CSE 512 Spring 2021 - Machine Learning - Homework 3

Your Name: Irfan Ahmed
Solar ID : 113166464
NetID Email: irfan.ahmed@stonybrook.edu

1.) 1.1.1.) For a binary classification problem,

$$r^*(x) = \min(\text{loss}(\hat{y}=0 \mid x=1)\, P(x=1),\ \text{loss}(\hat{y}=1 \mid x=0)\, P(x=0))$$

$P(x=1) = \eta(x)$, $P(x=0) = 1-\eta(x)$

$\text{loss}(\hat{y}=1 \mid x=0) = \alpha$, $\text{loss}(\hat{y}=0 \mid x=1) = 1$

$$r^*(x) = \min(1 \cdot (\eta(x)),\ \alpha(1-\eta(x)))$$

Optimal Bayes risk $r^*(x) = \min(\eta(x),\ \alpha(1-\eta(x)))$

1.1.2.) asymptotic risk:

$$r(x) = \underset{\text{Prob of +ve class}}{\eta(x)}\ \underbrace{\alpha(1-\eta(x))}_{\substack{\text{Cost of false positive} \\ \times \text{ prob of -ve class}}} + \underbrace{(1-\eta(x))}_{\text{Prob of -ve class}}\ \underbrace{\eta(x)}_{\text{Prob of +ve class}}.$$

1.1.3.) $r(x) = \alpha\, \eta(x)(1-\eta(x)) + (1-\eta(x))\,\eta(x)$.

$r(x) = (1+\alpha)\,\eta(x)(1-\eta(x)) \quad \rightarrow ①$

$r(x) \leq (1+\alpha)\,\alpha\, \eta(x)(1-\eta(x)) \quad (\because \alpha > 1)$

$r(x) \leq (1+\alpha)\,\eta(x)\,\alpha(1-\eta(x))$

$r(x) \leq (1+\alpha)\, r^*(x)(1-r^*(x))$ (From symmetry)

1.1.4) $R = E(r(x))$, $R^* = E(r^*(x))$

$$r(x) \leq (1+\alpha)\, r^*(x)(1 - r^*(x))$$

$$r(x) \leq (1+\alpha)\left(r^*(x) - (r^*(x))^2\right)$$

$$E(r^{*2}(x)) = (E(r^*(x)))^2 + \mathrm{Var}(r^*(x)) \geq E^2(r^*(x)) \Rightarrow E((r^*(x))^2) \geq E^2(r^*(x))$$

$$E(r(x)) \leq (1+\alpha)\left(E(r^*(x)) - E((r^*(x))^2)\right)$$

$$R \leq (1+\alpha)\left(R^* - (R^*)^2\right)$$

$$R \leq (1+\alpha)\, R^*(1 - R^*)$$

From ①

$$R \leq \frac{1+\alpha}{\alpha}\left(R^* - R^{*2}\right). \quad \underline{\underline{\alpha > 1}}$$

$$R \leq (1+\alpha) R^*(1 - R^*)$$

1.2.1) $r(x) = \eta(x)(1 - g(\eta, k)) + g(\eta, k)(1 - \eta(x))$.

1.2.2) We know, $g(\eta, k) = \sum_{i=\frac{k+1}{2}}^{k} {}^{k}C_i\, \eta^i (1-\eta)^{k-i}$

$$g(\eta, k) + g(1-\eta, k) = 1 \quad \rightarrow \left((\eta + 1 - \eta)^k = 1^k = 1\right)$$

$$r^*(x) = \min\left(\eta(x)(1 - g(\eta, k)),\ g(\eta, k)(1 - \eta(x))\right) \text{ — ①}$$

$(\because g(\eta, k) = 1 - g(1-\eta, k)$ & vice versa..$)$

$$r^*(x) = \min\left(\eta(x)\, g(1-\eta, k),\ g(\eta, k)(1 - \eta(x))\right) \text{ — ②}$$

$$r^*(x) = \min(\eta, 1-\eta).$$

$$r(x) = \eta(x) + g(\eta, k) - 2\eta(x)\, g(\eta, k). \quad (\because \text{From ①})$$

①

$$r(x) = \eta(x)\, g(1-\eta, k) + g(\eta, k)(1-\eta(x))$$

$$= \eta(x)\, g(1-\eta, k) + (1-g(1-\eta, k))(1-\eta(x))$$

$$= \eta(x) g(1-\eta, k) + 1 - \eta(x) - g(1-\eta, k) + \eta(x) g(1-\eta, k)$$

$$= 1-\eta(x) - g(1-\eta, k) + 2\eta(x) g(1-\eta, k)$$

$$= 1-\eta(x) + g(1-\eta, k) + 2\eta(x) g(1-\eta, k) - 2g(1-\eta, k)$$

$$r(x) = 1-\eta(x) + g(1-\eta, k) - 2(1-\eta(x)) g(1-\eta, k). \quad (\text{From } ②) \quad ②$$

$\Rightarrow$ $r(x)$ is same for $r^*(x)$ being $\eta(x)$ or $1-\eta(x)$.

