257 HW 2 01181878

Q 10

2.1 $\delta = 0.03$, $\varepsilon(M,N,\delta) = \sqrt{\frac{1}{2N}\ln\frac{2M}{\delta}}$

(a) $M = 1$; $\varepsilon \leq 0.05$

Given $\varepsilon(M,N,\delta) = \sqrt{\frac{1}{2N}\ln\frac{2M}{\delta}}$

$\Rightarrow$ $$N = \frac{1}{2(\varepsilon)^2}\ln\frac{2M}{\delta}$$

Substitut

$$N \geq \frac{1}{2(0.05)^2}\ln\frac{2(1)}{0.03}$$

($\varepsilon \leq 0.05$) divi when ~~multiph~~ on both side sign change.

$$N \geq \frac{1}{0.005}\ln(66.666)$$

$$N \geq \frac{4.199}{0.05}$$

$$N \geq 839.941$$

011818781

b) from a $N = \frac{1}{2(\varepsilon)^2} \ln \frac{2M}{\delta}$

$M = 100$

$\varepsilon \leq 0.05$

$$N \geq \frac{1}{2 \times (0.05)^2} \ln \left(\frac{200}{0.03} \right)$$

$$\geq \frac{1}{0.005} \ln (6{,}666)$$

$$\geq \frac{8.8048}{0.005}$$

$$\boxed{N \geq 1{,}760.976}$$

c) $M = 10{,}000$

$\varepsilon \leq 0.05$

$$N \geq \frac{1}{0.005} \ln \left(\frac{20{,}000}{0.03} \right)$$

$$N \geq \frac{13.41004}{0.005}$$

$$\boxed{N \geq 2{,}682.009}$$

Q2 01181878I

2.3

(a) $$\mathcal{H} \to f(a) = \begin{cases} +1 & [a, \infty) \\ +1 & (-\infty, a) \end{cases}$$

for growth function of position = $m_{\mathcal{H}}(N) = N+1$

" " " " negate = $m_{\mathcal{H}}(N) = N-1$

So in total $m_{\mathcal{H}}(N) = (N+1)+(N-1)$

$= 2N$

(b) $$\mathcal{H} \to f(a,b) = \begin{cases} +1 & [a,b] \\ -1 & \text{everywhere.} \end{cases}$$

&

$$f(a,b) = \begin{cases} -1 & (a,b) \\ +1 & \text{everywhere} \end{cases}$$

growth function for '+' interval = $\binom{N+1}{2} + 1 = \frac{1}{2}N^2 + \frac{1}{2}N + 1$

" " " '-' num = $N-2$

so max $m_{\mathcal{H}}(N) = \frac{1}{2}N^2 + \frac{1}{2}N + 1 + N - 2$

$= \frac{1}{2}N^2 + \frac{3}{2}N - 1$

c) two concentric spheres in $\mathbb{R}^d$: $\mathcal{H} \to f(a, b) \to$ '+1' for $a \le \sqrt{x_1^2 + \dots + x_d^2} + b$

Ans)

growth function of concentric circles mapping is req

So, map ϕ can be defined as

$$\phi(x_1, \dots x_d) \Rightarrow r = \sqrt{x_1^2 + \dots + x_n^2}$$

which is similar to positive interval of previous part. $\therefore$ $m_{\mathcal{H}}(N) = \frac{N^2}{2} + \frac{N}{2} + 1$

Q3
2.8

~~plot both $m_H(N)$ for 2.5, 2.6~~

~~for $d_{vc} = 2$, $d_{vc} = 5$~~

Ans)

2 growth function are present

(i) d_{vc} is finite & $m_{\mathcal{H}}(N) = 1$ is bounded $b(N^{d_{vc}} + 1)$

(ii) $d_{vc} = +\infty$ and $m_{\mathcal{H}}(N) = 2^N$

hence $m_{\mathcal{H}}(N) = N + 1 \Rightarrow$ Possible growth function
$d_{vc} = 1$

01818781

(ii) $1 + N + \frac{N(N-1)}{2}$ $\Rightarrow$ for $N = 2 \Rightarrow 1 + 2 + \frac{2(2-1)}{2} = 2^2$

