

## Νευρωνικά Δίκτυα, 8<sup>ο</sup> εξάμηνο (Neural Networks)

### Εργασία #2

Έστω το σύνολο δεδομένων plants<sup>1</sup> που μπορείτε να βρείτε στη βιβλιοθήκη UCI Machine Learning Repository<sup>2</sup>. Το σύνολο δεδομένων αφορά φυτά και σε ποιες πολιτείες των ΗΠΑ αυτά συναντώνται.

Πραγματοποιείστε ομαδοποίηση (clustering) πάνω στα δεδομένα χρησιμοποιώντας τον αλγόριθμο k-means και αξιολογείστε τα αποτελέσματα που πήρατε χρησιμοποιώντας διάφορες μετρικές.

Για την υλοποίηση της εργασίας μπορείτε να χρησιμοποιήσετε τη γλώσσα προγραμματισμού python και τη βιβλιοθήκη scikit-learn<sup>3</sup>.

Η υποβολή μπορεί να γίνει μέσω google drive. Θα πρέπει να ανεβάσετε την εργασία σας σε κάποιο φάκελο και να στείλετε link με δικαιώματα ανάγνωσης στον διδάσκοντα (στο [yrefanid@uom.edu.gr](mailto:yrefanid@uom.edu.gr)).

Η εργασία βαθμολογείται μέχρι έναν (1) βαθμό.

---

<sup>1</sup> <http://archive.ics.uci.edu/ml/datasets/Plants>

<sup>2</sup> <http://archive.ics.uci.edu/ml/index.php>

<sup>3</sup> <https://scikit-learn.org/stable/>