

DAMA502 ΠΕΡΙΓΡΑΜΜΑ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ

1. ΓΕΝΙΚΑ

ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ		
ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ	DATA SCIENCE AND MACHINE LEARNING		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΜΕΤΑΠΤΥΧΙΑΚΟ (επίπεδο 7)		
ΚΩΔΙΚΟΣ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ	DAMA502	ΕΞΑΜΗΝΟ	1/2/3
ΤΙΤΛΟΣ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ	Statistics and Optimization		
ΑΥΤΟΤΕΛΕΙΣ ΕΚΠΑΙΔΕΥΤΙΚΕΣ 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 φόρτου και το σύνολο των πιστωτικών μονάδων</i>		ΩΡΕΣ	ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ
Εβδομαδιαίες ώρες φόρτου: 32-33 ώρες x 13 εβδομάδες		420	15 ECTS
ΤΥΠΟΣ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ <i>Υποχρεωτική/Επιλογής/Κατ' επιλογήν υποχρεωτική</i>	Υποχρεωτική/Επιλογής		
ΠΡΟΑΠΑΙΤΟΥΜΕΝΕΣ ΘΕΜΑΤΙΚΕΣ ΕΝΟΤΗΤΕΣ:	Για την παρακολούθηση της DAMA502 δεν απαιτείται η ταυτόχρονη παρακολούθηση ή ολοκλήρωση άλλης ΘΕ του DAMA.		
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:	ΑΓΓΛΙΚΗ		
Η ΘΕΜΑΤΙΚΗ ΕΝΟΤΗΤΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	ΝΑΙ		
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ (URL)	Κάθε ΘΕ έχει τον δικό της χώρο στον ψηφιακό χώρο εκπαίδευσης του ΕΑΠ (http://study.eap.gr , http://courses.eap.gr), με ελεγχόμενη πρόσβαση (χρήση κωδικού) για φοιτητές και διδακτικό προσωπικό.		

2. ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

<p>Μαθησιακά Αποτελέσματα <i>Περιγράφονται τα μαθησιακά αποτελέσματα της ΘΕ, οι συγκεκριμένες γνώσεις, δεξιότητες και ικανότητες καταλλήλου επιπέδου που θα αποκτήσουν οι φοιτητές μετά την επιτυχή ολοκλήρωση της ΘΕ.</i> <i>Συμβουλευτείτε το Παράρτημα Α</i></p> <ul style="list-style-type: none"> • Περιγραφή του Επιπέδου των Μαθησιακών Αποτελεσμάτων για κάθε ένα κύκλο σπουδών σύμφωνα με Πλαίσιο Προσόντων του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης • Περιγραφικοί Δείκτες Επιπέδων 6, 7 & 8 του Ευρωπαϊκού Πλαισίου Προσόντων Διά Βίου Μάθησης και Παράρτημα Β • Περιληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων
<p>Γνώσεις: Με την επιτυχή ολοκλήρωση της θεματικής ενότητας, οι φοιτήτριες/φοιτητές θα μπορούν:</p> <ul style="list-style-type: none"> - Να αναγνωρίζουν ότι βασικοί πυλώνες για τη Μηχανική Μάθηση είναι η θεωρία πιθανοτήτων, η στατιστική και η βελτιστοποίηση, και να εφαρμόζουν αναλυτικά (analytical) και υπολογιστικά (computational) εργαλεία. - Να συνοψίζουν τις ιδιότητες της μονομεταβλητής και πολυμεταβλητής κανονικής κατανομής (Gaussian distribution), να υπολογίζουν τις περιθώριες (marginal) και υπό συνθήκη (conditional) κατανομές, καθώς και μετασχηματισμούς της Gaussian συνάρτησης. - Να εστιάζουν στη διωνυμική (binomial) και Bernoulli κατανομή και να παρουσιάζουν αναλυτικά την Beta κατανομή. - Να συνοψίζουν τις συζυγείς προτεραιότητες (conjugate priors) που συνδέονται διαμέσου

του Θεωρήματος Bayes.

