Week 4: Assignment 4

The due date for submitting this assignment has passed.

Due on 2024-08-21, 23:59 IST.

Assignment submitted on 2024-08-21, 21:38 IST

1) Which one of the following statements is true: ***1 point***

- Weight change criterion is a method of 'early stopping' that checks whether or not the error is dropping over epochs to decide whether to continue training or stop.
- L_2 norm tends to create more sparse weights than L_1 norm.
- During the training phase, for each iteration, Dropout ignores a random fraction, p, of nodes, and accounts for it in the test phase by scaling down the activations by a factor of p.
- A single McCulloch-Pitts neuron is capable of modeling AND, OR, XOR, NOR, and NAND functions

Yes, the answer is correct.
Score: 1
Accepted Answers:
During the training phase, for each iteration, Dropout ignores a random fraction, p, of nodes, and accounts for it in the test phase by scaling down the activations by a factor of p.

2) For a neural network f, let w_{ij} be the weight connecting neurons a_i in hidden layer-1 to b_j in adjacent hidden layer-2. Consider the following statements: ***1 point***

Statement-1 : $\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial b_j} a_i$, where L is the loss function of f.

Statement-2 : w_{ij} is not the only weight that is connecting neurons a_i and b_j

Choose the most appropriate answer:

- Statement-1 and Statement-2 are false
- Statement-1 and Statement-2 are true
- Statement-1 is true but Statement-2 is false
- Statement-1 is false but Statement-2 is true

Yes, the answer is correct.
Score: 1
Accepted Answers:
Statement-1 is true but Statement-2 is false

3) Which of the following statements are **true**? (Select all that apply) ***1 point***

- $Sigmoid$ activation function $\sigma(\cdot)$ can be represented in terms of $\tanh$ activation function as below:

$$\sigma(x) = (\tanh(x/2) - 1)/2$$

- The derivative of the sigmoid activation function is symmetric around the y-axis
- Gradient of a $sigmoid$ neuron vanishes at saturation.
- $Sigmoid$ activation is centered around 0 whereas $\tanh$ activation is centered around 0.5

No, the answer is incorrect.
Score: 0
Accepted Answers:
The derivative of the sigmoid activation function is symmetric around the y-axis
Gradient of a $sigmoid$ neuron vanishes at saturation.

4) Consider two 3×3 images x_1 and x_2 such that $x_1 = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 7 & 12 \\ 1 & 7 & 6 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 10 & 6 & 2 \\ 8 & 12 & 8 \\ 2 & 3 & 6 \end{bmatrix}$. Their corresponding one-hot encoded label vectors are $y1 = [0, 1, 0]$ and $y2 = [0, 0, 1]$. Perform mixup data augmentation between x_1 and x_2 given that $\lambda = 0.4$. ***1 point***

- $\tilde{x} = \begin{bmatrix} 6.8 & 4.8 & 2.8 \\ 6.4 & 10.0 & 9.6 \\ 1.6 & 4.6 & 6.0 \end{bmatrix}; \tilde{y} = [0, 0.6, 0.4]$
- $\tilde{x} = \begin{bmatrix} 5.2 & 4.2 & 3.2 \\ 5.6 & 9.0 & 10.4 \\ 1.4 & 5.4 & 6.0 \end{bmatrix}; \tilde{y} = [0, 0.6, 0.4]$
- $\tilde{x} = \begin{bmatrix} 6.8 & 4.8 & 2.8 \\ 6.4 & 10.0 & 9.6 \\ 1.6 & 4.6 & 6.0 \end{bmatrix}; \tilde{y} = [0, 0.4, 0.6]$
- $\tilde{x} = \begin{bmatrix} 5.2 & 4.2 & 3.2 \\ 5.6 & 9.0 & 10.4 \\ 1.4 & 5.4 & 6.0 \end{bmatrix}; \tilde{y} = [0, 0.4, 0.6]$

Yes, the answer is correct.
Score: 1
Accepted Answers:
$\tilde{x} = \begin{bmatrix} 6.8 & 4.8 & 2.8 \\ 6.4 & 10.0 & 9.6 \\ 1.6 & 4.6 & 6.0 \end{bmatrix}; \tilde{y} = [0, 0.4, 0.6]$

5) Consider the following statements P and Q regarding AlexNet and choose the correct option: ***1 point***

(P) In AlexNet, Response Normalization Layers were introduced to emulate the competitive nature of real neurons, where highly active neurons suppress the activity of neighboring neurons, creating
competition among different kernel outputs.

(Q) Convolutional layers contain only about 5% of the total parameters hence account for the least computation.

