	Problem 2.1 8-= (1) (1-) &	1 [-	: (Cas	01,00						
	S = wo + wixi + wzxz (H) (1)			n) _s s	, ,	4				
	SI = 4 + 2(-1) + 1(-1) = 11 10	-12 J	-tele	المار		(
4	S2 = 4 + 2(-1) + 1(+1) = 3	The second secon	Military and the second second			1				
1 :						1				
1 7					3	1				
1+	1.0 1- 1+ P.1-=(1-)(1) +	$(\varepsilon-)(i)$	+ 1.8	s pe						
	The Logic Function that this									
	performing is a <u>Tautology</u>									
	Problem 2.2	5	7 0	-Vidage	}					
	in the standard atou which are	- /	- Lampigne of							
ha	WK+1 = WK + 1 Dgn (dk) XI	<u> </u>	A STATE	Land	Santa es presidente					
1	Assume, a talk delia at the author was									
	Wo = 2.1, W1 = 1, W2 = 3,	r = 2 1	1100	- 64						
	6 = 2W (6 = 10 , 8 = 10)	0.00		- e A						
	S ₁ = 2.1 + (1)(-1)+(3)(-1) = -1.9	χ	1 3	(2	5	4				
6	3 8 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	15 W 4-1	X (G) +	Su : 3)	+1				
1.	W, (new) = 1 + 2 (1)(-1) = -1	-	1	11		+1				
1-	wa (new) = 3+2(1)(-1)=1	4	MARIE CONTRACTOR	-1		+1				
1-	8-1-1+1-	+	1 .	11		-1				
1+	S1 = 2.1 + (-1)(-1) + (1)(-1)									
100	Z= 2:1 1- 1+									
T 14	So = 2.1 + (-1)(-1)+ (1)(1) = 1	+•\								
17	S= = 2.1 + (-1)(1) + (1)(-1) = 0	4.1								
1+	c 21 1 (-1)(1) + (1)(1) - 2.	1								