- Να εξηγούν τι είναι οι επαρκείς στατιστικοί δείκτες (sufficient statistics) και να περιγράφουν την εκθετική οικογένεια κατανομών (exponential family).
- Να πραγματοποιούν αλλαγή τυχαίων μεταβλητών (change of random variables) και να βρίσκουν τη νέα συνάρτηση κατανομής.
- Να απαριθμούν βασικές τεχνικές στατιστικής ανάλυσης.
- Να διενεργούν ελέγχους υποθέσεων (hypothesis testing).
- Να εντοπίζουν και να ερμηνεύουν ακραίες τιμές (outliers) σε σύνολα δεδομένων.
- Να ανακαλούν τον τρόπο εύρεσης ελαχίστων για συναρτήσεις μιας μεταβλητής.
- Να συνοψίζουν τη διαδικασία εύρεσης ελαχίστου πολυμεταβλητής συνάρτησης μέσω του αλγορίθμου καθόδου κλίσης (gradient descent).
- Να εξηγούν τη λειτουργία του стоχαστικού αλγορίθμου καθόδου κλίσης (stochastic gradient descent), καθώς και τα πλεονεκτήματα και τους περιορισμούς του σε σύγκριση με τον απλό gradient descent.
- Να περιγράφουν τους πολλαπλασιαστές Lagrange (Lagrange multipliers) και τη χρήση τους σε περιορισμένα προβλήματα βελτιστοποίησης.
- Να παρουσιάζουν την κυρτή βελτιστοποίηση (convex optimization).
- Να χρησιμοποιούν το SageMath για:
 - ο Εύρεση ελαχίστων πολυμεταβλητών συναρτήσεων.
 - ο Ελαχιστοποίηση συναρτήσεων με περιορισμούς.
 - ο Να αξιολογούν την απόδοση μοντέλων (model performance).
 - ο Να εξάγουν χρήσιμες πληροφορίες από την οπτικοποίηση δεδομένων (data visualization).
 - ο Να απεικονίζουν μονοδιάστατα (1D) και δισδιάστατα (2D) δεδομένα χρησιμοποιώντας binning, διαγράμματα πυκνότητας (density plots), διαγράμματα διασποράς (scatter plots) και κουτιά (box plots).

Δεξιότητες:

Με την επιτυχή ολοκλήρωση της θεματικής ενότητας, οι φοιτήτριες/φοιτητές θα μπορούν:

- Να πραγματοποιούν αλλαγές μεταβλητών σε κατανομές πιθανοτήτων και να συναγουν τις αντίστοιχες συναρτήσεις κατανομής.
- Να διενεργούν ελέγχους υποθέσεων με τυποποιημένες στατιστικές μεθόδους.
- Να εντοπίζουν και να ερμηνεύουν ακραίες τιμές (outliers) σε σύνολα δεδομένων.
- Να χρησιμοποιούν το SageMath για την ελαχιστοποίηση πολυμεταβλητών συναρτήσεων και την επίλυση περιορισμένων προβλημάτων βελτιστοποίησης.
- Να αξιολογούν την απόδοση μοντέλων μηχανικής μάθησης χρησιμοποιώντας μετρικές βελτιστοποίησης και στατιστικές μετρικές.
- Να δημιουργούν κατάλληλες οπτικοποιήσεις (π.χ. ιστογράμματα, διαγράμματα πυκνότητας, διαγράμματα διασποράς) για 1D και 2D δεδομένα με τα κατάλληλα εργαλεία.

Ικανότητες:

Με την επιτυχή ολοκλήρωση της θεματικής ενότητας, οι φοιτήτριες/φοιτητές θα μπορούν:

- Να επιλέγουν και να εφαρμόζουν κατάλληλες μεθόδους πιθανοτήτων, στατιστικής και βελτιστοποίησης για την ανάλυση και επίλυση προβλημάτων στη μηχανική μάθηση.
- Να συνδυάζουν στατιστική σκέψη με υπολογιστικά εργαλεία για την ερμηνεία πραγματικών δεδομένων και τη λήψη τεκμηριωμένων αποφάσεων.
- Να αξιολογούν κριτικά την καταλληλότητα αλγορίθμων βελτιστοποίησης (π.χ. SGD vs. GD) για την εκπαίδευση συγκεκριμένων μοντέλων μηχανικής μάθησης.
- Να επιλύουν πρακτικά προβλήματα βελτιστοποίησης με περιορισμούς (π.χ. Lagrange multipliers) στο πλαίσιο της μηχανικής μάθησης.
- Να ερμηνεύουν και να επικοινωνούν τα ευρήματα από οπτικοποιήσεις δεδομένων για την υποστήριξη συμπερασμάτων βασισμένων σε δεδομένα.
- Να εργάζονται αυτόνομα με μαθηματικά λογισμικά (π.χ. SageMath) για την ανάλυση

δεδομένων και την επικύρωση αλγορίθμων μάθησης.

