

ΠΕΡΙΓΡΑΜΜΑ ΜΑΘΗΜΑΤΟΣ

(1) ΓΕΝΙΚΑ

ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ		
ΤΜΗΜΑ	ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΠΡΟΠΤΥΧΙΑΚΟ		
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	ΕΠ25	ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	7^ο
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	Ενισχυτική Μηχανική Μάθηση και Στοχαστικά Παίγνια		
ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ <i>σε περίπτωση που οι πιστωτικές μονάδες απονέμονται σε διακριτά μέρη του μαθήματος π.χ. Διαλέξεις, Εργαστηριακές Ασκήσεις κ.λπ. Αν οι πιστωτικές μονάδες απονέμονται ενιαία για το σύνολο του μαθήματος αναγράψτε τις εβδομαδιαίες ώρες διδασκαλίας και το σύνολο των πιστωτικών μονάδων</i>		ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ
Διαλέξεις		4	6
<i>Προσθέστε σειρές αν χρειαστεί. Η οργάνωση διδασκαλίας και οι διδακτικές μέθοδοι που χρησιμοποιούνται περιγράφονται αναλυτικά στο (4).</i>			
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ <i>γενικού υποβάθρου, ειδικού υποβάθρου, ειδίκευσης γενικών γνώσεων, ανάπτυξης δεξιοτήτων</i>	Προαιρετικό Μάθημα (ΠΜ)		
	K	E1	E2
	E3	E4	E5
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ:			
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:		ΕΛΛΗΝΙΚΑ/ΑΓΓΛΙΚΑ	
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS		ΝΑΙ	
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)		https://eclass.uoa.gr/courses/DI729/	

(2) ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Μαθησιακά Αποτελέσματα

Περιγράφονται τα μαθησιακά αποτελέσματα του μαθήματος οι συγκεκριμένες γνώσεις, δεξιότητες και ικανότητες καταλλήλου επιπέδου που θα αποκτήσουν οι φοιτητές μετά την επιτυχή ολοκλήρωση του μαθήματος.

