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?

   > **Solution:** $h(x) = \begin{cases} 0 & \text{if } x \leq 1.25, 1.75 \leq x < 2.25, 2.75 \leq x \\ 1 & \text{otherwise} \end{cases}$

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

   > **Solution:** False.

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

   > **Solution:**
   > 1. KD-tree
   > 2. Locality-Sensitive Hashing

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 $h(x)$?

   > **Solution:** $h(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$

   (b) What is the Bayes risk?

   > **Solution:**
   >
   > $$\begin{aligned} h^* &= \int_{-1}^{-1} P(\text{error}|x)P(x)dx \\ &= \int_{-1}^{0} (0.4 \times 0.2)dx + \int_{-1}^{0} (0.1 \times 0.8)dx \\ &= 0.08 + 0.08 = 0.16 \end{aligned}$$

(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?

Solution: We decide class 1 when $R(1|x) > R(0|x)$.
When $x < 0$, $5 \times P(Y = 0|x) = 5 \times 0.6 = 3.0$ and $P(Y = 1|x) = 0.4$. When $x \geq 0$, $5 \times P(Y = 0|x) = 5 \times 0.1 = 0.5$ and $P(Y = 1|x) = 0.9$. We get the same decision function:

$$h(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

3. Distance Function

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

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

Solution: S is PSD.

4. Covariance

Let X and Y be random variables.

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

Solution:

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

When X and Y are independent, $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= 0 \end{aligned}$$

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

Solution: Let X be a random variable that takes on values -1 and 1 with probabilities 0.5. Let $Y = 2X$. Clearly these variables are not independent.

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] \end{aligned}$$

XY is also a random variable that has mean 0. Therefore, the covariance is zero even though the variables are correlated.

5. Find the spectral decomposition of the following matrices.

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

Solution:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

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

Solution:

$$\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$