

Please check your email address for a confirmation link

Resend confirmation email

• LLM Course 

Search documentation  3,519

0. SETUP

1. TRANSFORMER MODELS

Introduction
Natural Language Processing and Large Language Models

Transformers, what can they do?

How do Transformers work?

How 😊 Transformers solve tasks

Transformer Architectures

Quick quiz

Inference with LLMs

Bias and limitations

Summary

Certification exam

2. USING 🤗 TRANSFORMERS

3. FINE-TUNING A PRETRAINED MODEL

4. SHARING MODELS AND TOKENIZERS

5. THE 📁 DATASETS LIBRARY

6. THE 🎮 TOKENIZERS LIBRARY

7. CLASSICAL NLP TASKS

8. HOW TO ASK FOR HELP

9. BUILDING AND SHARING DEMOS

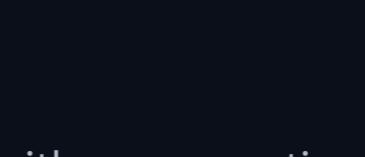
10. CURATE HIGH-QUALITY DATASETS

11. FINE-TUNE LARGE LANGUAGE MODELS

12. BUILD REASONING MODELS

COURSE EVENTS

Ungraded quiz



So far, this chapter has covered a lot of ground! Don't worry if you didn't grasp all the details, but it's to reflect on what you've learned so far with a quiz.

This quiz is ungraded, so you can try it as many times as you want. If you struggle with some questions, follow the tips and revisit the material. You'll be quizzed on this material again in the certification exam.

1. Explore the Hub and look for the roberta-large-mnli checkpoint. What task does it perform?

- Summarization
- Text classification

Correct! More precisely, it classifies if two sentences are logically linked across three labels (contradiction, neutral, entailment) — a task also called *natural language inference*.

- Text generation

Submit You got all the answers!

2. What will the following code return?

```
from transformers import pipeline  
  
ner = pipeline("ner", grouped_entities=True)  
ner("My name is Sylvain and I work at Hugging Face in Brooklyn.")
```

- It will return classification scores for this sentence, with labels "positive" or "negative".
- It will return a generated text completing this sentence.
- It will return the words representing persons, organizations or locations.

Correct! Furthermore, with `grouped_entities=True`, it will group together the words belonging to the same entity, like "Hugging Face".

Submit You got all the answers!

3. What should replace ... in this code sample?

```
from transformers import pipeline  
  
filler = pipeline("fill-mask", model="bert-base-cased")  
result = filler("...")
```

- This <mask> has been waiting for you.
- This [MASK] has been waiting for you.

Correct! This model's mask token is [MASK].

- This man has been waiting for you.

Submit You got all the answers!

4. Why will this code fail?

```
from transformers import pipeline  
  
classifier = pipeline("zero-shot-classification")  
result = classifier("This is a course about the Transformers library")
```

- This pipeline requires that labels be given to classify this text.
- This pipeline requires several sentences, not just one.
- The 😊 Transformers library is broken, as usual.
- This pipeline requires longer inputs; this one is too short.

Submit You got all the answers!

5. What does "transfer learning" mean?

- Transferring the knowledge of a pretrained model to a new model by training it on the same dataset.
- Transferring the knowledge of a pretrained model to a new model by initializing the second model with the first model's weights.

Correct! When the second model is trained on a new task, it *transfers* the knowledge of the first model.

- Transferring the knowledge of a pretrained model to a new model by building the second model with the same architecture as the first model.

Submit You got all the answers!

6. True or false? A language model usually does not need labels for its pretraining.

- True

Correct! The pretraining is usually *self-supervised*, which means the labels are created automatically from the inputs (like predicting the next word or filling in some masked words).

- False

Submit You got all the answers!

7. Select the sentence that best describes the terms "model", "architecture", and "weights".

- If a model is a building, its architecture is the blueprint and the weights are the people living inside.
- An architecture is a map to build a model and its weights are the cities represented on the map.
- An architecture is a succession of mathematical functions to build a model and its weights are those functions parameters.

Correct! The same set of mathematical functions (architecture) can be used to build different models by using different parameters (weights).

Submit You got all the answers!

8. Which of these types of models would you use for completing prompts with generated text?

- An encoder model
- A decoder model

Correct! Decoder models are perfectly suited for text generation from a prompt.

- A sequence-to-sequence model

Submit You got all the answers!

9. Which of those types of models would you use for summarizing texts?

- An encoder model
- A decoder model
- A sequence-to-sequence model

Correct! Sequence-to-sequence models are perfectly suited for a summarization task.

Submit You got all the answers!

10. Which of these types of models would you use for classifying text inputs according to certain labels?

- An encoder model

Correct! An encoder model generates a representation of the whole sentence which is perfectly suited for a task like classification.

- A decoder model

- A sequence-to-sequence model

Submit You got all the answers!

11. What possible source can the bias observed in a model have?

- The model is a fine-tuned version of a pretrained model and it picked up its bias from it.

Correct! When applying Transfer Learning, the bias in the pretrained model used persists in the fine-tuned model.

- The data the model was trained on is biased.

Correct! This is the most obvious source of bias, but not the only one.

- The metric the model was optimizing for is biased.

Correct! A less obvious source of bias is the way the model is trained. Your model will blindly optimize for whatever metric you chose, without any second thoughts.

Submit

You got all the answers!

↳ Update on GitHub

← Transformer Architectures

Inference with LLMs →

Ungraded quiz

1. Explore the Hub and look for the roberta-large-mnli checkpoint. What task does it perform?

2. What will the following code return?

3. What should replace ... in this code sample?

4. Why will this code fail?

5. What does "transfer learning" mean?

6. True or false? A language model usually does not need labels for its pretraining.

7. Select the sentence that best describes the terms "model", "architecture", and "weights".

8. Which of these types of models would you use for completing prompts with generated text?

9. Which of those types of models would you use for summarizing texts?

10. Which of these types of models would you use for classifying text inputs according to certain labels?

11. What possible source can the bias observed in a model have?