Pattern Recognition and Machine Learning, January 22, 2020

Name:

- 1. (5 points) Describe the K-means algorithm (1pt). Prove that K-means converges to a local minimum (2pts). Describe the K-means++ algorithm (1pt). Give an example when employing K-means++ leads, on average, to better solution than K-means (1pt).
- 2. (5 points) Formulate the logistic regression problem (1 point). Derive the gradient descent learning method for this task (2.5 points).
 Given the training set T = {(-1, -1), (-2, -1), (0, 1), (2, 1)}, starting point w = (0, 0), and an initial step size of 1, do the first iteration of gradient descent. (1.5 points).
- 3. (5 points) Parameter Estimation. The reliability of light bulbs is defined by the probability density function

$$p(t) = \frac{1}{\theta \left(\frac{t}{\theta} + 1\right)^2}, \ t \in [0, \infty],$$

where $p(t)\delta$ is the probability that a bulb fails in a short time interval $(t, t + \delta)$. That is, the probability P(T) that the bulb fails in interval (0, T), is

 $P(T) = \int_0^T p(t) dt = 1 - \frac{1}{(\frac{T}{\theta} + 1)}$. What is the maximum likelihood estimate of the reliability parameter θ if in an experiment with two bulbs the following lifetimes have been observed: $t_1 = 1.0, t_2 = 4.0$? (4pts)

For this setting, what is the probability that a light bulb is still OK at time 10? (1pt)

- 4. (5 points) Decision tree classifiers.
 - (a) (2 points) Describe the simplest top-down tree-building algorithm that selects the decision rule at a node by maximization of information gain.
 - (b) (3 points) Discuss advantages and disadvantages of decision trees.