

ΙΔΡΥΜΑ	ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ									
ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ									
ΤΜΗΜΑ	ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ									
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΜΕΤΑΠΤΥΧΙΑΚΟ - ΔΠΜΣ ΓΛΩΣΣΙΚΗ ΤΕΧΝΟΛΟΓΙΑ									
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	M906 Προγραμματισμός για Γλωσσική Τεχνολογία II									
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	M906	Εξάμηνο	2ο	ECTS	6					
ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΘΕΩΡ.	2	ΦΡΟΝΤ.	ΕΡΓΑΣΤ.	1					
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ	<p>Επιλέξτε ένα από τα ακόλουθα και διαγράψτε τα υπόλοιπα Βασικό (B) / Επιλογής (E)</p> <table border="1"> <tr> <td>B</td> <td>E</td> </tr> <tr> <td>X</td> <td></td> </tr> </table>						B	E	X	
B	E									
X										
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	https://eclass.uoa.gr/courses/DI523/									
ΠΡΟΕΡΑΙΤΙΚΑ / ΣΥΝΙΣΤΩΜΕΝΑ ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ	M901 Προγραμματισμός για Γλωσσική Τεχνολογία I M904 Εισαγωγή στην Φωνολογία και στην Ψηφιακή Επεξεργασία Σήματος									
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:	ΕΛΛΗΝΙΚΑ									
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	ΟΧΙ									

ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ
<i>Συμπληρώστε (λεκτική περιγραφή) το περιεχόμενο του μαθήματος αναφέροντας τη δομή και τα θέματα που καλύπτονται.</i>
<p>Python και NLTK, Πρόσβαση σε σώματα κειμένων και λεξικούς πόρους, Επεξεργασία μη επισημειωμένου κειμένου, Ανάπτυξη δομημένων προγραμμάτων, Επισημείωση λέξεων, Ταξινόμηση κειμένου, Εξαγωγή πληροφορίας από κείμενο, Ανάλυση προτασιακής δομής, Γραμματικές Χαρακτηριστικών, Ανάλυση προτασιακής σημασίας, Διαχείριση γλωσσικών δεδομένων, Python και Essentia, είσοδος/έξοδος αρχείων ήχου (wav, mp3, ogg, flac, etc.), Signal processing blocks (FFT, DCT, frame cutter, windowing, envelope, smoothing), Φίλτρα FIR & IIR (low/high/band pass, band reject, DC removal, equal loudness), στατιστική (διάμεσος, μέσος όρος, διασπορά, power means, raw and central moments, spread, kurtosis, skewness, flatness), Time-domain descriptors (διάρκεια, loudness, zero-crossing-rate, log attack time and other signal envelope descriptors) Spectral descriptors (Bark/Mel/ERB bands, MFCC, GFCC, LPC, spectral peaks).</p>

ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ
<i>Περιγράψτε τους στόχους ή/και τα μαθησιακά αποτελέσματα του μαθήματος. Μπορείτε να αναφερθείτε στις επιμέρους κατηγορίες των μαθησιακών αποτελεσμάτων σε επίπεδο γνώσεων, δεξιοτήτων (νοητικών, πρακτικών) και ικανοτήτων. Συμβουλευτείτε το Παράρτημα «Περληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων» για τη συγγραφή των μαθησιακών αποτελεσμάτων.</i>
<p>Με την επιτυχή ολοκλήρωση του μαθήματος οι φοιτητές/φοιτήτριες θα είναι σε θέση να:</p> <ul style="list-style-type: none"> • Γνωρίζουν τις λειτουργίες και βιβλιοθήκες που παρέχει το εργαλείο NLTK. • Μπορούν να στήσουν και να εκτελέσουν NLP pipelines σε σώματα κειμένων. • Μπορούν να αναπτύξουν μια NLP εφαρμογή (π.χ. spam filter, named entity recognizer, summarizer, sentiment analyzer) με την χρήση αλγορίθμων μηχανικής μάθησης και του NLTK. • Αναγνωρίζουν τις ομάδες ηχητικών χαρακτηριστικών που είναι χρήσιμες για την αντιμετώπιση αντίστοιχων ομάδων προβλημάτων, • Κατανοούν διαισθητικά ποια ηχητικά στοιχεία της ομιλίας απεικονίζονται σε διαφορετικά ηχητικά χαρακτηριστικά που μπορούν να εξαχθούν από ψηφιακό σήμα, • Σχεδιάζουν βασικά πρωτότυπα προγράμματα επεξεργασίας ήχου και στατιστικής ανάλυσης σε γλώσσα python για γρήγορο έλεγχο ερευνητικών υποθέσεων.

ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ**ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ**

Δια ζώσης ή εξ αποστάσεως

ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ

Υποστήριξη μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας e-class, συγκεκριμένα: Παροχή υλικού – ανάρτηση παρουσιάσεων και βοηθητικού υλικού, Συζητήσεις, Ανακοινώσεις, Ανάθεση-Κατάθεση εργασιών, Επικοινωνία μέσω ηλεκτρονικού ταχυδρομείου.

ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ – ΔΙΔΑΚΤΙΚΕΣ ΤΕΧΝΙΚΕΣ

Περιγράφονται αναλυτικά ο τρόπος και τεχνικές διδασκαλίας και αναγράφονται αναλυτικά οι ώρες μελέτης του φοιτητή για κάθε μαθησιακή δραστηριότητα καθώς και οι ώρες μη καθοδηγούμενης μελέτης ώστε ο συνολικός φόρτος εργασίας σε επίπεδο εξαμήνου να αντιστοιχεί στα standards του ECTS (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)

Περιγράφονται αναλυτικά ο τρόπος και τεχνικές διδασκαλίας, αριθμός ατόμων ανά ομάδα κτλ. Συμπληρώνεται αναλόγως και ο παρακάτω πίνακας.

Δραστηριότητα	Φόρτος (ώρες)
Διαλέξεις	26
Εργαστήριο	13
Ατομική εργασία σε μελέτη περίπτωσης	31
Μικρές ατομικές εργασίες εξάσκησης	40
Αυτοτελής Μελέτη	40
Σύνολο ωρών	150

ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ

Περιγραφή της διαδικασίας αξιολόγησης

Μέθοδοι αξιολόγησης (Διαμορφωτική ή/και Τελική), Εργαλεία Αξιολόγησης (Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, Επίλυση Προβλημάτων, Γραπτή Εργασία, Εκθεση/Αναφορά, Προφορική Εξέταση, Δημόσια Παρουσίαση, Εργαστηριακή Εργασία, Άλλη/άλλες,), Παροχή ανατροφοδότησης (περιγραφική, μέσω κλίμακας διαβαθμισμένων κριτηρίων)

Περιγράφονται ρητά οι μέθοδοι, τα εργαλεία αξιολόγησης και η παρεχόμενη ανατροφοδότηση αποτελεσμάτων. Συμπληρώνεται αναλόγως και ο παρακάτω πίνακας.

Αξιολόγηση	Αριθμός	Ποσοστό
Πρόοδοι	2	20%
Ασκήσεις	4	30%
Τελική εργασία	1	50%

ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

- <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.