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 η ΘΕ;

Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών

Προσαρμογή σε νέες καταστάσεις

Λήψη αποφάσεων

Αυτόνομη εργασία

Ομαδική εργασία

Εργασία σε διεθνές περιβάλλον

Εργασία σε διεπιστημονικό περιβάλλον

Παράγωγή νέων ερευνητικών ιδεών

Σχεδιασμός και διαχείριση έργων

Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα

Σεβασμός στο φυσικό περιβάλλον

Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας

και ευαισθησίας σε θέματα φύλου

Άσκηση κριτικής και αυτοκριτικής

Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης

Οι γενικές ικανότητες που οι φοιτητές θα αποκτήσουν είναι οι παρακάτω:

- Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών με χρήση των απαραίτητων τεχνολογιών
- Προσαρμοστικότητα σε νέες καταστάσεις
- Λήψη αποφάσεων
- Αυτοδύναμη εργασία
- Ομαδική εργασία
- Σχεδιασμός και διαχείριση έργων
- Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου
- Παραγωγή ελεύθερης, δημιουργικής και επαγωγικής σκέψης

3. ΠΕΡΙΕΧΟΜΕΝΟ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ

Σκοπός της Ενότητας

Οι φοιτητές θα μάθουν τα βασικά μαθηματικά εργαλεία που είναι απαραίτητα για τη Μηχανική Μάθηση (Machine Learning, ML). Αυτά περιλαμβάνουν βασικές έννοιες από τη θεωρία πιθανοτήτων (Probability Theory), την εισαγωγική στατιστική (Introductory Statistics) και την κυρτή βελτιστοποίηση (Convex Optimization). Επιπλέον, ο φοιτητής θα αποκτήσει βασικές τεχνικές οπτικοποίησης (Visualization) μονοδιάστατων (1D) και δισδιάστατων (2D) δεδομένων. Συνολικά, ένας φοιτητής χωρίς προϋπάρχουσες γνώσεις σε αυτούς τους μαθηματικούς τομείς θα δημιουργήσει το υπόβαθρο για να κατανοήσει τεχνικές Μηχανικής Μάθησης, ενώ όσοι έχουν ήδη υπόβαθρο σε μαθηματικές γνώσεις θα μπορούν να εμβαθύνουν περισσότερο στην εφαρμογή των μαθηματικών στη Μηχανική Μάθηση. Η μαθηματική μελέτη θα υποστηρίζεται από υπολογιστικό λογισμικό (computational software) που θα επιτρέπει τόσο αναλυτικές (analytical) όσο και αριθμητικές (numerical) αξιολογήσεις.

Τα κύρια θέματα της ενότητας είναι η Θεωρία Πιθανοτήτων και η Στατιστική (Probability Theory and Statistics), η Κυρτή Βελτιστοποίηση (Convex Optimization) και η Οπτικοποίηση (Visualization).

4. ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ <i>Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.</i>	- Τρόπος παράδοσης: Εξ αποστάσεως εκπαίδευση με τρεις (3) Ομαδικές Συμβουλευτικές Συναντήσεις (ΟΣΣ) διάρκειας 4 ωρών κατά τη διάρκεια του ακαδημαϊκού εξαμήνου, τα Σαββατοκύριακα. - Προσωπική επικοινωνία και ανατροφοδότηση (συμβουλευτικός ρόλος των συνεργαζόμενων διδασκόντων).
ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ <i>Χρήση Τ.Π.Ε. στη Διδασκαλία, στην</i>	Στις Ομαδικές Συμβουλευτικές Συναντήσεις (ΟΣΣ) ή/και κατά τη διδασκαλία χρησιμοποιούνται: - Εργαλεία απομακρυσμένων συναντήσεων (webex, Teams)

