo Sostenibile (SDGs)** | Istruzione di qualità |