

<b>INSTITUTION</b>	NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS					
<b>SCHOOL</b>	SCHOOL OF SCIENCE					
<b>DEPARTMENT</b>	INFORMATICS AND TELECOMMUNICATIONS					
<b>COURSE LEVEL</b>	GRADUATE					
<b>COURSE TITLE</b>	Programming for Language Technology II					
<b>COURSE CODE</b>	<b>C28</b>	<b>Semester</b>	2	<b>ECTS</b>	6	
<b>TEACHING HOURS per week</b>	<b>THEORY</b>	2	<b>SEMINAR.</b>	0	<b>LABORATORY</b>	0
<b>URL</b>	<a href="https://eclass.uoa.gr/courses/DI523/">https://eclass.uoa.gr/courses/DI523/</a>					

<b>COURSE CONTENT</b>
<p>Python and NLTK, Access to text corpora and lexical resources (e.g. Wordnet), Processing of raw text, Develop structured programs, Text/Word annotation (e.g.POS tagging), Text classification, Information extraction from text, Extract sentence structure , Grammatical analysis. Language models, word embeddings, n-grams, evaluation with precision, recall, F-measure and accuracy.</p> <p>Data processing, Python , entry / exit audio files (wav, mp3, ogg, flac, etc.), Signal processing blocks (FFT, DCT, frame cutter, windowing, envelope, smoothing), Filters FIR &amp; IIR (low / high / band pass, band reject, DC removal, equal loudness), statistics (median, mean, standard deviation, power means, raw and central moments, spread, kurtosis, skewness, flatness), Time-domain descriptors (duration, loudness, zero-crossing-rate, log attack time and other signal envelope descriptors) Spectral descriptors (Bark / Mel / ERB bands, MFCC, GFCC, LPC, spectral peaks).</p>

<b>STUDENT LEARNING OBJECTIVES</b>
<p>Upon successful completion of the course the student will be able to:</p> <ul style="list-style-type: none"> <li>- mention functions and libraries provided by the NLTK framework</li> <li>- set up and execute NLP pipelines on text corpora</li> <li>- develop an NLP application (e.g. spam filter, named entity recognizer, summarizer, sentiment analyzer) using machine learning algorithms (scikit-learn) and NLTK</li> <li>- Recognize groups of sound features that are useful for dealing with corresponding problem groups</li> <li>- Identify intuitively which audio elements of speech are reflected in different audio features that can be output from a digital signal</li> <li>- design basic prototype audio processing and statistical analysis programs in python for quick testing of research hypotheses</li> </ul>

<b>TEACHING AND LEARNING METHODS - ASSESSMENT</b>	
<b>TEACHING METHOD</b>	Remote - online
<b>USE OF INFORMATION AND COMMUNICATION TECHNOLOGIES</b>	Support for the learning process through the electronic platform e-class, namely: Provision of material - posting presentations and supporting material, Discussions, Announcements, Assignment-Submission of assignments. Communication via email.

**TEACHING ORGANIZATION**

Describe in detail the way and methods of teaching:

Enhanced Lectures,  
 Online Lectures,  
 Seminars,  
 Tutorial,  
 Laboratory,  
 Laboratory Exercise,  
 Study & analysis of literature,  
 Practice (Positioning),  
 Interactive teaching,  
 Developing a project,  
 Individual / group work  
 Telework (reference to tools) etc.

Details of the student's study hours for each learning activity and hours of non-guided study are shown to ensure that the total workload at the semester corresponds to the ECTS

Activity	Student Workload (hours)
Lectures	26
Laboratory	13
Teamwork in a case study	31
Small individual exercises	40
Independent Study	40
<b>Total Course (25 hours of workload per unit of credit)</b>	<b>150</b>

**ASSESSMENT OF STUDENTS**

Description of the assessment process

Assessment Methods, Formative or Concluding, Multiple Choice Test, Quick Response Questions, Test Development Questions, Problem Solving, Written Work, Report / Report, Oral Examination, Public Presentation, Laboratory Work, Other / Other

Fully defined evaluation criteria are mentioned and if and where they are accessible to students.

Describe explicitly methods, evaluation tools and provided feedback.

The table below is supplemented accordingly.

Assessment methods	Number	Percentage
Progress	2	20%
Exercises	4	30%
Final work	1	50%

**LITERATURE AND STUDY MATERIALS / READING LIST**

- <https://www.nltk.org/>
- <https://www.nltk.org/book/>
- Speech and Language Processing, Dan Jurafsky and James H. Martin: <https://web.stanford.edu/~jurafsky/slp3/>
- McLoughlin, I. (2009). *Applied speech and audio processing: with Matlab examples*. Cambridge University Press.
- Sen, S., Dutta, A., & Dey, N. (2019). *Audio Processing and Speech Recognition: Concepts, Techniques and Research Overviews*. Springer.
- Yu, D., & Deng, L. (2016). *AUTOMATIC SPEECH RECOGNITION*. Springer london limited.
- Kamath, U., Liu, J., & Whitaker, J. (2019). *Deep learning for nlp and speech recognition* (Vol. 84). Springer.
- Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer.
- Eyben, F. (2015). *Real-time speech and music classification by large audio feature space extraction*. Springer.
- Ogunfunmi, T., Togneri, R., & Narasimha, M. (Eds.). (2015). *Speech and audio processing for coding, enhancement and recognition*. Springer New York.