Συμβουλευτείτε το Παράρτημα Α

• Περιγραφή του Επιπέδου των Μαθησιακών Αποτελεσμάτων για κάθε ένα κύκλο σπουδών σύμφωνα με το Πλαίσιο Προσόντων του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης • Περιγραφικοί Δείκτες Επιπέδων 6, 7 & 8 του Ευρωπαϊκού Πλαισίου Προσόντων Διά Βίου Μάθησης και το Παράρτημα Β • Περιληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων																	
• Απόκτηση γνώσεων και δεξιοτήτων πάνω σε μια σημαντική περιοχή της μηχανικής μάθησης που έχει ισχυρούς δεσμούς με θεωρία βέλτιστου ελέγχου, στοχαστική βελτιστοποίηση, προσεγγιστικές δομές (νευρωνικά δίκτυα) και θεωρία παιγνίων, με εξαιρετικές επιτυχίες τα τελευταία χρόνια. • Κατά την ολοκλήρωση του μαθήματος, ο φοιτητής θα είναι σε θέση να: • μοντελοποιεί ορθώς προβλήματα ενισχυτικής μηχανικής μάθησης χρησιμοποιώντας των κατάλληλο φορμαλισμό ανά περίπτωση, • υλοποιεί εξ'ολοκλήρου αλγορίθμους βέλτιστου ελέγχου και στοχαστικής βελτιστοποίησης «πολλαπλών ληστών» και «Μαρκοβιανών διαδικασιών» βασισμένων στις εξισώσεις Bellman, • κατανοεί, γνωρίζει τις διαφορές και χρησιμοποιεί αλγορίθμους βαθιάς ενισχυτικής μηχανικής μάθησης τόσο σε θεωρητικό επίπεδο όσο και σε προγραμματιστικά περιβάλλοντα, • σχεδιάζει πειραματικές μεθόδους εφαρμογής αλγορίθμων ενισχυτικής μηχανικής μάθησης μέσω αριθμητικών/προγραμματιστικών προσομοιώσεων και να αξιολογεί τα αποτελέσματά τους.																	
Γενικές Ικανότητες Λαμβάνοντας υπόψη τις γενικές ικανότητες που πρέπει να έχει αποκτήσει ο πτυχιούχος (όπως αυτές αναγράφονται στο Παράρτημα Διπλώματος και παρατίθενται ακολούθως) σε ποια / ποιες από αυτές αποσκοπεί το μάθημα; <table> <tr> <td>Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών</td><td>Σχεδιασμός και διαχείριση έργων</td></tr> <tr> <td>Προσαρμογή σε νέες καταστάσεις</td><td>Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα</td></tr> <tr> <td>Λήψη αποφάσεων</td><td>Σεβασμός στο φυσικό περιβάλλον</td></tr> <tr> <td>Αυτόνομη εργασία</td><td>Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου</td></tr> <tr> <td>Ομαδική εργασία</td><td>Άσκηση κριτικής και αυτοκριτικής</td></tr> <tr> <td>Εργασία σε διεθνές περιβάλλον</td><td>Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης</td></tr> <tr> <td>Εργασία σε διεπιστημονικό περιβάλλον</td><td>.....</td></tr> <tr> <td>Παραγωγή νέων ερευνητικών ιδεών</td><td>Άλλες...</td></tr> </table>		Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών	Σχεδιασμός και διαχείριση έργων	Προσαρμογή σε νέες καταστάσεις	Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα	Λήψη αποφάσεων	Σεβασμός στο φυσικό περιβάλλον	Αυτόνομη εργασία	Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου	Ομαδική εργασία	Άσκηση κριτικής και αυτοκριτικής	Εργασία σε διεθνές περιβάλλον	Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης	Εργασία σε διεπιστημονικό περιβάλλον		Παραγωγή νέων ερευνητικών ιδεών	Άλλες...
Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών	Σχεδιασμός και διαχείριση έργων																
Προσαρμογή σε νέες καταστάσεις	Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα																
Λήψη αποφάσεων	Σεβασμός στο φυσικό περιβάλλον																
Αυτόνομη εργασία	Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου																
Ομαδική εργασία	Άσκηση κριτικής και αυτοκριτικής																
Εργασία σε διεθνές περιβάλλον	Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης																
Εργασία σε διεπιστημονικό περιβάλλον																	
Παραγωγή νέων ερευνητικών ιδεών	Άλλες...																
Προσαρμογή σε νέες καταστάσεις Λήψη αποφάσεων Αυτόνομη εργασία Εργασία σε διεπιστημονικό περιβάλλον Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης																	

(3) ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

Το μάθημα εξετάζει την ακόλουθη ερώτηση: Πώς μπορεί ένας πράκτορας (agent) να μάθει να ενεργεί για την επίτευξη κάποιου σκοπού σε ένα δυναμικό, αβέβαιο και μερικώς παρατηρήσιμο περιβάλλον, πιθανώς υπό την παρουσία και άλλων δρώντων πρακτόρων; Το ερώτημα αυτό κι οι ειδικότερες εκδοχές του (ένας πράκτορας, γνώση του περιβάλλοντος, στοχαστικότητα, μάθηση με χρήση πρότερης γνώσης, κοινός σκοπός και συντονισμός ενεργειών των μονάδων) έχει κεντρικό ενδιαφέρον σε ένα μεγάλο φάσμα επιστημονικών περιοχών και εφαρμογών, όπως η υπολογιστική όραση, ρομποτική, αυτόνομη οδήγηση, επεξεργασία φυσικής γλώσσας,

παιχνίδια (επιτραπέζια, παιχνίδια με τράπουλες, βίντεο), και στα ασύρματα συστήματα επικοινωνιών και δίκτυα.

