ASSIGNMENT 1

Group no 57 - Akatsuki

Kakumanu Vamsee Krishna - 22111065
Sanket Sanjay Kale - 22111052
Pratik Mahipal Patil - 22111047
Shrawan Kumar - 22111080

1 Question 1

Given, m: is a function that maps {0,1} $\rightarrow \{-1,+1\}$
and
f: is a function that maps $\{-1,+1\} \rightarrow \{0,1\}$

Aim is to map {0,1} $\rightarrow \{-1,+1\}$ and vice-versa
so that
$\Rightarrow XOR(b_1, b_2, \cdots, b_n) = f(\prod_{i=n}^{n} m(b_i)$ for any $n \in N$
$i.e.$
$\Rightarrow XOR(b_1, b_2, \cdots, b_n) = f(m(b_1).m(b_2)\cdots.m(b_n))$

s$_i \in \{-1,+1\}$ and $b_i \in \{0,1\}$

m(b$_i$) : $s_i = 1 - 2b_i$
found through line joining (0,+1) and (+1,-1)
$m : \left\{ \substack{+1; b_i=0 \\ -1; b_i=1} \right\}$

f(s$_i$) : $b_i = (1 - s_i)/2$
found through line joining (+1,0) and (-1,+1)
$f : \left\{ \substack{0; s_i=+1 \\ 1; s_i=-1} \right\}$

say B(b$_1, b_2 \cdots, b_n) = (10101001)_2$

lets see if XOR(B) = f(m(b$_1).m(b_2).\cdots.m(b_n))$

$\Rightarrow XOR(1,0,1,0,1,0,0,1) = f(m(1).m(0).m(1).m(0).m(1).m(0).m(0).m(1))$
$\Rightarrow 0 = f((-1).(+1).(-1).(+1).(-1).(+1).(+1).(-1))$
$\Rightarrow 0 = f(+1)$
$\Rightarrow 0 = 0$
$LHS = RHS$

Hence it is proved that there exists a way to map {0,1} $\rightarrow \{-1,+1\}$ and viceversa
so that $XOR(b_1, b_2, \cdots, b_n) = f(m(b_1).m(b_2).\cdots.m(b_n))i.e.$
we can represent XOR as product by functions m and f

2 Question 2

Given,

$\prod_{i=1}^{n}(sign(r_i)) = sign(\prod_{i=1}^{n}(r_i)) \rightarrow EQ1$

sign(0) = 0
+ve is short term for positive and -ve is short term for negative

"We will Prove this by using Mathematical Induction"

For i = 1
$\Rightarrow sign(r_1) = sign(r_1)$
$\Rightarrow LHS = RHS$

For i = 2
$\Rightarrow sign(r_1).sign(r_2) = sign(r_1.r_2)$

If r_1 $is + ve\; then$
$\Rightarrow sign(r_2) = sign(r_2)$
$\Rightarrow LHS = RHS$

If r_1 $is - ve\; then$
$\Rightarrow -sign(r_2) = sign(-(r_2))$
$\Rightarrow -sign(r_2) = -sign(r_2)$
$\Rightarrow LHS = RHS$

Lets assume EQ1 is valid for i = k
i.e.
Our Hypothesis is *sign(*$r_1).sign(r_2).\cdots.sign(r_k) = sign(r_1.r_2.\cdots.r_k)$

Lets prove that EQ1 is also true for i = k+1
i.e. sign($r_1).sign(r_2).\cdots.sign(r_k).sign(r_{k+1}) = sign(r_1.r_2.\cdots.r_k.r_{k+1}) \rightarrow EQ2$

rewrite EQ2 as
$\Rightarrow [sign(r_1).sign(r_2).\cdots.sign(r_k)].sign(r_{k+1}) = sign((r_1.r_2.\cdots.r_k).r_{k+1})$

Now lets prove by analysing possible cases of "sign(r_{k+1})"

CASE 1: If sign($r_{k+1}) = 0$ then
$\Rightarrow [sign(r_1).sign(r_2).\cdots.sign(r_k)].0 = sign((r_1.r_2.\cdots.r_k).0)$
as from given $sign(0) = 0$
$\Rightarrow 0 = sign(0)$
$\Rightarrow 0 = 0$
$\Rightarrow LHS = RHS$

CASE 2: If sign(r_{k+1}) is $+ve$ then
$\Rightarrow [sign(r_1).sign(r_2).\cdots.sign(r_k)].(+ve) = sign((r_1.r_2.\cdots.r_k).(+ve))$

