# ESE 5420 Homework 4

Mohammed Raza Syed - Penn ID: 37486255

## Problem 1

### b) Written Questions

#### (i) Random splitting better than sequential splitting in our case

Random splitting is preferred over sequential splitting because it ensures that the model is trained and evaluated on a representative and unbiased dataset. Sequential splitting may lead to biased subsets if the data has an inherent order. The reasons are as follows:

- Representative Distribution: Random splitting ensures that both the training and test sets have a similar distribution, reducing the risk of bias or skewness.
- Avoids Bias: Sequential splitting can introduce bias if the data is ordered (e.g., temporally or by features), which might result in poor generalization to unseen data.
- Breaks Sequential Patterns: Random splitting prevents the model from overfitting to patterns specific to the order of the data, such as trends in time-ordered datasets.
- Improves Generalization: By mixing the data randomly, the model learns to generalize better, as the test set provides a more accurate representation of unseen scenarios.

#### (ii) Derive the classification rule for the threshold 0.5

To classify using logistic regression with a threshold of 0.5, we derive the rule as follows:

The logistic regression model outputs the probability $P(Y = 1|X)$, given by:

$$P(Y = 1|X) = \frac{1}{1 + \exp(-(\beta_0 + \beta^T X))}.$$

The classification rule is to classify as $+1$ if $P(Y = 1|X) \geq 0.5$. Substituting the formula:

$$\frac{1}{1 + \exp(-(\beta_0 + \beta^T X))} \geq 0.5.$$

Simplify this inequality:

$$\begin{aligned}
1 &\geq 0.5 \cdot (1 + \exp(-(\beta_0 + \beta^T X))) \\
1 &\geq 0.5 + 0.5 \exp(-(\beta_0 + \beta^T X)) \\
0.5 &\geq 0.5 \exp(-(\beta_0 + \beta^T X)) \\
1 &\geq \exp(-(\beta_0 + \beta^T X)).
\end{aligned}$$

Take the natural logarithm of both sides:

$$\ln(1) \geq -(\beta_0 + \beta^T X).$$

Since $\ln(1) = 0$:

$$0 \geq -(\beta_0 + \beta^T X).$$

Multiply through by $-1$, reversing the inequality:

$$\beta_0 + \beta^T X \geq 0.$$

Thus, the final classification rule is:

$$\hat{y} = \begin{cases} +1 & \text{if } \beta_0 + \beta^T X \geq 0, \\ -1 & \text{otherwise.} \end{cases}$$

# Problem 2

## a) Designing $P(Y = y|X = x)$

To satisfy the given requirements, we need to design the conditional probability $P(Y = y|X = x)$ as follows:

- $X$ is uniformly distributed on the interval $[0, 1]$.
- $Y$ is a binary variable, taking values 0 or 1.
- The conditions are:
  - $P(Y = 0) = P(Y = 1) = 0.5$ (balanced classes),
  - Maximum achievable accuracy of any classifier should be less than or equal to 0.9,
  - The Bayes optimal classifier accuracy should be at least 0.8.

To meet these conditions, we define $P(Y = 1|X = x)$ as a piecewise function:

$$P(Y = 1|X = x) = \begin{cases} 0.9 & \text{if } x > 0.5 \\ 0.1 & \text{if } x \leq 0.5 \end{cases}$$

$$P(Y = 0|X = x) = 1 - P(Y = 1|X = x) = \begin{cases} 0.1 & \text{if } x > 0.5 \\ 0.9 & \text{if } x \leq 0.5 \end{cases}$$

**Verification of Requirements**

- **Balanced Classes:** Since $X$ is uniformly distributed, we calculate the overall probability of each class $Y$:

$$P(Y = 1) = 0.9 \cdot 0.5 + 0.1 \cdot 0.5 = 0.5$$

$$P(Y = 0) = 0.1 \cdot 0.5 + 0.9 \cdot 0.5 = 0.5$$

  Thus, $P(Y = 0) = P(Y = 1) = 0.5$, satisfying the balanced classes requirement.

- **Maximum Classifier Accuracy:** For any given $x$, the maximum probability of $Y = y$ is 0.9, so no classifier can achieve an accuracy greater than 0.9.

- **Bayes Optimal Classifier Accuracy:** The Bayes optimal classifier selects the most probable class for each $x$. Thus:

  - For $x > 0.5$, it predicts $Y = 1$ with probability 0.9,
  - For $x \leq 0.5$, it predicts $Y = 0$ with probability 0.9.

  The overall accuracy of the Bayes optimal classifier is:

$$0.9 \cdot 0.5 + 0.9 \cdot 0.5 = 0.9$$

  which is greater than the minimum required accuracy of 0.8.

