Università degli Studi di Roma "Tor Vergata"

Facsimile Scheda Insegnamento

Mod. Scheda Insegnamento v. 1.1 -

Docente responsabile dell'insegnamento/attività formativa

Nome	Claudia Roberta
Cognome	Combei

Denominazione insegnamento/attività formativa

Italiano	Linguistica Applicata
Inglese	Applied Linguistics

Informazioni insegnamento/attività formativa

A.A.	2026/2027
	◉ L ○ LM ○ LM CU
CdS	Lettere; Lingue e Letterature Moderne
Codice	804003452; 804001769
Canale	
CFU	6
Lingua	italiano

Docente del modulo didattico (compilare solo per attività formative articolate in moduli)

Nome	Claudia Roberta
Cognome	Combei

Denominazione modulo didattico (compilare solo per attività formative articolate in moduli)

Italiano	Linguistica Applicata
Inglese	Applied Linguistics

Obiettivi formativi e risultati di apprendimento attesi

Italiano	OBIETTIVI FORMATIVI: Il corso si propone di offrire una panoramica introduttiva sugli strumenti per l ' analisi dei corpora e sui modelli di base per il trattamento automatico delle lingue e del linguaggio. CONOSCENZA E CAPACITÀ DI COMPRENSIONE: Comprendere i concetti fondamentali della linguistica computazionale, comprese le tecniche di base dell ' elaborazione del linguaggio naturale. CAPACITÀ DI APPLICARE CONOSCENZA E COMPRENSIONE: Usare i corpora e gli strumenti di base per l ' elaborazione e l ' analisi computazionale dei dati e dei fenomeni linguistici. AUTONOMIA DI GIUDIZIO: Valutare criticamente le applicazioni di base del trattamento automatico del linguaggio e i loro limiti. ABILITÀ COMUNICATIVE: Descrivere e spiegare i concetti di base della linguistica computazionale e il funzionamento dei principali metodi del trattamento automatico del linguaggio. CAPACITÀ DI APPRENDIMENTO: Acquisire gli strumenti necessari per approfondire in autonomia aspetti teorici e pratici della linguistica computazionale e per applicarli a nuovi contesti di ricerca.
Inglese	LEARNING OUTCOMES: The course aims to provide an introductory overview of tools for corpus analysis and basic models for automatic language processing. KNOWLEDGE AND UNDERSTANDING: Understand the main concepts of computational linguistics, including basic techniques of natural language processing. APPLYING KNOWLEDGE AND UNDERSTANDING: Use corpora and the main tools for computational processing and analysis of linguistic data and phenomena. MAKING JUDGEMENTS: Evaluate basic applications of natural language processing and their limitations. COMMUNICATION SKILLS: Describe and explain the basic concepts of computational linguistics and the functioning of main methods in automatic language processing. LEARNING SKILLS: Acquire the necessary tools to independently deepen both the theoretical and practical aspects of computational linguistics and to apply them to new research contexts.

Prerequisiti

Italiano	È richiesta una comprensione dei concetti di base della linguistica generale. Non sono richieste conoscenze pregresse di programmazione in Python o di sviluppo di modelli e applicazioni NLP. Durante il corso verranno fornite alcune basi di programmazione, utili per affrontare le attività previste dal syllabus.
Inglese	A basic understanding of general linguistics concepts is required. No prior knowledge of Python programming or the development of NLP models and applications is required. Some basic programming concepts will be introduced during the course, as they will be useful for completing the activities outlined in the syllabus.

Programma

Italiano	Il corso si propone di fornire un ' introduzione teorica e pratica alla linguistica computazionale e al trattamento automatico del linguaggio. Nella prima parte del corso si illustreranno le metodologie per la raccolta, la compilazione e la gestione dei corpora. Verranno approfondite le tecniche di (pre-)elaborazione del testo (ad es., tokenizzazione, stemming, lemmatizzazione, ecc.) e di annotazione linguistica (ad es., POS-tagging, annotazione semantica, ecc.). L ' analisi dei dati e dei fenomeni linguistici sarà affrontata attraverso lo studio delle co-occorrenze, delle analisi frequentiste e dell ' estrazione dei pattern dai corpora. Saranno esplorati i modelli di base per la rappresentazione del testo, come il Bag-of-Words e il TF-IDF. Saranno inoltre introdotti alcuni modelli di base per la classificazione supervisionata, come il Naive Bayes e la regressione. Infine, il corso offrirà un ' introduzione pratica all ' uso di Python per il trattamento automatico del linguaggio,
Inglese	The course aims to provide a theoretical and practical introduction to computational linguistics and natural language processing. In the first part, methodologies for corpus collection, compilation, and management will be presented. Techniques for text (pre-) processing (e.g., tokenization, stemming, lemmatizatio, etc.) and linguistic annotation (e.g., POS tagging, semantic annotation, etc.) will be explored in depth. The analysis of linguistic data and phenomena will be addressed through the study of co-occurrences, frequency-based analyses, and pattern extraction from corpora. Fundamental models for text representation, such as Bag-of-Words and TF-IDF, will be taught. Basic models for supervised classification, such as Naive Bayes and regression, will be introduced. Finally, the course will provide a hands-on introduction to the use of Python for automatic language processing, with a particular focus on the NLTK and spaCy libraries. During practical activities, students will have the opportunity to apply the acquired knowledge by working