Το μάθημα επικεντρώνει στη θεωρία, τις μεθόδους και τους αλγορίθμους της ενισχυτικής μηχανικής μάθησης (reinforcement learning) που απαντούν στο παραπάνω ερώτημα. Στοχεύοντας σε πλείστες περιπτώσεις όπου η αλληλεπίδραση μεταξύ των πρακτόρων και του περιβάλλοντος στο οποίο αυτοί ενεργούν είναι συνεχής, η βασική θεώρηση του μαθήματος είναι ότι οι πράκτορες παρατηρούν την κατάσταση του περιβάλλοντος μέσω σχετικών μετρήσεων και έμμεσα αντλούν πληροφορίες για τις ενέργειες των άλλων πρακτόρων. Κάποια στοιχεία πληροφορίας είναι κοινά σε όλους τους πράκτορες, άλλα συνιστούν ιδιωτική πληροφορία. Με βάση τις παρατηρήσεις που έχει στη διάθεσή του, κάθε πράκτορας επιλέγει μια ενέργεια και την εκτελεί. Η κατάσταση του περιβάλλοντος μεταβάλλεται χρονικά και εν γένει με αβέβαιο (πιθανοτικό) τρόπο, ανάλογα με τη μέχρι εκείνη τη στιγμή εξέλιξή της και με βάση τις ενέργειες των υπολοίπων πρακτόρων του συστήματος. Οι αποφάσεις κάθε πράκτορα λαμβάνονται σε αλληλέπληλες χρονικές στιγμές και στηρίζονται σε κανόνες προτέρας εκμάθησης, κατά συνέπεια η μηχανική μάθηση είναι ακολουθιακή. Ο κύκλος αυτός επαναλαμβάνεται συνεχώς επιτρέποντας στους πράκτορες να μαθαίνουν το τρόπο ενεργειών τους ώστε να βελτιώνουν, όχι μόνο τη στιγμιαία ανταμοιβή, αλλά το συνολικό μακροπρόθεσμο όφελος/απόδοσή τους.

Οι θεματικές ενότητες του μαθήματος:

- Προβλήματα πολλαπλών ληστών (με ή χωρίς περιεχόμενο) που απλοποιούν τη διαδικασία αποφάσεων.
- Μαρκοβιανές διαδικασίες απόφασης (Markov decision processes - MDPs) με γνωστά (πιθανώς στοχαστικά) μοντέλα συστήματος, υπό τη σκοπιά του δυναμικού προγραμματισμού και των εξισώσεων Bellman, επεκτείνοντας τα υπό μελέτη προβλήματα σε περιπτώσεις μερικώς-παρατηρήσιμων MDPs (partially observable MDPs - POMDPs).
- Αλγόριθμοι εύρεσης βέλτιστων πολιτικών (Q-learning, Monte Carlo/REINFORCE) σε περιπτώσεις απουσίας γνώσης του μοντέλου περιβάλλοντος.
- Αλγόριθμοι βαθιάς ενισχυτικής μηχανικής μάθησης που μαθαίνουν προσεγγιστικά βέλτιστες αποφάσεις με τη χρήση τεχνητών νευρωνικών δικτύων.
- Επέκταση αλγορίθμων βαθιάς ενισχυτικής μηχανικής μάθησης ως «μοντέλα κόσμου» καθώς και σε περιβάλλοντα πολλαπλών πρακτόρων με συναδελφικότητα ή αντιπαλότητα.
- Προγραμματιστικά περιβάλλοντα υλοποίησης αλγορίθμων βαθιάς ενισχυτικής μηχανικής μάθησης σε Python/Pytorch με τη χρήση βιβλιοθηκών Gymnasium και Stable Baselines.

