Your grade: 100%

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

Next item

1. Which tokenization method generates a smaller vocabulary but increases input dimensionality and computational needs? 1 / 1 point

- Word-based tokenization
- WordPiece tokenization
- (selected) Character-based tokenization
- SentencePiece tokenization

Correct
Character-based tokenization generates a smaller vocabulary as each character is treated as a separate token. However, since each character becomes a unique token, it increases the input dimensionality and computational needs due to the larger number of tokens.

2. Imagine you are training a sentiment analysis model where the input consists of user reviews. After tokenization, you find that the sequences have varying lengths. Which concept will you employ to address the issue of varied lengths while using data loaders? 1 / 1 point

- Shuffling
- Iteration
- Batching
- (selected) Padding

Correct
You would use padding to ensure that each sample in a data loader is of the same length.

3. Fill in the blank. 1 / 1 point

In subword-based tokenization, the _____ indicates that the word should be attached to the previous word without a space.

- (selected) ## symbol
- <eos> special token
- <pad> token
- Underscore symbol

Correct
In subword-based tokenization, the ## symbol indicates that the word should be attached to the previous word without a space.

4. Identify an advantage of word-based tokenization. 1 / 1 point

- It evaluates the benefits and drawbacks of splitting and merging two symbols
- It creates smaller vocabulary
- (selected) It preserves the semantic meaning
- It breaks down infrequent words to meaningful subwords

Correct
In word-based tokenization, the text is divided into individual words, each word considered a token. An advantage of word-based tokenization is that it preserves the semantic meaning.

5. Which input provided during data loader creation helps prevent the model from learning patterns based on the order of the data? 1 / 1 point

- (selected) The shuffle argument
- The data set

- ○ The padding value
- ○ The batch size

Correct
You can mention the shuffle argument as true. This shuffling is particularly useful for training deep learning models, as it prevents the model from learning patterns based on the order of the data.