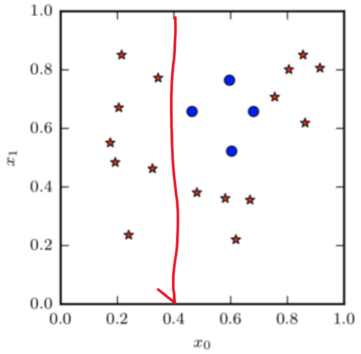
End Sem, CSE 542, Winter 2020
Duration: 2 hours, Total marks # 20

Instructions: Write all necessary steps that are needed. Submit scanned copy by 1pm.

Q1. Assume a binary classification setup with stars and bubbles as two classes shown in figure below.



A cut is made at $x_0 = 0.4$ as shown by the red line. Compute Gini index for the left region and the right region. [3]

Q2. Compute out of bag error for following. Suppose you are fitting a function to data distributed according to sin(x). Let the training samples be D = {(0, sin(0)), (45, sin(45)), (90, sin(90)), (225, sin(225)), (270, sin(270))}. Suppose you try to apply bagging to learn decision tree.
Let there be two sampled datasets from D. D is in form (x,y) where y is the label for x.
D1 = {(0, sin(0)), (45, sin(45)), (90, sin(90)), (225, sin(225)), (225, sin(225))}
D2 = {(0, sin(0)), (0, sin(0)), (90, sin(90)), (225, sin(225)), (225, sin(225))}
For both D1 and D2, let the learnt decision stump make a cut at x = 91. Compute OOB error for both D1 and D2. [3]

Q3. Consider following 2-D samples {(-1,-1), (-2,-2), (0,0), (2,2)}. Assuming that k-means with k=2 is to be performed and initialized with cluster centers as {(-1,-1),(0,0)}, show:

a. Which points are assigned to which cluster? You must show all the steps [2]
b. After assignment of points to cluster, compute cluster mean, within cluster covariance and between cluster covariance. [3]

Q4. Consider the following training data
x = {-1, 0, 2} and f(x) = {-4, 1, 4}. Apply Gaussian process regression to find mean and variance at a test sample z=1. Assume a Gaussian kernel of the form
$k(x,z) = \sigma^2 \exp\left[-\frac{(x-z)^2}{2l^2}\right]$, where $\sigma = l = 1$. [3]

Q5. The PCA objective formulation is given by

$$\max_w w^T S w \;\text{ s.t } w^T w = 1$$

Show that the solution to this formulation is the eigenvector of S with maximum eigenvalue. Assume standard definition of symbols. [2]

Q6. Assuming W to be weights and X to be input, let the output of a neuron be $\sigma(W^T X)$, where σ denotes sigmoid. Give an expression for X in terms of W such that the $\sigma(W^T X)$ attains the maximum possible value. Also assume that $X^T X = 1$. [2]

Q7. Arrange in the increasing order of performance

a. Single decision tree
b. Bagged decision tree
c. Random forest
d. Give reasons. [2]