Mini Project 2

CS536, Spring 2015

Due: Feb 22, 2015

1 Problem

Consider the family of concepts in a 2D Euclidean plane $X = \Re^2$ consisting of concentric circles, $c = \{(x, y) : x^2 + y^2 \leq r^2\}$ for some $r \in \Re$. Show that this class can be (ϵ, δ)-PAC-learned from training data of size $N \geq (1/\epsilon)\log(1/\delta)$.

2 Problem

For important questions, President Mouth relies on expert advice. He selects an appropriate advisor from a collection of $H = 2,800$ experts.

1. Assume that laws are proposed in a random fashion independently and identically according to some distribution P determined by an unknown group of senators. Assume that President Mouth can find and select an expert senator out of H who has consistently voted with the majority for the last $N = 200$ laws. Give a bound on the probability that such a senator incorrectly predicts the global vote for a future law. What is the value of the bound with 95% confidence?

2. Assume now that President Mouth can find and select an expert senator out of H who has consistently voted with the majority for all but $N' = 20$ of the last $N = 200$ laws. What is the value of the new bound?

3 Problem

Consider the following decision rule for a 2-category one-dimensional problem: Decide ω_1 when $x > \theta$, otherwise decide ω_2.

1. Show that the probability of error of this rule is

$$P(error) = P(\omega_1)\int_{-\infty}^{\theta} p(x|\omega_1)dx + P(\omega_2)\int_{\theta}^{\infty} p(x|\omega_2)dx.$$

2. By differentiating, show that a necessary condition to minimize $P(error)$ is that θ satisfies

$$p(\theta|\omega_1)P(\omega_1) = p(\theta|\omega_2)P(\omega_2).$$

3. Does this equation define θ uniquely?

4. Give an example where a value of θ satisfying the above condition actually maximizes the probability of error!

4 Problem

Suppose we replace the deterministic decision rule $\alpha(x)$ by the stochastic rule, namely the one giving the probability $P(\alpha_i|x)$ of taking α_i upon observing x.

1. Show that the resulting risk is given by

$$R(\alpha) = \int \left[\sum_{i=1}^{a} R(\alpha_i|x) P(\alpha_i|x) \right] p(x) dx.$$

2. Show that R is minimized by choosing $P(\alpha_i|x) = 1$ for the action α_i associated with the minimum conditional risk $R(\alpha_i|x)$, thereby showing that no benefit can be gained from randomizing the best decision rule.

3. Can one gain from the randomizing a suboptimal rule?

5 Problem

It is often useful (and only feasible) to establish bounds on error instead of computing the explicit Bayes error rate.

1. Show that for any two non-negative numbers a and b

$$\min[a, b] \leq \sqrt{ab}.$$

2. Use this to show the error rate for a binary Bayes classifier must satisfy

$$P(error) \leq \sqrt{P(\omega_1)P(\omega_2)}\rho \leq \frac{1}{2}\rho,$$

where ρ is the so-called *Bhattacharyya coefficient*

$$\rho = \int p(x|\omega_1)^{1/2} p(x|\omega_2)^{1/2} dx.$$

3. Calculate the Bhattacharyya error bound for two multivariate Gaussian densities with the unit covariance and uniform priors.