ACCUSATABLE DE CONTROL DE LA C

电影的 经股份的 医克里克氏性 医克里克氏性 医克里克氏性 医克里克氏性 医克里克氏性

		ω,	(neu) = -	1 +	5 (-1) (1)	= -3	1	5 1	roldoz9		
-	E	$\omega_1 (new) = -1 + 2(-1)(1) = -3$ $\omega_2 (new) = 1 + 2(-1)(+1) = -1 + 12(10) + 12(10)$										
	1	Calculating 5 with the new weigh										
	1					(-1)(-1)		-	\ x_2		1 4	
	!	1				+ (-1)(1)					+1	
						+ (-1)(-1				4.1	+1	
					Company of the Compan	(1)(-1)			-1		*1	
						it tot	· · · · · · · · · · · · · · · · · · ·		1+1	-1.9	-	
						1280104.					-	
						T. B. Tool,	7		your m	ish as	-	
		Park	3 6							11 9		
		I SUN								30 1 - 60 1		
		To	Jigires.	uring	2 m	najority model	with a	unction bio	ality s of	. I'm	and	
		To Con	e jigu	lement uring setting equal	a no she she she she	model the we	vote g with a nights the	unction bid had i	aluty	ZELO ZELO	and	
ť	0	Con easy W:	imp fight	lement uring setting equal wo =	dhe dhe to	model when the week other was a wing a distribution of the way a distribution of the way and the way a distribution of the way and the way a distribution of the way and the w	vote guith a hights the	biou biou in that in	ality	· I'm zero	and	
	0	Con Less W:	drip fight	lement uring betting equal wo =	e n e she to	model when when the we cach other wing a distribution of the wing a distrib	vote g with a hights the ex	had in	ality	. I'm 3000 , w2 0	and und	
1+	3	Con con w:	unipi fight	lement uring betting equal wo =	2 m Hhe Sall to	model when the war each oth w, = a, w	vate grith a hights the a	bion had in was =	ality a of a.	I'm zero	and und	
1+	0	Con Lon W: Asso	dight	lement uring betting equal wo =	and the to	model when the war each oth w, = 2, w	vote grights to a 2 = 2 , (1-)(1) & 7,	bion bion was =	ality ality ality ality ality ality ality ality ality	I'm zero	and and	
	0	To con sees.	sign dign dign dign dign dign dign dign d	lement uring setting equal wo =	and the to	model when the war each oth w, = a, w	vote grights to a 2 = 2 , (1-)(1) & 7,	bious bious was =	ality ality ality ality ality ality ality ality ality	I'm zero	and und	
1+	0	Con Lon W: Asso	sign dign dign dign dign dign dign dign d	lement uring betting equal wo =	and the to	model when the war each oth w, = 2, w	vote grights to a 2 = 2 , (1-)(1) & 7,	was = 1 x2 +1	2. 23 -1 (1)	I'm zero	and and	
1+	0	To con sees.	sign dign dign dign dign dign dign dign d	lement uring setting equal wo =	and the to	model when the war each oth w, = 2, w	vote & suights the sights to the sight t	bious bious was =	2. 2. 2. 2. 2. 2. 2. 4. 4. 4.	I'm zero	and unal	
1+		To con sees.	sign dign dign dign dign dign dign dign d	lement uring setting equal wo =	and the to	model when the war each oth w, = 2, w	vate & sights the sights the sights the sights the sights the sight sigh	was = 1 x2 +1 +1 +1	2. 23 -1 41 -1 +1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	. I'm 3000 3000 S -2 -2	and und	
1+		To con sees.	sign dign dign dign dign dign dign dign d	lement uring setting equal wo =	and the to	model when the war each oth w, = 2, w	vote & stith a sights the sights the sights the sights the sight s	wation bia hat is wa = 12	2. 23 -1 41 41 41 41 41 41 41 41 41 41 41 41 41	. I'm 3000 , w2 0	and und y -1 -1 +1	
1+		To con sees.	sign dign dign dign dign dign dign dign d	lement uring setting equal wo =	and the to	the we each oth with a	vate & stath a sights the sights the sights the sights the sight and sight a	was = 1 x2 +1 +1 +1	2. 23 -1 41 -1 +1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	. I'm 300 , w2 0 , -2 -2 -2 -2	and und y -1 -1 +1 -1	

THE PROPERTY OF THE PARTY OF TH

Paob	lem	2.4
1 5+	Ada	line

Wo = 0.5, W1 = -1, W2 = 1

(ex-#) 1 (e	CH	X2 -	5	y	
S1 = 0.5 + (-1)(-1) + (1) (-1) =0.5	-1	=1	0.5	+1	
Sa = 0.5 + (-1)(-1) + (1)(1) = 2.5	ion lia	+1 5	9.7	+ 1	
S3 = 0.5 + (-1)(1) + (1)(-1)=-1.5	+1,00	-1: 8	-1.5	-1	
Sq = 0.5 + (-1) (1) + (1)(1) = 0.5	+1	+1	0.5	+1	
x 6x3 x 6x5	16	- 16	(;)		
	sx6	,wE			

1.5 meller1

$\omega_0 = 0.5$, $\omega_1 = 1$, $\omega_2 (= \times - 14)$	x, 1	212	S	9	
S ₁ = 0.5 + (-1)(1) + (-1)(-1) = 0.5			0.5		
Sa = 10.5 + (-1)(1) + ((-1) (1) = -1.5 x	(- 1 _	+1	-1.5	-1	
S3 = 0.5 + (1)(1) + (-1)(-1) = 2.5	+1	-46	2.5	+1	
Sy = 0.5 + (1)(1) + (-1)(1) = 0,5	+1 =	+1	0.5	+1	

