

# Εξεταστέα ύλη μαθήματος «Διαχείριση Μεγάλων Δεδομένων» Ακαδημαϊκό Έτος 2021-2022

Η εξεταστέα ύλη του μαθήματος «Διαχείριση Μεγάλων Δεδομένων» για το ακαδημαϊκό έτος 2020-2021 έχει ως εξής:

- 1) Όλες οι σημειώσεις και πηγαίοι κώδικες σε R και Python διαθέσιμες στις ενότητες LAB-R Εγκατάσταση και παραδείγματα χρήσης, LAB-Python Εγκατάσταση και παραδείγματα χρήσης, Lecture 1 έως και Lecture 6 στην ιστοσελίδα του μαθήματος στο eclass ( <https://eclass.upatras.gr/courses/ECON1332/> ) .  
ΣΗΜ: Από το σύνολο διαφανειών στο αρχείο “Lecture4.1.AlternativeClassificationMethods.pdf” στην ενότητα «Lecture 4: Classification» εντός ύλης είναι **ΜΟΝΟ** οι αλγόριθμοι:
  - i) *k*-Nearest Neighbors (διαφάνειες#25 έως και #73)
  - ii) *Naïve Bayes* (διαφάνειες από #78 έως και τη διαφάνεια #108).
- 2) Τα ζητούμενα στις 3 (τρεις) εργασίες που έχετε εκπονήσει.
- 3) Από το βιβλίο Βερούκιος, Σταυρόπουλος, Κοτσιαντής, Τζαγκαράκης: *Η ΕΠΙΣΤΗΜΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ Βασικές Αρχές, Θεωρία, & Εφαρμογές με τη Γλώσσα R*, Εκδόσεις Νέων Τεχνολογιών, ISBN-13: 978-960-578-043-2, 2019, οι εξής ενότητες:

**Κεφάλαιο 3:** Ολόκληρο.

**Κεφάλαιο 4:** Μόνοι ενότητα 4.6

**Κεφάλαιο 6:** Ολόκληρο (εκτός της ενότητας 6.7.1)

**Κεφάλαιο 7:** Από 7.1 έως και 7.3

**Κεφάλαιο 8:** Από 8.1 έως και 8.4.1 και ενότητα 8.5

**Κεφάλαιο 9:** Από 9.1 έως και 9.3

- 4) Από το βιβλίο *Rang-Ning Tan, Michael Steinbach, Vipin Kumar: “Εισαγωγή στην Εξόρυξη Δεδομένων”*, Εκδόσεις ΤΖΙΟΛΑ, 2010 (ελληνική έκδοση), οι εξής ενότητες:

**Κεφάλαιο 1:** Ολόκληρο

**Κεφάλαιο 2:** Ολόκληρο

**Κεφάλαιο 3:** Ολόκληρο

**Κεφάλαιο 4:** Ολόκληρο εκτός των ενότητων: 4.5 και 4.6

**Κεφάλαιο 6:** Από 6.1 έως και 6.4, 6.7 (εκτός 6.7.2 και 6.7.3)

**Κεφάλαιο 8:** Από 8.1 έως και 8.3