<p>Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές</p>	<p>Λογισμικό παρουσίασης (PowerPoint, εκπαιδευτικά βίντεο – animations) - Εξειδικευμένο λογισμικό και βάσεις δεδομένων για τα υπό εκπαίδευση αντικείμενα (π.χ. SageMath κ.ά.)</p> <p>Επιπλέον, οι φοιτητές χρησιμοποιούν εργαλεία αυτοματισμού γραφείου, φυλλομετρητές ιστού (web browsers) και e-reader για ψηφιακά βιβλία.</p> <p>Η επικοινωνία με τους φοιτητές υποστηρίζεται από: - Την ψηφιακή πλατφόρμα του ΕΑΠ (https://courses.eap.gr/login/index.php / https://study.eap.gr/login/index.php) για πληροφορίες ενότητας, αναρτήσεις εκπαιδευτικού υλικού, ανακοινώσεις, μηνύματα, αποτελέσματα εξετάσεων, ομάδες χρηστών, φόρουμ συζητήσεων κ.ά. - Email και μηνύματα (messaging).</p>												
<p>ΟΡΓΑΝΩΣΗ ΕΚΠΑΙΔΕΥΣΗΣ Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι εκπαίδευσης. Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη & ανάλυση βιβλιογραφίας, Φροντιστήριο, Πρακτική (Τοποθέτηση), Κλινική Άσκηση, Καλλιτεχνικό Εργαστήριο, Διαδραστική διδασκαλία, Εκπαιδευτικές επισκέψεις, Εκπόνηση μελέτης (project), Συγγραφή εργασίας / εργασιών, Καλλιτεχνική δημιουργία, κ.λπ.</p> <p>Αναγράφονται 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</p>	<table border="1"> <thead> <tr> <th>Δραστηριότητα</th> <th>Φόρτος Εργασίας Έτους</th> </tr> </thead> <tbody> <tr> <td>3 ΟΣΣ (x 4 ώρες)</td> <td>12</td> </tr> <tr> <td>3 εξαμηνιαίες εργασίες (x 27 ώρες)</td> <td>81</td> </tr> <tr> <td>Ατομική μελέτη (25 ώρες x 13 εβδομάδες)</td> <td>325</td> </tr> <tr> <td>Τελικές Εξετάσεις</td> <td>3</td> </tr> <tr> <td>Σύνολο φόρτου ΘΕ (ώρες)</td> <td>421</td> </tr> </tbody> </table>	Δραστηριότητα	Φόρτος Εργασίας Έτους	3 ΟΣΣ (x 4 ώρες)	12	3 εξαμηνιαίες εργασίες (x 27 ώρες)	81	Ατομική μελέτη (25 ώρες x 13 εβδομάδες)	325	Τελικές Εξετάσεις	3	Σύνολο φόρτου ΘΕ (ώρες)	421
Δραστηριότητα	Φόρτος Εργασίας Έτους												
3 ΟΣΣ (x 4 ώρες)	12												
3 εξαμηνιαίες εργασίες (x 27 ώρες)	81												
Ατομική μελέτη (25 ώρες x 13 εβδομάδες)	325												
Τελικές Εξετάσεις	3												
Σύνολο φόρτου ΘΕ (ώρες)	421												
<p>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ Περιγραφή της διαδικασίας αξιολόγησης</p> <p>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</p> <p>Αναφέρονται ρητά προσδιορισμένα κριτήρια αξιολόγησης και εάν και που είναι προσβάσιμα από τους φοιτητές.</p>	<p>Αξιολόγηση Φοιτητών – Διαμόρφωση Τελικής Βαθμολογίας</p> <p>α. Τρεις (3) Εργασίες Εξαμήνου (A₁, A₂, A₃), η κάθε μία με συντελεστή 10%, συνολικά 30%.</p> <p>Η βαθμολόγηση των εργασιών ενεργοποιείται μόνον εάν ο φοιτητής επιτύχει συνολικό ποσοστό ≥50% στην τελική ή επαναληπτική εξέταση.</p> <p>β. Τελική ή επαναληπτική εξέταση (E) με συντελεστή 70%.</p> <p>Η τελική βαθμολογία της ενότητας υπολογίζεται ως εξής (με μέγιστο βαθμό το 10):</p> <p>Τελική Βαθμολογία = A₁×10% + A₂×10% + A₃×10% + E×70%</p> <p>Γλώσσα αξιολόγησης: Αγγλικά</p>												

5. ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

-Προτεινόμενη Βιβλιογραφία :

- Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. Mathematics for Machine Learning. Cambridge University Press, 2020
- Paul Zimmermann, Alexandre Casamayou, Nathann Cohen, Guillaume Connan, Thierry Dumont, Laurent Fousse, François Maltey, Matthias Meulien, Marc Mezzarobba, Clément Pernet, Nicolas M. Thiéry, Erik Bray, John Cremona, Marcelo Forets, Alexandru Ghitza, Hugh Thomas. Computational Mathematics with SageMath. SIAM, 2018
- Michael R. Berthold, Christian Borgelt, Frank Höppner and Frank Klawonn. (2010). Guide to Intelligent Data Analysis. How to Intelligently Make Sense of Real Data. Springer.
- EMC, E. S., & EMC, E. S. (Eds.). (2015). Data science and big data analytics: Discovering, analyzing, visualizing and presenting data. Wiley.

Επιπλέον ψηφιακό (και πολυμεσικό) υλικό θα διατεθεί online.