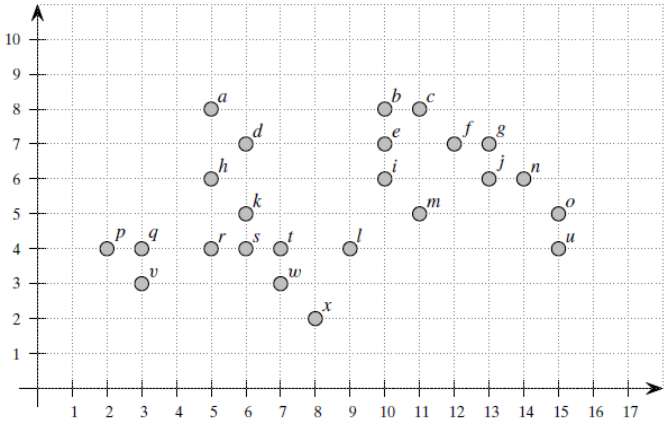
Alexandria University
Faculty of Engineering
Computer and Communication Program

Due: Friday 21st of May 2021
CCE: Pattern Recognition

Sheet #4 Clustering Evaluation

1. Using K-means: set K=2,3,4,5,6. Report different clustering results.
2. K-ways normalized: cut k=2,3,4,5,6. Use Similarity graph as the {3,5}-NN graph. Where Sim (xi, xj) = 1 iff xj is one of the nearest three points to xi (or vise-versa). Report different clustering results.
3. Assume the ground truth clustering results is T1= {p, q, v},
   T2 ={a, d, h, k, r, s, t, l, w, x} and T3={b, c, e, i, m, f, g, j, n, a, u}.
   i. Compute the external measures we studied such as
      1. Conditional Entropy
      2. Purity
      3. Pairwise measures (Jaccard and Rand index)
      4. Max matching when number of clusters =3.
      5. F-Measure
   ii. Compute the internal measures we studied. You will need the proximity matrix before proceeding.
      1. BetaCV
      2. Normalized-Cut



Notes

- Make sure you did everything on your own.
- Marks are put on trial and effort not on correct answer, so cheating will be severely penalized.
- Deliver the solution in PDF format on the email:

  patternssp2021@gmail.com with subject [Sheet4][ID]

Prof. Dr. Marwan Torki

Eng. Abdelrahman Wael
Eng. Ismail El-Yamany