CS771 Introduction to Machine Learning
Assignment 1

Bharat
22111074
bharat22@iitk.ac.in

Pradeep Chalotra
22111045
pchalotra22@iitk.ac.in

Pulkit Sharma
22111048
Pulkits22@iitk.ac.in

Sarthak Neema
22111079
sarthakn22@iitk.ac.in

Sumit Kumar Chaudhary
22111060
sumitkc22@iitk.ac.in

Problem 1:-

Given that, we have to map the binary digits 0 , 1 to sign -1 , +1 and it can be shown by function as follows :

$$m\{0,1\} \rightarrow \{-1,+1\}$$

To find the equation of above function, lets assume that

$$m(x) = ax + b \tag{1}$$

$$\text{m}\{0{,}1\} \rightarrow \{-1,1\}$$

Mapping values of m we get,

m(0) = -1 & m(1) = 1

$$m(0) = a * 0 + b = -1$$
$$\therefore b = -1 \tag{2}$$
$$m(1) = a * 1 + b = +1$$
$$a + b = 1$$
$$a - 1 = 1$$
$$\therefore a = 2 \tag{3}$$

From eq. (1) , (2) and (3 , We get , m(x) = 2x-1

The equation of function $m\{0,1\} \rightarrow \{-1,+1\}$ is

$$m(x) = 2x - 1$$

Now, we have to map the signs to the bits (0 & 1) which is shown by function below,

F{-1 , +1} $\rightarrow \{0,1\}$

Let us assume the equation of function F,

$$F(x) = ax + b \tag{4}$$

F{-1,+1} $\rightarrow \{0,1\}$

Mapping values of F we get,

F(-1) = 0 & F(1) = 1

$$F(-1) = a * (-1) + b = 0$$
$$\therefore a = b \tag{5}$$
$$m(1) = a * 1 + b = +1$$
$$a + b = 1$$
$$a + a = 1 \qquad [usingeq5]$$
$$\therefore a = \frac{1}{2} \tag{6}$$

From eq. (4) , (5) and (6 , We get ,

$$F(x) = \frac{x}{2} + \frac{1}{2}$$
$$F(x) = \frac{x+1}{2}$$

Problem 2:-

$\prod_{i=1}^{n} sign(r_i) = sign(\prod_{i=1}^{n} r_i)$

Now we will take LHS of the equation that we have to prove,

$$\prod_{i=1}^{n} sign(r_i) = sign(r_1) * sign(r_2) \cdots * sign(r_n)$$

We know that,

sign(x) = x/|x|

So,

$$\prod_{i=1}^{n} sign(r_i) = \frac{r_1}{|r_1|} * \frac{r_2}{|r_2|} \cdots * \frac{r_n}{|r_n|}$$

Also we know that,

$$|x * y| = |x| * |y|$$

Since,

$$|xy| = \sqrt{x^2 * y^2}$$
$$\Rightarrow \quad \sqrt{x^2} * \sqrt{y^2}$$
$$\Rightarrow \quad |x| * |y|$$

So,

$$\prod_{i=1}^{n} sign(r_i) = \frac{r_1 * r_2 * r_3 ------r_n}{|r_1 * r_2 * r_3 \cdots * r_n|} \tag{7}$$

Now we will take RHS of the equation,

$$sign(\prod_{i=1}^{n} r_i) = sign(r_1 * r_2 * r_3 \cdots * r_n) \Rightarrow \frac{r_1 * r_2 * r_3 ----r_n}{|r_1 * r_2 * r_3 ----- r_n|}$$

Thus,

LHS = RHS

Hence,

$$\prod_{i=1}^{n} sign(r_i) = sign(\prod_{i=1}^{n} r_i)$$

Therefore,

$$sign(\bar{u}^T \bar{x} * \bar{v}^T \bar{x} * \bar{w}^T \bar{x}) = sign(\bar{u}^T \bar{x}) * sign(\bar{v}^T \bar{x}) * sign(\bar{w}^T \bar{x})$$

Problem 3:-

$$(\tilde{\mathbf{u}}^T \tilde{\mathbf{x}}) \cdot (\tilde{\mathbf{v}}^T \tilde{\mathbf{x}}) \cdot (\tilde{\mathbf{w}}^T \tilde{\mathbf{x}})$$

