① می‌دانیم $S=\{(x_i,f(x_i))\}_{i=1}^{m} \subseteq (\mathbb{R}^d \times \{0,1\})^m$ و همچنین می‌دانیم $\Leftarrow$ $\exists P_s,\ h_s(x)=1 \Leftrightarrow P_s(x)\geq 0$

حال چند جمله‌ای P_s را به شکل زیر تعریف می‌کنیم (h_s به ERM انتخاب شده است) :

$$A=\{x_i\in S \mid f(x_i)=1\}$$

$$P_s(x) = -(x-x_1)\times\cdots\times(x-x_i) \quad \text{که در آن } x_i\in A$$

حال به وضوح در چند جمله‌ای $P_s(x)$ در مجموعه‌ی A برابر 0 است و $h_s(x)$ در A برابر 1 است و P_s در سایر نقاط برابر با $n<0$ است.

② با توجه به خاصیت خطی بودن امید ریاضی داریم :

$$E_{S\sim D^m}[L_s(h)] = E_{S\sim D^m}\left[\frac{1}{m}\sum_{i=1}^{m} 1_{[h(x_i)\neq f(x_i)]}\right]$$

$$=\frac{1}{m}\sum_{i=1} E_{x_i\sim D}\left[1_{[h(x_i)\neq f(x_i)]}\right] = \frac{1}{m}\times m\times L_{(D,f)}(h)$$

$$=L_{(D,f)}(h)$$

③ فرض می‌کنیم که A کوچکترین مستطیل است که تمام نمونه‌ها مثبت t_s را در برمی‌گیرد. پس اگر A^* را کوچکتر انتخاب کنیم تعدادی از نمونه‌های مثبت در بر گرفته نمی‌شوند پس $L_s(A)<L_s(A^*)$ و اگر A^{**} بزرگتر انتخاب کنیم در نتیجه نمونه‌های مثبت بیشتری را در برمی‌گیرند که داخل t_s برچسب 0 دارند پس $L_s(A)<L_s(A^{**})$ و در نتیجه A، ERM است.

۳) ۲) با استفاده از راهنمای سوال تعریف می کنیم $R^* = R(R_1^*, R_2^*, R_3^*, R_4^*)$ همچنین R_1 و R_2 و R_3 و R_4 را نیز به شکل $R(a_1^*, a_1, a_2^*, b_2^*)$ و $R(b_1, b_1^*, a_2^*, b_2^*)$ و $R(a_1^*, b_1^*, a_2^*, a_2)$ و $(a_1^*, b_1^*, b_2, b_2^*)$ تعریف می کنیم که احتمال انتخاب هرکدام $\epsilon/4$ است. حال $R(S)$ را با الگوریتم $A(S)$ (قسمت الف) فرض کنیم حال $R(S) \subseteq R^*(S)$ برای هر S است. پس در نتیجه داریم که

$$L_{(D,f)}(R(S)) = D(R^* - R(S))$$

حال طبق فرض می دانیم احتمال نمونه ها در R_i ها $\epsilon/4$ است. حال F_i ها را به شکل زیر می نویسیم:

$$F_i = \{ S|_m : S|_m \cap R_i = \emptyset \}.$$

i from 1 to 4

طبق شکل راهنمای سوال (شکل ۲.۲) واضح است که $R(S) \subseteq \cup F_i$ پس داریم:

$$D^m(\{S : L_{(D,f)}(A(S)) > \epsilon\}) \leq D^m(\cup F_i) \leq \sum_{1}^{4} D^m(F_i)$$

می دانستیم طبق فرض وجود نمونه در R_i، $\epsilon/4$ است پس وجود نمونه در F_i، $1-\epsilon/4$ است

پس داریم:

$$\sum_{1}^{4} D^m(F_i) = 4 \times (1-\epsilon/4)^m = 4e^{-m\epsilon/4}$$

$$\Rightarrow D^m(\{S : L_{(D,f)}(A(S)) > \epsilon\}) \leq 4e^{-m\epsilon/4} < \delta =$$

$$\Rightarrow 4e^{-m\epsilon/4} < \delta \Rightarrow 4e^{\frac{m\epsilon}{4}} > \frac{1}{\delta} \Rightarrow e^{\frac{m\epsilon}{4}} > \frac{4}{\delta} \Rightarrow \boxed{m > \frac{4\log\frac{4}{\delta}}{\epsilon}}$$

Scanned with CamScanner