Assume (Ex-3) = LEX-3] = 0.9 = 16

ω₀ = 2, σω₁₌~-1=, ω₂ = -2 ε×6 6x8

	21,	212	5	4	
S1 = 2 + (1)(-1) + (1)(-2)==1	21 53	+1	5-1	-1	
Sa = 2 + (1)(-1) + (-1)(-2) = 3	34 P	-1 14	3	+1	1
Sy = 2+ (-1)(-1) + (-2)(1)=1	78	+1.x		+1	
Sh = St (-1)(1) + (-5)(1) = -1	+1)	+/160	-1	-1	

1. dt = (28-E) (wg) | X, e 4, w, +bi

		Problem St.	
		Problem 3.1	
		1 - 1 - 1 - 1 - 2 - 2 - 2 - 2	
		L= 1 (+-x3) (+-x3)	
1 1 -1	-	10 10 (1-) (1) + (1-)(1) + (1-)(1-) + (1-)(1-)	
1 1 5	-72		
1		0	
1	1	1-x2 = Wax2 + b2-=(1-)(1) + (1)(1) + (1)(1)	
,	1	2 1 1 1 1 3.0 = (1)(1) + (1)(1) + 3.0 + 3	
	(;)		
/		The state of the s	
1	5	A	
1-7	2.0		
1-	7.7-		
1-1-	7.5	9x3 1994 9x3 (E)(E)(+ (1941) + 200 + 20	
1+	7.6	1 + = 14 1 x 2 p = (P()-) + (D()) + 20) + 2	
		dL = p. d [t-x3] = (t-x3) (-1)	
		dx3 dx3 == x3 = t. 6 = 0	
1 1	1	E J.K J.R	
1 1-		1 11 111	
17		$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
1+		20.33	
		1 2 x2 = x1 ex, w, +b1) (2-) + (1-) (1-) + 2 = 2	
1 -		90011	
		· / ~ 1) (ω) [- χ,ω,+b,]	
		$\frac{\partial L}{\partial \omega_1} = \left(x_3 - E \right) \left(\omega_{\alpha} \right) \left[\frac{\chi_1}{\left(1 + e^{\chi_1 \omega_1 + b_1} \right)^2} \right]$	
		QW1 (1+ e /) / Δ	

ii)
$$\frac{\partial L}{\partial \omega_{2}} = \frac{\partial L}{\partial w_{3}} \times \frac{\partial w_{3}}{\partial \omega_{2}}$$

$$= (4 \times 3 - t)(x_{2})$$
iii) $\frac{\partial L}{\partial b_{1}} = \frac{\partial L}{\partial w_{3}} \times \frac{\partial x_{3}}{\partial x_{2}} \times \frac{\partial x_{3}}{\partial b_{1}}$

$$= (x_{3} - t)(\omega_{3}) \times \left[\frac{e^{x_{1}\omega_{1} + b_{1}}}{(1 + e^{x_{1}\omega_{1} + b_{1}})^{2}}\right]$$
iv) $\frac{\partial L}{\partial b_{2}} = \frac{\partial L}{\partial x_{3}} \times \frac{\partial x_{3}}{\partial b_{2}}$

$$= \frac{(x_{3} - t)}{(x_{3} - t)}$$

$$\frac{\partial L}{\partial \omega_{1}} = \frac{\partial L}{\partial \omega_{2}} \times \frac{\partial x_{3}}{\partial \omega_{2}} \times \frac{\partial x_{3}}{\partial \omega_{2}}$$

$$= \frac{(x_{3} - t)}{(x_{3} - t)(w_{3})(x_{1})} \times \frac{(x_{3} + b_{1})}{(x_{3} - t)(x_{2})}$$

$$\frac{\partial L}{\partial \omega_{2}} = (x_{3} - t)(x_{2}).$$

$$\frac{\partial L}{\partial \omega_{2}} = (x_{3} - t)(x_{2}).$$

$$\frac{\partial L}{\partial b_1} = \begin{cases} (x_3 - t) & (\omega_3), & (\omega_3 x_1 + b_1 > 0, 16) \\ 0 & (\omega_3 - t) \end{cases}$$