Thus, the final design for $P(Y = y|X = x)$ is:

$$P(Y = 1|X = x) = \begin{cases} 0.9 & \text{if } x > 0.5 \\ 0.1 & \text{if } x \leq 0.5 \end{cases}$$

$$P(Y = 0|X = x) = \begin{cases} 0.1 & \text{if } x > 0.5 \\ 0.9 & \text{if } x \leq 0.5 \end{cases}$$

This design meets all the specified requirements.

# Problem 4

## a) Write down the expected regret $E[R_T]$ of the algorithm in terms of $E[T_2]$ and $\Delta$

The regret is the difference between the total reward if the optimal arm (arm 1) was always pulled and the reward obtained by the algorithm. Let:

$T_1$ be the number of times arm 1 is pulled, and $T_2$ be the number of times arm 2 is pulled.

The total reward of the algorithm is:

$$\text{Total Reward} = T_1\mu_1 + T_2\mu_2.$$

The reward from always pulling the optimal arm (arm 1) is:

$$\text{Optimal Reward} = T\mu_1,$$

where $T$ is the total number of rounds.

Thus, the regret is:

$$R_T = \text{Optimal Reward} - \text{Total Reward}.$$

Substitute the rewards:

$$R_T = T\mu_1 - (T_1\mu_1 + T_2\mu_2).$$

Simplify:

$$R_T = (T - T_1)\mu_1 - T_2\mu_2.$$

Since $T_1 + T_2 = T$, this reduces to:

$$R_T = T_2(\mu_1 - \mu_2).$$

Let $\Delta = \mu_1 - \mu_2$. Then:

$$R_T = T_2\Delta.$$

Taking the expected value:

$$\mathbf{E}[\mathbf{R_T}] = \mathbf{E}[\mathbf{T_2}] \cdot \boldsymbol{\Delta}.$$

## b) Use the Hoeffding inequality to show that $P[\hat{\mu}_2 > \hat{\mu}_1] \leq 2e^{-m\Delta^2/2}$

We want to bound the probability that $\hat{\mu}_2 > \hat{\mu}_1$, which can be written as:

$$P[\hat{\mu}_2 > \hat{\mu}_1] = P[\hat{\mu}_2 - \hat{\mu}_1 > 0].$$

**Step 1: Decompose $\hat{\mu}_2 - \hat{\mu}_1$:**

Let $\hat{\mu}_2$ and $\hat{\mu}_1$ represent the empirical mean estimates of arms 2 and 1, respectively. Then:

$$\hat{\mu}_2 - \hat{\mu}_1 = (\hat{\mu}_2 - \mu_2) + (\mu_2 - \mu_1) + (\mu_1 - \hat{\mu}_1).$$

Substitute $\mu_2 - \mu_1 = -\Delta$:

$$\hat{\mu}_2 - \hat{\mu}_1 = (\hat{\mu}_2 - \mu_2) - \Delta + (\mu_1 - \hat{\mu}_1).$$

The event $\hat{\mu}_2 > \hat{\mu}_1$ implies:

$$(\hat{\mu}_2 - \mu_2) + (\mu_1 - \hat{\mu}_1) > \Delta.$$

**Step 2: Bound Each Term Using Hoeffding's Inequality:**

Hoeffding's inequality states that for independent random variables $X_i \in [0,1]$, the sample mean $\hat{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$ satisfies:

$$P[|\hat{X} - E[X]| \geq \epsilon] \leq 2e^{-2m\epsilon^2}.$$

Apply Hoeffding's inequality to:

- $P[\hat{\mu}_2 - \mu_2 > \Delta/2]$:
  Here, the deviation $\epsilon$ is bounded by $\Delta/2$, so:

$$P[\hat{\mu}_2 - \mu_2 > \Delta/2] \leq e^{-2m(\Delta/2)^2}.$$

- $P[\mu_1 - \hat{\mu}_1 > \Delta/2]$:
  Similarly, the deviation $\epsilon$ is also $\Delta/2$, so:

$$P[\mu_1 - \hat{\mu}_1 > \Delta/2] \leq e^{-2m(\Delta/2)^2}.$$

**Step 3: Combine Bounds:**

The event $(\hat{\mu}_2 - \mu_2) + (\mu_1 - \hat{\mu}_1) > \Delta$ requires either $\hat{\mu}_2 - \mu_2 > \Delta/2$ or $\mu_1 - \hat{\mu}_1 > \Delta/2$. Using the union bound:

$$P[\hat{\mu}_2 > \hat{\mu}_1] \leq P[\hat{\mu}_2 - \mu_2 > \Delta/2] + P[\mu_1 - \hat{\mu}_1 > \Delta/2].$$

