Your grade: 100%

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

Next item

1. You are working as a software developer in an MNC and have been assigned a project on natural language processing (NLP) that involves implementing a self-attention mechanism. In your team, you have team members who are new to the self-attention mechanism. What is the primary purpose of the self-attention mechanism that you will explain when you kick off a meeting with your team members? 1 / 1 point

- ○ Remove irrelevant words from the input sentence
- ◉ Encode contextual information from surrounding words to represent them
- ○ Generate alternative text based on the input sequence
- ○ Perform part of speech tagging on the individual words

Correct
The self-attention mechanism helps NLP translate and summarize text, encode the contextual information from the surrounding words and present it. It also allows the model to capture dependencies and relationships among words in the input sequence.

2. What is the specific purpose of the dimension index parameter in positional encoding? 1 / 1 point

- ◉ Generates a unique sine or cosine wave for each embedding
- ○ Represents the position of the sine wave over time
- ○ Represents the number of words in the sequence
- ○ Identifies the position of each word embedding within the sequence

Correct
It controls the number of oscillations for each wave. Each one of these waves is added to a different dimension in the word embedding.

3. Imagine you are using the attention mechanism formula for translating French words to English. What do the query, key, and value vectors signify? Select the answer that correctly describes all three. 1 / 1 point

- ○ Query vectors: Represent word embeddings of all the words in the French language
  Key vectors: Represent the word embeddings of the words in the English language
  Value vector: Represents the translated English word
- ○ Query vectors: Represent word embeddings from the French language
  Key vector: Represents the word embedding of the French word to be translated
  Value vectors: Represent word embeddings from the English language
- ◉ Query vector: Represents the word embedding of the French word to be translated
  Key vectors: Represent word embeddings from the French language
  Value vectors: Represent word embeddings from the English language
- ○ Query vectors: Represent all the word embeddings from the English language
  Key vector: Represents the word embedding of the French word to be translated
  Value vector: Represent the word embedding of the translated English word

Correct
The query vector represents the word embedding of the French word you want to translate, key vectors represent the embeddings of the French words, and value vectors represent the embeddings of the English words in the target vocabulary.

4. Select the parameters you will provide when initializing an instance of the **nn.TransformerEncoderLayer class** in PyTorch. 1 / 1 point

- ○ Batch size and sequence length
- ◉ Embedding dimension and number of heads

- Sequence length and embedding dimension
- Number of layers and batch size

Correct
Embedding dimension and number of heads are important parameters that define the structure and behavior of the transformer encoder layer. The embedding dimension indicates the dimensionality of the input embeddings; however, the number of heads indicates the attention heads in the multi-head attention mechanism.

5. When using transformer-based models for text classification, the model is created after establishing the text pipeline. Identify the missing step (step number 2) from the following sequence of steps in creating the model. 1 / 1 point

Steps for creating the model:

1. Instantiate the embedding layer
2. ?
3. Apply the transformer encoder layers
4. Use the classifier layer to predict the label

- (•) Add positional encoding
- ( ) Record cumulative losses
- ( ) Construct a vocabulary
- ( ) Generate tokens

Correct
After instantiating the embedding layer, positional encoding embeds sequence order into word embeddings.