Boston University
Department of Electrical and Computer Engineering

ENG EC 414 Introduction to Machine Learning

HW 2

Issued: Tue 8 Sept 2020 **Due:** 4:55pm Tue 15 Sept 2020 on Gradescope

---

Important: Before you proceed, please read the documents pertaining to *Homework formatting and submission guidelines* in the Homeworks section of Blackboard.

Note: Problem difficulty = number of coffee cups ☕

Problem 2.1 [8pts] *(max, argmax, min, argmin)* Let $f(x) := (2 + \sin(2\pi x))$ and $\mathcal{A} := [0, 2]$.

(a) [2pts] Compute: $\max_{x\in\mathcal{A}} f(x)$ and $\operatorname*{argmax}_{x\in\mathcal{A}} f(x)$.

(b) [2pts] Compute: $\min_{x\in\mathcal{A}} f(x)$ and $\operatorname*{argmin}_{x\in\mathcal{A}} f(x)$.

(c) [2pts] Compute: $\operatorname*{argmax}_{x\in\mathcal{A}} \left[2 \cdot e^{-f(x)+5} - 7\right]$.

(d) [2pts] Compute: $\operatorname*{argmin}_{x\in\mathcal{A}} [11 - \ln(2 \cdot f(x) + 5)]$.

Problem 2.2 [16pts] *(Training errors for constant predictors)* A dataset has $n = 12$ examples with feature vectors in $\mathbb{R}^{10}$ and the following labels $\{y_1 = 2.0, y_2 = 1.0, y_3 = 5.0, y_4 = 3.0, y_5 = 2.0, y_6 = 4.0, y_7 = 1.0, y_8 = 4.0, y_9 = 6.0, y_{10} = 4.0, y_{11} = 2.0, y_{12} = 6.0\}$. We consider constant predictors with different loss functions.

(a) [4pts] *(Empirical average zero-one loss)*
Let $F_{\text{error}}(y) := \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}[y \neq y_j]$, where $\mathbf{1}[A]$ is 1 if A is true and 0 otherwise. Compute the argmin.

(b) [5pts] *(Empirical average squared-error loss)*
Let $F_{\text{MSE}}(y) := \frac{1}{n}\sum_{j=1}^{n}(y - y_j)^2$. Compute the argmin.

(c) [7pts] ☕☕ *(Empirical average absolute-error loss)*
Let $F_{\text{MAE}}(y) := \frac{1}{n}\sum_{j=1}^{n} |y - y_j|$. Compute the argmin.

Problem 2.3 [12pts] *(Argmin for random variables)*

(a) [5pts] ☕ *(Expected squared-error loss given* **x***)* Suppose $p(y|\mathbf{x})$ denotes the conditional pdf/pmf of random variable Y (modeling labels) given that random vector $\mathbf{X}$ (modeling feature vectors) equals $\mathbf{x}$. Let $L_{\text{MSE}}(y|\mathbf{x}) := E[(Y - y)^2 \mid \mathbf{X} = \mathbf{x}]$, i.e., the conditional expectation of the squared-error loss given $\mathbf{X} = \mathbf{x}$. Express $h(\mathbf{x}) := \operatorname*{argmin}_{y\in\mathbb{R}} L_{\text{MSE}}(y)$ in terms of $p(y|\mathbf{x})$.

(b) [7pts] ☕☕ *(Expected absolute-error loss given* **x***)* Suppose $p(y|\mathbf{x})$ denotes the conditional pdf/pmf of random variable Y (modeling labels) given that a random vector $\mathbf{X}$ (modeling feature vectors) equals $\mathbf{x}$. Let $L_{\text{MAE}}(y|\mathbf{x}) := E[|Y - y| \mid \mathbf{X} = \mathbf{x}]$, i.e., the conditional expectation of the absolute-error loss given $\mathbf{X} = \mathbf{x}$. Express $h(\mathbf{x}) := \operatorname*{argmin}_{y\in\mathbb{R}} L_{\text{MAE}}(y)$ in terms of $p(y|\mathbf{x})$.