Substitute the bounds:

$$P[\hat{\mu}_2 > \hat{\mu}_1] \leq e^{-2m(\Delta/2)^2} + e^{-2m(\Delta/2)^2}.$$

Simplify:

$$\mathbf{P}[\hat{\boldsymbol{\mu}}_\mathbf{2} > \hat{\boldsymbol{\mu}}_\mathbf{1}] \leq \mathbf{2e^{-m\Delta^2/2}}.$$

## c) Use part (b) to show that $E[T_2] \leq m + 2(T - 2m)e^{-m\Delta^2/2}$

The algorithm pulls arm 2 during:

- **The exploration phase:** Both arms are pulled $m$ times.
- **The exploitation phase:** The arm with the higher estimated mean is pulled for $T - 2m$ rounds. Arm 2 is pulled in this phase only if $\hat{\mu}_2 > \hat{\mu}_1$.

Thus:

$$E[T_2] = m + E[\text{Exploitation Phase Pulls for Arm 2}].$$

From part (b), the probability $P[\hat{\mu}_2 > \hat{\mu}_1]$ is bounded by:

$$P[\hat{\mu}_2 > \hat{\mu}_1] \leq 2e^{-m\Delta^2/2}.$$

During the $T-2m$ exploitation rounds, the expected number of pulls for arm 2 is:

$$E[\text{Exploitation Phase Pulls}] = (T-2m)\cdot P[\hat{\mu}_2 > \hat{\mu}_1].$$

Substitute the bound for $P[\hat{\mu}_2 > \hat{\mu}_1]$:

$$E[\text{Exploitation Phase Pulls}] \le (T-2m)\cdot 2e^{-m\Delta^2/2}.$$

Combine with the exploration phase:

$$\mathbf{E[T_2] \le m + 2(T-2m)e^{-m\Delta^2/2}}.$$

## d) Using the bounds, show that $E[R_T] \le m\Delta + 2T\Delta e^{-m\Delta^2/2}$

From part (a), the expected regret is:

$$E[R_T] = E[T_2]\cdot\Delta.$$

Substitute the bound for $E[T_2]$ from part (c):

$$E[T_2] \le m + 2(T-2m)e^{-m\Delta^2/2}.$$

This gives:

$$E[R_T] \le \Delta\cdot\left(m + 2(T-2m)e^{-m\Delta^2/2}\right).$$

Distribute $\Delta$:

$$E[R_T] \le m\Delta + 2\Delta(T-2m)e^{-m\Delta^2/2}.$$

Simplify:

$$E[R_T] \le m\Delta + 2T\Delta e^{-m\Delta^2/2} - 4m\Delta e^{-m\Delta^2/2}.$$

Since $2T\Delta e^{-m\Delta^2/2}$ dominates $-4m\Delta e^{-m\Delta^2/2}$

Since $T$ (the total number of rounds) is typically much larger than $m$ (the number of exploration rounds), the term $2T\Delta e^{-m\Delta^2/2}$ dominates $-4m\Delta e^{-m\Delta^2/2}$. Thus, we simplify to:

$$\mathbf{E[R_T] \le m\Delta + 2T\Delta e^{-m\Delta^2/2}}.$$

## e) Find the $m$ that minimizes this regret bound and show why it is optimal

The regret bound is:

$$E[R_T] \le m\Delta + 2T\Delta e^{-m\Delta^2/2}.$$

Let:
$$f(m) = m\Delta + 2T\Delta e^{-m\Delta^2/2}.$$

Take the derivative:
$$f'(m) = \Delta - T\Delta^3 e^{-m\Delta^2/2}.$$

Set $f'(m) = 0$:
$$\Delta = T\Delta^3 e^{-m\Delta^2/2}.$$

Simplify:
$$e^{-m\Delta^2/2} = \frac{1}{T\Delta^2}.$$

Take the natural logarithm:
$$-\frac{m\Delta^2}{2} = \ln\left(\frac{1}{T\Delta^2}\right).$$

Solve for $m$:
$$\mathbf{m} = \frac{\mathbf{2}\ln(\mathbf{T\Delta^2})}{\mathbf{\Delta^2}}.$$

Verify convexity with the second derivative:
$$f''(m) = T\Delta^5 e^{-m\Delta^2/2}.$$

Since $f''(m) > 0$ for all $m > 0$, $f(m)$ is convex, and the solution is a global minimum.

## Bonus: Show that for the optimal $m$, we have the regret bounds

We aim to show the following:
$$E[R_T] \leq \frac{C_1}{\Delta} + \frac{C_2 \log(T\Delta^2)}{\Delta},$$
where $C_1$ and $C_2$ are constants. Furthermore, by maximizing the second term, we will show:
$$E[R_T] \leq 1 + C\sqrt{T},$$
where $C$ is some constant.

