Στατιστική Μηχανική Μάθηση *(Statistical Machine Learning)*

ΔΙΔΑΣΚΟΝΤΕΣ: Ι.ΚΟΣΜΙΔΗΣ – Ξ.ΠΕΝΤΕΛΗ

Γενικά Στοιχεία Μαθήματος

Κωδικός: 61232
Τύπος: Υποχρεωτικό Ομάδας Μαθημάτων 2
Επίπεδο: Μεταπτυχιακό
Έτος σπουδών: Α'
Εξάμηνο σπουδών: 2ο
ECTS: 7,5
Γλώσσα διδασκαλίας: Αγγλική

Περιεχόμενο Μαθήματος

Το μάθημα περιλαμβάνει ένα μεγάλος εύρος μεθόδων Στατιστικής και Μηχανικής Μάθησης για επιβλεπόμενη (supervised) και μη επιβλεπόμενη (unsupervised) μάθηση. Συγκεκριμένα θα διδαχθούν μέθοδοι μη επιβλεπόμενης μάθησης που αφορούν την ομαδοποίηση (clustering) όπως, η έννοια της απόστασης, ιεραρχική ομαδοποίηση, ο αλγόριθμος K-means και οι παραλλαγές τους, ομαδοποίηση με τη χρήση μοντέλων και άλλες μέθοδοι όπως επίσης και μέθοδοι μείωσης διάστασης. Μέθοδοι αξιολόγησης της ομαδοποίησης, επιλογή μεταβλητών. Από τις μεθόδους επιβλεπόμενης μάθησης θα διδαχθούν μέθοδοι ταξινόμησης όπως, διακριτική ανάλυση, μέθοδος κοντινότερων γειτόνων, δέντρα αποφάσεων και τυχαία δέντρα, μηχανές διανυσμάτων στήριξης, αφελής Bayes ταξινομητής και άλλοι. Μέθοδοι cross-validation. Μέθοδοι για μεγάλο όγκο δεδομένων, Κανονικοποιήση. Στατιστικά μοντέλα για δίκτυα. Τεχνικές εξομάλυνσης στην παλινδρόμηση.

Προαπαιτούμενα

Στατιστική συμπερασματολογία

Επιδιωκόμενα Μαθησιακά Αποτελέσματα

Μετά την επιτυχή ολοκλήρωση του μαθήματος ο φοιτητής θα διαθέτει τη γνώση και την τεχνική ικανότητα

- να υλοποιήσει τις μεθοδολογίες που απαντούν σε προβλήματα ταξινόμησης,
- να υλοποιήσει τις μεθοδολογίες που απαντούν σε προβλήματα μείωσης της διάστασης του προβλήματος
- να υλοποιήσει τις μεθοδολογίες που απαντούν σε προβλήματα ομαδοποίησης.
- να ερμηνεύει τα αποτελέσματα και να αξιολογεί τις διάφορες μεθόδους μεταξύ τους.
- να αντιμετωπίζει προβλήματα εξαιτίας του μεγάλου όγκου δεδομένων.

Συνιστώμενη Βιβλιογραφία

- T. Hastie, R. Tibshirani and R. Friedman (2009) Elements of Statistical Learning, Springer.
- J. Witten, J, Hastie, T. and Tibshirani, R. (2011) Introduction to Statistical Learning with applications in R, Springer
- C. Giraud (2015). Introduction to High-Dimensional Statistics. Philadelphia: Chapman and Hall/CRC.
- E. D. Kolaczyk (2014) Statistical Analysis of Network Data with R. Springer

Διδακτικές και Μαθησιακές Μέθοδοι

Δια ζώσης διδασκαλία, η οποία καλύπτει θεωρία και πρακτική εξάσκηση. Η πρακτική εξάσκηση γίνεται με τη βοήθεια της R.

Μέθοδοι Αξιολόγησης και Βαθμολόγησης

Γραπτή εξέταση (80%) και εργασίες (20%) .