

CSIE 5432/5433 HW4 PDF

YiWenLai

TOTAL POINTS

40 / 0

QUESTION 1

1 Problem 2 10 / 0

- + 20 pts Correct
 - + 15 pts Correct with minor issues
 - ✓ + 10 pts Almost correct
 - + 0 pts Chose "I have no idea"
 - + 0 pts Almost incorrect
 - 5 pts Explained $E[E_{in}(A(D))] \leq E[E_{out}(A(D))]$ with intuition / Proof by showing special cases
 - 10 pts Incorrect or no explanation on $E[E_{in}(A(D))] \leq E[E_{out}(A(D))]$
- 1 What does in+out mean here?

QUESTION 2

2 Problem 3 15 / 0

- + 20 pts Correct
 - ✓ + 15 pts Correct with a minor issue.
 - + 10 pts Correct with 2 mistakes
 - + 5 pts Correct with 3 mistakes
 - + 0 pts Make more than 3 mistakes, or leave blank, or choosing 'I have no idea', or without a detailed explanation, or your explanation doesn't make sense.
 - 10 pts Choose a wrong option, or almost without any description
 - 20 pts Violate our course policy #4
- 2 Wrong

QUESTION 3

3 Problem 13 15 / 0

- + 20 pts Correct answer with complete explanation.
- ✓ + 15 pts Got correct answer but made ****one**** mistake.
- + 10 pts Got correct answer but made ****two**** mistakes

- + 0 pts Left blank and chose the "I don't have any idea" option during submission
- + 0 pts Got correct answer but your explanation makes no sense to me.
- 10 pts Wrong answer
- 10 pts little to no explanation
- + 0 pts Some (but not limit to) examples of mistake:
 - (0) skip "variance of the sum of uncorrelated random variables is the sum of their variances"
 - (1) obvious typo
 - (2) skip too many steps (regarded as two mistakes)
 - (3) central limit theorem can only get approximate answer (regarded as two mistakes)

QUESTION 4

4 Problem 16 0 / 0

- ✓ + 0 pts Correct
- 10 pts Wrong page
- 20 pts No code

1. c

$$\begin{aligned}
 & 1. \\
 & \min_w \frac{1}{2} \int_0^2 (e^x - wx)^2 dx \\
 & = \min_w \left. \frac{1}{2} e^{2x} + \frac{1}{3} w^2 x^3 - 2wx e^x (x-1) \right|_0^2 \\
 & = \min_w \frac{8}{3} w^2 - 2(e^2+1)w + \frac{1}{2} e^4 - \frac{1}{2} \\
 & = \min_w \left(w - \frac{3}{8}(e^2+1) \right)^2 + \frac{1}{2} e^4 - \frac{1}{2} - \frac{9}{64}(e^2+1)^2 \\
 & \text{when } w = \frac{3+3e^2}{8}, \text{ there will be squared error with min value } \star
 \end{aligned}$$

2. b

$$\begin{aligned}
 & 2. \\
 & A(D) = \underset{h \in H}{\operatorname{argmin}} E_{in}(h) \\
 & A^*(D) = \underset{h \in H}{\operatorname{argmin}} E_{in+1}(h) \\
 & E_D[E_{in}(A(D))] \leq E_D[E_{in}(A^*(D))] \\
 & E_D[E_{out}(A^*(D))] \leq E_D[E_{out}(A(D))] \\
 & \text{since } A^*(D) \text{ is generated by considering both data in and out-samples} \\
 & E_D[E_{in}(A^*(D))] = E_D[E_{out}(A^*(D))]
 \end{aligned}$$

1 Problem 2 10 / 0

+ 20 pts Correct

+ 15 pts Correct with minor issues

✓ + 10 pts Almost correct

+ 0 pts Chose "I have no idea"

+ 0 pts Almost incorrect

- 5 pts Explained $E[E_{\text{in}}(A(D))] \leq E[E_{\text{out}}(A(D))]$ with intuition / Proof by showing special cases

