

# Αλγοριθμική Επιστήμη Δεδομένων

Εισαγωγή και διαδικαστικά

Άνοιξη 2021

Πρόγραμμα Διδακτορικών Σπουδών ΣΗΜΜΥ ΕΜΠ  
ΔΠΜΣ Επιστήμη Δεδομένων και Μηχανική Μάθηση (ε.δε.μ<sup>2</sup>) ΕΜΠ  
ΔΠΜΣ Εφαρμοσμένες Μαθηματικές Επιστήμες (ΕΜΕ) ΕΜΠ  
ΔΠΜΣ Αλγόριθμοι, Λογική και Διακριτά Μαθηματικά (ΑΛΜΑ) ΕΚΠΑ-ΕΜΠ

---

# Περιεχόμενο μαθήματος

- Εντοπισμός κανόνων συσχέτισης και συνόλων στοιχείων που εμφανίζονται συχνά σε μεγάλα δεδομένα.
- Ροές δεδομένων, δειγματοληψία και sketching σε ροές δεδομένων.
- Εντοπισμός παρόμοιων στοιχείων - locality sensitive hashing.
- Πολυδιάστατοι χώροι και μείωση διάστασης, τυχαίες προβολές και Johnson-Lindenstrauss λήμμα.
- Recommendation systems, singular value decomposition, εφαρμογές.

# Περιεχόμενο μαθήματος

- Το web ως γράφημα και ως αλυσίδα Markov, ανάλυση συνδέσμων, PageRank.
- Ανάλυση κοινωνικών δικτύων, ανίχνευση κοινοτήτων.
- Συσταδοποίηση δεδομένων (clustering), k-means, spectral clustering.
- Ηλεκτρονική διαφήμιση, δημοπρασίες, Generalized Second Price.
- Ασφάλεια και ιδιωτικότητα δεδομένων, differential privacy.

# Προαπαιτούμενα

- **Αλγόριθμοι και πολυπλοκότητα:** ανάλυση χρονικής πολυπλοκότητας αλγορίθμων, βασικές δομές δεδομένων, βασικοί αλγόριθμοι γράφων, NP-πληρότητα
- **Μαθηματικά εργαλεία:** πιθανότητες, concentration inequalities, Markov chains, γραμμική άλγεβρα, θεωρία γραφημάτων

# Διαδικαστικά

- ▶ Διδάσκοντες: [Α. Παγουρτζής](#), [Δ. Σούλιου](#), [Β. Νάκος](#)
- ▶ Βοηθοί διδασκαλίας: [Σ. Πετσαλάκης](#)
- ▶ Επικοινωνία: [ads@corelab.ntua.gr](mailto:ads@corelab.ntua.gr)
  
- ▶ Βαθμολογία (περυσινό σχήμα):
  - ▶ 8 μον. από ΤΔ («βάση»: 3)
  - ▶ 3 μον. από ασκήσεις
  
  - ▶ φέτος: παρουσιάσεις φοιτητών; προγραμματιστική άσκηση;

# Βιβλιογραφία

- ▶ **Mining of Massive Datasets**, Jure Leskovec, Anand Rajaraman, and Jeff Ullman
  - ▶ [www.mmds.org](http://www.mmds.org) (book, slides, videos, etc.)
- ▶ **Foundations of Data Science**, Avrim Blum, John Hopcroft, and Ravindran Kannan
  - ▶ <https://www.cs.cornell.edu/jeh/book.pdf>
- ▶ .... και πολλά ακόμη βιβλία, άρθρα και ιστοσελίδες!