RQ3: DL Refresher

- Due Jan 31 at 4:59pm
- Points 4
- Questions 10
- Time Limit 15 Minutes
- Allowed Attempts Unlimited

Take the Quiz Again

Attempt History

	Attempt	Time	Score
LATEST	**Attempt 1**	6 minutes	3.6 out of 4

⚠ Correct answers are hidden.

Score for this attempt: 3.6 out of 4
Submitted Jan 31 at 4:43pm
This attempt took 6 minutes.

⠿

Question 1
0.4 / 0.4 pts

Consider this simple PyTorch model for regression:

```
import torch
import torch.nn as nn

class SimpleMLP(nn.Module):
    def __init__(self):
        super(SimpleMLP, self).__init__()
        self.fc1 = nn.Linear(4, 3)   # from 4 inputs to 3 neurons
        self.relu = nn.ReLU()
        self.fc2 = nn.Linear(3, 1)   # from 3 neurons to 1 output

    def forward(self, x):
        x = self.fc1(x)
        x = self.relu(x)
        x = self.fc2(x)
        return x
```

How many trainable parameters (weights + biases) does this model contain in total?

Hint: For nn.Linear(a,b), there are a × b weights and b biases.

○ 12

○ 22

○ 19

○ 15

⋮⋮

Question 2

0.4 / 0.4 pts

You have the sentence: "I love to [MASK] every morning." and your BERT-based model must predict the missing word. Which statement correctly describes BERT's approach in this scenario?

○ BERT is an autoregressive model that generates tokens one at a time, each conditioned on previous tokens.

◉ BERT uses Masked Language Modeling, attending to all surrounding tokens to infer the [MASK].

○ BERT is an autoregressive model that generates tokens one at a time, each conditioned on previous tokens.

○ BERT ignores the context and focuses purely on final-layer hidden states.

⋮⋮

Question 3

0.4 / 0.4 pts

Neural networks are said to be "biologically inspired," yet they differ greatly from real brains. Which statement best explains one key difference?

◉ Biological neurons have more complex connectivity and signaling patterns than typical artificial neurons.

○ Biological neurons rely exclusively on matrix multiplication to transmit signals.

○ Artificial neurons use chemical neurotransmitters for activations, unlike real neurons.

○ Artificial networks are only trained with linear activations, whereas biology uses nonlinear signals.

⋮⋮

Question 4

0.4 / 0.4 pts

Some claim that "using only linear activations in multiple layers is not beneficial" because it collapses to a single linear operation. Why is stacking purely linear layers generally ineffective?

◉ Multiple linear layers combine into the same transform, providing no additional representation power.

○ Linear layers inherently cause vanishing gradients.

○ Purely linear networks can only handle binary classification.

○ Nonlinear layers slow down training, so fewer are better.

⋮⋮

Question 5

0.4 / 0.4 pts

You have two 300-dimensional word embeddings, v1 and v2. You want to measure how closely they match in meaning. Which measure is most commonly used to assess their similarity?

◉ Cosine similarity

○ Euclidean distance

○ Mean squared error

○ Cross-entropy loss

⋮⋮

Question 6

0.4 / 0.4 pts

You discover that your model can be fooled by carefully crafted inputs, causing nonsensical outputs. In LLM terminology, such deceptive or misleading inputs are known as:

Hint: These are maliciously designed prompts or inputs to break or confuse the model

◉ Adversarial examples

○ Instruction-tuned prompts

○ Autoregressive prompts

○ Vocabulary collisions

⋮⋮

Question 7

0.4 / 0.4 pts

You are building a neural network to predict housing prices (a continuous value). Which two activation functions are generally suitable at the output layer for a continuous prediction task?

☐ Sigmoid

☑ Linear

☐ Softmax

☑ ReLU

⋮⋮

Question 8

0.4 / 0.4 pts

In a typical feed-forward neural network architecture (with an input layer, one or more hidden layers, and an output layer), which of the following is not a usual term for a component in that architecture?

○ Neurons

◉ Axons

○ Layers

○ Activation function

⋮⋮

Question 9

0.4 / 0.4 pts

You have a large pre-trained language model and want it to follow specific user instructions (e.g., "turn these notes into a summary"). Which method refers to training a model explicitly with instruction-oriented examples?

Hint: One approach uses tasks posed as explicit instructions (e.g., in a Q&A or summarization format).

- ○ Increasing the Context Window
- ○ Fine-Tuning on unstructured text
- ○ Prompt Engineering
- ◉ Instruction Tuning

⋮⋮

IncorrectQuestion 10

0 / 0.4 pts

Assume a feed-forward network with 4 input features, then a single hidden layer of 5 neurons, and an output layer of size 1 for regression. All hidden neurons use ReLU. How many weight parameters (not counting biases) connect the input layer to the hidden layer?

- ○ 4 × 5 = 20
- ○ 4 + 5 = 9
- ◉ 4 × 5 + 5 = 25
- ○ 5 × 1 = 5

Quiz Score: 3.6 out of 4