Week 8 Exercise: Unsupervised Learning

Note: An indicative mark is in front of each question. The total mark is 10. You may mark your own work when we release the solutions.

[1] 1. We have a 24-bit colour image of size 100×100. How many possible images of this size and bit depth?

[3] 2. An alternative to derive PCA is to minimize the reconstruction error (Slide 26) for all N data samples $\mathbf{x}^{(i)}, i = 1, \cdots, N$, assuming that the mean $\boldsymbol{\mu} = \sum_i \mathbf{x}^{(i)}$ is zero. Take this approach to derive the first principal component (as the first eigenvector of the data matrix).

[1] 3. In k-means clustering, how could we determine k if it is not given?

[2] 4. For the graph on Slide 39, compute the normalised cut $Ncut(A, B)$.

[3] 5. In spectral clustering, show that the smallest eigenvalue for the formulated generalized eigenvalue problem on Slide 41 is 0 with the corresponding generalized eigenvector $\mathbf{y} = \mathbf{1}$, hence the same "representation/embedding" for all nodes.