CSE 250B: Section 4 - Sharad Vikram

1. Nearest Neighbor Practice

   You're given the following (X, y) pairs: $(1.5, 1), (2.5, 1), (1, 0), (2, 0), (3, 0)$

   (a) What is the 1-nearest neighbor decision rule?

   (b) **True/False**: The 1-nearest neighbor algorithm is consistent.

   (c) Name two methods that can speed up nearest neighbor queries.

2. Statistical Learning Theory

   Let $\eta(x) = \begin{cases} 0.4 & \text{if } x < 0 \\ 0.9 & \text{if } x \geq 0 \end{cases}$, and $\mu(x) = \begin{cases} 0 & \text{if } x < -1 \\ 0.2 & \text{if } -1 \leq x < 0 \\ 0.8 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x > 1 \end{cases}$.

   (a) What is the decision rule $h(x)$?

   (b) What is the Bayes risk?

   (c) We define a cost sensitive risk function: $R(h) = 5 \times Pr(Y = 0, h(X) = 1) + P(Y = 1, h(X) = 0)$. What is our decision rule with this cost function?

3. Distance Function

   Under what conditions of S is the following function a valid distance metric?

   $$d(x, y) = (x - y)^T S(x - y)$$

4. Covariance

   Let X and Y be random variables.

   (a) Show that if X is of independent Y, $\text{Cov}(X, Y) = 0$.

   (b) Show that if $\text{Cov}(X, Y) = 0$, X is not necessarily independent of Y.

5. Find the spectral decomposition of the following matrices.

   (a) $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

   (b) $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$