

MACHINE LEARNING FOUNDATIONS AND APPLICATIONS

Assignment 2

Due Date: 6 November, 2019

[Instructions: Please show all the steps and calculations with proper explanation. Numerical accuracy is less important than methodological accuracy. You can submit through either Moodle or Piazza. You can work out on paper and submit scanned copies, or you can directly work out using LATEX. The submission filename should be <roll number>.pdf. Copying will result in 0 marks.]

1. You are given a set of 10 data-points (2D) below. Attempt to partition them into 2 clusters using a) K-means, b) Kernelized K-means using any Kernel function you want (except linear kernel). Show 3 complete iterations in each case. **[5 + 5 = 10 marks]**

$[(-1,1), (5,-6), (0,-1), (-4,4), (1,0), (0,-2), (6,4), (-6,5), (-2,-1), (-5,-7)]$

2. Derive maximum likelihood estimate for the parameters of a) Geometric Distribution, b) Binomial distribution, c) Multinomial distribution. In each case, show all the calculations precisely. **[3+3+4=10 marks]**

3. Derive MAP estimate for the parameters of the following distributions: a) Poisson (using Gamma as prior), b) Categorical (using Dirichlet as prior), c) Multivariate Gaussian with known covariance matrix, with multivariate normal as prior for the mean vector. In each case, show all the calculations precisely. **[5+5+10=20 marks]**

4. You have the following 1D observations, and you wish to fit a Gaussian Mixture Model on it. Looking at the data, decide how many Gaussian components you want to use. Then estimate the Gaussian parameters using E-M algorithm. **[2+8=10 marks]**

$[2.3, 4.7, -5.5, -4.8, 9.1, 3.5, 10.4, -4.3, 11.2, 1.9, 10.8, 3.4]$