۶.۲) الف (a) ثابت می‌کنیم که $VCdim(H_{=k}) = \min\{k, |X|-k\}$

ابتدا باید نشان دهیم که $VCdim(H) \le k$ ، فرض می‌کنیم C مجموعه‌ای با اندازه $k+1$ باشد، آنگاه $h \in H$ وجود ندارد که به ازای همه‌ی $x \in C$ ، $h(x)=1$ باشد

به عبارت دیگر برای هر C ، $k+1$ عضوی، $h \in H$ ای وجود دارد که C shatter نمی‌شود

به طور مشابه اگر C مجموعه‌ای به اندازه‌ی $|X|-k+1$ باشد، هیچ $h \in H$ وجود ندارد که C را shatter کند، به عبارت دیگر به ازای تمامی $x \in C$ ، هیچ $h \in H$ وجود ندارد که $h(x)=0$

در حالت بعدی که $VCdim(H) \ge \min\{k, |X|-k\}$ ، فرض می‌کنیم $C=\{x_1, \dots, x_m\}$ مجموعه‌ای با اندازه m باشد که $(y_1, \dots, y_m)$ را label کند، آنگاه $h \in H$ وجود دارد به طوری که $h(x_i)=y_i$ ، که حداکثر k بار اشتراک (?) وجود دارد که label یک (?) را می‌گیرند

پس چون $m \le VCdim(H)$ در نتیجه ثابت می‌شود که $VCdim(H) = \min\{k, |X|-k\}$

۶.۲) b) ثابت می‌کنیم که $VCdim(H)=k$ ، ابتدا باید نشان دهیم که $VCdim(H) \le k$

فرض می‌کنیم $C \subseteq X$ ، مجموعه‌ای به اندازه $k+1$ باشد، به ازای تمامی $x \in C$ هیچ $h \in H$ وجود ندارد که $h(x)=1$ باشد، به طور مشابه باید نشان داد که $VCdim(H) \ge k$.

فرض می‌کنیم $C=\{x_1, \dots, x_m\}$ ، $C \subseteq X$ ، $m \le k$ ، به ازای تمامی $x \in C$ ها $h \in H$ وجود ندارد که $h(x)=1$ ، به عبارت دیگر $h \in H$ وجود ندارد که C را shatter کند.

اکنون برای $VCdim(H) \ge k$ می‌توان $h \in H$ را پیدا کرد که $h(x_i)=y_i$ و برای هر $x_i \in C$ ، $h(x)=0$ در نتیجه $VCdim(H)=k$

۶.۶) a) برای هر $h \in H$، d تا متغیر وجود دارد که می‌توانند x_i، $\bar{x}_i$ و یا داده‌ی گمشده باشند، بنابراین 3^d حالت ممکن وجود دارد در نتیجه:

$$|H^d_{con}| \le 3^d + 1$$

۶.۶) b) می‌توانیم یک کران بالا برای $vcdim(H)$ در نظر بگیریم

فرض می‌کنیم $vcdim(H) = Q$، اگر H، یک مجموعه Q عضوی را shatter کند

با توجه به فرضیات داریم:

$$\log(|H|) \ge Q \Rightarrow \log(|H|) \ge vcdim(H)$$

$$\Rightarrow vcdim(H^d_{con}) \le \lfloor \log(|H^d_{con}|) \rfloor \le \log(3^d+1) \le 3\log d$$

$$|H^d_{con}| \le 3^d + 1$$

۶.۹) برای $vcdim(H)$ دو حالت وجود دارد

$$\left.\begin{array}{ll} \text{①} \quad H \text{، } c \text{ را shatter می‌کند} \quad |c| = d & \exists c \\ \text{②} \quad H \text{، } c \text{ را shatter نمی‌کند} \quad |c| = d+1 & \forall c \end{array}\right\}$$

حال یک نقطه، ۲ نقطه، ۳ نقطه را بررسی می‌کنیم، با در نظر گرفتن تمامی حالات در $c=\{c_1\}$، $c=\{c_1, c_2\}$، $c=\{c_1, c_2, c_3\}$، می‌بینیم که H، c را shatter می‌کند

پس $vcdim(H) \ge 3$، اگر اندازه $c=\{c_1, c_2, c_3, c_4\}$ باشد، با بررسی تمامی حالات متوجه می‌شویم که حالت $(+,-,+,-)$ را هیچ تابعی نمی‌تواند برگرداند، پس در اینجا H، c را shatter نمی‌کند

