






Εργασία 8

(Προθεσμία: 7 Φεβρουαρίου 2021)

UNSUPERVISED CLUSTERING

Το Fisher's IRIS data set (δες http://en.wikipedia.org/wiki/Iris_flower_data_set) περιέχει μετρήσεις της μορφής: (μήκος σέπαλου, πλάτος σέπαλου, μήκος πετάλου, πλάτος πετάλου) σε cm για 150 φυτά iris (είδος κρίνου, αγριόκρino). Από αυτά τα 150 φυτά, 50 είναι Iris Setosa (ω_1), 50 είναι Iris Versicolour (ω_2) και 50 είναι Iris Virginica (ω_3). Γνωρίζουμε ότι μόνο η μία (Iris Setosa) από τις άλλες δυο κλάσεις είναι γραμμικά διαχωρίσιμη και ότι η Versicolour με την Virginica «περιπλέκονται» μεταξύ τους αρκετά.

Iris Setosa	Iris Versicolour	Iris Virginica
		

1. Χρησιμοποιώντας τον αλγόριθμο K-means προσπαθήστε να ταξινομήσετε τα φυτά σε 3 ομάδες. Ποιό είναι το λάθος που βρίσκετε?
2. Χρησιμοποιώντας τον αλγόριθμο fuzzy C-means προσπαθήστε να ταξινομήσετε τα φυτά σε 3 ομάδες. Ποιό είναι το λάθος που βρίσκετε?
3. Χρησιμοποιώντας τον αλγόριθμο ISODATA προσπαθήστε να ταξινομήσετε τα φυτά σε ομάδες. Πόσες ομάδες βρίσκετε? Ποιό είναι το λάθος που βρίσκετε?

Ξάνθη, 16/12/2020