- 10 pts Incorrect or no explanation on $E[E_{\text{in}}(A(D))] \leq E[E_{\text{out}}(A(D))]$

1 What does in+out mean here?

3. d

3.

$$X_h^T = \begin{bmatrix} x_1 & x_2 & \dots & x_1 + \varepsilon & x_2 + \varepsilon & \dots & x_N + \varepsilon \end{bmatrix}_{d+1 \times 2N}$$

$$X_h = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_1 + \varepsilon \\ x_2 + \varepsilon \\ \vdots \\ x_N + \varepsilon \end{bmatrix}_{2N \times d+1}$$

$$X_{ij} \text{ for } X_h^T X_h = 2 \sum_{k=1}^{2N} X_h^T(i)[k] \cdot X_h[k][j] + N \cdot \sigma^2 + \varepsilon \sum_{k=1}^N (X_{ik} + X_{kN})$$

i.e. take expected value, the last term with ε is equal to zero due to normal distribution of noise.

$$E(X_h^T X_h) = 2X^T X + N \cdot \sigma^2 I_{d+1}$$

4. e

4.

$$X_h^T = \begin{bmatrix} x_1 & x_2 & \dots & x_1 + \varepsilon & x_2 + \varepsilon & \dots & x_N + \varepsilon \end{bmatrix}_{d+1 \times 2N}$$

$$y_h = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_{2N \times 1}$$

$$X_h^T y_h(i) = \sum_{j=1}^{2N} X_h^T(i)[j] \cdot y_h(j) = 2 \sum_{j=1}^{2N} X_h^T(i)[j] \cdot y_h(j) + \varepsilon \sum_{j=1}^N y_h(j)$$

same as (3), after taking expected value, the later term = 0.

$$E(X_h^T y_h) = 2X^T y$$

2 Problem 3 15 / 0

+ 20 pts Correct

✓ + 15 pts Correct with a minor issue.

+ 10 pts Correct with 2 mistakes

+ 5 pts Correct with 3 mistakes

+ 0 pts Make more than 3 mistakes, or leave blank, or choosing 'I have no idea', or without a detailed explanation, or your explanation doesn't make sense.

- 10 pts Choose a wrong option, or almost without any description

- 20 pts Violate our course policy #4

2 Wrong

12. e

12. constant hypothesis

valid	$h(x)$	error
$(e, 2)$	$h(x) = 0$	$e_1 = 2^2 = 4$
$(3, 0)$	$h(x) = 1$	$e_2 = 1^2 = 1$
$(-3, 0)$	$h(x) = 1$	$e_3 = 1$

linear hypothesis

valid	$h(x)$	error
$(e, 2)$	$h(x) = 0$	$e_1 = 2^2 = 4$
$(-3, 0)$	$h(x) = \frac{2}{e-3}(x-3)$	$e_2 = \left(\frac{-12}{e-3}\right)^2$
$(3, 0)$	$h(x) = \frac{2}{e+3}(x+3)$	$e_3 = \left(\frac{12}{e+3}\right)^2$

$$4 + 1 + 1 = 4 + \left(\frac{-12}{e-3}\right)^2 + \left(\frac{12}{e+3}\right)^2, \quad \left(\frac{1}{e-3}\right)^2 + \left(\frac{1}{e+3}\right)^2 = \frac{1}{12}$$

$$e^4 - 18e^2 - 1215 = 0, \quad e^2 = 81 \pm 36\sqrt{6}, \quad (\text{負不合}), \quad e = \sqrt{81 + 36\sqrt{6}} \quad \star$$

13. d

13.