$\Rightarrow 3\ Subcases$

$\Rightarrow Subcase\ 1 : If (sign(r_1).sign(r_2).\cdots.sign(r_k))\ is + ve$ then
$sign((r_1.r_2.\cdots.r_k))$ is also +ve, from hypothesis
$So, \Rightarrow [+ve].(+ve) = sign((+ve).(+ve))$
$\Rightarrow +ve = sign(+ve)$
$\Rightarrow +ve = +ve$
$\Rightarrow LHS = RHS$

$\Rightarrow Subcase\ 2 : If (sign(r_1).sign(r_2).\cdots.sign(r_k))\ is - ve$ then
$sign((r_1.r_2.\cdots.r_k))$ is also -ve, from hypothesis
$So, \Rightarrow [-ve].(+ve) = sign((-ve).(+ve))$
$\Rightarrow -ve = sign(-ve)$
$\Rightarrow -ve = -ve$
$\Rightarrow LHS = RHS$

$\Rightarrow Subcase\ 3 : If (sign(r_1).sign(r_2).\cdots.sign(r_k))\ is$ 0 (due to any $r_i(i : 1$ to $k) = 0)$
then $sign((r_1.r_2.\cdots.r_k))$ is also 0 ,from hypothesis
$\Rightarrow [0].(+ve) = sign((0).(+ve))$
$\Rightarrow 0 = sign(0)$
$\Rightarrow 0 = 0$
$\Rightarrow LHS = RHS$

CASE 3: If sign(r_{k+1}*)* $is - ve$ then
$\Rightarrow [sign(r_1).sign(r_2).\cdots.sign(r_k)].(-ve) = sign((r_1.r_2.\cdots.r_k).(-ve))$

$\Rightarrow 3\ Subcases$

$\Rightarrow Subcase\ 1 : If (sign(r1).sign(r2).\cdots.sign(rk))\ is + ve$ then
$sign((r1.r2.\cdots.rk))$ is also $+\ ve$, from hypothesis
$So, \Rightarrow [+ve].(-ve) = sign((+ve).(-ve))$
$\Rightarrow -ve = sign(-ve)$
$\Rightarrow -ve = -ve$
$\Rightarrow LHS = RHS$

$\Rightarrow Subcase\ 2 : If (sign(r1).sign(r2).\cdots.sign(rk))\ is - ve$ then
$sign((r1.r2.\cdots.rk))$ is also $-\ ve$, from hypothesis
$So, \Rightarrow [-ve].(-ve) = sign((-ve).(-ve))$
$\Rightarrow +ve = sign(+ve)$
$\Rightarrow +ve = +ve$
$\Rightarrow LHS = RHS$

$\Rightarrow Subcase\ 3 : If (sign(r1).sign(r2).\cdots.sign(rk))\ is$ 0 (due to any r_i (1 to k) = 0)
then $sign((r1.r2.\cdots.rk))$ is also 0, from hypothesis
$So, \Rightarrow [0].(-ve) = sign((0).(-ve))$
$\Rightarrow 0 = sign(0)$
$\Rightarrow 0 = 0$
$\Rightarrow LHS = RHS$

From the above cases i.e. for $r_{k+1} = \{-ve, 0, +ve\}$
It is proven that,
$sign(r_1).sign(r_2).\cdots.sign(r_k).sign(r_{k+1}) = sign(r_1.r_2.\cdots.r_k.r_{k+1})$; $k \in N$

3 Question 3

Aim is to prove that there exists a way to map 9 dimensional vectors to D dimensional vectors as
$\phi : R^9 -> R^D$
$such\ that\ for\ any\ triple\ (\tilde{u}, \tilde{v}, \tilde{w}),\ there\ always\ exists\ a\ vector\ W \in R^D$
$such\ that\ for\ every\ \tilde{x} \in R^9,\ we\ have$
$(\tilde{u}^T\tilde{x}).(\tilde{v}^T\tilde{x}).(\tilde{w}^T\tilde{x}) = W^T.\phi(\tilde{x}) \rightarrow EQ1$

If we expand the terms on the LHS of EQ1
$\Rightarrow (\tilde{u}^T\tilde{x}).(\tilde{v}^T\tilde{x}).(\tilde{w}^T\tilde{x}) = (\sum_{i=1}^{9} \tilde{u}_i\tilde{x}_i)(\sum_{j=1}^{9} \tilde{u}_j\tilde{x}_j)(\sum_{k=1}^{9} \tilde{w}\tilde{x}_k)$

we get $\sum_{i=1}^{9}\sum_{j=1}^{9}\sum_{k=1}^{9} \tilde{u}_i\tilde{u}_j\tilde{u}_k\tilde{x}_i\tilde{x}_j\tilde{x}_k$

