

DAMA510 ΠΕΡΙΓΡΑΦΜΑ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ

1. ΓΕΝΙΚΑ

ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ		
ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ	DATA SCIENCE AND MACHINE LEARNING		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΜΕΤΑΠΤΥΧΙΑΚΟ (επίπεδο 7)		
ΚΩΔΙΚΟΣ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ	DAMA510	ΕΞΑΜΗΝΟ	2
ΤΙΤΛΟΣ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ	Machine Learning		
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		
ΠΡΟΑΠΑΙΤΟΥΜΕΝΕΣ ΘΕΜΑΤΙΚΕΣ ΕΝΟΤΗΤΕΣ:	Για την παρακολούθηση της DAMA510 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</p> <p>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.</p> <p>Συμβουλευτείτε το Παράρτημα Α</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:</p> <p>Με την επιτυχή ολοκλήρωση της θεματικής ενότητας, οι φοιτήτριες/φοιτητές θα μπορούν:</p> <ul style="list-style-type: none"> - Εξηγούν τις βασικές φάσεις της διαδικασίας Επιστήμης Δεδομένων (data science process) και τον ρόλο του επιστήμονα δεδομένων (data scientist). <p>Δεξιότητες:</p> <p>Με την επιτυχή ολοκλήρωση της θεματικής ενότητας, οι φοιτήτριες/φοιτητές θα μπορούν:</p> <ul style="list-style-type: none"> - Αξιολογούν την ποιότητα και τα χαρακτηριστικά των εισαγόμενων δεδομένων, όπως τύπους δεδομένων (data types), ελλείπουσες τιμές (missing values) και ακραίων τιμών (outliers). - Εφαρμόζουν τεχνικές προεπεξεργασίας δεδομένων (data preprocessing), όπως καθαρισμό (data cleaning), μετασχηματισμό (transformation) και κλιμάκωση χαρακτηριστικών (feature scaling), χρησιμοποιώντας κατάλληλα εργαλεία και γλώσσες.

- Πραγματοποιούν μείωση διαστάσεων (dimensionality reduction) για τη μείωση της πολυπλοκότητας σε δεδομένα υψηλής διάστασης.
- Υπολογίζουν μέτρα ομοιότητας και απόστασης (similarity and distance measures) για αριθμητικά και κατηγορικά γνωρίσματα.
- Εφαρμόζουν αλγόριθμους συσταδοποίησης (clustering algorithms) για την ανακάλυψη ομάδων σε μη επισημασμένα δεδομένα.
- Εφαρμόζουν εξόρυξη συχνών στοιχειοσυνόλων (frequent itemset mining) και μάθηση κανόνων συσχέτισης (association rule learning) για την εξαγωγή προτύπων από συναλλακτικά δεδομένα.
- Εφαρμόζουν μοντέλα παλινδρόμησης (regression) και ταξινόμησης (classification) σε επισημασμένα σύνολα δεδομένων.
- Αναλύουν τα χαρακτηριστικά ενός συνόλου δεδομένων και το προετοιμάζουν για επιβλεπόμενη μάθηση, αντιμετωπίζοντας την ανισοροπία (imbalance), επιλέγοντας γνωρίσματα και κωδικοποιώντας μεταβλητές.
- Επιλέγουν σχετικά γνωρίσματα με τεχνικές μείωσης διαστάσεων και επιλογής χαρακτηριστικών (feature selection) για τη βελτίωση της προγνωστικής ακρίβειας.
- Εφαρμόζουν μάθηση με Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines) και Νευρωνικά Δίκτυα (neural networks) σε εργασίες ταξινόμησης, αναλύοντας την απόδοσή τους και τη συμπεριφορά τους.

Ικανότητες:

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

- Επικοινωνούν αποτελεσματικά τα ευρήματα και τα αποτελέσματα, χρησιμοποιώντας κατάλληλα εργαλεία οπτικοποίησης (visualization tools).
- Αξιολογούν την καταλληλότητα διαφορετικών παραδειγμάτων συσταδοποίησης (clustering paradigms) για συγκεκριμένα προβλήματα.
- Αξιολογούν μοντέλα κανόνων συσχέτισης βάσει μετρικών επικύρωσης και συνάφειας με το αντικείμενο (domain relevance).
- Αξιολογούν την απόδοση μοντέλων (model performance) χρησιμοποιώντας ποικιλία μετρικών.

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
- Προσαρμογή σε νέες καταστάσεις
- Λήψη αποφάσεων
- Αυτοδύναμη εργασία
- Ομαδική εργασία
- Σχεδιασμός και διαχείριση έργων
- Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου
- Παράγωγή ελεύθερης, δημιουργικής και επαγωγικής σκέψης

3. ΠΕΡΙΕΧΟΜΕΝΟ ΘΕΜΑΤΙΚΗΣ ΕΝΟΤΗΤΑΣ

Σκοπός της Ενότητας

Οι φοιτητές θα αποκτήσουν στέρεο υπόβαθρο σχετικά με τους αλγοριθμικούς άξονες και τις υπολογιστικές απαιτήσεις των βασικών προσεγγίσεων στην Επιστήμη Δεδομένων (Data Science) και τη Μηχανική Μάθηση (Machine Learning). Θα μάθουν θεμελιώδεις έννοιες και αρχές που διέπουν τις τεχνικές εξαγωγής γνώσης από δεδομένα, θα εξοικειωθούν με πρακτικές παραμέτρους που αφορούν την ανάλυση και την ερμηνεία των δεδομένων, την αξιολόγηση της ποιότητας των εισαγόμενων δεδομένων και την εξαγωγή συμπερασμάτων από τα αποτελέσματα της εξόρυξης. Με την ολοκλήρωση της ενότητας θα μπορούν να εφαρμόζουν τη θεωρία και να χρησιμοποιούν γλώσσες, αλγορίθμους και εργαλεία για την επίλυση προβλημάτων του πραγματικού κόσμου, καθώς και να ερμηνεύουν και να επικοινωνούν τα ευρήματα σε κάθε είδους κοινό.