$$\frac{\partial L}{\partial b_2} = (x_3 - t)$$

$$\frac{\partial L}{\partial b_3} = (x_3 - t) = (x_3 - t)$$

$$\frac{\partial L}{\partial \omega_2} = \begin{pmatrix} x_3 - t \end{pmatrix} x_2$$

$$\begin{pmatrix} -2 \\ -5 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$\frac{2}{3} \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$\frac{\partial L}{\partial b_1} = \begin{bmatrix} -a \\ -a \end{bmatrix} \begin{bmatrix} 1 & -a \\ -3 & 1 \end{bmatrix}$$

$$\frac{\partial A}{\partial b_1} = \begin{bmatrix} -a \\ -a \end{bmatrix} \begin{bmatrix} 1 & -a \\ -a \end{bmatrix}$$

$$\frac{dL}{db_2} = (x_3 - t) = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

ECE 661 Homework-1

Problem 1:

True/False Questions

- 1. **False**. To enhance a neural network's ability to generalize, one can either augment the training dataset by increasing the number of samples or reduce the model's complexity by decreasing the number of parameters. Increasing the number of parameters can make the model overfit the noise.
- 2. **False.** The final set of weights got after training depends on the initialization values.
- 3. **True.** Using a full batch size provides a more accurate gradient estimate, reducing the likelihood of becoming trapped in local minima during training.
- 4. **False.** An if else statement can split the model into two parts where each part is differentiable, for example a condition which says n>0, the backpropagation training algorithm cannot be applied to the entire model, only part of the model depending on the condition.
- 5. **False.** If the padding size and stride size are both set to one, enlarging the height and width of the kernel size will result in a smaller output feature map size.

Problem 4:

2D convolution:

Input matrix:

```
([[ 0., 0., -1., 0., 0., 0., 1., 0., 0.],
        [ 0., -1., -1., -1., 0., 1., 1., 1., 0.],
        [-1., -1., -1., -1., 0., 1., 1., 1., 1.],
        [ 0., -1., -1., -1., 0., 1., 1., 1., 0.],
        [ 0., 0., -1., 0., 0., 0., 1., 0., 0.]])
```

Kernel:

```
([[ 0.0000, -0.5000, 0.0000],
[-0.5000, 1.0000, -0.5000],
[ 0.0000, -0.5000, 0.0000]])
```

Output matrix:

```
0.
                              -1.
       0.
                   0.
                         0.
                               0.
                                    -1.
                                           0.
                                                -1. ]
                                                 0.5]
-0.5
       1.
             1.
                   0.5
                         0.
                              -0.5 -1.
                                          -1.
                   0.
                         0.
                               0.
       1.
[ 0.
            -0.5
                         0.
                              -1.
                                     0.5 -1.
```

Applying Kernel to an image

Original Image



Feature Map After Convolution

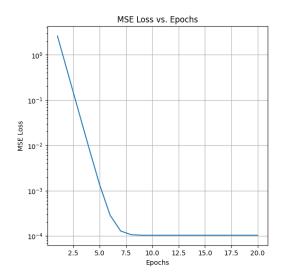


When the kernel is applied to the image of the dog, results in inversion. This happens because the kernel contains both positive and negative values, and during the convolution operation, these values interact with the pixel values of the image. Specifically, the positive central value amplifies the importance of the central pixel, while the negative values subtract from the surrounding pixels. This combination of operations results in the image appearing inverted or flipped.

