

20233_csci_544_30249:
Applied Natural
Language Processing

Announcements

Syllabus

Content

Assignments

Tools

USC Code of Ethics

Review Test Submission: Quiz4

User	Kayvan Shah
Course	20233_csci_544_30249: Applied Natural Language Processing
Test	Quiz4
Started	9/22/23 7:35 PM
Submitted	9/22/23 7:42 PM
Due Date	9/23/23 5:50 PM
Status	Completed
Attempt Score	90 out of 100 points
Time Elapsed	6 minutes out of 10 minutes
Results Displayed All Answers, Submitted Answers, Correct Answers	

Question 1

10 out of 10 points

Given a one-layer perceptron with weights of [1,-2,3,4], bias of 1 and activation function of ReLU, what is the output given the input [-1,4,2,-1]?

Selected Answer: ☒ 0

Correct Answer: ☒ 0

Answer range +/- 0 (0 - 0)

Question 2

0 out of 10 points

For the back propagation of RNN, we should update the weights based on the gradient descent computed by the last time-step's loss.

Selected Answer: ☒ True

Answers: ☐ True
☒ False

Question 3

10 out of 10 points

Which of the following statement(s) is/are true about training RNN?

Selected Answers: ☒ Training it is computationally expensive
☒ It might experience gradient vanishing.
☒ It might experience gradient explosion.

Answers: ☒ Training it is computationally expensive
☐ It should take inputs of the same size
☒ It might experience gradient vanishing.
☒ It might experience gradient explosion.

Question 4

10 out of 10 points

Word2Vec word embedding method considers the word occurrence at context-level.

Selected Answer: ☒ True

Answers: ☒ True
☐ False

Question 5

10 out of 10 points

In RNN, as weights (W) occur at each time step, each path from W to L (loss) should be seen as a separate dependency for differentiation.

Selected Answer: ☒ True

Answers: ☒ True
☐ False

Question 6

10 out of 10 points

Which of the following statement(s) is/are true about the limitation of Feedforward Neural Network?

Selected Answers: ☒ It is not efficient to catch the long-term dependency.
☒ All input instance should have the same length
☐ The solution of the network is just an approximation and is not perfect

- Answers:
- ✓ The solution of the network is just an approximation and is not perfect.
 - ✓ It is not efficient to catch the long-term dependency.
 - ✓ All input instance should have the same length
 - ✓ The weight of the parameters are shared across time.
 - ✓ The solution of the network is just an approximation and is not perfect.

Question 7

10 out of 10 points

Skip-Gram uses ____ word to predict ____ word, while CBOW uses ____ word to predict ____ word.

Selected Answer: ✓ center, context, context, center

- Answers:
- ✓ center, context, context, center
 - context, center, center, context
 - context, context, center, center
 - center, center, context, context

Question 8

10 out of 10 points

What are the training samples of Skip-Gram given a sentence "I love natural language processing", and the word "natural" is the center word?

Selected Answer: ✓ (natural, I), (natural love), (natural, language), (natural, processing)

- Answers:
- (natural, language), (natural, processing)
 - (natural, I), (natural, love), (natural, natural), (natural language), (natural, processing)
 - (natural, I) (natural love)
 - ✓ (natural, I), (natural love), (natural, language), (natural, processing)

Question 9

10 out of 10 points

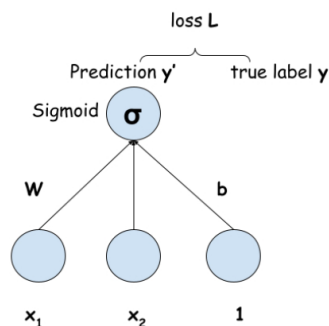
Which of the following statement(s) is/are true about neural network?

- Selected Answer: ✓ ReLU activation function can efficiently solve the gradient vanishing issue.
- Answers:
- ✓ MLP can be used to approximate most of the functions given the proper values of the weights.
 - ✓ ReLU activation function can efficiently solve the gradient vanishing issue.
 - RNN is good at capturing temporal information from input text because it is designed to match the unidirectional property of the natural language.
 - ✓ MLP can be used to approximate most of the functions given the proper values of the weights.
 - One-layer perceptron can be used to solve XOR problem.

Question 10

10 out of 10 points

As shown in the following figure, which back propagation computation for $\frac{\partial L}{\partial w_i}$ is correct for 1-layer feedforward neural network given input x , weight matrix \mathbf{W} and vector b , function Sigmoid, and true label y . The forward computation is $y' = \text{Sigmoid}(z)$, $z = \mathbf{W}x + b$ and the loss is L .



Selected Answer: ✓ $\frac{\partial L}{\partial y'} \frac{\partial y'}{\partial z} \frac{\partial z}{\partial w_i}$

- Answers:
- $\frac{\partial L}{\partial y} \frac{\partial y}{\partial z} \frac{\partial z}{\partial w_i}$
 - ✓ $\frac{\partial L}{\partial y'} \frac{\partial y'}{\partial z} \frac{\partial z}{\partial w_i}$
 - $\frac{\partial L}{\partial y} \frac{\partial y}{\partial z}$

$$\frac{\partial y}{\partial x} \frac{\partial z}{\partial x}$$

$$\frac{\partial L}{\partial y} \frac{\partial y}{\partial z} \frac{\partial z}{\partial b}$$

Saturday, January 20, 2024 11:39:33 PM PST

← OK