We have 9 dimensional vectors mapped to D dimensional vectors as,

$$\phi(x) : \mathbb{R}^9 \rightarrow \mathbb{R}^D$$

and for any triple $(\tilde{u}, \tilde{v}, \tilde{w})$, there always exists a vector $W \in \mathbb{R}^D$, where D depends only on the number of PUFs.
Now,

$\Rightarrow \quad (\tilde{u}^T \tilde{x}) \cdot (\tilde{v}^T \tilde{x}) \cdot (\tilde{w}^T \tilde{x})$

$\Rightarrow \quad (\sum_{i=1}^{n} \tilde{u}_i \tilde{x}_i)(\sum_{i=1}^{n} \tilde{v}_i \tilde{x}_i)(\sum_{i=1}^{n} \tilde{w}_i \tilde{x}_i)$

We can also write it as follows :

$\Rightarrow \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \tilde{u}_i \tilde{v}_j \tilde{w}_k \tilde{x}_i \tilde{x}_j \tilde{x}_k$

$\Rightarrow \tilde{u}_1 \tilde{v}_1 \tilde{w}_1 \tilde{x}_1 \tilde{x}_1 \tilde{x}_1 + \tilde{u}_1 \tilde{v}_1 \tilde{w}_2 \tilde{x}_1 \tilde{x}_1 \tilde{x}_2 + \ldots \ldots \cdots + \tilde{u}_9 \tilde{v}_9 \tilde{w}_9 \tilde{x}_9 \tilde{x}_9 \tilde{x}_9$

$$\Rightarrow [u_1 v_{w1} \quad u_1 v_1 w_2 \quad \ldots \quad u_1 v_9 w_9 \quad \ldots \quad u_9 v_9 w_9] \begin{bmatrix} x_1 x_1 x_1 \\ x_1 x_1 x_2 \\ \cdots \\ x_1 x_9 x_9 \\ \cdots \\ x_9 x_9 x_9 \end{bmatrix}$$

$\Rightarrow \mathrm{w}_0 x_0 + w_1 x_1 + \ldots \cdots + w_{729} x_{729}$

$$\therefore (\tilde{u}^T \tilde{x}) \cdot (\tilde{v}^T \tilde{x}) \cdot (\tilde{w}^T \tilde{x}) = W^T \phi(\tilde{x})$$

where, W = $(\tilde{u}_1 \tilde{v}_1 \tilde{w}_1, \tilde{u}_2 \tilde{v}_2 \tilde{w}_2, \ldots \ldots, \tilde{u}_9 \tilde{v}_9 \tilde{w}_9)$

Problem 5:-

Hyperparameters :

$\rightarrow$ Learning Rate : 0.01

With a constant learning rate, each step the algorithm takes toward the minimum is of the same length. For this reason, the algorithm gets close to but never converges on the true value. With a time decaying learning rate (the right hand plot), the steps get shorter as the algorithm continues to iterate. As the steps get smaller and smaller, the algorithm converges closer and closer to the true value. Centering and scaling data beforehand is another important step for convergence.

$\rightarrow \quad \lambda : 0.0001$

We tried regularizer to be $[1, 0.1, 0.001, 0.0001, 0.00001, 0.0000001]$ and we see that $\lambda = 10^{-6}$ got the highest accuracy.

Problem 6:-

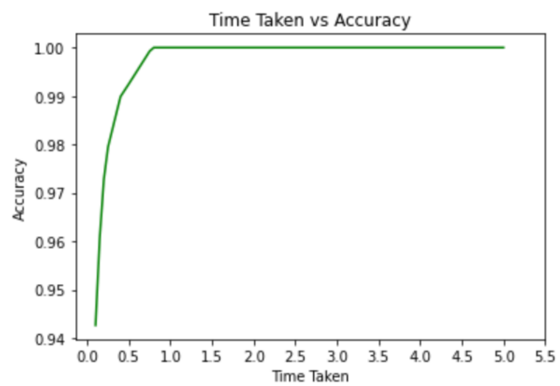

Figure 1: Convergence graph