Alexandria University
Faculty of Engineering
Computer and Communications Program

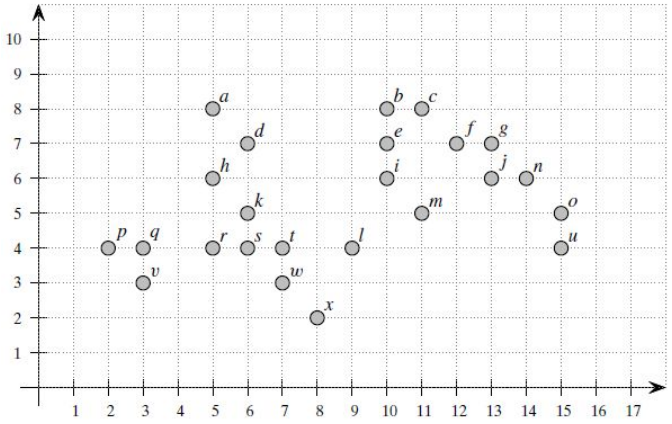
Due : Sunday 17/3/2019

CCE :Pattern Recognition

Sheet#4 Clustering Evaluation

Submit a report Report is essential. Detailed steps are required. Final answers will not be marked.

1. **Perform clustering on the following data**
   a. Using Kmeans: set K=2,3,4,5,6. Report different clustering results.
   b. K-ways normalized: cut k=2,3,4,5,6
      i. Use RBF kernel with gamma = {0.01,0.1}. Report the Report different clustering results.
      ii. 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.
   c. 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
   d. Summarize your finding using graphs, tables and comment on what you obtain
      i. Compare parameter setting for every algorithm
      ii. Compare between different algorithms results



Dr.Marwan Torki

Eng. Mai Ibrahim
Eng.Nada Osman