ELEG/CISC 867 Advanced Machine Learning

Homework 3 University of Delaware
Handout: May 9, 2019 Due: May 16, 2019

PROBLEM 1 (STABILITY AND LEARNABILITY; 20 PTS). Consider a learning algorithm A. We say that A

- is on-average-replace-one-stable with rate $\epsilon(n)$ if

$$\mathbb{E}_{Z^{n+1}\sim P^{n+1}, I\in U[1:n]}[\ell(A(Z^n_{I/n+1}), Z_I) - \ell(A(Z^n), Z_I)] \leq \epsilon(n), \quad \forall P;$$

- is an AERM (Asymptotic Empirical Risk Minimizer) over hypothesis class $\mathcal{H}$ with rate $\epsilon(n)$ if

$$\mathbb{E}_{Z^n\sim P^n}\left[L(A(Z^n), Z^n) - \min_{h\in\mathcal{H}} L(h, Z^n)\right] \leq \epsilon(n), \quad \forall P;$$

- learns hypothesis class $\mathcal{H}$ with rate $\epsilon(n)$ if

$$\mathbb{E}_{Z^n\sim P^n}\left[L(A(Z^n), P) - \min_{h\in\mathcal{H}} L(h, P)\right] \leq \epsilon(n), \quad \forall P.$$

Show that if learning algorithm A is on-average-replace-one-stable with rate $\epsilon_1(n)$ and is an AERM over hypothesis class $\mathcal{H}$ with rate $\epsilon_2(n)$ then it learns class $\mathcal{H}$ with rate $\epsilon_1(n)+\epsilon_2(n)$.

PROBLEM 2 (MARGIN AND PERCEPTRON; 20 PTS). Assume the training dataset $z^n = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ is separable by homogenous hyperplanes $\mathbf{w}^T\mathbf{x}, \mathbf{w} \in \mathbb{R}^d$.

(a) Write down the expression of the margin γ associated with a homogenous separating hyperplane $\mathbf{w}$. [10 pts]

(b) Show that if there exists a homogenous separating hyperplane with margin γ and all the instances from the training dataset are within a ball of radius ρ, then the Batch Perceptron algorithm, when applied to z^n, will run at most $(\rho/\gamma)^2$ steps. [10 pts]

PROBLEM 3 (KERNEL RIDGE REGRESSION; 30 PTS). Solve Problem 16.3 on Pages 188-189 of the textbook.