(4) ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ <i>Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.</i>	Στην τάξη (Πρόσωπο με πρόσωπο)	
ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ <i>Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές</i>	Υποστήριξη μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας e-class. Ειδικότερα: Περιγραφή Μαθήματος, Παροχή Υλικού, Ανακοινώσεις, Ημερολόγιο, Ανάθεση και Κατάθεση Εργασιών, Συζητήσεις για Εργασίες, Ερωτηματολόγια, Εξωτερικοί Σύνδεσμοι) Επικοινωνία μέσω ηλεκτρονικού ταχυδρομείου	
ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ <i>Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι διδασκαλίας.</i> <i>Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη & ανάλυση βιβλιογραφίας, Φροντιστήριο, Πρακτική (Τοποθέτηση), Κλινική Άσκηση, Καλλιτεχνικό</i>	Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου
	Διαλέξεις (φυσική παρουσία)	39
	Φροντιστήριο (φυσική παρουσία)	13
	Μία (1) Ατομική Εργασία	78

Εργαστήριο, Διαδραστική διδασκαλία, Εκπαιδευτικές επισκέψεις, Εκπόνηση μελέτης (project), Συγγραφή εργασίας / εργασιών, Καλλιτεχνική δημιουργία, κ.λπ.	Μελέτη για Προφορική Εξέταση	20									
Αναγράφονται οι ώρες μελέτης του φοιτητή για κάθε μαθησιακή δραστηριότητα καθώς και οι ώρες μη καθοδηγούμενης μελέτης σύμφωνα με τις αρχές του ECTS	Σύνολο Μαθήματος	150									
ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ Περιγραφή της διαδικασίας αξιολόγησης Γλώσσα Αξιολόγησης, Μέθοδοι αξιολόγησης, Διαμορφωτική ή Συμπερασματική, Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμών, Επίλυση Προβλημάτων, Γραπτή Εργασία, Έκθεση / Αναφορά, Προφορική Εξέταση, Δημόσια Παρουσίαση, Εργαστηριακή Εργασία, Κλινική Εξέταση Ασθενούς, Καλλιτεχνική Ερμηνεία, Άλλη / Άλλες Αναφέρονται ρητά προσδιορισμένα κριτήρια αξιολόγησης και εάν και που είναι προσβάσιμα από τους φοιτητές.	Οι φοιτητές αξιολογούνται με μία (1) υποχρεωτική ατομική εργασία και μια (1) τελική προφορική εξέταση βασισμένη στην εργασία και την ύλη του μαθήματος. Η εργασία χωρίζεται σε υποερωτήματα που καλύπτουν το σύνολο της ύλης και περιέχει προγραμματιστική υλοποίηση αλγορίθμων και προβλημάτων, καθώς και θεωρητικές ερωτήσεις πάνω στο υλοποιημένο υλικό. Συνοδεύεται από εκτεταμένο κείμενο αναφοράς και σχολιασμού των αποτελεσμάτων. Η προφορική εξέταση είναι συνδεδεμένη με την αναφορά της εργασίας και σκοπός της είναι να καλύψει την κατανόηση του αντικειμένου των υλοποιημένων τμημάτων μέσω περαιτέρω θεωρητικών ερωτήσεων πάνω στην ύλη του μαθήματος. <table border="1" data-bbox="638 835 1474 938"> <thead> <tr> <th>Αξιολόγηση</th><th>Αριθμός</th><th>Ποσοστό</th></tr> </thead> <tbody> <tr> <td>Ατομική Εργασία</td><td>1</td><td>40%</td></tr> <tr> <td>Προφορική εξέταση</td><td>1</td><td>60%</td></tr> </tbody> </table>		Αξιολόγηση	Αριθμός	Ποσοστό	Ατομική Εργασία	1	40%	Προφορική εξέταση	1	60%
Αξιολόγηση	Αριθμός	Ποσοστό									
Ατομική Εργασία	1	40%									
Προφορική εξέταση	1	60%									

(5) ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

1. "Reinforcement Learning: An Introduction," R. S. Sutton and A. G. Barto, MIT Press, Cambridge, MA, 2018.
2. "Dynamic Programming and Optimal Control," D. P. Bertsekas, Athena Scientific, 4th edition, 2007.
3. "Reinforcement Learning: State-of-the-Art," M. Wiering and M. Otterlo, Springer Berlin, Heidelberg, 2012.