

**Instructor: Dr. Nizar Bouguila**

1. Provide an R or python implementation of the expectation maximization algorithm for learning multivariate Gaussian mixture models. Provide also two testing examples (a Data set and a final result), one where you supposed that you know the number of clusters in advance and one where you determine the number of clusters via MDL and AIC. You should discuss the obtained results.
2. Explain and discuss the DBSCAN approach. Implement the DBSCAN technique proposed in <https://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=50DAF4112C4EC341A0B98A770A669A2C?doi=10.1.1.121.9220&rep=rep1&type=pdf>  
Provide also a testing example (a Data set and a final result).