در نتیجه:

$$\boxed{vcdim(H) = 3}$$

۶.۱۰) a) فرض می‌کنیم $C=\{c_1, \dots, c_d\}$، همچنین می‌دانیم که H، C را shatter می‌کند، حال با توجه به تمرین ۵.۳ داریم:

$$\exists f: X \longrightarrow \{0,1\} : L_D(f)=0$$

$$E_{S\sim D^m}[L_D(A(S))] \geq \frac{1}{2} - \frac{1}{2k}$$: $k = \frac{|C|}{m} = \frac{d}{m}$

$$E_{S\sim D^m}[L_D(A(S))] \geq \underbrace{\frac{1}{2} - \frac{m}{2d}}_{=\frac{d-m}{2d}}$$ چون $k = \frac{d}{m}$ ؛

$$E_{S\sim D^m}[L_D(A(S))] \geq 0 + \frac{d-m}{2d}$$

در نتیجه با توجه به realizability assumption ($\exists f: L_D(f)=0$)

$$E_{S\sim D^m}[L_D(A(S))] \geq \min_{h\in H} L_D(f) + \frac{d-m}{2d}$$

۶.۱۰) b) با استفاده از برهان خلف فرض می‌کنیم $vcdim(H)=\infty$، با توجه به قضیه free lunch، در هر training set به اندازه‌ی m، یک مجموعه مانند C به اندازه‌ی $d=2m$ وجود دارد که shatter می‌گردد، بنابراین:

$$E_{S\sim D^m}[L_D(A(S))] \geq 0 + \frac{d-m}{2d} = \frac{1}{4}$$

حال طبق قضیه مارکوف:

$$P(Z>a) \geq \frac{E(Z)-a}{1-a}$$

اگر $a=1/8$

$$P_{S\sim D^m}[L_D(A(S)) > 1/8] \geq \frac{E(L_D(A(S)))-1/8}{1-1/8} = 1/7$$

و این بدان معناست که خطای حقیقی با احتمال بیشتر از $1/7$ از $1/8$ بزرگ‌تر است، اگر $\epsilon \leq 1/8$ و $\delta \leq 1/7$، در تناقض با PAC learning بودن H است، در نتیجه فرض خلف باطل و حکم برقرار است.

۶.۱۱) ۱) فرض کنیم $C=\{c_1,\dots,c_k\}$, $h\in H_i$ اگر $vcdim(H_i)=d\geq 3$

یک k به طوریکه $(k<d)$ انتخاب می‌کنیم که C، $\cup H_i$ را shatter می‌کند

حال با توجه به اینکه $\tau_{\cup H}(k)=2^k$ $(k<d)$

$$\tau_{\cup H}(k)=\max_{C\subseteq X:|C|=k}|H_C|=\max\left|\bigcup_{i=1}^{r}H_{iC}\right|\leq\sum_{i=1}^{r}\max|H_{iC}|$$

$$\Rightarrow \boxed{\tau_{\cup H_i}(k)\leq\sum_{i=1}^{r}\tau_{H_i}(k)}\ *$$

با توجه به Sauer's Lemma:

$$\tau_{H_i}(k)\leq\left(\frac{ek}{d}\right)^d\leq\left(\frac{ek}{3}\right)^d\leq k^d\ **$$

با توجه به * و ** داریم:

$$\tau_{\cup H_i}(k)\leq\sum_{i=1}^{r}k^d=rk^d\rightarrow\tau_{\cup H_i}(k)\leq rk^d$$

همانطور که داشتیم $\tau_{\cup H_i}(k)=2^k$، با توجه به رابطه قبلی:

$$2^k\leq rk^d\Rightarrow k\leq\log_2 r+d\log_2 k$$

با توجه به لم A.2

$$k\leq\log_2 r+d\log_2 k\Rightarrow k\leq 2\log_2 r+4d\log_2 2d$$

چون $|C|=k$, C توسط $\cup H_i$ shatter می‌شود پس:

$$vcdim(\cup H_i)\leq|C|=k$$

$$\Rightarrow \boxed{vcdim(\cup H_i)\leq k\leq 2\log_2 r+4\log_2 2d}$$

۱۱.۶) ط) اگر $H=H_1 \cup H_2$ ، $k \geq 2d+2$ ، $\tau_H(k) < 2^k$

داریم:

$$\Rightarrow \tau_H(k) \leq \tau_{H_1}(k) + \tau_{H_2}(k)$$

$$\leq \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{i}$$

$$= \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{k-i}$$

$$= \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=k-d}^{k} \binom{k}{i}$$

$$\leq \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=d+2}^{k} \binom{k}{i}$$

$$< \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=d+1}^{d} \binom{k}{i}$$

$$= \sum_{i=0}^{k} \binom{k}{i} = \boxed{2^k} \Rightarrow \tau_H(k) < 2^k$$

پس H، مجموعه‌ای با اندازه‌ی $2d+2$ را shatter نمی‌کند، در نتیجه:

$$\underline{vcdim(H) \leq 2d+1}$$