Now, if we write a $9^3 = 729\ dimensional\ function\ that\ maps$
$\Rightarrow \tilde{x} = (\tilde{x}_1, ..., \tilde{x}_9)\ \rightarrow\ \phi(\tilde{x}) = (\tilde{x}_1\tilde{x}_1\tilde{x}_1, \tilde{x}_1\tilde{x}_1\tilde{x}_2, \cdots, \tilde{x}_1\tilde{x}_1\tilde{x}_9, \tilde{x}_1\tilde{x}_2\tilde{x}_1, \cdots, \tilde{x}_9\tilde{x}_9\tilde{x}_9)$
$i.e.$
$\Rightarrow\ \phi(\tilde{x}) = \tilde{x}.\tilde{x}.\tilde{x} \rightarrow cartesian\ product\ of\ vectors$
$\Rightarrow cardinality\ of\ vector(\phi(x)) = 729$

so that we get $(\tilde{u}^T\tilde{x}).(\tilde{v}^T\tilde{x}).(\tilde{w}^T\tilde{x}) = W^T.\phi(\tilde{x})\ by\ taking$
$\Rightarrow W = \tilde{u}.\tilde{v}.\tilde{w} \rightarrow cartesian\ product\ of\ vectors$
$\Rightarrow W = (\tilde{u}_1\tilde{v}_1\tilde{w}_1, \tilde{u}_1\tilde{v}_1\tilde{w}_2, \cdots, \tilde{u}_1\tilde{v}_1\tilde{w}_9, \tilde{u}_1\tilde{v}_2\tilde{w}_1, \tilde{u}_1\tilde{v}_2\tilde{w}_9, \cdots, \tilde{u}_9\tilde{v}_9\tilde{w}_9)$
$\Rightarrow cardinality\ of\ vector(W) = 729$

i.e.
$\Rightarrow (\tilde{u}^T\tilde{x}).(\tilde{v}^T\tilde{x}).(\tilde{w}^T\tilde{x}) =$
$(\tilde{u}_1\tilde{v}_1\tilde{w}_1, \cdots, \tilde{u}_1\tilde{v}_1\tilde{w}_9, \tilde{u}_1\tilde{v}_2\tilde{w}_1, \tilde{u}_1\tilde{v}_2\tilde{w}_9, \cdots, \tilde{u}_9\tilde{v}_9\tilde{w}_9)^T.(\tilde{x}_1\tilde{x}_1\tilde{x}_1, \cdots, \tilde{x}_1\tilde{x}_1\tilde{x}_9, \tilde{x}_1\tilde{x}_2\tilde{x}_1, \cdots, \tilde{x}_9\tilde{x}_9\tilde{x}_9)$
Hence it is proved that we can map 9 dimensional to D(729) dimensional vectors,
there always exists a vector $W \in R^D\ for\ any\ triple\ (\tilde{u}, \tilde{v}, \tilde{w})$
$such\ that\ for\ every\ \tilde{x} \in R^9,$
$we\ have\ (\tilde{u}^T\tilde{x}).(\tilde{v}^T\tilde{x}).(\tilde{w}^T\tilde{x}) = W^T.\phi(\tilde{x})$

4 Question 5

The hyperparameters we used are C and η
$C\ in\ \nabla W = W + C(\tilde{x}.y)$
$\eta\ in\ W = W - \eta(\nabla W)$

How we have set the final values of C and η is by following the technique of obtaining highest accuracy possible,
with less punishment(C) for loss, a modable step-length(η) which doesn't lead to over-shooting around global minima.

From experimental history we have initialized η with 0.01 and C with 1
on tinkering with η with C = 1
we checked η for various values less than 0.01 like 0.008, 0.004, 0.002 and finally
arrived @$\eta = 0.001$ for which we had an accuracy of 98.53% with less over-shooting around global minima.
Now with η = 0.001 we tuned C for which we can get near to 100% accuracy, the progression of accuracy
with different C was like this.

C = 1	Accuracy = 98.53%
C = 3	Accuracy = 95.83%
C = 5	Accuracy = 99.36%
C = 5.5	Accuracy = 97.58%
C = 5.7	Accuracy = 100%
C = 6	Accuracy = 100%
C = 7	Accuracy = 100%
C = 30	Accuracy = 100%

Finally we concluded $\eta = 0.001\ and\ C = 5.7$ as sweet-spot for our hyperparameters based on train data.

5 Question 6