

MKhalusova [Docs] Model_doc structure/clarity improvements (#26876)  

5964f82 · 6 months ago

 History 

57 lines (34 loc) · 2.38 KB



main

[transformers](#) / [docs](#) / [source](#) / [en](#) / [model_doc](#) / [flan-t5.md](#) Top

Preview

Code

Blame

Raw



FLAN-T5

Overview

FLAN-T5 was released in the paper [Scaling Instruction-Finetuned Language Models](#) - it is an enhanced version of T5 that has been finetuned in a mixture of tasks.

One can directly use FLAN-T5 weights without finetuning the model:

```
>>> from transformers import AutoModelForSeq2SeqLM, AutoTokenizer

>>> model = AutoModelForSeq2SeqLM.from_pretrained("google/flan-t5-small")
>>> tokenizer = AutoTokenizer.from_pretrained("google/flan-t5-small")

>>> inputs = tokenizer("A step by step recipe to make bolognese pasta:", return_tensors="pt")
>>> outputs = model.generate(**inputs)
```



```
>>> print(tokenizer.batch_decode(outputs, skip_special_tokens=True))  
['Pour a cup of bolognese into a large bowl and add the pasta']
```

FLAN-T5 includes the same improvements as T5 version 1.1 (see [here](#) for the full details of the model's improvements.)

Google has released the following variants:

- [google/flan-t5-small](#)
- [google/flan-t5-base](#)
- [google/flan-t5-large](#)
- [google/flan-t5-xl](#)
- [google/flan-t5-xxl](#).

The original checkpoints can be found [here](#).

Refer to [T5's documentation page](#) for all API reference, code examples and notebooks. For more details regarding training and evaluation of the FLAN-T5, refer to the model card.