Lab 1: LMS Algorithm

Problem 5(d):

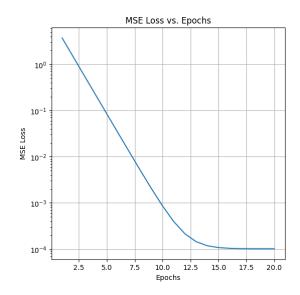
MSE vs Epochs when learning rate=0.01



Optimal weight when learning rate=0.01

[[1.00074855] [1.00082859] [-2.00068123]]

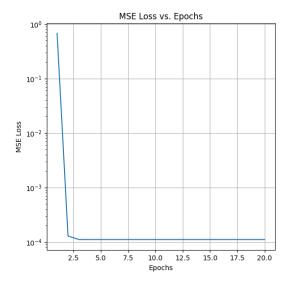
MSE vs Epochs when learning rate=0.005



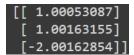
Optimal weight when learning rate=0.005

[[1.00068274] [1.0006024] [-2.00033003]]

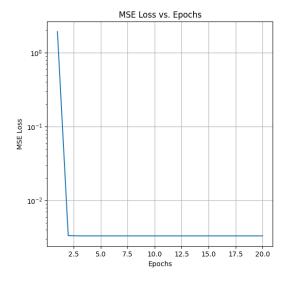
MSE vs Epochs when learning rate=0.05



Optimal weight when learning rate=0.05



MSE vs Epochs when learning rate=0.5



Optimal weight when learning rate=0.5

[[0.97969496] [0.98520802] [-1.9666911]]

As the learning rate increases, convergence occurs more rapidly, but it must be carefully chosen, as an excessively high learning rate can cause the model to overshoot the minimum.

Lab 2: Simple NN

Problem 6(a)

```
# Create the neural network module: LeNet-5
class SimpleNN(nn.Module):
    def __init__(self):
        super(SimpleNN, self).__init__()
        # Layer definition
        self.conv1 = CONV(in_channels=3, out_channels=16, kernel_size=5, stride=1, padding=2)#Your code here
        self.conv2 = CONV(in_channels=16, out_channels=16, kernel_size=3, stride=1, padding=2)#Your code here
        self.conv3 = CONV(in_channels=16, out_channels=32, kernel_size=7, stride=1, padding=2)#Your code here
        self.fc1 = FC(in_features=288, out_features=32)#Your code here
        self.fc2 = FC(in_features=32, out_features=10)#Your code here
```

```
def forward(self, x):
   # Conv 1
   out = F.relu(self.conv1(x))
   out = F.max_pool2d(out, 4, 2)
   # Conv 2
   out = F.relu(self.conv2(out))
   # MaxPool
   out = F.max_pool2d(out, 3, 2)
   # Conv 3
   out = F.relu(self.conv3(out))
   out = F.max_pool2d(out, 2, 2)
    # Flatten
   out = out.view(out.size(0), -1)
   # FC 1
   out = F.relu(self.fc1(out))
   # FC 2
   out= F.relu(self.fc2(out))
   return out
```

```
device = 'cuda' if torch.cuda.is_available() else 'cpu'
if device =='cuda':
    print("Run on GPU...")
else:
    print("Run on CPU...")

# Model Definition
net = SimpleNN()
net = net.to(device)

# Test forward pass
data = torch.randn(5,3,32,32)
data = data.to(device)

# Forward pass "data" through "net" to get output "out"
out = net(data) #Your code here

# Check output shape
assert(out.detach().cpu().numpy().shape == (5,10))
print("Forward pass successful")

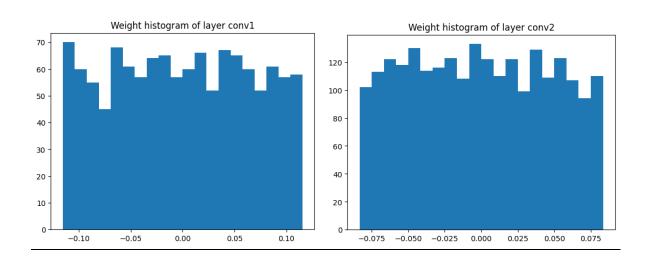
Run on CPU...
Forward pass successful
```

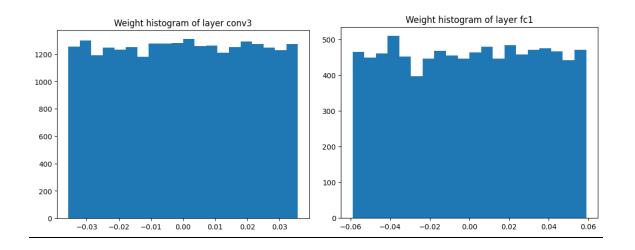
Problem 6(b)

