CHENNAI MATHEMATICAL INSTITUTE

FML

Date: Sep 11, 2024
Due: 23 Sep, 2024

(1) Show for a non negative random variable X

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X \geq \nu) d\nu$$

(2) Define the KL divergence between two Bernoulli distributions to be

$$KL^+(p||q) = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q}$$

Show that $KL^+(p||q) \geq 2(p-q)^2$.

(3) Let $f : X \to \{0,1\}$ where $X \subseteq \mathbb{R}^2$. Suppose the hypothesis class is the set of all concentric circles centred at the origin. Is this class agonstic-PAC learnable? What about PAC learnability? The same question for concentric spheres?

(4) Consider learning the concept class of conjunctions of at most n-boolean literals, $x_1, \ldots, x_n, \overline{x_1}, \ldots, \overline{x_n}$. A positive example for f a conjunction of literals is a setting of the n-variables $x_1, \ldots, x_n$ which makes f true. Is the hypothesis class of conjunction of Boolean literals is PAC learnable? Agnostic PAC learnable? What is the sample complexity of learning this class?

(5) Consider learning axis parallel rectangles when the input data is subject to noise. Points which are labeled 0 are not subject to noise. Points which are labeled 1 are subject to noise - the label is flipped to 0 with probability $\eta < 1/2$. η is not known to the learner but the learner knows an upper bound η', with $\eta < \eta' < 1/2$. Show that we can still agnostic PAC learn in this situation.

(6) Let $\mathcal{C}$ be a concept class consisting of 10 concepts $c_0, \ldots, c_9$, Let the domain $X = \{1, 2, \ldots, 999\}$. n is in concept class c_i if the decimal representation of n has the integer i. What is the VC dimension of this concept class?

(7) Let $X_1, X_2, \ldots, X_m$ be i.i.d random variables which take values in $[0,1]$ with mean μ. Define $\overline{X} = \frac{1}{m}\sum_{i=1}^{i=m} X_i$. For every ϵ with $\mu + \epsilon < 1$ show that

$$\mathbb{P}(\overline{X} \geq \mu + \epsilon) \leq e^{-mKL^+(\mu+\epsilon||\mu)}$$

Deduce Hoeffdings bound from this.

(8) Show that when q is fixed, $KL^+(p||q)$ is a convex function of p and when p is fixed it is a convex function of q.

(9) Let p and q be two functions on a hypothesis space $\mathcal{H}$ such that $p(h) \in (0,1)$ and $q(h) \in (0,1)$. Let Q be a distribution on the

hypothesis space. Show that

$$KL^+(\mathbb{E}_{h\sim Q}[p(h)]||E_{h\sim Q}[q(h)]) \leq \mathbb{E}_{h\sim Q}[KL^+(p(h)||q(h))]$$