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 point

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

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

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

- ( ) 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 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
- ( ) 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 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

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

- ( ) Sequence length and embedding dimension
- ( ) Number of layers and batch size
- (●) Embedding dimension and number of heads
- ( ) Batch size and sequence length

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 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

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

Upgrade to submit