H.W #6.

1. (a) $m_0^u = \lceil \frac{4}{\varepsilon}((3)\cdot \log_2 \frac{12}{\varepsilon} + \log_2 \frac{2}{\delta}) \rceil$, $\varepsilon = 0.05$, $\delta = 0.05$

$= \lceil \frac{4}{0.05} (3 \cdot \log_2 \frac{12}{0.05} + \log_2 \frac{2}{0.05}) \rceil$

$= \lceil 80 \cdot (3 \cdot \log_2 240 + \log_2 40) \rceil$

$= 2323.4$

$\fallingdotseq 2324$.

(b) $m_0^L = \lceil \frac{1-\varepsilon}{\varepsilon} \ln \frac{1}{\delta} \rceil$

$= \lceil \frac{1-0.05}{0.05} \ln \frac{1}{0.05} \rceil = \lceil 56.9189 \rceil = 57$

$*$ $m_0^L = \max \{ \frac{1}{\varepsilon} \ln \frac{1}{\delta}, \frac{VCD(C)-1}{32\varepsilon} \}$ $\delta \leq 0.01$ and $\varepsilon \leq \frac{1}{8}$

2. <sol> $\ell = 1000$, $\delta = 0.05$, $d = 3$

$$\varepsilon = \frac{2}{\ln 2} \cdot \frac{3(1+\ln \frac{2000}{3}) - \ln \frac{0.05}{2}}{1000} = 0.0756$$

3. We have $P_r[\sup_{\alpha \in \Lambda} R(\alpha) - R_{emp}(\alpha) > \varepsilon] \leq 2\, \Pi_H(2\ell)\, 2^{-\varepsilon\ell/2B}$

Let $\delta = 2 \cdot \Pi_H(2\ell)\, 2^{-\varepsilon\ell/2B}$. $\leq \left(\frac{e \cdot 2\ell}{d}\right)^d$

$\Rightarrow \frac{\delta}{2} \leq \left(\frac{e \cdot 2\ell}{d}\right)^d \cdot 2^{-\frac{\varepsilon\ell}{2B}}$

$\Rightarrow \ln \frac{\delta}{2} \leq d\left(1 + \ln \frac{2\ell}{d}\right) - \frac{\varepsilon\ell}{2B} \ln 2$

$\Rightarrow \varepsilon \cdot \frac{\ell}{2B} \cdot \ln 2 \leq d\left(1 + \ln \frac{2\ell}{d}\right) - \ln \frac{\delta}{2}$

$\therefore \varepsilon \leq \frac{2B}{\ln 2} \cdot \frac{d(1+\ln\frac{2\ell}{d}) - \ln \frac{\delta}{2}}{\ell}$

$\therefore$ With prob. at least $1-\delta$,

$R(\alpha) \leq R_{emp}(\alpha) + \varepsilon$, $\forall \alpha \in \Lambda$.

4. For the ANNs.

$O(\omega \ln h) \le VCD(H) \le O(\omega^2 h^2)$ where

$\begin{cases} \omega : \text{the number of total parameters} \\ h : \text{the num. of hidden units} \end{cases}$

Assume $\exists$ positive constants k_1, k_2 s.t.

$k_1 \omega \ln h \le VCD(H) \le k_2 \omega^2 h^2$

Then

$$G(2l) \le d\left(1 + \ln \frac{2l}{d}\right)$$
$$\le k_2 \omega^2 h^2 \left(1 + \ln \frac{2l}{k_1 \omega \ln h}\right)$$

(1) For the bounded loss functions

$$R(\alpha) \le R_{emp}(\alpha) + \frac{\varepsilon^2}{2}\left(1 + \sqrt{1 + \frac{4R_{emp}(\alpha)}{\varepsilon^2}}\right) \quad \text{where}$$

$$\varepsilon^2 = 4(B-A)\,\frac{d\left(1 + \ln \frac{2l}{d}\right) - \ln \frac{\delta}{4}}{l}$$

$$\therefore\ \varepsilon^2 \le 4(B-A) \cdot \frac{k_2 \omega^2 h^2 \left(1 + \ln \frac{2l}{k_1 \omega \ln h}\right) - \ln \delta/4}{l}$$

$$\varepsilon^2 = O\left\{(B-A)\,\frac{\omega^2 h^2 \ln \frac{l}{\omega \ln h} - \ln \delta}{l}\right\}$$

(2) For the non-negative loss functions

$p > 2$, $$R(\alpha) \le \frac{R_{emp}(\alpha)}{(1 - \tau^* a(p)\varepsilon)_+} \quad \text{where} \quad \sup_{\alpha \in \Lambda} \frac{\sqrt[p]{\int Q^p(z,\alpha)\, p(z)\, dz}}{\int Q(z,\alpha)\, p(z)\, dz} = \tau_p < \tau^*,$$

$$a(p) = \sqrt[p]{\frac{1}{2}\left(\frac{p-1}{p-1}\right)^{p-1}}, \quad \varepsilon = \sqrt{4 \cdot \frac{G(2l) - \ln \delta/4}{l}}$$

So, $$\varepsilon \le \sqrt{4 \cdot \frac{k_2 \omega^2 h^2 \left(1 + \ln \frac{2l}{k_1 \omega \ln h}\right) - \ln \frac{\delta}{4}}{l}}$$

$$\therefore\ \varepsilon = O\left(\sqrt{\frac{\omega^2 h^2 \ln \frac{l}{\omega \ln h} - \ln \delta}{l}}\right).$$

5. $V(l) = r_n + T_n \left(\frac{h_n \ln l}{l} \right)^{\frac{1}{2}}$

$= \frac{1}{n} + T_n \left(\frac{n \cdot \ln l}{l} \right)^{\frac{1}{2}}$

$\frac{\partial V(l)}{\partial n} = -\frac{1}{n^2} + T_n \cdot \frac{1}{2} \left(\frac{n \cdot \ln l}{l} \right)^{\frac{1}{2}} \cdot \frac{\ln l}{l} = 0$

$\frac{1}{n^2} = \frac{T_n}{2} \left(\frac{l}{n \cdot \ln l} \right)^{1/2} \frac{\ln l}{l}$

$= \frac{T_n}{2} \left(\frac{\ln l}{n l} \right)^{1/2}$

$n^2 = \frac{2}{T_n} \left(\frac{n l}{\ln l} \right)^{1/2}$

$n^{\frac{3}{2}} = \frac{2}{T_n} \left(\frac{l}{\ln l} \right)^{1/2}$

$\therefore n = \left(\frac{2}{T_n} \right)^{\frac{2}{3}} \left(\frac{l}{\ln l} \right)^{1/3}$

$\therefore n \propto \left(\frac{l}{\ln l} \right)^{1/3}$