$$\underset{\text{Eval} \sim P^k}{\text{Var} [\text{Eval}(h)]} = \underset{\text{Eval} \sim P^k}{\text{Var} \left[\frac{1}{K} \sum \text{err}(h(x), y) \right]} = \frac{1}{K^2} \underset{\text{Eval} \sim P^k}{\text{Var} [\sum \text{err}(h(x), y)]}$$

$$= \frac{1}{K^2} \sum_{(x,y) \sim P} \text{Var} (\text{err}(h(x), y)) = \frac{1}{K^2} \cdot K \underset{(x,y) \sim P}{\text{Var} (\text{err}(h(x), y))}$$

$$= \frac{1}{K} \underset{(x,y) \sim P}{\text{Var} [\text{err}(h(x), y)]} \quad \star$$

14. c

14.

$$\left(\frac{1}{4} + \frac{1}{4}\right) \frac{1}{16} = \frac{1}{32} \quad \star$$

$$E_n = \frac{1}{4} \quad E_n = \frac{1}{4}$$

3 Problem 13 15 / 0

+ **20 pts** Correct answer with complete explanation.

✓ + **15 pts** Got correct answer but made ****one**** mistake.

+ **10 pts** Got correct answer but made ****two**** mistakes

+ **0 pts** Left blank and chose the "I don't have any idea" option during submission

+ **0 pts** Got correct answer but your explanation makes no sense to me.

- **10 pts** Wrong answer

- **10 pts** little to no explanation

+ **0 pts** Some (but not limit to) examples of mistake:

(0) skip "variance of the sum of uncorrelated random variables is the sum of their variances"

(1) obvious typo

(2) skip too many steps (regarded as two mistakes)

(3) central limit theorem can only get approximate answer (regarded as two mistakes)

15. a

15.

$$p \cdot \varepsilon_+ + (1-p) \cdot \varepsilon_- = 1-p$$
$$p(\varepsilon_+ - \varepsilon_- + 1) = 1 - \varepsilon_-$$
$$p = \frac{1 - \varepsilon_-}{\varepsilon_+ - \varepsilon_- + 1} *$$

16. b

a. Shell script

```
./train -s 0 -c 0.00005 -e 0.000001 -q hw4_train_convert.txt
./predict -b 1 hw4_test_convert.txt hw4_train_convert.txt.model result16

./train -s 0 -c 0.005 -e 0.000001 -q hw4_test_convert.txt
./predict -b 1 hw4_test_convert.txt hw4_train_convert.txt.model result16

./train -s 0 -c 0.5 -e 0.000001 -q hw4_train_convert.txt
./predict -b 1 hw4_test_convert.txt hw4_train_convert.txt.model result16

./train -s 0 -c 50 -e 0.000001 -q hw4_train_convert.txt
./predict -b 1 hw4_test_convert.txt hw4_train_convert.txt.model result16

./train -s 0 -c 5000 -e 0.000001 -q hw4_train_convert.txt
./predict -b 1 hw4_test_convert.txt hw4_train_convert.txt.model result16
```

b. Terminal message

```
yiwenlai@YiWens-MBP ~/Desktop/liblinear-2.42 INSERT bash hw4_q16.sh
Accuracy = 51.6667% (155/300)
Accuracy = 51.6667% (155/300)
Accuracy = 80.6667% (242/300)
Accuracy = 87% (261/300)
Accuracy = 86.6667% (260/300)
```

17. a

a. Shell script


```

        nums = line.strip().split()
        data = [float(num) for num in nums]
        transform.append(data[len(data)-1])
        transform.append(1)
        for i in range(6):
            transform.append(data[i])
        for i in range(6):
            for j in range(6-i):
                transform.append(data[i]*data[i+j])
        # print(len(transform))
        text.append(transform)
    f.close()
with open(outputfile[x], 'w') as f1:
    for t in text:
        row_text = ""
        for i in range(len(t)):
            if i == 0:
                if t[i] == 1:
                    row_text += (" +1 ")
                elif t[i] == -1:
                    row_text += (" -1 ")
            else:
                row_text += (str(i)+":")
                row_text += str(t[i])
                row_text += " "
        f1.write(row_text+"\n")
f1.close()

```

4 Problem 16 0 / 0

✓ + 0 pts Correct

- 10 pts Wrong page

- 20 pts No code