Università degli Studi di Roma "Tor Vergata"

Facsimile Scheda Insegnamento

Modalità di valutazione

- [x] Prova scritta
- [ ] Prova orale
- [ ] Valutazione in itinere
- [x] Valutazione di progetto
- [ ] Valutazione di tirocinio
- [ ] Prova pratica
- [ ] Prova di laboratorio

Descrizione delle modalità e dei criteri di verifica dell'apprendimento

Italiano	MODALITA' DI VALUTAZIONE: L ' esame si compone di due parti: un progetto (elaborato scritto), concordato con la docente e incentrato su uno degli argomenti del corso, e una prova scritta. Il progetto va consegnato almeno una settimana prima della data dell ' appello a cui si è iscritti/e. Chi non consegna il progetto non può sostenere la prova scritta. CRITERI DI VERIFICA DI APPRENDIMENTO: Durante la prova scritta, gli studenti e le studentesse risponderanno a domande sul programma d ' esame. L ' esame coprirà sia i temi trattati a lezione sia quelli lasciati allo studio individuale. Saranno oggetto di esame anche i testi di riferimento indicati nel presente syllabus. Le domande saranno di tipo teorico e/o argomentativo. Il voto finale terrà conto del progetto (50%) e dell ' accuratezza e delle completezza delle risposte date alla prova scritta (50%), considerando anche correttezza terminologica e adeguatezza espositiva. La valutazione finale dell ' esame
Inglese	ASSESSMENT METHODS: The exam consists of two parts: a project, agreed upon with the instructor and focused on one of the course topics, and a written examination. The project must be submitted at least one week before the scheduled exam date. Those who do not submit the written project cannot take the written exam. CRITERIA FOR ASSESSMENT: During the written exam, students will answer questions on the course material. The reference texts listed in this syllabus will also be part of the exam. The exam will cover both the topics addressed in class and those assigned for independent study. Questions will be theoretical and/or argumentative in nature. The final grade will be based on the project (50%) and the accuracy and completeness of the responses of the written exam (50%), considering also the correct use of terminology and clarity of presentation. The final exam is graded on a scale of thirty. The exam will be assessed according to the following criteria: Fail: Significant gaps and inaccuracies in

Testi adottati

Italiano	Nissim, Malvina & Pannitto, Ludovica. (2022). Che cos ’ è la linguistica computazionale. Roma: Carocci. [tutto] Lenci, Alessandro, Montemagni, Simonetta, & Pirrelli, Vito (2016). Testo e computer: Elementi di linguistica computazionale. Roma: Carocci. [capitoli: 1, 2, 4, 7, 8, 9]
Inglese	Nissim, Malvina & Pannitto, Ludovica. (2022). Che cos ’ è la linguistica computazionale. Roma: Carocci. [all chapters] Lenci, Alessandro, Montemagni, Simonetta, & Pirrelli, Vito (2016). Testo e computer: Elementi di linguistica computazionale. Roma: Carocci. [chapters: 1, 2, 4, 7, 8, 9]

Bibliografia di riferimento

Italiano	I due manuali indicati sotto la voce “ Testi adottati ” , le slides ed eventuale altro materiale fornito dalla docente.
Inglese	The two textbooks listed under "Required Manuals", the slides, and any other material provided by the instructor.

Modalità di svolgimento

☒ Modalità in presenza
☐ Modalità a distanza

Descrizione della modalità di svolgimento e metodi didattici adottati

Italiano	Lezioni frontali ed esercitazioni pratiche.
Inglese	Lectures and practical exercises.

Modalità di frequenza

◉ Frequenza facoltativa
○ Frequenza obbligatoria

Descrizione della modalità di frequenza

Italiano	La frequenza è facoltativa e il programma è identico per frequentanti e non-frequentanti. Gli studenti e le studentesse dovranno contattare la docente con congruo anticipo per ottenere le istruzioni necessarie alla realizzazione del progetto (elaborato scritto).
Inglese	Attendance is optional, and the syllabus is the same for both attending and non-attending students. Students are required to contact the instructor well in advance to receive the necessary instructions for completing the written project (written work).