Περιεχόμενο Θεματικής Ενότητας

Προεπεξεργασία δεδομένων (Data preprocessing), Μηχανική χαρακτηριστικών (Feature engineering), Ανίχνευση εκτός τιμών (Outlier detection), Μείωση διαστάσεων (Dimensionality reduction), Συσταδοποίηση (Clustering), Εξόρυξη συχνών συνόλων αντικειμένων (Frequent itemsets), Κανόνες συσχέτισης (Association rules), Δέντρα αποφάσεων (Decision Trees), Παλινδρόμηση (Regression), Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines), Νευρωνικά Δίκτυα (Neural Networks)

4. ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

<p>ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ <i>Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.</i></p>	<p>- 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).</p>				
<p>ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ <i>Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές</i></p>	<p>Στις Ομαδικές Συμβουλευτικές Συναντήσεις (ΟΣΣ) ή/και κατά τη διδασκαλία χρησιμοποιούνται: - Εργαλεία απομακρυσμένων συναντήσεων (webex, Teams) Λογισμικό παρουσίασης (PowerPoint, εκπαιδευτικά βίντεο – animations) - Εξειδικευμένο λογισμικό και βάσεις δεδομένων για τα υπό εκπαίδευση αντικείμενα (π.χ. SageMath κ.ά.) Επιπλέον, οι φοιτητές χρησιμοποιούν εργαλεία αυτοματισμού γραφείου, φυλλομετρητές ιστού (web browsers) και e-reader για ψηφιακά βιβλία. 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 <i>Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι εκπαίδευσης. Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη & ανάλυση</i></p>	<table border="1"> <thead> <tr> <th data-bbox="695 1787 1117 1854">Δραστηριότητα</th> <th data-bbox="1125 1787 1352 1854">Φόρτος Εργασίας Έτους</th> </tr> </thead> <tbody> <tr> <td data-bbox="695 1860 1117 1894">3 ΟΣΣ (x 4 ώρες)</td> <td data-bbox="1125 1860 1352 1894">12</td> </tr> </tbody> </table>	Δραστηριότητα	Φόρτος Εργασίας Έτους	3 ΟΣΣ (x 4 ώρες)	12
Δραστηριότητα	Φόρτος Εργασίας Έτους				
3 ΟΣΣ (x 4 ώρες)	12				

<p>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>	5 Εκπαιδευτικές δραστηριότητες (x 4 ώρες)	20
	2 εξαμηνιαίες εργασίες (x 30 ώρες)	60
	Ατομική μελέτη (25 ώρες x 13 εβδομάδες)	325
	Τελικές Εξετάσεις	3
	Σύνολο φόρτου ΘΕ (ώρες)	420
<p>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ Περιγραφή της διαδικασίας αξιολόγησης</p> <p>Γλώσσα Αξιολόγησης, Μέθοδοι αξιολόγησης, Διαμορφωτική ή Συμπερασματική, Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, Επίλυση Προβλημάτων, Γραπτή Εργασία, Εκθεση / 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 Ενότητας</p> <p>α. Πέντε (5) Εκπαιδευτικές Δραστηριότητες τύπου quiz πολλαπλής επιλογής (Q), οι οποίες συνεισφέρουν εξίσου στον τελικό βαθμό με ποσοστό 2 % η καθεμία.</p> <p>β. Δύο (2) Εργασίες Εξαμήνου (A), οι οποίες συνεισφέρουν εξίσου στον τελικό βαθμό με ποσοστό 10 % η καθεμία.</p> <p>Η βαθμολόγηση των εργασιών ενεργοποιείται μόνον εάν ο φοιτητής επιτύχει συνολικό ποσοστό $\geq 50\%$ στην τελική ή επαναληπτική εξέταση.</p> <p>γ. Τελική ή επαναληπτική εξέταση (E) με συντελεστή 70%.</p> <p>Η τελική βαθμολογία της ενότητας υπολογίζεται ως εξής (με μέγιστο βαθμό το 10):</p> $\text{Τελικός Βαθμός} = (Q_1 \times 2 \%) + (Q_2 \times 2 \%) + (Q_3 \times 2 \%) + (Q_4 \times 2 \%) + (Q_5 \times 2 \%) + (A_1 \times 10 \%) + (A_2 \times 10 \%) + (E \times 70 \%)$ <p>Γλώσσα αξιολόγησης: Αγγλικά.</p>	

5. ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

- Προτεινόμενη Βιβλιογραφία :

- 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.
- W. N. Venables, D. M. Smith and the R Core Team (2022). An Introduction to R.
- H. Wickham, M. Cetinkaya-Rundel, G. Grolemund (2023). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data (2nd edition). O' Reilly Media.

Επιπλέον ψηφιακό (και πολυμεσικό) υλικό θα διατεθεί online.

- Related scientific Journals:

- Journal of Machine Learning Research (<http://www.jmlr.org>)

- Machine Learning (<https://www.springer.com/journal/10994>)