By symmetry.

$$r(x) = r^*(x) + g(r^*, k) - 2r^*(x)\, g(r^*, k)$$

$$r(x) = r^*(x) + (1-2r^*(x))\, g(r^*(x), k).$$

1.2.3.) Hoeffding inequality,

$$P(H(n) \geq (p+\varepsilon)n) \leq \exp(-2\varepsilon^2 n).$$

Here $n = k$

$p = \eta$

we need atleast $\frac{k+1}{2}$ +ve $\Rightarrow$ $P(H(n) \geq \frac{k+1}{2}) = g(\eta, k)$

$\Rightarrow$ $(p+\varepsilon)n = \frac{k+1}{2}$

$(\eta+\varepsilon)k = \frac{k+1}{2} \Rightarrow \varepsilon = \frac{1}{2} - \eta + \frac{1}{2k}$

If $\frac{1}{2k} \to 0 \Rightarrow k \to \infty$ i.e. $\eta \to r^*$

$$g(r^*(x), k) \leq \exp(-2\varepsilon^2 n)$$

$$\varepsilon = \frac{1}{2} - r^*(x)$$

$$\therefore \; g(r^*(x), k) \leq \exp(-2(0.5 - r^*(x))^2 k).$$

1.2.4.) $r(x) = r^*(x) + (1 - 2r^*(x))\, g(r^*(x), k)$

$$r(x) - r^*(x) = (1-2r^*)\, g(r^*, k) \leq (1-2r^*)\, e^{-2(1/2 - r^*)^2 k}$$

$$r(x) - r^*(x) \leq (1-2r^*)\, e^{-(1-2r^*)^2 k/2}$$

let's find maximum of $u e^{-u^2 k/2}$ w.r.t u. (Minimum is 0)

Derivative $= 0 \Rightarrow$ $e^{-u^2 k/2} + u\, e^{-u^2 k/2} (-2uk/2) = 0$

$$1 - u^2 k = 0 \Rightarrow u^* = 1/\sqrt{k}$$

$$\mathrm{Max}(u e^{-u^2 k/2}) = \frac{1}{\sqrt{k}} \cdot e^{-u^2 k/2} = \frac{1}{\sqrt{k}} e^{-1/2} = \frac{1}{\sqrt{ke}}$$

$$\therefore \; r(x) - r^*(x) \leq \mathrm{Max}((1-2r^*)\, e^{-(1-2r^*)^2 k/2}$$

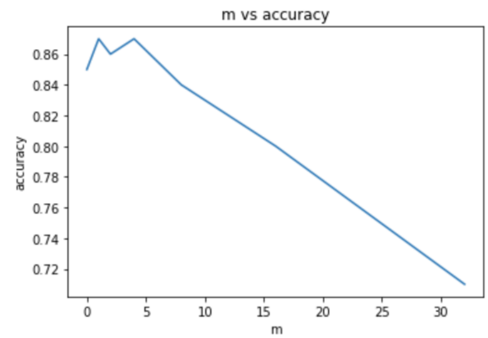
$$r(x) - r^*(x) \leq \frac{1}{\sqrt{ke}} \leq \frac{1}{\sqrt{2k}} \qquad (e \approx 2.81)$$

$$r(x) \leq r^*(x) + \frac{1}{\sqrt{2k}}$$

As $k \to \infty$, $r(x) \to r^*(x)$

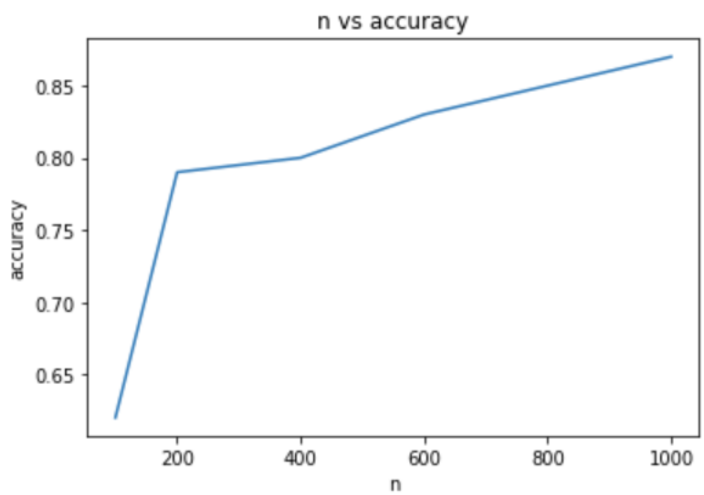
Bayes risk.

```
[0, 1, 2, 4, 8, 16, 32] [0.85 0.87 0.86 0.87 0.84 0.8  0.71]
```



2.2.1) With increasing k value, the accuracy is getting lowered. In the beginning with increasing k value, accuracy was around 0.85 and then instantaneously decreased for m = 16 and kept decreasing for greater values of m.

```
[100, 200, 400, 600, 800, 1000] [0.62 0.79 0.8  0.83 0.85 0.87]
```



2.2.2) With increasing training data, the accuracy is increasing and settles at 0.87.

```
0.83
```

2.2.3) The accuracy is lower than accuracy using euclidean distance. The accuracy obtained is 0.83 whereas using euclidean we get 0.87 for k=3.

```
2.2.4)
```

```
------------
```


```
------------
```


```
------------
```