Machine Learning

**Question 1.** Consider a 1D dataset with four examples: -3, -1, 2, 4. By hand, apply \( k \)-means clustering until convergence, assuming the initial cluster centres are -4 and 0. For each iteration give the assignment of examples to clusters and the new values of the cluster centres.

**Question 2.** The \( k \)-means algorithm will in general converge to a local optimum rather than a global one. Given this, how would you adapt the algorithm to increase the chances of finding a good solution?

**Question 3.** Sketch a dataset on which \( k \)-means would work poorly but a Gaussian Mixture Model with the same number of clusters would do well. Describe why \( k \)-means would not work well.

**Question 4.** Describe a Gaussian Mixture Model and the Expectation-Maximisation algorithm used to train it.

**Question 5.** Explain how the number of components/clusters in a Gaussian Mixture Model can affect over/under fitting of the data. Describe how to use cross-validation to select the number of components/clusters.