Advanced Machine Learning: Homework Problem Set I

Guidelines: You have to submit hardcopy of the solutions (printed or handwritten) by February 14, 2018 begining of lecture class. Write your name and roll number clearly on top of the solution. Be clear and precise in your solution.

Problem 1:

We consider a two distribution variant of the PAC model in which the learning algorithm may explicitly request positive examples and negative examples, but must find a hypothesis that performs well on both the distributions of positive and negative examples.

We say that an algorithm A PAC learns a hypothesis class $\mathcal{H}$ in the two distribution variant of the PAC model if for any target concept $c \in \mathcal{H}$, for any distribution $\mathcal{D}^+$ over the positively labeled instances, distribution $\mathcal{D}^-$ over the negatively labeled instances, and for any given (ϵ, δ), if A is given access sufficiently large number (finite) of positive and negative examples that are i.i.d. in $\mathcal{D}^+$, $\mathcal{D}^-$, then A outputs a hypothesis $h \in \mathcal{H}$ such that with probability at least $1-\delta$, $Pr_{x\sim\mathcal{D}^+}[h(x)=0] \leq \epsilon$ and $Pr_{x\sim\mathcal{D}^-}[h(x)=1] \leq \epsilon$.

(a) Prove that if $\mathcal{H}$ is PAC learnable using the basic (one distribution) model, then it is also PAC learnable using the two distribution model.

(b) Let h_0 be a function that always outputs 0, and h_1 be a function that always outputs 1. Prove that if a hypothesis class $\mathcal{H}$ is PAC learnable using the two distribution model, then the hypothesis class $\mathcal{H} \cup \{h_0, h_1\}$ is PAC learnable in the basic one distribution model.

Problem 2:

Let $X = \mathbb{R}^2$ be the domain and $Y = \{0, 1\}$ be the label set of a learning problem. Let $\mathcal{H} = \{h_r, r \in \mathbb{R}_+\}$ be the set of hypothesis corresponding to all concentric circles in the plane that classify as

$$h_r(x) = \begin{cases} 1 & ||x||_2 \leq r \\ 0 & \text{otherwise} \end{cases}$$

Prove that under realizability assumption $\mathcal{H}$ is PAC learnable with sample complexity

$$m_{\mathcal{H}}(\epsilon, \delta) \leq \frac{1}{\epsilon} \log \frac{1}{\delta}$$

.