**ΠΕΡΙΓΡΑΜΜΑ ΜΑΘΗΜΑΤΟΣ**

1. **ΓΕΝΙΚΑ**

|  |  |
| --- | --- |
| **ΣΧΟΛΗ** | ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗΣ |
| **ΤΜΗΜΑ** | ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ |
| **ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ**  | ΠΡΟΠΤΥΧΙΑΚΟ |
| **ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ** | ΣΑΘΕΜΕΔ-17 | **ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ** | **8ο**  |
| **ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ** | ΘΕΜΑΤΑ ΕΠΙΣΤΗΜΗΣ ΔΕΔΟΜΕΝΩΝ |
| **ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ** *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* | **ΕΒΔΟΜΑΔΙΑΙΕΣΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ** | **ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ** |
|  | 3 | 6 |
|  |  |  |
|  |  |  |
| *Προσθέστε σειρές αν χρειαστεί. Η οργάνωση διδασκαλίας και οι διδακτικές μέθοδοι που χρησιμοποιούνται περιγράφονται αναλυτικά στο (δ).* |  |  |
| **ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ***γενικού υποβάθρου, ειδικού υποβάθρου, ειδίκευσης* *γενικών γνώσεων, ανάπτυξης δεξιοτήτων* | Ειδίκευσης γενικών γνώσεων |
| **ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ:** | Η ομαλή παρακολούθηση του μαθήματος απαιτεί βασικές γνώσεις απειροστικού λογισμού, γραμμικής άλγεβρας, στατιστικής, προγραμματισμού υπολογιστών, και διαχείρισης δεδομένων. Γνώσεις αριθμητικής ανάλυσης και επιχειρησιακής έρευνας (βελτιστοποίησης) είναι επιθυμητές. |
| **ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:** | ΕΛΛΗΝΙΚΗ |
| **ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS**  | ΟΧΙ |
| **ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)** | https://eclass.unipi.gr/courses/SAE203/ |

1. **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*
 |
| Η επιστήμη δεδομένων (ΕΔ) συνίσταται στη συνέργεια της διαχείρισης δεδομένων, της μηχανικής μάθησης και της εξόρυξης δεδομένων και αποσκοπεί στην ανάλυση (μεγάλου όγκου) δεδομένων, για τα οποία δεν διατίθεται (ακριβές) υπόδειγμα, με σκοπό την εξαγωγή πληροφορίας/γνώσης απ’ αυτά. Τα δεδομένα μπορεί να είναι αδόμητα ή δομημένα, ελλιπή, χρονικά μεταβαλλόμενα, ετερογενή ή/και ασαφή. Το μάθημα αυτό αποσκοπεί στο να εισάγει το φοιτητή σ’ αυτήν την νεοφανή επιστήμη, παρέχοντας τις βασικές γνώσεις σε θέματα προπαρασκευής και διαχείρισης δεδομένων με τα παραπάνω χαρακτηριστικά, μεθόδων εξόρυξης γνώσης απ’ αυτά (π.χ. εξόρυξης κανόνων συσχέτισης) και διαχείρισης συνόλων δεδομένων μεγάλου όγκου. Το θεωρητικό-αλγοριθμικό μέρος του μαθήματος θα συνοδεύεται από πρακτική εξάσκηση στο εργαστήριο, με τη χρήση μοντέρνων εργαλείων λογισμικού για ΕΔ.Με την επιτυχή ολοκλήρωση του μαθήματος, οι φοιτητές θα μπορούν να:α) Κατονομάζουν και κατανοούν το ρόλο και τα οφέλη της εφαρμογής της ΕΔ σε μοντέρνες πρακτικές εφαρμογέςβ) Κατονομάζουν και κατανοούν τις κύριες κατηγορίες προβλημάτων και είναι σε θέση να εφαρμόζουν αντιπροσωπευτικές μεθόδους επίλυσηςγ) Κατανοούν τις δυσκολίες που προκύπτουν σε δεδομένα μεγάλης κλίμακας και τρόπους αντιμετώπισής τουςδ) Εργάζονται ανεξάρτητα με μοντέρνα εργαλεία λογισμικού για να επιλύουν προβλήματα ΕΔ. |
| **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;.* |
| *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**……**Άλλες…**…….* |
| Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών. Αυτόνομη εργασία. Σχεδιασμός και διαχείριση έργων.Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης. |

