Join the Hugging Face community

and get access to the augmented documentation experience

Sign Up   to get started

○ **Pytorch**   🌀 TensorFlow   🔵 Ask a question

End-of-chapter quiz

Test what you learned in this chapter!

1. The emotion dataset contains Twitter messages labeled with emotions. Search for it in the Hub , and read the dataset card. Which of these is not one of its basic emotions?

☐   Joy

☐   Love

☑   Confusion

> **Correct!** Correct! Confusion is not one of the six basic emotions.

☐   Surprise

Submit   **You got all the answers!**

2. Search for the ar_sarcasm dataset in the Hub . Which task does it support?

☑   Sentiment classification

> **Correct!** That's right! You can tell thanks to the tags.

☐   Machine translation

☐   Named entity recognition

☐ Question answering

Submit   **You got all the answers!**

3. How does the BERT model expect a pair of sentences to be processed?

☐ Tokens_of_sentence_1 [SEP] Tokens_of_sentence_2

☐ [CLS] Tokens_of_sentence_1 Tokens_of_sentence_2

☑ [CLS] Tokens_of_sentence_1 [SEP] Tokens_of_sentence_2 [SEP]

> **Correct!** That's correct!

☐ [CLS] Tokens_of_sentence_1 [SEP] Tokens_of_sentence_2

Submit   **You got all the answers!**

4. What are the benefits of the Dataset.map() method?

☐ The results of the function are cached, so it won't take any time if we re-execute the code.

☑ It can apply multiprocessing to go faster than applying the function on each element of the dataset.

> **Correct!** This is a neat feature of this method, but it's not the only one!

☐ It does not load the whole dataset into memory, saving the results as soon as one element is processed.

Submit   **You didn't select all the correct answers, there's more!**

5. What does dynamic padding mean?

☐ It's when you pad the inputs for each batch to the maximum length in the whole dataset.

☑ It's when you pad your inputs when the batch is created, to the maximum length of the sentences inside that batch.

> **Correct!** That's correct! The "dynamic" part comes from the fact that the size of each batch is determined at the time of creation, and all your batches might have different shapes as a result.

It's when you pad your inputs so that each sentence has the same number of tokens as the previous one in the dataset.

Submit   **You got all the answers!**

6. What is the purpose of a collate function?

☐ It ensures all the sequences in the dataset have the same length.

☑ It puts together all the samples in a batch.

> **Correct!** Correct! You can pass the collate function as an argument of a `DataLoader`. We used the `DataCollatorWithPadding` function, which pads all items in a batch so they have the same length.

☐ It preprocesses the whole dataset.

☐ It truncates the sequences in the dataset.

Submit   **You got all the answers!**

7. What happens when you instantiate one of the AutoModelForXxx classes with a pretrained language model (such as bert-base-uncased) that corresponds to a different task than the one for which it was trained?

☐ Nothing, but you get a warning.

☑ The head of the pretrained model is discarded and a new head suitable for the task is inserted instead.

> **Correct!** Correct. For example, when we used `AutoModelForSequenceClassification` with `bert-base-uncased`, we got warnings when instantiating the model. The pretrained head is not used for the sequence classification task, so it's discarded and a new head is instantiated with random weights.

☐ The head of the pretrained model is discarded.

☐ Nothing, since the model can still be fine-tuned for the different task.

Submit   **You got all the answers!**

8. What's the purpose of TrainingArguments ?

☑ It contains all the hyperparameters used for training and evaluation with the `.Trainer`

> **Correct!** Correct!

- [ ] It specifies the size of the model.
- [ ] It just contains the hyperparameters used for evaluation.
- [ ] It just contains the hyperparameters used for training.

Submit   **You got all the answers!**

9. Why should you use the 🤗 Accelerate library?

- [ ] It provides access to faster models.
- [ ] It provides a high-level API so I don't have to implement my own training loop.
- [x] It makes our training loops work on distributed strategies.

> **Correct!** Correct! With 🤗 Accelerate, your training loops will work for multiple GPUs and TPUs.

- [ ] It provides more optimization functions.

Submit   **You got all the answers!**

<> Update on GitHub

← Fine-tuning, Check!   ✓ Complete Chapter