

Υπολογιστική Νοημοσύνη

Εργαστηριακές Ασκήσεις ακαδημαϊκού έτους 2022-23

Διδάσκων καθηγητής: Α. Λύκας

Μέλη ομάδας:

Όνομ/πώνυμο	ΑΜ	Email
Θεοδωρίδης Χαράλαμπος	4674	cs04674@uoi.gr
Μπονίτσης Παντελής	4742	cs04742@uoi.gr
Σιδηρόπουλος Γεώργιος	4789	cs04789@uoi.gr

Άσκηση 2

Πρόγραμμα ομαδοποίησης (ΠΟ) με M ομάδες βασισμένο στον αλγόριθμο k-means.

Γενικές πληροφορίες

- Κατάλογος άσκησης : Exercise2
- Εντολή μεταγλώττισης : « javac *.java »
- Εντολή εκτέλεσης : « java Engine »
- Γλώσσα Προγραμματισμού : Java
- Το πρόγραμμα αποτελείται συνολικά από τέσσερις (4) κλάσεις και ένα (1) πρόγραμμα σε γλώσσα Python.
 - Κλάση FileConstructor : Υλοποιεί την δημιουργία δύο (2) αρχείων txt. Το ένα περιέχει τα παραδείγματα (συντεταγμένες σημείων) και το άλλο τις συντεταγμένες των κέντρων.
 - Κλάση FileReader : Υλοποιεί το διάβασμα του αρχείου txt που περιέχει τα παραδείγματα.
 - Κλάση Algorithm : Υλοποιεί τα βήματα και τις λειτουργίες που απαιτούνται για τον αλγόριθμο k-means.
 - Κλάση Engine : Υλοποιεί τις 15 εκτελέσεις του αλγορίθμου k-means.

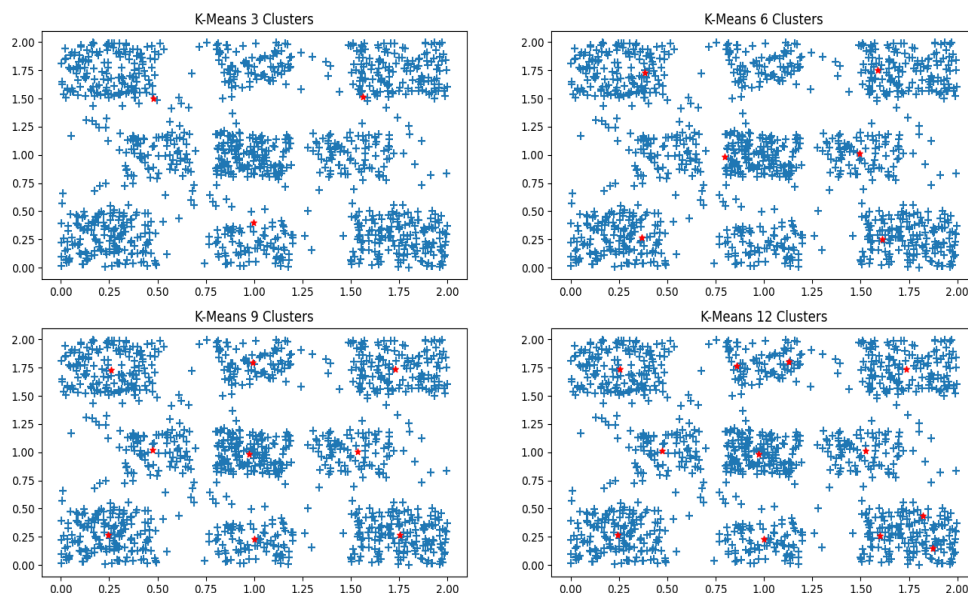
- Πρόγραμμα PlotConstructor.py : Υλοποιεί την δημιουργία του σχήματος που περιέχει τα παραδείγματα και τις θέσεις των κέντρων στο τέλος της εκτέλεσης του βασικού προγράμματος.

Ερώτημα α)

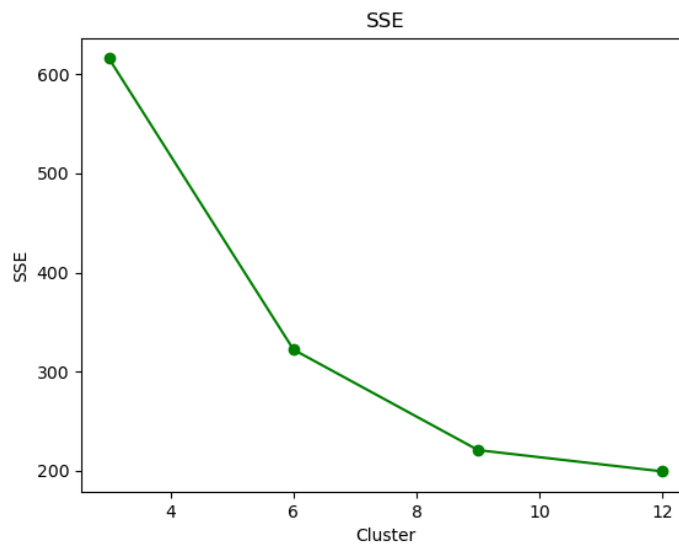
Τα 15 τρεξίματα του αλγορίθμου k-means με τυχαία επιλεγμένα αρχικά κέντρα καθώς και η αποθήκευση της λύσης με το μικρότερο σφάλμα ομαδοποίησης υλοποιείται μέσα στην κλάση Engine και συγκεκριμένα στις γραμμές 28 – 35. Η επιλογή των τυχαίων αρχικών κέντρων υλοποιείται από την μέθοδο randomCentroid() (κλάση Algorithm στην γραμμή 48) και η εκτέλεση του αλγορίθμου k-means πραγματοποιείται από την μέθοδο kMeans() (κλάση Algorithm στην γραμμή 206). Τέλος, ο υπολογισμός του σφάλματος ομαδοποίησης πραγματοποιείται από την συνάρτηση calculateError() (κλάση Algorithm στην γραμμή 192) .

Ερώτημα β)

Παρακάτω παρουσιάζονται τα σχήματα που περιέχουν τα παραδείγματα και τις θέσεις των κέντρων ανάλογα με τον αριθμό των ομάδων (3, 6, 9, 12).



Παρακάτω παρουσιάζεται το διάγραμμα που δείχνει πώς μεταβάλλεται το σφάλμα ομαδοποίησης με βάση τον αριθμό των ομάδων (3, 6, 9, 12).



- Το σφάλμα ομαδοποίησης δεν μπορεί χρησιμοποιηθεί για να εκτιμηθεί ο πραγματικός αριθμός των ομάδων διότι στην περίπτωση που χρησιμοποιηθεί αρκετά μεγάλος αριθμός ομάδων το σφάλμα θα προσεγγίσει το μηδέν.