## Εξόρυξη δεδομένων και αλγόριθμοι μάθησης

**Εαρινό εξάμηνο 2019  
Project Python**

Στην παρούσα εργασία θα χρησιμοποιήσουμε

1. Το εργαλείο [Anaconda](https://www.anaconda.com/what-is-anaconda/) που μας επιτρέπει να διαχειριζόμαστε packages και περιβάλλοντα (environments) ώστε να χρησιμοποιήσουμε την Python. Εφόσον εγκαταστήσετε το εργαλείο, θα χρησιμοποιήσετε το Anaconda Navigator ώστε να δημιουργήστε ένα περιβάλλον python έκδοσης 3.5 όπου και θα εγκαταστήσετε τις σχετικές βιβλιοθήκες που χρειάζονται τα project.
2. Στο αντίστοιχο περιβάλλον θα εγκαταστήσετε το [Jupyter Notebook](http://jupyter.org/) που είναι μια web εφαρμογή που μας επιτρέπει να συνδυάσουμε text, code και visualizations σε ένα αρχείο. Τα τρία (3) projects είναι σε αρχεία Jupyter Notebooks.

## Παραδοτέα

Ως παραδοτέα της παρούσας εργασίας ορίζονται τα αντίστοιχα τρία (3) αρχεία Jupyter Notebook

* 1\_Decision Tree.ipynb
* 2\_Naive Bayes.ipynb
* 3\_K-Means.ipynb

με συμπληρωμένο τον κώδικα και τα σχόλια σας ώστε να επιλύεται το κάθε ξεχωριστό πρόβλημα.

## Προφορική εξέταση

Η εξέταση θα αφορά την κατανόηση της θεωρίας ως προς τα

1. Decision Trees (supervised learning)
2. Naïve Bayes Theorem (supervised learning)
3. K-Means (unsupervised learning)

καθώς και του υλικού που αναφέρεται εντός των Notebooks.

Οι εργασίες έχουν επιλεγεί με γνώμονα τις τρέχουσες τάσεις του Machine Learning κι αποτελούν μια εισαγωγή της χρήσης της Python σε επίλυση συναφών προβλημάτων. Η προσωπική ενασχόληση είναι απαραίτητη για την αξιολόγηση σας. Θετικά θα αξιολογηθεί οποιαδήποτε βελτιστοποίηση των λύσεων.

Για οποιαδήποτε απορία σας μπορείτε να επικοινωνήσετε στο [sergiang@ceid.upatras.gr](mailto:sergiang@ceid.upatras.gr) .

Καλή επιτυχία!!!