$\Rightarrow 4 = 4$

and equation is a polynom with

possible growth funtion bounded by $N^2 + 1$

(iii) 2^N. which imp d_{vc} can be $\infty = N$

There can be a ~~possible~~ growth funtion but

bounde by $N^N + 1$

(iv) $2^{\lfloor \sqrt{N} \rfloor}$ $\Rightarrow d_{vc} = 1$ can be bounded by

$N^1 + 1$ but for $N = 25$ $2^5 = 32$

but bound is 26 which is not possible

(v) $2^{\frac{N}{2}}$ when $d_{vc} = 0$ so $N^0 + 1 = 2$

which cannot have a growth funch

(vi) $1 + N + \frac{N(N-1)(N-2)}{6}$ $d_{vc} = 1$, bound $= N+1$

for $N = 4$; 9 9 bound $= 5$

not possible

5

Q4

2.12 $\cancel{\Omega}\,\cancel{\varepsilon}(N, \mathcal{H}, \delta) = \sqrt{\frac{8}{N} \ln\left(\frac{4 m_{\mathcal{H}}(2N)}{\delta}\right)}$

$\delta = 0.05$

$d_{vc} = 10$

(or)

$$\Omega(N, \mathcal{H}, \delta) \leq \sqrt{\frac{8}{N} \ln\left(\frac{4(2N)^{d_{vc}} + 1}{\delta}\right)}$$

$\Omega = 95\%$ of conf

$= 5\%$ gen

$= \frac{5}{100}$.

$$(0.05)^2 \leq \frac{8}{N} \ln\left[\frac{4[(2N)^{10}] + 1}{0.05}\right]$$

$$N \geq \frac{8}{(0.05)^2} \times \ln\left(\frac{4 (2000)^{10} + 1}{0.05}\right)$$

$$\boxed{N \geq 2.5725 \times 10^5}$$

6

Q5

2.22

011818781

$$E_D\left[E_{out}(g^{(D)})\right] = \sigma^2 + bias + var.$$

$$\Rightarrow \quad E_D\left[E_{xy}\left[g^{(D)}(x) - y(x)\right]^2\right]$$

$$= E_{xy}\, E_D\left[\left[g^D(x) - y(x)\right]^2\right]$$

$$= E_{xy}\left[E_D\, g^D(x)^2 + y(x)^2 - 2g^D(x)y(x)\right]$$

$$= E_{xy}\left[E_D\, g^D(x)^2 - 2g^D(x)f(x) - 2g^D(x)\varepsilon + f(x)^2 + 2f(x)\varepsilon x + \varepsilon^2\right]$$

$$= \left[E_D\, g^D(x)^2 - \bar{g}(x)^2\right] + \left(\bar{g}(x)^2 - 2\bar{g}(x)f x + f(x)^2\right) - 2\varepsilon\left(\bar{g}(x) - f(x)\right) + \varepsilon^2$$

$$= E_D\left[(g^d(x)^2) - 2g^{(D)}(x)\bar{g}(x) + \bar{g}(x)^2\right] + (\bar{g}(x) - f(x))^2 + \varepsilon^2 - 2\left(\bar{g}(x) - f(x)\right)\varepsilon$$

$$= Var(x) + bia(x) + \varepsilon^2 - 2\varepsilon\left(\bar{g}(x) - f(x)\right)$$

$\rightarrow$

$\therefore \quad E_D\left[E_D f(y)^D\right] = E_x\left[E_D\left[g^D(x)^2\right] - 2\bar{g}(x)\,y(x) + y(x^2)\right]$

$= E_{xy}\left[\mathrm{Var}(x)\right] + E_{xy}\left[\mathrm{bia}(x)\right] + E_{xy}\left[\varepsilon^2\right] - 2E_{xy}\left(g(x) - f(x)\right] \cdot \varepsilon)$

$= E_x\left[\mathrm{Var}(x)\right) + E_x\left[\mathrm{bia}(x)\right] + E_x\left[E_\varepsilon(\varepsilon^2)\right] - 2E_x\left[\bar{g}(x) - f(x)\right]E_\varepsilon(\varepsilon)\right]$

