

Data Mining Lab 3: Graph Spectra

Gibson Chikafa
chikafa@kth.se

Akhil Yerrapragada
akhily@kth.se

1 Assignment Goals

In this assignment we implement the spectral clustering algorithm as described in [1].

2 Solution overview

We implement the solution in Python programming language. We use the *networkx* library for creation, manipulation and study of structure of our graphs. We use the *numpy* library for the matrix operations of our graphs e.g computing eigenvalues and matrices. For clustering with Kmeans we use *sklearn*.

3 Results

Figure 1 and Figure 2 shows results of unclustered and clustered graphs for *example1.dat* file respectively. As seen the graphs is clustered into 4 partitions which agrees to the data i.e the original graph consists of 4 disjoint sets.

Figure 3 and Figure 4 shows results of unclustered and clustered graphs for *example2.dat* file respectively. The original graph has been clustered into 4 clusters.

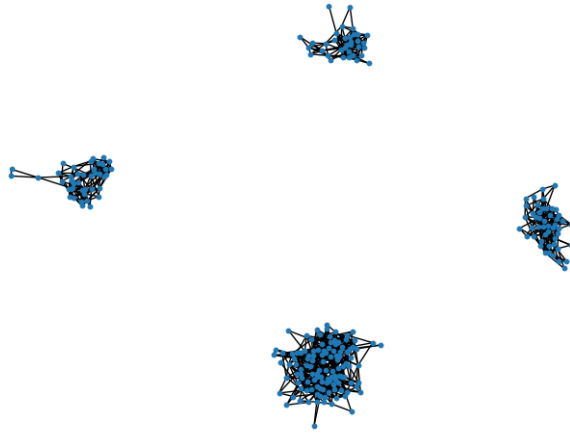


Figure 1: Unclustered graph for example1.dat

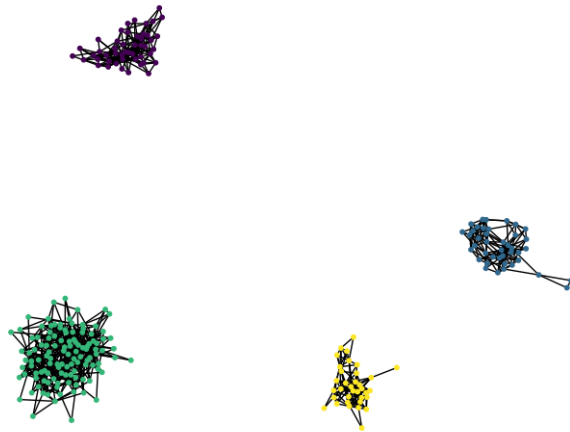


Figure 2: Clustered graph for example1.dat

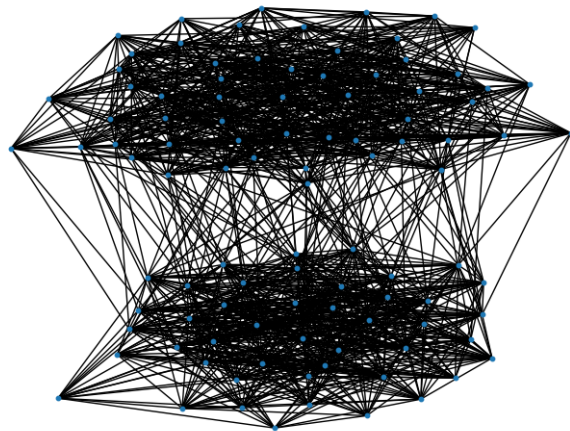


Figure 3: Unclustered graph for example2.dat

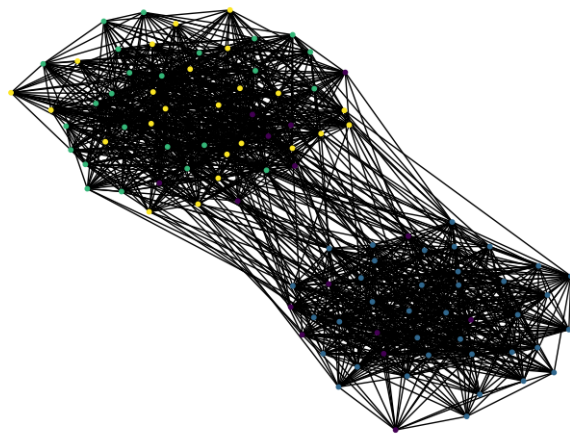


Figure 4: Clustered graph for example2.dat

References

- [1] A.Y. Ng, Michael Jordan, and Y Weiss. On spectral clustering: Analysis and an algorithm. volume 2, 11 2001.