

Q1 Instructions

1 Point

Open notes: The quiz is open notes. You are free to use any content from the course website or from your own personal notes.

No communication: ANY communication with other students about the quiz content is strictly forbidden and will result in a failing grade for the whole class (not just this quiz).

No partial credit: Every question is all or nothing credit. Thus, you must get the answer exactly right to get credit for the question (including SELECT ALL questions). No partial credit will be given on quizzes.

Number Format: When giving numbers as short answers, please give in standard decimal notation with preceding "0." for decimals if needed but no trailing 0s (e.g., "0.15", "2.9", "0.001", "100" but NOT "0.15000" NOR ".15" NOR ".001" NOR "6.0").

Honor Pledge: I assert that I have not received any information about this quiz and will not share any quiz content with anyone else. I understand that any violation of this will result in a failing grade for the whole class (not just this quiz).

Yes, I understand the policies above and assert the honor pledge.

No

Q2

1 Point

Which of the following tasks can be accomplished using self-supervised learning? Select all the correct answers.

- ✓ Image classification
- ✓ Image colorization (from grayscale to color image)
- ✓ Predicting missing words in a sentence

✓ ~~Answer~~ Speech recognition

Image colorization (from grayscale to color image), Predicting missing words in a sentence

Q3

1 Point

RNNs can only be used for problems where the input and output are sequences.

True

False

Q4

1 Point

Which of the following distributions are valid generative models for count-valued data? (Select all)

Gaussian

✓ Multinomial

Mixture of Gaussians

✓ Mixture of Multinomials

Q5

1 Point

The ELBO objective is always a ____ bound on the marginal log likelihood $\log p_g(x)$ regardless of the choice of q_f .

Lower

Tight

Upper

Q6

2 Points

The LDA joint distribution can be formulated as

$$p(\theta, \beta, Z, W) = p_{\eta}(\beta) \prod_{i=1}^M p_{\alpha}(\theta_i) \prod_{\ell=1}^{L_i} P(z_{i,\ell}|\theta_i) P(w_{i,\ell}|\beta_{z_{i,\ell}}),$$

where β are the topic-word distributions, θ_i are the sample-specific topic distributions, W are the words in the training data, and Z are the topic assignments for each corresponding words in W .

Q6.1

1 Point

Which of the following is the posterior distribution of LDA?

$$p(\theta|\beta, Z, W)$$

$$p(\theta, \beta, Z, W)$$

$$p(\theta, \beta, Z|W)$$

$$p(\theta, \beta, Z)$$

Q6.2

1 Point

What is the prior distribution of θ_i ?

$$p(\theta_i)$$

$$p(z_{i,1}, z_{i,2}, \dots, z_{i,L_i}|\theta_i)$$

$$p(\theta_i|z_{i,1}, z_{i,2}, \dots, z_{i,L_i})$$

$$p(\theta_i, z_{i,1}, z_{i,2}, \dots, z_{i,L_i})$$

Q7
1 Point

In the word2vec task for predicting the target word given context words using the CBOW architecture, the order of the words in the context does not matter.

- True
- False

Q8
1 Point

What are some limitations with the windowing approach to handling sequences?

- ✓ The window size is fixed.
- ✓ Lacks long-range dependencies.

The model's predictions of different windows are very dependent.

Q9
1 Point

Latent semantic indexing (LSI) is a non-probabilistic topic model that can be solved by applying SVD to a word-count matrix.

- True
- False

Q10
1 Point

In Gibbs sampling for LDA, sampling the topics of all the words one time is sufficient to learn the parameters.

True

False

Q11

1 Point

Let X and Y be independent continuous random variables. If we know that $p(X = 2|Y = 1) = 3$, $p(Y = 2) = 3$, and $p(Y = 1|X = 3) = 4$, what is $p(X = 2, Y = 1)$? (Insufficient information [-1] is a valid answer.)

9

Quiz 9	● Graded
Student	
Paloma Arellano	
Total Points	
10 / 12 pts	
Question 1	
Instructions	1 / 1 pt
Question 2	
(no title)	0 / 1 pt
+ 1 pt	Correct
✓ + 0 pts	Incorrect
Question 3	
(no title)	1 / 1 pt
Question 4	

(no title)	1 / 1 pt
Question 5 (no title)	1 / 1 pt
Question 6 (no title)	2 / 2 pts
6.1 (no title)	1 / 1 pt
6.2 (no title)	1 / 1 pt
Question 7 (no title)	1 / 1 pt
Question 8 (no title)	1 / 1 pt
Question 9 (no title)	1 / 1 pt
Question 10 (no title)	1 / 1 pt
Question 11 (no title)	0 / 1 pt