1. Classification allows you to identify similarity between two things while siamese networks allow you to categorize things. 1 / 1 point

- ○ True
- ◉ False

Correct
Correct.

2. Do the two subnetworks in a siamese network share the same parameters? 1 / 1 point

- ○ No
- ◉ Yes

Correct
Correct.

3. When training a siamese network to identify duplicates, which pairs of questions from the following questions do you expect to have the highest cosine similarity ? 1 / 1 point

Is learning NLP useful for me to get a job? (ANCHOR)

What should I learn to get a job? (POSITIVE)

Where is the job? (NEGATIVE)

- ◉ Anchor, Positive
- ○ Anchor, Negative
- ○ Negative, Positive

Correct
Correct.

4. In the triplet loss function below, will decreasing the hyperparameter alpha from 0.5 to 0.2 require more, or less, optimization during training ? 1 / 1 point

$$\text{diff} = \text{s}(A, N) - \text{s}(A, P)$$

$$\mathcal{L}(A, P, N) = \max(diff + \alpha, 0)$$

- (●) Less
- ( ) More.

Correct
Correct. Alpha is the margin, so the smaller it is the less you have to optimize.

5. The orange square below corresponds to the similarity score of question duplicates? 1 / 1 point

0.7	-0.6	-0.4
-0.6	0.4	0.1
-0.4	0.1	0.5

- (●) False
- ( ) True

Correct
Correct. They correspond to non question duplicates.

6. What is the closest negative in this set of numbers assuming a duplicate pair similarity of 0.6? 1 / 1 point

[-0.9,-0.4,0.4, 0.8]

- ( ) -0.9
- ( ) -0.4
- (●) 0.4

0.8

Correct
Correct.

7. In one shot learning, is any retraining required when new classes are added? For example, a new bank customer's signature. 1 / 1 point

- Yes
- (•) No

Correct
Correct.

8. During training, you have to update the weights of each of the subnetworks independently. 1 / 1 point

- (•) False.
- True.

Correct
Correct. You update the same weight.

9. The mean negative is defined as the closest off-diagonal value to the diagonal in each row (excluding the diagonal). 1 / 1 point

- (•) False
- True

Correct
Correct.

10. In what order are Siamese networks performed in lecture? 1 / 1 point

- (•)
  1. Convert each input into an array of numbers
  2. Feed arrays into your model
  3. Compare $\boldsymbol{v}1$, $\boldsymbol{v}2$ using cosine similarity
  4. Test against a threshold

Correct. You update the same weight.

9. The mean negative is defined as the closest off-diagonal value to the diagonal in each row (excluding the diagonal). 1 / 1 point

- (●) False
- ( ) True

Correct
Correct.

10. In what order are Siamese networks performed in lecture? 1 / 1 point

- (●)
  1. Convert each input into an array of numbers
  2. Feed arrays into your model
  3. Compare $\boldsymbol{v}1$, $\boldsymbol{v}2$ using cosine similarity
  4. Test against a threshold
- ( )
  1. Convert each input into an array of numbers
  2. Feed arrays into your model
  3. Run logistic regression classifier
  4. Classify by using the probability
- ( )
  1. Convert each input into an array of numbers
  2. Feed arrays into your model
  3. Run soft-max classifier for all classes
  4. Take the arg-max of the probabilities
- ( )
  1. Convert each input into an array of numbers
  2. Feed arrays into your model
  3. Compare $\boldsymbol{v}1$, $\boldsymbol{v}2$ using euclidean distance
  4. Test against a threshold

Correct
Correct.