$= \mathrm{Var} + \mathrm{bia} + E_x\left[\varepsilon - E_\varepsilon(\varepsilon)^2\right] - 2E_x\left[\bar{g}(x)\,g(x)\right.$

$$\boxed{= \mathrm{Var} + \mathrm{bia} + \sigma^2}$$

//

Q6 $\prod_{n=1}^{N} P(y_n|x_n)$ is equavalent to minimizin

$$E_{in}(w) = \frac{1}{N} \sum_{n=1}^{N} \ln(1+e^{-y_n w^T x_n})$$

$\Rightarrow$ from the linea classifica $h(x) = \text{sig}(w^T x)$

linea regu use no sign

$$h(x) = w^T x$$

four logistic regressia it neede to be in b/w above 2

$$h(x) = \theta(w^T x) \quad \because \theta(s) = \frac{e^s}{1+e^s}$$

four our target function

$$P(y_n|x_n) = \begin{cases} h(x) & y = +1 \\ 1 - h(x) & y = -1 \end{cases}$$

we know the $1-\theta(s) = \theta(-s)$

$$\therefore \quad P(y|x) = \begin{cases} \cancel{h(x)}\ \theta(y w^T x) \\ \theta(-y w^T x) \end{cases}$$

$\Rightarrow$ we can use $-\frac{1}{N}\ln(\;)$ in a decreas funct

Since likelihood selects the hypo h which maximizes

$$-\frac{1}{N}\ln\left(\prod_{n=1}^{N} P(y_n|x_n)\right) = \frac{1}{N}\sum_{n=1}^{N}\ln\left(\frac{1}{P(y_n|x_n)}\right)$$

$$= \frac{1}{N}\sum_{n=1}^{N}\ln\left[\frac{1}{\theta(y_n w^T x_n)}\right]$$

Since we are minimizing quality error each case use error

$$E_{in}(w) = \frac{1}{N}\sum_{n=1}^{N}\ln\left[\frac{1+e^{y_n w^T x_n}}{e^{y_n w^T x_n}}\right] \qquad \theta(s)=\frac{e^s}{1+e^s}$$

$$E_{in}(w) = \frac{1}{N}\sum_{n=1}^{N}\ln\left[1+e^{-(y_n w^T x_n)}\right]$$

Q7

A)

in Sample Err

$$E_{in}(w) = \frac{1}{N} \sum_{n=1}^{n} \ln\left(1 + e^{-y_n w^T x_n}\right)$$

$$\nabla E_{in}(w_t) = \frac{d}{dw}\left[E_{in}(w)\right]$$

$$= \frac{1}{N} \sum_{n=1}^{n} \frac{d}{du} \ln\left(1 + e^{-y_n w^T x_n}\right)$$

$$= \frac{1}{N} \sum_{n=1}^{n} \frac{1}{1+e^{-y_n w^T x_n}} \times \left(0 + e^{-y_n w^T x_n} (-y_n x_n)\right)$$

$$\nabla E_{in}(w) = \frac{1}{N} \sum_{n=1}^{N} \frac{-y_n x_n}{e^{y_n w^T x_n} + 1}$$

$$\Rightarrow \nabla E_{in}(w) = \frac{-1}{N} \sum_{n=1}^{N} x_n y_n \times \theta(-y_n \cdot w^T x_n)$$

Q. 8

A) $(x-3)^2 + x_2 = 1$

$\Rightarrow \quad x_1^2 - 6x + 8 + x_2 = 0$

$8 - 6x + x_2 + x_1^2 = 0$

which can be Represented in Z when $Z = \phi(x)$

$$h(x) = \cancel{\text{sign}} \left\{ [8 \; -6 \; 1 \; 1] \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \end{pmatrix} \right\} - Z$$

$$\phi(x) = (1, x_1, x_2, x_1^2)$$

b) line $\quad (x_1-3)^2 + (x_2-4)^2 = 1$

$x^2 - 6x + x_2^2 - 8x_2 + 24 = 0$

$24 - 6x_1 - 8x_2 + x_1^2 + x_2^2 =$

$$[24 \; -6 \; -8 \; 1 \; 1] \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \end{bmatrix} \Big\} \; Z$$

$$\phi(x) = (1.\, x_1 x_2 x_1^2\, x_2^2)$$

(12)

c)

$$2(x_1-3)^2+(x_2-4)^2=1$$

$$\Rightarrow \quad 2x_1^2 - 12x_1 + 18 + x_2^2 - 8x_2 + 15 = 0$$

$$33 - 12x_1 - 8x_2 + 2x_1^2 + x_2^2 = 0$$

$$\cancel{h(x) = (w_0, w_1, \ldots w_n)}$$

$$\left.\begin{bmatrix} 33 & -12 & -8 & 2 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \end{pmatrix}\right\} z$$

$$\phi(x) = (1, x_1, x_2, x_1^2, x_2^2)$$