1. **ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ**

|  |
| --- |
| * **Εισαγωγή στην επιστήμη δεδομένων** (1 βδομάδα)
* **Eξόρυξη γνώσης από δεδομένα** (3 βδομάδες)
	+ Κατηγοριοποίηση: k-nearest neighbor (k-NN), δένδρα απόφασης
	+ Συσταδοποίηση: διαμεριστικοί αλγόριθμοι, αλγόριθμοι βάσει πυκνότητας, online/αυξητικοί αλγόριθμοι, αλγόριθμοι για μεγάλες βάσεις δεδομένων, αλγόριθμοι για κατηγορικά δεδομένα
	+ Εξόρυξη κανόνων συσχέτισης: αλγόριθμοι εύρεσης συχνών item-sets (Apriori, αλγόριθμοι διαμερισμού, δειγματοληψίας)
* **Θέματα διαχείρισης μεγάλων βάσεων δεδομένων** (5 βδομάδες)
	+ Παράλληλα/κατανεμημένα/in-memory συστήματα διαχείρισης βάσεων δεδομένων)
	+ Δεδομένα στο υπολογιστικό νέφος
	+ Ρεύματα δεδομένων
	+ Το κατανεμημένο σύστημα αρχείων Hadoop, δημιουργία αντιγράφων, ανοχή σε σφάλματα, υψηλός βαθμός ρυθμαπόδοσης στην ανάγνωση.
	+ Το Apache Hadoop ως μια υλοποίηση του MapReduce.
	+ Περιορισμοί του Hadoop.
	+ Σχεδίαση εργασιών MapReduce.
	+ Παράλληλη επεξεργασία, επεξεργασία στην κύρια μνήμη, τα Dataframes στο Spark, αποθήκευση κατά στήλες και κατά γραμμές, παράδειγμα χρήσης.
	+ Resilient Distributed Datasets (RDDs)
	+ Σύγκριση μεταξύ Spark και Hadoop.
	+ Δηλωτική επεξεργασία επερωτήσεων, η Spark SQL, προγραμματισμός με Dataframes, η μηχανή επεξεργασίας του Spark, κατάτμηση δεδομένων, δουλεύοντας με δεδομένα JSON.

**Εργαστήρια με χρήση Python (scikit-learn), SPARK** (4 βδομάδες) |

1. **ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ**

|  |  |
| --- | --- |
| **ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ***Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.* | Διαλέξεις στην αίθουσα και στο εργαστήριο |
| **ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ***Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές* | e-classSlidesΕργαστηριακή εκπαίδευση |
| **ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ***Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι διδασκαλίας.**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.**Αναγράφονται οι ώρες μελέτης του φοιτητή για κάθε μαθησιακή δραστηριότητα καθώς και οι ώρες μη καθοδηγούμενης μελέτης σύμφωνα με τις αρχές του ECTS* |

|  |  |
| --- | --- |
| ***Δραστηριότητα*** | ***Φόρτος Εργασίας Εξαμήνου*** |
| Διαλέξεις | 39 |
| Εργαστηριακή άσκηση | 10 |
| Φροντιστήριο | 3 |
| Μελέτη & ανάλυση βιβλιογραφίας | 54 |
| Εκπόνηση μελέτης | 24 |
| Συγγραφή εργασίας | 20 |
|  |  |
|  |  |
|  |  |
| Σύνολο Μαθήματος  | **150** |

 |
| **ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ** *Περιγραφή της διαδικασίας αξιολόγησης**Γλώσσα Αξιολόγησης, Μέθοδοι αξιολόγησης, Διαμορφωτική ή Συμπερασματική, Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, 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 πού είναι προσβάσιμα στους φοιτητές.* | Γλώσσα αξιολόγησης: ΕλληνικήΗ αξιολόγηση περιλαμβάνει:* Εργασία με εξέταση της σχετικής έκθεσης/αναφοράς (40%)
* Γραπτή εξέταση (60%), που περιλαμβάνει ερωτήσεις σύντομης απάντησης και επίλυση προβλημάτων.

Οι φοιτητές ενημερώνονται για τα κριτήρια αξιολόγησης προφορικά και μέσω eclass. |
|  |  |

1. **ΣΥΝΙΣΤΩΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ**

|  |
| --- |
| Προτεινόμενη Βιβλιογραφία:1. G. Joel (2021). Επιστήμη Δεδομένων: Βασικές Αρχές και Εφαρμογές με

Python, 2η έκδοση. Παπασωτηρίου.1. H. Wickam, G. Grolemund (2022). Προγραμματισμός σε R για την

Επιστήμη των Δεδομένων. Κλειδάριθμος.1. Β. Βερύκιος, Β. Καγκλής, Η. Σταυρόπουλος (2015). Η Επιστήμη των

Δεδομένων μέσα από τη Γλώσσα R. [ηλεκτρ. βιβλ.] Αθήνα: ΣύνδεσμοςΕλληνικών Ακαδημαϊκών Βιβλιοθηκών.1. M. J. Zaki, W. Meira, Jr. (2017). Εξόρυξη και Ανάλυση Δεδομένων:

Βασικές Έννοιες και Αλγόριθμοι. Κλειδάριθμος. |