- Only statement P is true
- Only statement Q is true
- Both statements are true
- None of the statements is true

Yes, the answer is correct.
Score: 1
Accepted Answers:
Only statement P is true

6) Given an input image of shape (10,10,3), you want to use one of the two following layers: ***1 point***

- Fully connected layer with 2 neurons, with biases
- Convolutional layer with three 2x2 filters (with biases) with 0 padding and a stride of 2.

If you use the fully-connected layer, the input volume is "flattened" into a column vector before being fed into the layer. What is the difference in the number of trainable parameters between these two layers?

- The fully connected layer has 566 fewer parameters
- The convolutional layer has 518 fewer parameters
- The convolutional layer has 570 fewer parameters
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:

Accepted Answers:
None of the above

7) Which of the following statements is false? ***1 point***

- For a fixed padding, the bigger the size of the kernel, the smaller is the output after convolution.
- To get the output with the same size as that of the input, padding used is $\lfloor \frac{k}{2} \rfloor$ where $k \times k$ is the kernel used.
- The number of feature maps obtained after a convolution operation depends on the depth of the input but not on the number of filters.
- Stride is a hyper-parameter

Yes, the answer is correct.
Score: 1
Accepted Answers:
The number of feature maps obtained after a convolution operation depends on the depth of the input but not on the number of filters.

8) Compute the value for the following expression **ELU(tanh(x))** where $x = -1.3$ and $\alpha = 0.3$ (Round decimal point till 2 places).

-0.17

Yes, the answer is correct.
Score: 1
Accepted Answers:
(Type: Range) -0.19,-0.15

1 point

9) Using RMSProp-based Gradient Descent, find the new value of parameter θ_{t+1}, given that the old value $\theta_t = 1.2$, aggregated gradient $\Delta\theta_t = 0.85$, gradient accumulation $r_{t-1} = 0.7$, learning rate $\alpha = 0.9$, decay rate $\rho = 0.3$ and small constant $\delta = 10^{-7}$ (Round decimal point till 3 places).

0.516

No, the answer is incorrect.
Score: 0
Accepted Answers:
(Type: Range) 0.286,0.306

1 point

If we convolve a feature map of size 32 × 32 × 6 with a filter of size 7 × 7 × 3, with a stride of 1 across all dimensions and a padding of 0, the width of the output volume is ___A___,the height of the output volume is ___B___ and the depth of the output volume is ___C___

10) A _____

26

No, the answer is incorrect.
Score: 0
Accepted Answers:
(Type: Numeric) 122

0.33 points

11) B _____

26

No, the answer is incorrect.
Score: 0
Accepted Answers:
(Type: Numeric) 84

0.33 points

12) C _____

3

No, the answer is incorrect.
Score: 0
Accepted Answers:
(Type: Numeric) 4

0.34 points

Assume that the feature map given below is generated from a convolution layer in CNN, after which a 2 × 2 Max Pooling layer with a stride 2 is applied to it.

156 (1,1)	**40** (1,2)	**179** (1,3)	**240** (1,4)
99 (2,1)	**106** (2,2)	**51** (2,3)	**5** (2,4)
149 (3,1)	**88** (3,2)	**87** (3,3)	**71** (3,4)
62 (4,1)	**254** (4,2)	**20** (4,3)	**14** (4,4)

While backpropagation, we get the following gradient for the pooling layer.

8	**10**
14	**2**

Assign the appropriate gradient value for the locations at feature map.

13) Location (1,1):-

8

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 8

0.17 points

14) Location (1,4):

10

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 10

0.17 points

15) Location (2,2):

0

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 0

0.17 points

16) Location (3,1):

0

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 0

0.17 points

17) Location (3,3):

2

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 2

0.17 points

18) Location (4,3):

0

Yes, the answer is correct.
Score: 0.15
Accepted Answers:
(Type: Numeric) 0

0.15 points

For the same previous question, assign the appropriate gradient value for the locations at feature map but use **Average Pooling** layer instead of Max Pooling layer.

19) Location (1,1):

2

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 2

0.17 points

20) Location (1,4):

2.5

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 2.5

0.17 points

21) Location (2,2):

2

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 2

0.17 points

22) Location (3,1):

3.5

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 3.5

0.17 points

23) Location (3,3):

0.5

Yes, the answer is correct.
Score: 0.17
Accepted Answers:
(Type: Numeric) 0.5

0.17 points

24) Location (4,3):

0.5

Yes, the answer is correct.
Score: 0.15
Accepted Answers:
(Type: Numeric) 0.5

0.15 points