HW 2 – Machine Learning Introduction

All code is present in the GitHub repo
https://github.com/somesh-bagadiya/CMPE257-Fall23-Somesh-Bagadiya.git
Folder Name is HW 2

Task 1 HP 1

Task 1, Hp1

→ $\prod_{n=1}^{N} = P(y_n | x_n)$ is the likelihood for Linear Regⁿ

Now, Let's maximize it.

Maximizing the likelihood is equivalent to Maximizing the below mentioned equation.

$$\frac{1}{N} \ln \left(\prod_{n=1}^{N} \theta (y_n \, w^T x_n) \right) \quad [max]$$

We are using $\frac{1}{N}(\ln)$ because maximizing the ~~above~~ Likelihood will have monotonically increasing graph. and the above mentioned eqⁿ will also have monotonically increasing graph

But we don't want to maximize the in sample error, we want to minimize it.

∴ New eqⁿ will be.

$$-\frac{1}{N} \ln\left(\prod_{n=1}^{N} \theta\left(y_n w^T x_n\right)\right)$$

$$\Rightarrow \frac{1}{N} \ln\left[\prod_{n=1}^{N} \left(\frac{1}{\theta(y_n w^T x_n)}\right)\right]$$

Since its Log/ln we can write above mentioned equation as

$$\Rightarrow \frac{1}{N} \sum_{n=1}^{N} \ln\left(\frac{1}{\theta(y_n w^T x_n)}\right)$$

Now, $\underset{\text{(signal)}}{\theta(y_n w^T x_n)} = \frac{1}{1 + e^{-y_n w^T x_n}}$

thus,

$$\Rightarrow \frac{1}{N} \sum_{n=1}^{N} \ln\left(1 + e^{-y_n w^T x_n}\right) = E_{in}(w)$$

We know that likelihood is directly proportional to signal while Ein is inversely proportional to signal. Hence, we can state that maximizing likelihood will minimize in sample error.

Task 1 HP 2

Task 1 , HP 2

$\rightarrow \nabla E_{in}(w(t)) = ?$

$$E_{in}(w) = \frac{1}{N} \sum_{n=1}^{N} \ln\left(1 + e^{-y_n w^T x_n}\right)$$

$$\nabla E_{in} = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{(1+e^{-y_n w^T x_n})} \cdot \left(1 + e^{-y_n w^T x_n}\right)'$$

$$= \frac{1}{N} \sum_{n=1}^{N} \frac{1}{(1+e^{-y_n w^T x_n})} \left[0 + \left(-y_n x_n\right) e^{-y_n w^T x_n}\right]$$

$$= \frac{1}{N} \sum_{n=1}^{N} \frac{1}{1+e^{-y_n w^T x_n}} \cdot \left(-y_n x_n e^{-y_n w^T x_n}\right)$$

$$= -\frac{1}{N} \sum_{n=1}^{N} y_n x_n \left(\frac{e^{-y_n w^T x_n}}{1+e^{-y_n w^T x_n}}\right)$$

$$\theta(s) = \frac{e^s}{1+e^s} = \frac{1}{1+e^{-s}}$$

$$\therefore \nabla E_{in}(w) = -\frac{1}{N} \sum_{n=1}^{N} y_n x_n \, \theta(y_n w^T x_n)$$

Task 4

PAGE No.
DATE / /

Task 4

	+1	-1
+1	0	Ca
-1	Cr	0

a) i) Cost Accept $= (1 - g(x))\, Ca$

complete classificat^n for Accept is True positive and False Positeve

$= g(x) \times TP + (1 - g(x)) \times FP$

$= g(x) \times 0 + (1 - g(x)) \times Ca = \underline{\underline{(1 - g(x))\, Ca}}$

~~b)~~ ii) Similarly

Cost Reject $= g(x) \times Cr + (1 - g(x)) \times 0$

$= \underline{\underline{(g(x) \cdot Cr)}}$

b) Threshold is K, and for the system to be fair Cost Accept = Cost Reject

$$(1 - g(x)) \cdot Ca = g(x) \cdot Cr$$

$$\left(\frac{1}{g(x)} - 1\right) = \frac{Cr}{Ca}$$

$$\frac{1}{g(x)} = \frac{Cr + Ca}{Ca} \Rightarrow g(x) = \frac{Ca}{Cr + Ca} = K$$

c) Super market, $C_u = 1$, $C_r = 10$

$$K = \frac{1}{1+10} = \frac{1}{11}$$

Super market will easyly accept fingerprints of the customer

CIA, $C_u = 1000$, $C_r = 1$

$$K = \frac{1000}{1000+1} = \frac{1000}{1001}$$

CIA will reject almost every fingerprint untill it is very confident about the agent