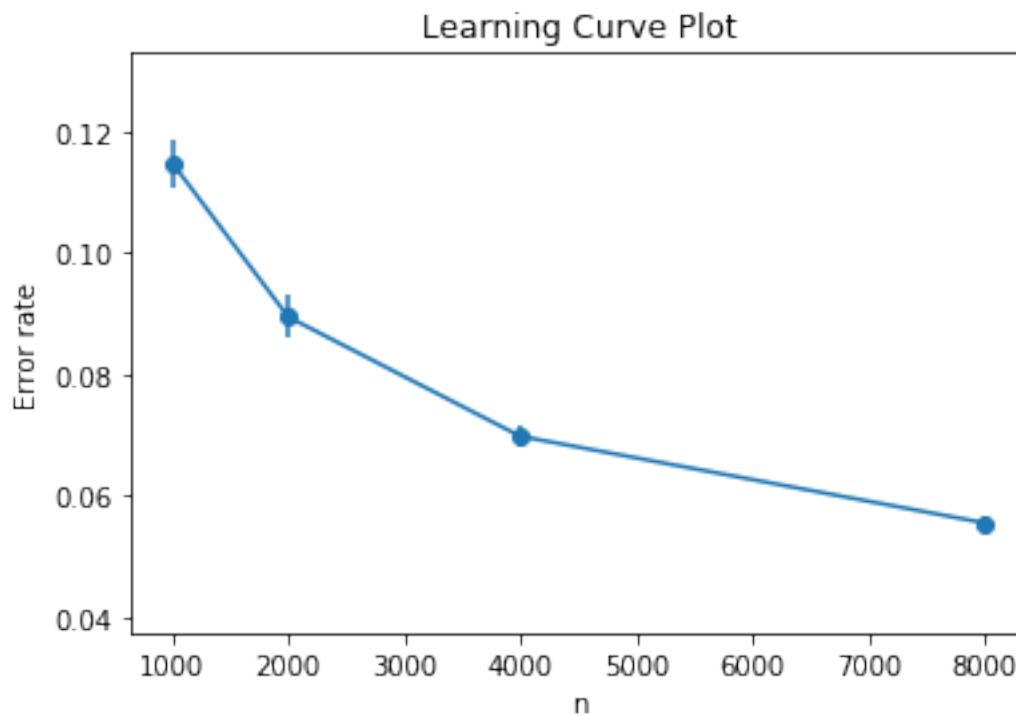# COMS 4771-2 Fall 2018 Homework 1

Zhuangyu Ren(zr2209)

September 27, 2018

Problem 1

(a) I separately counted the distance between the test points and all the data points, and then using "argmin" to count the min distance from every test point to the data.

(b) The curve is as below

(c) The training error rate will always be 0.

Problem 2

(a)

$$\int_0^{+\infty} cx^2 e^{-\frac{x}{\theta}} dx = 1$$

$$a = \frac{1}{2\theta^3}$$

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} \frac{1}{2\theta^3} x_i^2 e^{-\frac{x_i}{\theta}}$$

$$ln\mathcal{L}(\theta) = \sum_{i=1}^{n}(2lnx_i - 3ln\theta - \frac{x_i}{\theta} + 3ln2)$$

$$\theta = \frac{\sum_{i=1}^{n} x_i}{3n}$$

(b)

$$\int_{0}^{\theta} 1dx = 1$$

$$\theta = \frac{1}{x}$$

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \frac{1}{\theta^n} I(x_1, ...x_n)$$

$\varphi(\theta) = 0$ if $\theta < max(x_1, ...x_n)$ and $\varphi(\theta) = \frac{1}{\theta^n}$ if $\theta \geq max(x_1, ...x_n)$
therefore $\widehat{\theta} = max(x_1, ...x_n)$

(c)

$$\int ce^{-\frac{(x-\theta)^2}{2\theta^2}} = 1$$

$$c = \frac{1}{\theta\sqrt{2\pi}}$$

2

$$\ln f_\theta(x) = 20\ln(\frac{1}{\theta\sqrt{2\pi}}) - \frac{20\theta^2 - 42\theta + 41}{2\theta^2}$$

$$\frac{-20}{\theta} + \frac{41}{\theta^3} - \frac{21}{\theta^2} = 0$$

$\theta = 1$ or $\theta = -\frac{41}{20}$

(d)
$$\mathcal{L}(\theta) = (\frac{1}{2})^8 * e^{4\theta} * e^{7\theta} * (\frac{1}{2} - 3e^\theta)$$

$$\mathcal{L}(\theta) = \frac{1}{4}e^{11\theta} - \frac{2}{3}e^{12\theta}$$

$\theta < -\ln 6$, $\mathcal{L}(\theta)$ do not have a max value

Problem 3

(a)
$$\mu = \frac{\sum_{i=1}^n 1\{y_i = y\} x_{i,j}}{\sum_{i=1}^n 1\{y_i = y\}}$$

(b) 20 classes: training error rate: 0.216257
   20 classes: test error rate: 0.376016

(c) 2 classes: training error rate: 0.0577939
   2 classes: test error rate: 0.131383

(d) largest words are ['firearms', 'occupied', 'israelis', 'argic', 'serdar', 'melkonian', 'sahak',
   'appressian', 'ohanus', 'villages', 'cramer', 'armenia', 'cpr', 'sdpa', 'palestine', 'optilink',
   'handgun', 'firearm', 'budget', 'arabs']
   smallest words are ['prophet', 'mozumder', 'jesus', 'benedikt', 'ksand', 'alink', 'believ-
   ers', 'theology', 'scriptures', 'wpd', 'solntze', 'atheist', 'livesey', 'testament', 'revela-
   tion', 'teachings', 'clh', 'atheists', 'atheism', 'athos']