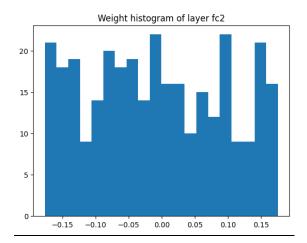
Layer	Input Shape	Output shape	Weight	# Param	#MAC
			shape		
Conv1	(1, 3, 32, 32)	(1, 16, 32, 32)	(16, 3, 5, 5)	1200	1228800
Conv2	(1, 16, 15, 15)	(1, 16, 17, 17)	(16, 16, 3, 3)	2304	665856
Conv3	(1, 16, 8, 8)	(1, 32, 6, 6)	(32, 16, 7, 7)	25088	903168
FC1	(1, 288)	(1, 32)	(32, 288)	9216	9216
FC2	(1, 32)	(1, 10)	(10,32)	330	330

Lab 3: Bonus Question

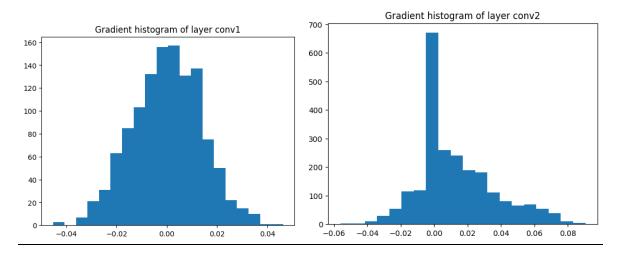
Weight Histogram:

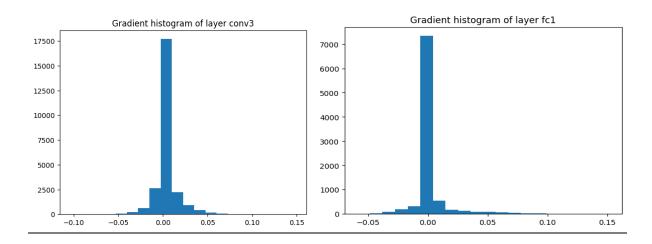


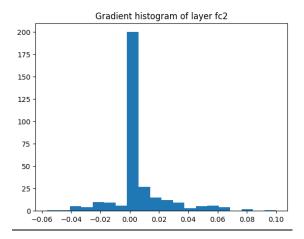




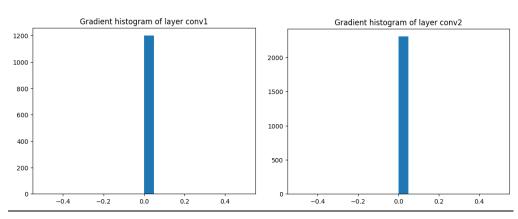
Gradient Histogram:

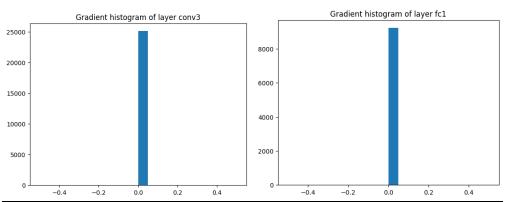


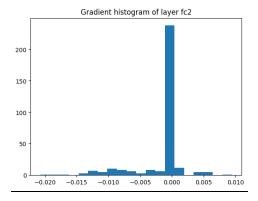




Zero Initialization:







In Convolutional Neural Networks (CNNs), the "dying ReLU" problem occurs when ReLU activation neurons become stuck at outputting 0 due to consistently negative weighted sums of their inputs. This issue can render a portion of the network 'dead,' leading to ineffective training as these neurons do not contribute to learning. It results in a loss of model capacity to capture intricate patterns in the data and can hinder the network's ability to adapt during training.