**Step 1: Expected regret bound for optimal $m$:**

From earlier analysis, the expected regret is:
$$E[R_T] \leq m\Delta + 2T\Delta e^{-m\Delta^2/2}.$$

For the optimal $m$, we have:
$$m = \frac{2\ln(T\Delta^2)}{\Delta^2}.$$

1. **First term ($m\Delta$):**
Substitute $m$ into $m\Delta$:

$$m\Delta = \frac{2\ln(T\Delta^2)}{\Delta^2} \cdot \Delta = \frac{2\ln(T\Delta^2)}{\Delta}.$$

This contributes to the term $\frac{C_2 \log(T\Delta^2)}{\Delta}$.
2. **Second term ($2T\Delta e^{-m\Delta^2/2}$):**
Substitute $m = \frac{2\ln(T\Delta^2)}{\Delta^2}$:

$$2T\Delta e^{-m\Delta^2/2} = 2T\Delta e^{-\ln(T\Delta^2)}.$$

Simplify $e^{-\ln(T\Delta^2)}$:

$$e^{-\ln(T\Delta^2)} = \frac{1}{T\Delta^2}.$$

Substitute back:

$$2T\Delta e^{-m\Delta^2/2} = 2T\Delta \cdot \frac{1}{T\Delta^2} = \frac{2}{\Delta}.$$

This contributes to the term $\frac{C_1}{\Delta}$.
**Combine both terms:**

$$E[R_T] \leq \frac{2}{\Delta} + \frac{2\ln(T\Delta^2)}{\Delta}.$$

Define constants $C_1 = 2$ and $C_2 = 2$, so:

$$E[R_T] \leq \frac{C_1}{\Delta} + \frac{C_2 \log(T\Delta^2)}{\Delta}.$$

**Step 2: Maximizing $\frac{\log(T\Delta^2)}{\Delta}$:**
Now we maximize the second term:

$$f(\Delta) = \frac{\log(T\Delta^2)}{\Delta}.$$

Take the derivative of $f(\Delta)$:

$$f'(\Delta) = \frac{d}{d\Delta}\left(\frac{\log(T\Delta^2)}{\Delta}\right).$$

Using the quotient rule:

$$f'(\Delta) = \frac{\Delta \cdot \frac{d}{d\Delta}[\log(T\Delta^2)] - \log(T\Delta^2)}{\Delta^2}.$$

The derivative of $\log(T\Delta^2)$ is:

$$\frac{d}{d\Delta}[\log(T\Delta^2)] = \frac{2}{\Delta}.$$

Substitute back:

$$f'(\Delta) = \frac{\Delta \cdot \frac{2}{\Delta} - \log(T\Delta^2)}{\Delta^2}.$$

Simplify:

$$f'(\Delta) = \frac{2 - \log(T\Delta^2)}{\Delta^2}.$$

Set $f'(\Delta) = 0$ to find the critical point:

$$2 - \log(T\Delta^2) = 0.$$

Solve for $\Delta$:

$$\log(T\Delta^2) = 2.$$

Exponentiate both sides:

$$T\Delta^2 = e^2.$$

Solve for $\Delta$:

$$\Delta = \frac{e}{\sqrt{T}}.$$

**Step 3: Substituting $\Delta = \frac{e}{\sqrt{T}}$ into $E[R_T]$:**
Substitute $\Delta = \frac{e}{\sqrt{T}}$ into the regret bound:

$$E[R_T] \leq \frac{C_1}{\Delta} + \frac{C_2 \log(T\Delta^2)}{\Delta}.$$

1. First term:

$$\frac{C_1}{\Delta} = \frac{C_1}{\frac{e}{\sqrt{T}}} = C_1 \frac{\sqrt{T}}{e}.$$

2. Second term: Substitute $T\Delta^2 = e^2$:

$$\log(T\Delta^2) = \log(e^2) = 2.$$

Substitute $\Delta = \frac{e}{\sqrt{T}}$:

$$\frac{C_2 \log(T\Delta^2)}{\Delta} = \frac{C_2 \cdot 2}{\frac{e}{\sqrt{T}}} = C_2 \frac{2\sqrt{T}}{e}.$$

Combine both terms:

$$E[R_T] \leq C_1 \frac{\sqrt{T}}{e} + C_2 \frac{2\sqrt{T}}{e}.$$

Factor out $\sqrt{T}$:

$$E[R_T] \leq (C_1 + 2C_2) \frac{\sqrt{T}}{e}.$$

Let $C = \frac{C_1 + 2C_2}{e}$, then:

$$E[R_T] \leq C\sqrt{T}.$$

For large $T$, the regret is dominated by the $\sqrt{T}$ term, so:

$$E[R_T] \leq 1 + C\sqrt{T}.$$