

## Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 80%. We keep your highest score.

[Next item →](#)

1. Consider the sentence, “she loves watching football.” In the word2vec model, what will be the context and target word(s) for “t=2”? (Take the window width as 1)

1 / 1 point

- ☐ Context: she, watching; Target word: loves  
☒ Context: loves, football; Target word: watching  
☐ Context: she, football; Target word: loves, watching  
☐ Context: loves, watching; Target word: she, football

✓ **Correct**

In word2vec, for the context and target words, with a window width of 1, the context includes the word at t minus 1 and t plus 1. The target word is at t. At t=1, the context consists of “she” and “watching.” The target word is “loves.”

As you move to t=2, the context includes “loves,” and “football,” and the target word is “watching.”

2. Consider the phrase “quick fox jumps over lazy dog.” Using the skip-gram model, find the context and target word for “t=3.” (Take the window width as 2)

1 / 1 point

- ☐ Target word: over; Context: lazy, dog  
☐ Target word: over; Context: jumps, lazy, dog  
☒ Target word: over; Context: fox, jumps, lazy, dog  
☐ Target word: over; Context: quick, fox, jumps

✓ **Correct**

Skip-gram model predicts surrounding context words from a specific target word. It aims to forecast words “t-2”, “t-1”, “t+1”, and “t+2” for a given target word t.

At t=1, with “fox” as the target, the model predicts “quick”, “jumps”, and “over”.

At t=2, when “jumps” is the target, it aims to predict “quick”, “fox”, “over”, and “lazy”.

Similarly, at t=3, when “over” is the target, it aims to predict “fox”, “jumps”, “lazy”, and “dog”.

3. Which generative AI model matches the description below?

1 / 1 point

“It is a type of simulated neural network that uses time series data. It is designed to remember past information.”

- ☒ Recurrent neural networks (RNNs)  
☐ Generative adversarial networks (GANs)  
☐ Word2vec’s neural network  
☐ Feedforward neural network

✓ **Correct**

RNN is a type of artificial neural network that uses sequential or time series data. As its name suggests, RNN is designed to remember past information and use it to influence future decisions.

4. Fill in the blank:

1 / 1 point

In general, \_\_\_(Blank 1)\_\_\_ models are more difficult to train compared to recurrent neural networks (RNNs). However, the aim is to minimize the \_\_\_(Blank 2)\_\_\_.

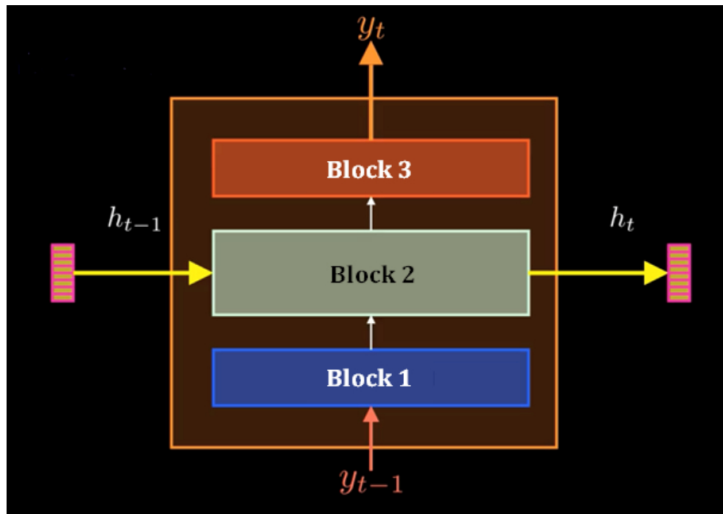
- ☐ Blank 1: N-gram; Blank 2: Cross-entropy loss  
☐ Blank 1: Sequence-to-sequence; Blank 2: Divergence  
☒ Blank 1: Sequence-to-sequence; Blank 2: Cross-entropy loss  
☐ Blank 1: Word2vec; Blank 2: Divergence

✓ Correct

In general, sequence-to-sequence models are more difficult to train than RNNs, and several factors come into play. However, typically, the aim is to minimize the cross-entropy loss by summing the output of the predicted outcomes with the actual labels.

5. Identify the blocks in the RNN decoder.

1 / 1 point



- ☒ Block 1: Embedding layer; Block 2: RNN cell; Block 3: Linear layer
- ☐ Block 1: Linear layer; Block 2: RNN cell; Block 3: Embedding layer
- ☐ Block 1: Linear layer; Block 2: Embedding layer; Block 3: RNN cell
- ☐ Block 1: Embedding layer; Block 2: Linear layer; Block 3: RNN cell

✓ Correct

The decoder is composed of many decoder units within which the predicted token is generated. Each unit has an embedding layer, where embedded vectors are created. These vectors then pass through the RNN cell to output the updated hidden state, which will be passed to the RNN cell of the subsequent decoder unit. Next, a linear layer maps the RNN output to the output dimension to generate the next token.

6. Which of the following expressions is used to calculate the F1 score?

1 / 1 point

- ☒  $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$
- ☐  $\frac{\text{CountMatch}}{\text{CountReference}} = \frac{\text{Number of common n-grams}}{\text{Number of n-grams in } R}$
- ☐  $e^{-\frac{1}{t} \sum_i P(\omega_i) \log(Q(\omega_i))}$
- ☐  $\frac{\text{CountMatch}}{\text{CountGenerated}} = \frac{\text{Number of common n-grams}}{\text{Number of n-grams in } H}$

✓ Correct

The F1 score is the harmonic mean of precision and recall that is used to judge the performance of a model based on both.