EPISTEMOLOGY AND THEORY OF MACHINE LEARNING

EXERCISE SET #1

The following is an adaptation of exercise 2.3 (items 1 and 2) in the textbook. Let the class of axis aligned rectangles, the algorithm A, and the rectangles R^* and $R(S)$ as defined there.

We want to show that for training sample of size $m \geq \frac{4\log(4/\delta)}{\epsilon}$, we have

$$\mathcal{D}^m(\{S|_x : L_{\mathcal{D},f}(R(S)) \leq \epsilon\}) \geq 1 - \delta.$$

(1) Show that A is an ERM.
(2) Show that for any training sample S, we have that $R(S) \subseteq R^*$.
(3) Suppose that $\mathcal{D}$ is such that $\mathcal{D}(R^*) < \epsilon$. Show that in that case, we trivially have the result.
(4) Thus suppose that $\mathcal{D}(R^*) \geq \epsilon$. Define $R_1, \ldots, R_4$ as in the textbook. Reason that if S contains an instance in *each* of the R_i, then we have

$$L_{\mathcal{D},f}(R(S)) \leq \epsilon.$$

(It might help to draw a picture.)
(5) Let $F_i = \{S|_x : S|_x \cap R_i = \emptyset\}$ be the event that the sample S contains *no* instances in R_i. Reason from (4) that

$$\mathcal{D}(\{S|_x : L_{\mathcal{D},f}(R(S)) > \epsilon\}) \leq \mathcal{D}^m(\cup_{i=1}^4 F_i).$$

(6) Rewrite this bound by upper bounding each F_i and using the union bound.
(7) Show how from this bound the desired result follows.