COURSE OUTLINE

GENERAL

SCHOOL	Sciences and Engineering		
ACADEMIC UNIT	Computer Science		
LEVEL OF STUDIES	1st Cycle		
COURSE CODE	COMP-348	**SEMESTER**	Fall/Spring
COURSE TITLE	Natural Language Processing		

INDEPENDENT TEACHING ACTIVITIES *if credits are awarded for separate components of the course, e.g. lectures, laboratory exercises, etc. If the credits are awarded for the whole of the course, give the weekly teaching hours and the total credits*	**WEEKLY TEACHING HOURS**	**CREDITS**
	2.5	6
Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).		

COURSE TYPE *general background, special background, specialised general knowledge, skills development*	Specialization
PREREQUISITE COURSES:	COMP-244
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	English
IS THE COURSE OFFERED TO ERASMUS STUDENTS	
COURSE WEBSITE (URL)	

LEARNING OUTCOMES

Learning outcomes

The course learning outcomes, specific knowledge, skills and competences of an appropriate level, which the students will acquire with the successful completion of the course are described.

Consult Appendix A

- *Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area*
- *Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B*
- *Guidelines for writing Learning Outcomes*

After completion of the course students are expected to be able to:

1. Comprehend various fundamental concepts of NLP: Text processing (normalization, lemmatization, stemming, etc.), language models (N-Grams), word representation (word embeddings), and text classification with Machine Learning.
2. Familiarise with known NLP tasks: Named Entity Recognition (NER), Part-of-Speech tagging (PoS), Dependency and Syntax parsing.
3. Employ Machine Learning (ML) techniques for text classification (e.g., Naive Bayes) and be able to properly apply the NLP feature engineering process.
4. Extend and integrate their knowledge with advanced methods in NLP and Deep Learning:

Word Vectors, Word2vec algorithm, BERT.

5. Apply their advanced knowledge on real-world research applications of NLP and recognize the societal impact in cases of misinformation and hate-speech identification.
6. Design, develop, and evaluate efficient and effective NLP solutions to a variety of problems, using state-of-the-art tools.

General Competences

Taking into consideration the general competences that the degree-holder must acquire (as these appear in the Diploma Supplement and appear below), at which of the following does the course aim?

Search for, analysis and synthesis of data and information, with the use of the necessary technology	*Project planning and management*
Adapting to new situations	*Respect for difference and multiculturalism*
Decision-making	*Respect for the natural environment*
Working independently	*Showing social, professional and ethical responsibility and sensitivity to gender issues*
Team work	*Criticism and self-criticism*
Working in an international environment	*Production of free, creative and inductive thinking*
Working in an interdisciplinary environment	*……*
Production of new research ideas	*Others…*
	…….

Search for, analysis and synthesis of data and information, with the use of the necessary technology
Adapting to new situations
Decision-making
Working independently
Production of new research ideas
Project planning and management
Criticism and self-criticism
Production of free, creative and inductive thinking
Working in an interdisciplinary environment

SYLLABUS

The content of the course is organized in the following groups:

<u>Part 1: Fundamental Knowledge, Concepts, and Techniques of NLP.</u>

- Introduction: A brief introduction of Natural Language Processing (NLP) in the matters of extracting information from language using methods such as Information Retrieval (IR), text classification with Machine Learning (ML) and sentiment analysis.
- Basic Text Processing: Fundamental applications of text processing using regular expressions, sentence segmentation, word tokenization and word normalisation using lemmatization and stemming. Also, introduction to string similarity techniques such as minimum edit distance (MED).
- Language Modelling: Introduction to probabilistic language models such as Unigram, Bigram and N-gram models. Methods of language model evaluation and generalisation.
- Text Classification: Use of ML algorithms (e.g., Naive Bayes) to classify text into categories. Identify distinguishable textual features using feature extraction and feature selection methods. Classification task evaluation measures of Precision, Recall and F-measure. Introduction to sentiment analysis task.
- Natural Language Processing Tasks: Introduction to widely known NLP tasks such as Part-of-Speech (PoS) tagging, Named Entity Recognition (NER), and Dependency and Constituency parsing.

<u>Part 2: Introduction to Deep Learning methods for NLP.</u>

- Word Vector Semantics: Introduction to vector space models and word distributional similarity. Analysis of the word2vec algorithm and word embeddings usage in NLP tasks.
- NLP, Deep Learning, and Transformers: Introduction to Deep Learning (DL) for NLP. Take a look into Neural Network structures used in NLP, such as Multi-Layer Perceptrons (MLP), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN). Introduction to recent DL advancements in NLP with Seq2Seq, followed by a detailed overview of how Transformers operate (e.g. encoder-decoder, self-attention). We will then introduce Large Language Models (LLMs), covering key concepts such as in-context learning , fine-tuning, and specific metrics.
- Future of NLP: Discuss on the future of NLP using examples from recent advancements in both industry and research. Also discuss the ethical issues that arise along with the advancement in NLP

TEACHING and LEARNING METHODS - EVALUATION

<table>
<tr><td>DELIVERY
Face-to-face, Distance learning, etc.</td><td>Face-to-face</td></tr>
<tr><td>USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Use of ICT in teaching, laboratory education, communication with students</td><td>Use of ICT in teaching / Χρήση ΤΠΕ
Communication with students / Επικοινωνία με Φοιτητές</td></tr>
<tr><td>TEACHING METHODS
The manner and methods of teaching are described in detail.
Lectures, seminars, laboratory practice, fieldwork, study and analysis of bibliography, tutorials, placements, clinical practice, art workshop, interactive teaching, educational visits, project, essay writing, artistic creativity, etc.

The student's study hours for each learning activity are given as well as the hours of non-directed study according to the principles of the ECTS</td><td>
<table>
<tr><th>Activity</th><th>Semester workload</th></tr>
<tr><td>Lectures</td><td>35</td></tr>
<tr><td>Preparation, homework, quizzes</td><td>60</td></tr>
<tr><td>Project</td><td>28</td></tr>
<tr><td>Exam preparation</td><td>25</td></tr>
<tr><td>Final Exam</td><td>2</td></tr>
<tr><td></td><td></td></tr>
<tr><td>Course total</td><td>150</td></tr>
</table>
</td></tr>
<tr><td>STUDENT PERFORMANCE EVALUATION
Description of the evaluation procedure

Language of evaluation, methods of evaluation, summative or conclusive, multiple choice questionnaires, short-answer questions, open-ended questions, problem solving, written work, essay/report, oral examination, public presentation, laboratory work, clinical examination of patient, art interpretation, other

Specifically-defined evaluation criteria are given, and if and where they are accessible to students.</td><td>- Mid-term exam
- Projects
- Final Examination
- Participation/Homework Assignments/Quizzes</td></tr>
</table>

ATTACHED BIBLIOGRAPHY

Required Textbooks / Readings:

Title	Author(s)	Publisher	Year	ISBN
Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition	Daniel Jurafsky, James H. Martin	Prentice Hall PTR, Upper Saddle River, NJ, United States	2023	9780130950697

Recommended Textbooks / Readings:

Title	Author(s)	Publisher	Year	ISBN
Natural Language Processing with Python	Steven Bird, Ewan Klein, Edward Loper	O'Reilly Media, Inc.	2009	9780596516499
Natural Language Processing with Transformers, Revised Edition	Lewis Tunstall, Leandro von Werra, Thomas Wolf	O'Reilly Media, Inc.	2022	978098136796